Genetic drift promotes and recombination hinders speciation on a holey fitness landscape

Ata Kalirad[1,2], Christina L. Burch[3] & Ricardo B. R. Azevedo[1]

[1]*Department of Biology and Biochemistry, University of Houston, Texas 77204-5001, USA*

[2]*Department for Integrative Evolutionary Biology, Max Planck Institute for Biology Tübingen, Max Planck Ring 9, 72076 Tübingen, Germany*

[3]*Department of Biology, University of North Carolina at Chapel Hill, North Carolina 27599-3280, USA*

Dobzhansky and Muller proposed a general mechanism through which microevolution, the substitution of alleles within populations, can cause the evolution of reproductive isolation between populations and, therefore, macroevolution. As allopatric populations diverge genetically, many combinations of alleles differing between them have not been tested by natural selection and may thus be incompatible. Such genetic incompatibilities often cause low fitness in hybrids between species. Furthermore, the number of incompatibilities grows with the genetic distance between diverging populations. However, what determines the rate and pattern of accumulation of incompatibilities remains unclear. Here we investigate this question by simulating evolution on a holey fitness landscape. We find that genetic incompatibilities accumulate at a rate that depends critically on genetic robustness: incompatibilities accumulate more slowly among genetically robust populations. We go on to identify two determinants of the rate of accumulation of genetic incompatibilities: recombination and

population size. In large populations with abundant genetic variation, recombination selects for increased genetic robustness and, as a consequence, incompatibilities accumulate more slowly. In small populations, genetic drift interferes with this process and promotes the accumulation of genetic incompatibilities. Our results suggest a novel mechanism by which genetic drift promotes and recombination hinders speciation.

The rates of origination of new species vary extensively across the tree of life. For example, some lineages of cichlids are estimated to produce new species approximately 70× faster than some lineages of hawks[1]. A major challenge for evolutionary biology is to explain this macroevolutionary pattern through the action of microevolutionary processes operating within populations[2,3]. Stanley[4] speculated that speciation is "a largely random process" and that, therefore, "macroevolution is decoupled from microevolution". Here we explore the alternative: that macroevolution and microevolution are, in fact, mechanistically coupled.

Speciation results from the build up of reproductive isolation (RI) between populations. Thus, variation in speciation rates should be caused by variation in the rates at which RI increases between populations as they diverge. As allopatric populations diverge, genetic incompatibilities accumulate between them and generate postzygotic RI[5,6], one of the causes of speciation. Orr[7] developed a combinatorial model to describe this process. He proposed that any pair of alleles at two divergent loci, not previously tested by natural selection, has a probability p of being incompatible. If two populations evolving in allopatry have diverged at k loci, then $k(k-1)/2$ combinations of alleles at pairs of these loci will not have been tested by natural selection (Fig. 1a). Combining

these two terms, the number of Dobzhansky–Muller incompatibilities (DMIs) between pairs of loci is expected to increase with k according to:

$$I(k) = \frac{pk(k-1)}{2} \approx \frac{pk^2}{2} \tag{1}$$

Orr[7] termed this pattern of faster than linear accumulation "snowballing". Empirical tests of this prediction confirmed that genetic incompatibilities accumulated faster than linearly for inviability in species of *Drosophila*[8,9], male sterility in subspecies of *Mus musculus*[10], and both seed and pollen sterility in species of *Solanum*[11,12]. Computational studies using models of RNA folding[13], transcription factor-DNA binding[14], and spatial gene regulation during development[15] also found that genetic incompatibilities snowballed. These empirical and computational studies were all conducted over relatively short evolutionary distances. Surprisingly, a study conducted over much longer genetic distances discovered that genetic incompatibilities did not snowball for animal proteins[16]; rather, they accumulated approximately linearly with genetic distance. F. Kondrashov et al.[16] suggested that a linear accumulation of genetic incompatibilities is not well explained by Orr's combinatorial model. We note, however, that the snowballing prediction in Equation 1 relies on particular assumptions[7,13,17], in particular, that the probability of incompatibility p does not, itself, evolve as populations diverge[13]. Relaxing this assumption could explain the linear accumulation pattern. For example, if p declines with k according to the relationship $p(k) = C/k$ (where C is a constant) then DMIs will accumulate linearly: $I(k) = C(k-1)$. The extent to which evolved differences in p account for the linear accumulation of genetic incompatibilities among animal proteins is unknown. Nonetheless, the possibility raises the question of what might cause p, a determinant of the rate and pattern of accumulation of genetic incompatibilities, to evolve.

Here we use a computational model of RNA secondary structure to investigate whether and why populations differ in the rate at which they accumulate genetic incompatibilities. We chose the folding of RNA sequences as a simple model of a genotype-phenotype map grounded in biophysical reality[18]. Crucially, this genotype-phenotype map naturally incorporates epistasis and has been used to study its evolutionary consequences (e.g., robustness[19,20], evolvability[21,22], molecular evolution[23]). As in Kalirad and Azevedo[13], we conduct evolution experiments that monitor the evolution of RNA secondary structures on a holey fitness landscape[24] (also known as a neutral network[25,26]), in which genotypes are either viable or inviable (Fig. 1c). Inviable genotypes carry combinations of alleles displaying DMIs (Fig. 1b). Thus, holey fitness landscapes provide the simplest modeling framework to study the evolution of reproductive isolation caused by genetic incompatibilities.

We follow the approach taken in the classic empirical tests of the combinatorial model[8,11] and monitor the accumulation of incompatible introgressions[13] (IIs) rather than DMIs. Because introgression is a perturbation of the genotype, we predict that the probability that an introgression is incompatible will be inversely related to the recipient's genetic robustness, its ability to maintain viability in the face of perturbations to the genotype. We test this prediction by experimentally manipulating population genetic parameters known to promote the evolution of genetic robustness.

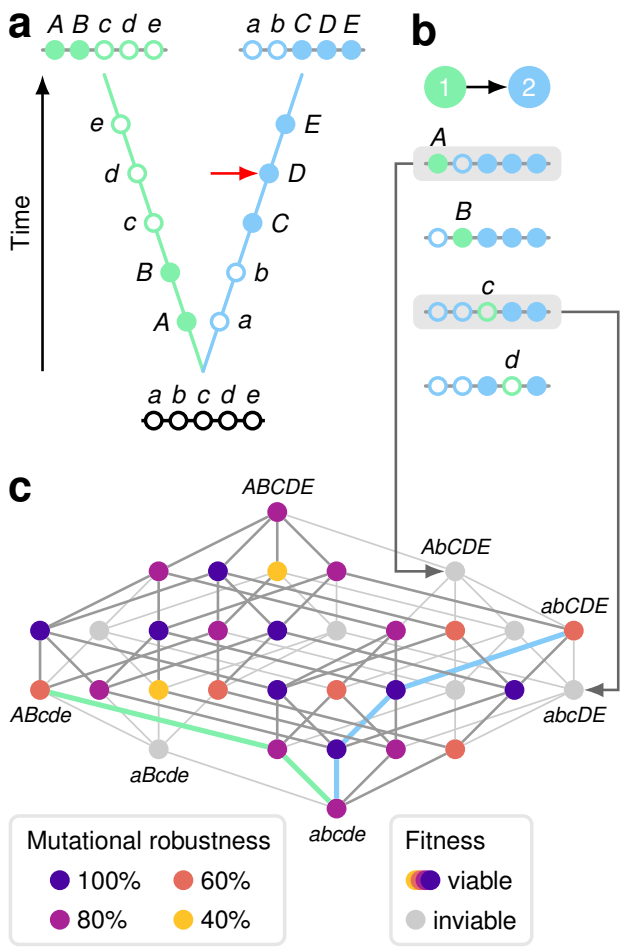


Fig. 1: Populations evolving on a holey fitness landscape accumulate genetic incompatibilities.

Fig. 1 (full caption): Populations evolving on a holey fitness landscape accumulate genetic incompatibilities. (a) Two populations diverge in allopatry. Both populations are initially fixed for lowercase alleles (open circles) at five loci, A–E. Derived alleles are indicated by uppercase letters (closed circles). The first two substitutions (closed circles) were of alleles *A* and *B* in population 1 (left, green); the next three substitutions were of alleles *C*, *D*, and *E* in population 2 (right, blue). **(b)** Genetic incompatibilities can be inferred by introgressing individual divergent alleles from one population into the other and counting the number of incompatible introgressions (IIs). This is illustrated by the introgression of the divergent alleles from population 1 (*ABcde*) into population 2 (*abCDE*) to create the four combinations that have not been tested by natural selection. We ignore the introgression of the *e* allele because it results in the transitional *abCDe* genotype in population 2, which is known to be viable (red arrow in (a)). Two alleles from population 1 (*A* and *c*) are deleterious in the genetic background of population 2 (grey boxes). We refer to these as IIs. Each II implies the existence of one or more DMIs. **(c)** Holey fitness landscape of the 32 genotypes at loci A–E. Nodes (circles) represent genotypes. Lines connect genotypes that can be reached by a mutation at a single locus. Each genotype has 5 mutational neighbors. 7 of the 32 genotypes (22%) are inviable, shown in light grey. The genotypes generated by the two IIs shown in (b) are marked by arrows. The remaining 25 genotypes (78%) are viable, shown in colors from yellow to blue, and define a neutral network. The different colors indicate the proportion ν of the mutational neighbors of a genotype that are viable, a measure of mutational robustness. The evolutionary trajectories of the populations shown in (a) are indicated by thick blue and green lines.

Experimental Design

Our experimental design capitalises on prior findings regarding how populations evolve on a holey fitness landscape (Figs. 1a and 1c). On a holey fitness landscape, all viable genotypes have the same viability but, in the presence of mutation and recombination, they do not all have the same fitness. Mutations are either lethal or neutral with respect to viability, so viable individuals can differ in their expected number of viable offspring depending on the proportion ν of their mutational neighbors that are viable, a measure of mutational robustness (Fig. 1c). Similarly, recombination between different pairs of genotypes may produce varying proportions of viable recombinants. Thus, populations evolving on a holey fitness landscape experience selection for higher genetic robustness[19,27–29]. The ability of populations to respond to this selection depends on factors that determine the relative strengths of selection on genetic robustness and genetic drift: mutation rate, recombination rate, and population size (Fig. 2; U, r, $N \rightarrow \nu$).

To test our prediction that the rate at which IIs accumulate between populations will be inversely related to their genetic robustness, we first experimentally evolved populations of RNA secondary structures using combinations of population sizes and recombination rates expected to result in different equilibrium levels of genetic robustness. Populations evolved under an individual-based Wright-Fisher model with a selection-recombination-mutation life cycle, with a constant population size (between $N = 10$ and $N = 500$) and recombination rate (between no recombination and free recombination). Our use of the Wright-Fisher model differs from most theoretical research into the combinatorial model[7,13,30,31] in that it does not impose a strong selection, weak

mutation (SSWM) regime[32]. Unlike SSWM, in which mutations arise and transit to fixation or loss sequentially, our implementation of the Wright-Fisher model uses a genome-wide mutation rate of $U = 0.1$, ensuring that multiple mutations segregate simultaneously ($NU \geqslant 1$ in all of our experimental populations). This difference in approach has two important consequences for this work. It enables 1) recombination and 2) selection for genetic robustness, neither of which occur in the SSWM regime[19].

Once these populations reached a mutation-recombination-selection-drift equilibrium, we made exact copies of each population to obtain pairs of identical ancestral populations. These population pairs were then allowed to diverge allopatrically, again evolving under a Wright-Fisher model with a selection-recombination-mutation life cycle and using the same population size and recombination rate used to generate the ancestral populations. We then investigated the effect that the differences in genetic robustness among the ancestral populations had on the rate at which genetic incompatibilities accumulated between the diverging population pairs.

We followed the standard approach for inferring genetic incompatibilities: introgression[8,11] (illustrated in Fig. 1b). Specifically, we introgressed each of the k divergent alleles from population i to population j, one at a time, and counted the number of the resulting introgressions that are incompatible, $\Pi_{ij}(k)$. For example, in the landscape depicted in Fig. 1c, $\Pi_{12}(5) = 2$. The number of IIs, like the number of DMIs, is also expected to snowball[13]. Because the introgressed alleles are single base pair substitutions, the probability that an introgression is incompatible with the genetic background of a recipient genotype should be inversely related to the mutational robustness, ν,

of the recipient genotype, that is, its ability to maintain viability in the face of mutation. Thus, ν, like p in the combinatorial model, determines the rate and pattern of accumulation of genetic incompatibilities.

We note that by counting IIs, we determine neither the order nor the number of the underlying DMIs[13,33] and, therefore, we cannot estimate p. In fact, p is impossible to measure in practice. The issue is that Equation 1 ignores all DMIs that involve more than a pair of loci, but such higher-order DMIs are known to exist[7,13,34]. In principle, the combinatorial model can be extended to higher-order DMIs[7,17] by introducing additional p parameters. However, there is no established method to infer the precise order of a DMI. Furthermore, the order of a DMI can change as populations diverge[13], something the combinatorial model cannot capture. Our focus on IIs and their relationship to ν eliminates these problems, enabling our investigation of the determinants of the rate at which genetic incompatibilities accumulate.

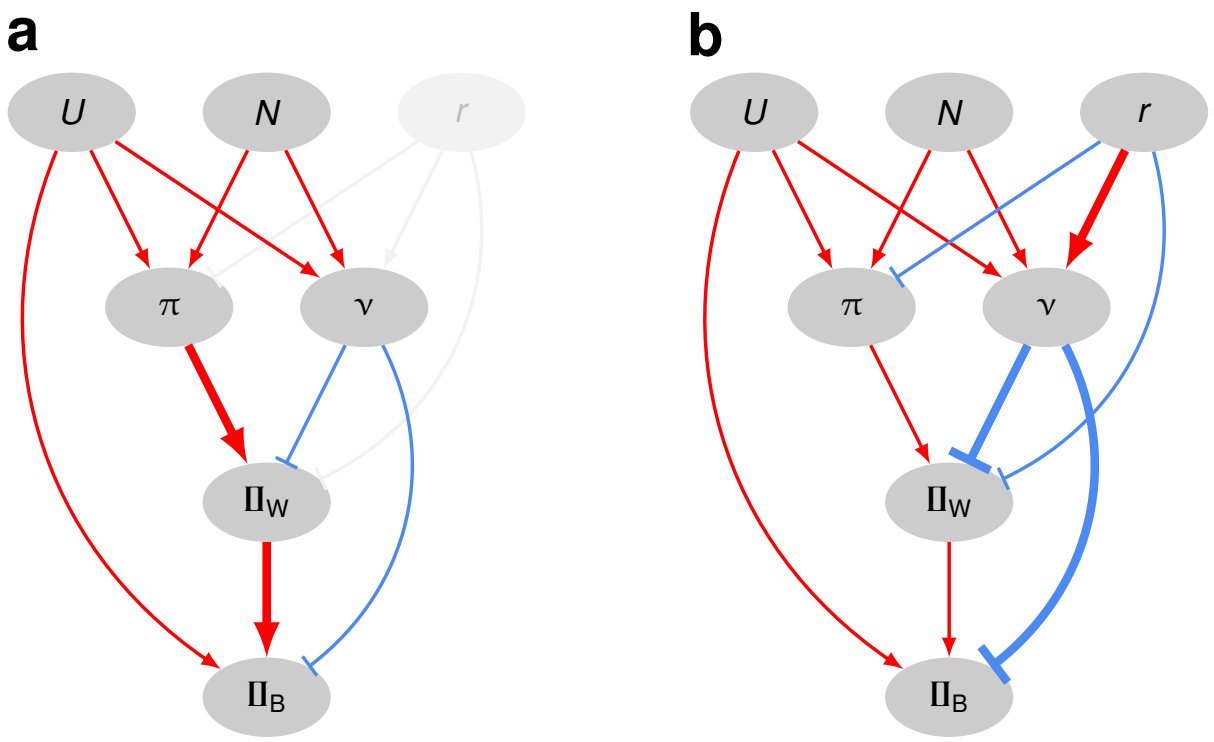


Fig. 2: Evolutionary causes of the accumulation of genetic incompatibilities between populations in our model. Effects are summarized for asexual **(a)** and sexual **(b)** populations. Ellipses contain the following properties of populations: U, genomic mutation rate; N, population size; r, recombination probability or rate; π, nucleotide diversity; ν, mutational robustness; Π_{W} and Π_{B}, genetic incompatibilities within and between populations, respectively. Light gray ellipses in **(a)** indicate that in asexual populations recombination does not operate $(r = 0)$. Red lines ending in arrows indicate positive effects; blue lines ending in bars indicate negative effects. The thickness of the line indicates the strength of the effect. Light gray lines in *a* indicate that the effect does not operate.

Results

Genetic drift promotes the accumulation of genetic incompatibilities in asexual populations. We allowed asexual populations of different sizes—between 10 and 500 individuals—to evolve until they reached an equilibrium genetic robustness. We then made a copy of the resulting population and allowed the two populations to diverge in allopatry. Both populations experience the same selection regime (see Methods). Thus, this is a mutation-order speciation[35] scenario. Fig. 3a shows the pattern of accumulation of IIs following the population split. The strength of selection for genetic robustness in these populations is of the order of the deleterious mutation rate[36] $U_d = U(1-\nu)$ (Fig. 2a: $U \to \nu$). For comparison, we subjected "populations" of $N = 1$ individual to a random walk, which is equivalent to the SSWM regime and does not impose selection for genetic robustness[19] (Fig. 3a, gray circles).

We modelled the accumulation of IIs using the following phenomenological equation

$$\mathrm{II} = a + bk^c \tag{2}$$

where a is the initial number of IIs within populations ($a = 0$ when $N = 1$), b is the rate of accumulation of IIs with genetic divergence, and c is a snowball exponent. Values of $c = 1$ and 2 correspond to linear and quadratic patterns of II accumulation, respectively. The SSWM simulations showed an overall snowball exponent of $c = 1.619$ (maximum likelihood, ML, estimate; 95% confidence interval, CI: 1.510, 1.736). Thus, IIs accumulated according to a snowball pattern $(c > 1)$.

Increasing population size caused the accumulation of IIs to slow down in asexual populations. Fitting Equation 2 to the data at each population size while keeping the snowball exponent constant at $c = 1.479$ (the ML estimate of c taking into account all simulated populations summarized in Figs. 3 and 4, from $N = 1$ to $N = 500$) resulted in a reduction of the slope b with population size (Fig. 3c, blue circles). For example, populations of $N = 10$ and 500 individuals accumulated IIs at 90.6% and 81.1% of the rate of populations of $N = 1$ individual, respectively. Increasing population size reduced the importance of genetic drift and was expected to allow selection for genetic robustness to operate more efficiently (Fig. 2a: $N \rightarrow \nu$). Indeed, larger populations evolved higher genetic robustness, ν, before populations began diverging from each other (Fig. 3d, blue circles). The level of genetic robustness in diverging populations explains $R^2 = 75.0\%$ of the variation in the rate of II accumulation in populations of different sizes.

Genetic drift promotes the accumulation of genetic incompatibilities in sexual populations. Next, we repeated the evolutionary simulations but allowed genotypes to mate at random and recombine freely. Fig. 3b shows the accumulation of IIs in populations of the same size as in Fig. 3a but with free-recombination. In populations of a single individual (i.e., the SSWM regime), recombination cannot operate. For populations with $N > 1$, the effect of population size on the accumulation of IIs was stronger than among the asexual populations (Fig. 3c, orange circles). Larger populations accumulated IIs much more slowly than smaller populations. Again, this pattern is well explained by variation in the level of genetic robustness in free-recombining populations of different sizes ($R^2 = 97.9\%$; Fig. 3d, orange circles).

What is the mechanism underlying the differences in genetic robustness among populations of different sizes? One possibility is that more robust genotypes are further from the holes in the fitness landscape. In all simulations reported here, a sequence is inviable if it is at a base-pair distance $\delta > 12$ from a reference structure (see Methods). As expected, genotypes from a population of $N = 10$ individuals were at an average distance of $\bar{\delta} = 11.32 \pm 0.10$ (95% CI); living on the edge of a hole made them more fragile. More abundant populations were further from the edge and, therefore, more robust: $\bar{\delta} = 10.18 \pm 0.16$ for $N = 50$ and $\bar{\delta} = 6.84 \pm 0.23$ for $N = 500$ individuals.

Recombination hinders the accumulation of genetic incompatibilities. Changing the recombination probability in populations of the same size $(N = 100)$ allows us to isolate the causal role of recombination in determining the rate of II accumulation (Fig. 4a). Populations experiencing more recombination both evolved higher genetic robustness and accumulated IIs more slowly (Fig. 2b: $r \to \nu$; Figs. 4b and 4c). For example, populations where $r = 10\%$ and 100% of matings experienced one crossover event accumulated IIs at 83.1% and 50.4% of the rate of asexual populations, respectively (again, assuming $c = 1.479$). Increasing the *amount* of recombination from one crossover to free-recombination caused a further reduction in the rate of II accumulation to 27.5% of the rate of asexual populations. The level of genetic robustness in diverging populations explains $R^2 = 98.8\%$ of the variation in the rate of II accumulation in populations experiencing different probabilities and amounts of recombination (Fig. 4b).

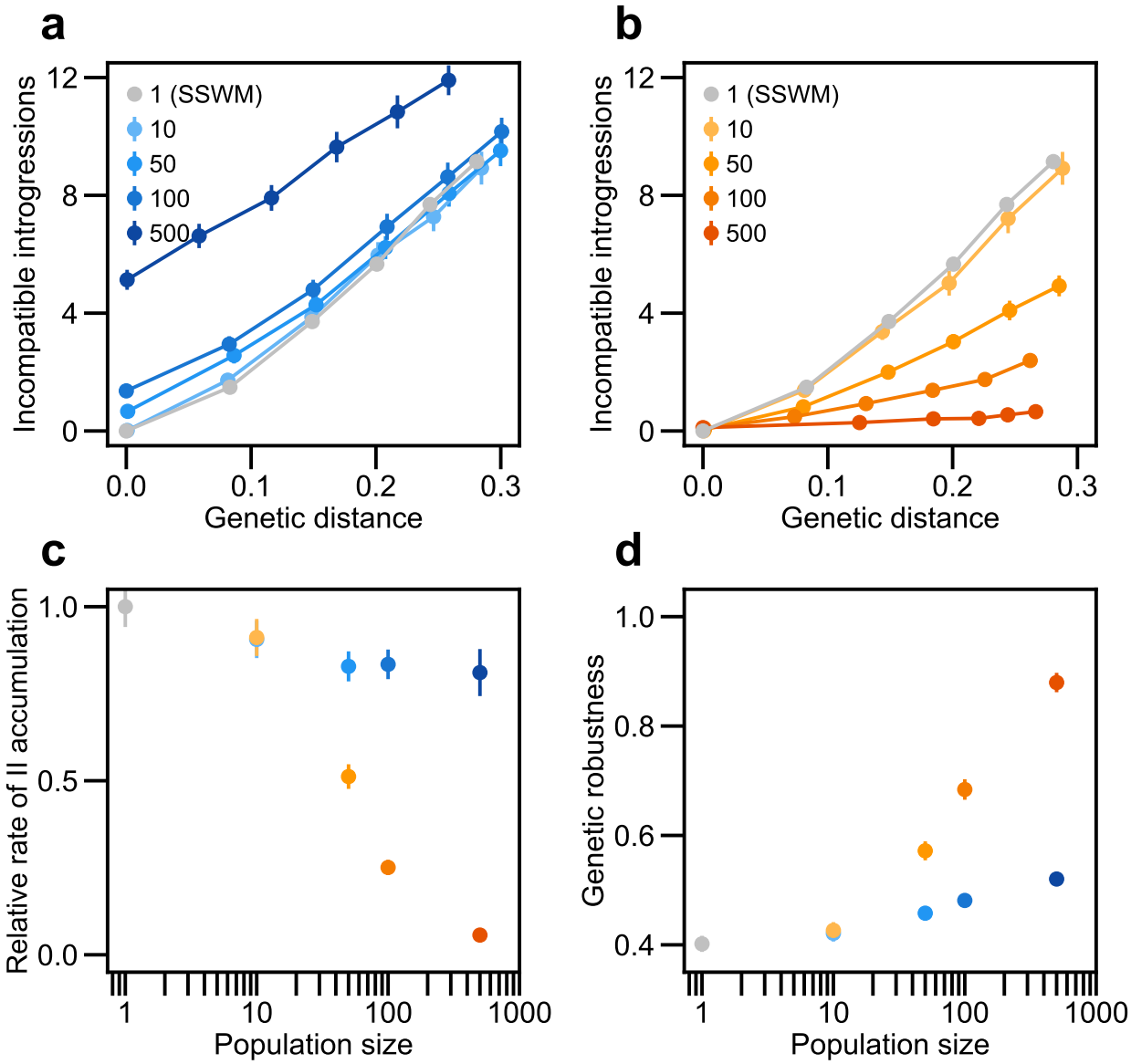


Fig. 3: Genetic drift promotes and recombination hinders the accumulation of genetic incompatibilities.

Fig. 3 (full caption): Genetic drift promotes and recombination hinders the accumulation of genetic incompatibilities. Incompatible introgressions (IIs) accumulate with genetic distance in asexual **(a)** and sexual **(b)** populations of different sizes ($N = 1, 10, \ldots 500$). Genetic distance (k) between divergent populations was measured by Jost's D[37]. In all simulations, genome length was $L = 100$ nucleotides and the mutation rate was $U = 0.1$ per genome per generation. Sexual populations experienced random mating and free recombination. **(c)** Relative rates of II accumulation (b_N/b_1) at different population sizes. For each N, the rate b_N was estimated by fitting Equation 2 to each simulation run, as described in the Methods. The average value of b_N was then divided by the average value under the SSWM regime (b_1). **(d)** Genetic robustness at different population sizes. The genetic robustness of a genotype was measured as the proportion ν of its mutational neighbors that are viable. The genetic robustness of each replicate population was measured every 100 ($N = 1, 10, 50, 100$) or 400 ($N = 500$) generations and averaged across every individual and timepoint. Note that these populations are at equilibrium before divergence begins, so that robustness does not evolve over the period shown. Plotted values are means of 200 replicate simulation runs at each population size. Error bars are 95% CIs.

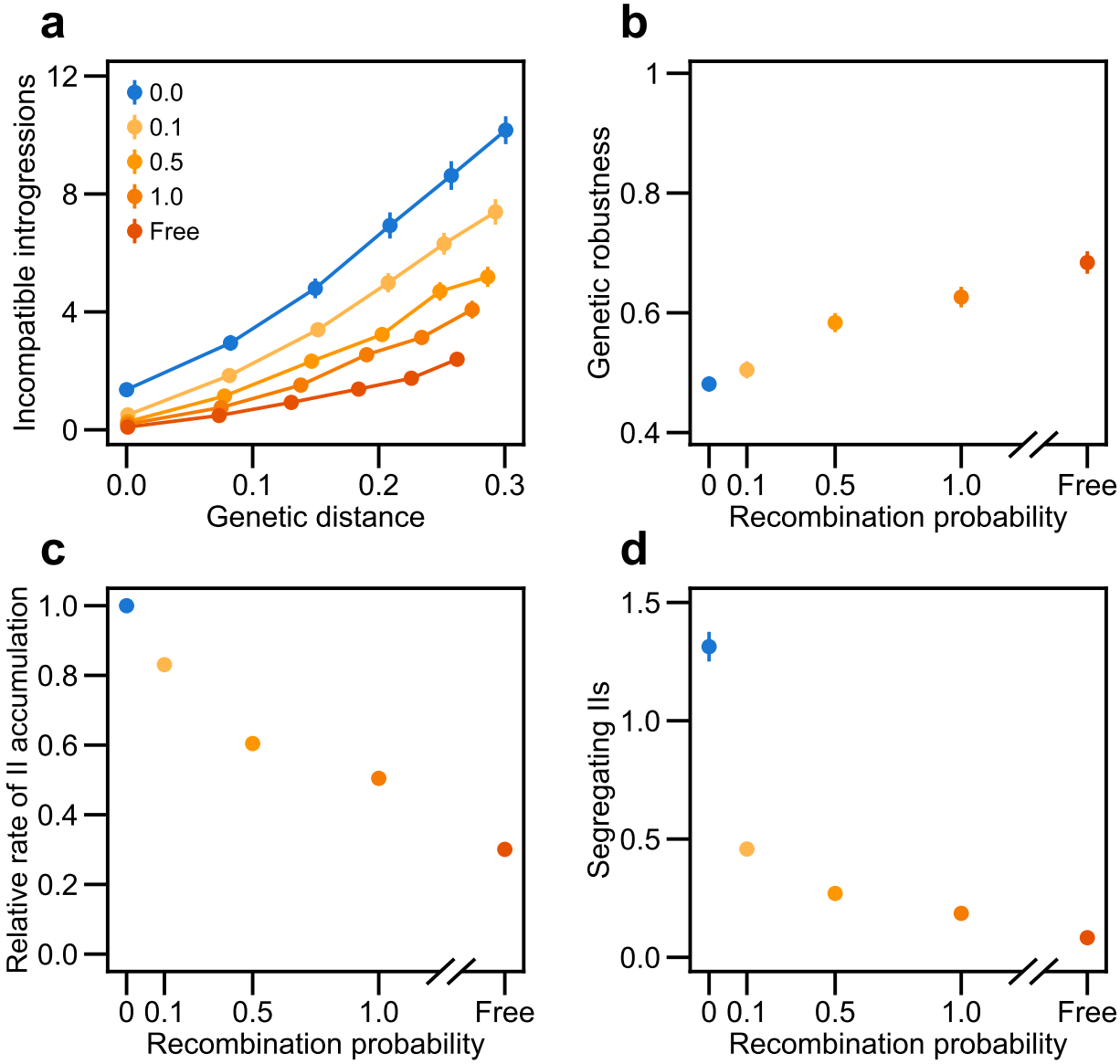


Fig. 4: Recombination hinders the accumulation of genetic incompatibilities.

Fig. 4 (full caption): Recombination hinders the accumulation of genetic incompatibilities. **(a)** IIs accumulate with genetic distance in populations of $N = 100$ individuals experiencing different recombination probabilities. Sexual populations $(r > 0)$ experienced random mating. The values of the recombination probability $(r = 0, 0.1, 0.5, 1)$ indicate the expected proportion of matings resulting in a recombinant progeny containing one crossover event. Under free recombination, every progeny is a recombinant with an expected 49.5 crossover events. **(b)** Genetic robustness at different recombination probabilities. **(c)** Relative rates of II accumulation (b_r/b_0) at different recombination probabilities. For each r, the rate b_r was estimated by fitting Equation 2 to each simulation run. The average value of b_r was then divided by the average value under the asexual regime (b_0). **(d)** Number of IIs segregating within populations at different recombination probabilities (see Methods). Values are means of 200 replicate simulation runs at each recombination probability calculated every 100 generations. Error bars are 95% CIs.

Recombination hinders the accumulation of genetic incompatibilities within populations. The previous results indicate that recombination hinders the accumulation of genetic incompatibilities. But how does recombination do that? To begin to answer that question we return to our expectation for the accumulation of IIs in the absence of recombination (Fig. 2a). In sufficiently large populations, genetic variation builds up. Because of epistastic interactions, some of these variants constitute within-population IIs (Fig. 2a: $\pi \rightarrow \Pi_{\mathrm{W}}$). In the absence of recombination, however, these segregating genetic incompatibilities will only rarely be expressed—and, therefore, be exposed to natural selection—because they can only occur in the same individual through multiple mutations, which are rare. Therefore, cryptic IIs are expected to accumulate *within* asexual populations. We observed this in our simulations: asexual populations of $N = 100$ and 500 individuals had accumulated 1.36 ± 0.17 and 5.13 ± 0.34 segregating IIs on average *before* divergence began (95% CIs; Fig. 3a, genetic distance of 0; Fig. 4d, $r = 0$).

We hypothesize that the extent to which a mutation is involved in an II *within* a population is an important determinant of whether that mutation will cause an II *among* populations if it spreads in the population (Fig. 2a: $\Pi_{\mathrm{W}} \rightarrow \Pi_{\mathrm{B}}$). To test this hypothesis we isolated from each asexual population the first mutation to rise to high frequency and cause an II among populations and the first mutation to rise to high frequency and not cause an II. We then went back to the generations in which these mutations arose, placed them in every individual present in the population at the time, and noted the proportion of these "backcrossed" individuals that remained viable (Fig. 5a). As predicted, we found that the proportion of viable backcrossed individuals was much lower for mutations that caused IIs (0.71 ± 0.04, mean and 95% CI; Fig. 5b, blue filled circle) than for

mutations that did not (0.96 ± 0.01; Fig. 5b, blue open circle). In other words, mutations that caused IIs among populations were likely to have started out causing IIs within populations.

We propose that recombination slows down the accumulation of IIs *between* populations because it suppresses the accumulation of IIs *within* populations. In the presence of recombination, segregating IIs can be expressed in recombinant individuals. This suppresses the accumulation of IIs within populations in two ways. First, it causes selection against mutations involved in IIs within populations (Fig. 2b: $r \dashv \Pi_{\mathrm{W}}$). Second, it generates recombination load, which in turn selects for genetic robustness[27,38–41]. Genetically robust individuals are less likely to express IIs (Fig. 2b: $r \rightarrow \nu \dashv \Pi_{\mathrm{W}}$). We observe this suppression of segregating IIs by recombination in our simulations (Fig. 4d). We believe the reduced supply of segregating IIs slows down the accumulation of IIs among populations.

We tested this hypothesis by conducting backcross simulations at different recombination rates (Fig. 5b). As recombination rate and robustness increased, we observed a higher proportion of viable backcrossed individuals for mutations that caused IIs among populations (Fig. 5b, orange filled circles). Thus, in the presence of recombination, even mutations that are ultimately destined to participate in genetic incompatibilities among populations are less likely to result in within-population genetic incompatibilities (Fig. 2b). That is because recombination selects for genotypes that are more compatible with the genetic backgrounds of other individuals in the population and, therefore, with mutations currently segregating in the population.

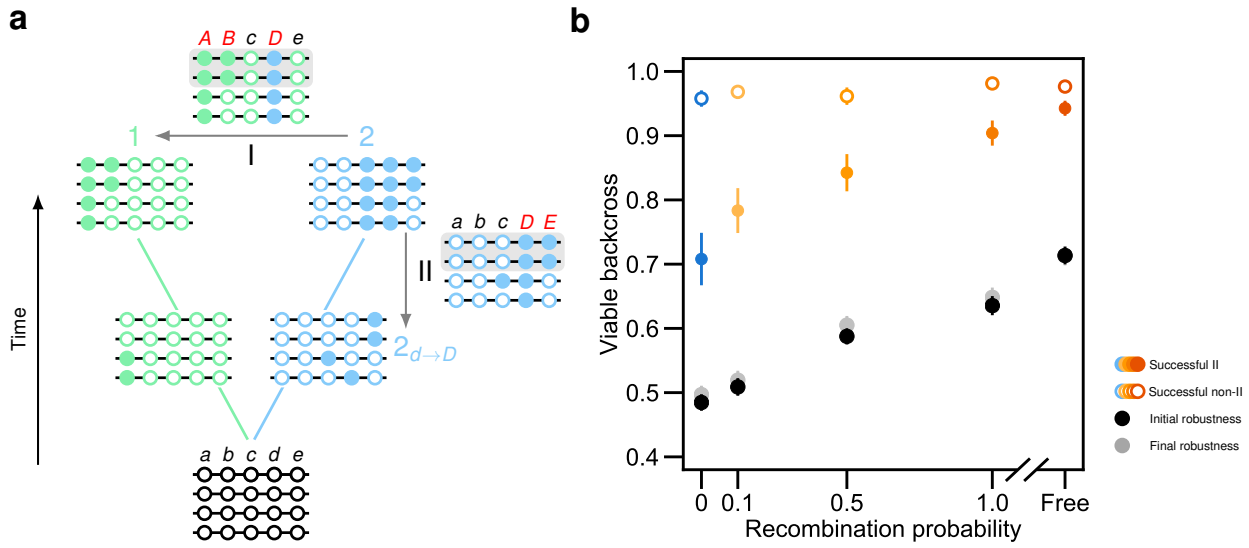


Fig. 5: Recombination selects for alleles that are more compatible with the genetic backgrounds of other individuals in the population.

Fig. 5 (full caption): Recombination selects for alleles that are more compatible with the genetic backgrounds of other individuals in the population. (a) Illustration of the experimental design used to generate data in **(b)**. Two populations of $N = 4$ individuals diverge in allopatry over the holey fitness landscape shown in Fig. 1c. After some time, population 2 (blue) fixes derived allele *D*. Introgressing *D* from population 2 to population 1 (arrow I) results in an inviable genotype (*ABcDe*), rendering 50% of introgressed sequences in population 2 inviable (gray box). Backcrossing *D* into every genotype present in population 2 (arrow II) when the *D* allele arose by mutation reveals an inviable genotype (*abcDE*), and renders 50% of the sequences in that population inviable too (gray box). Thus, *D* is involved in IIs both within and among populations. **(b)** We repeated the simulations summarized in Fig. 4 and isolated the first mutation to occur in a population, rise to a frequency of at least 50%, and be involved in an II with the other population. We then went back to the generation in which the mutation arose and backcrossed it into every individual present within its own population at that time. Panel **(b)** shows the proportion of viable backcrossed individuals made from populations experiencing different recombination probabilities (colored filled circles). As a control, we did the same calculation for the first mutation to rise to high frequency that does not cause an II among populations (open circles). Black and gray circles show the mean genetic robustness of the populations when the mutations arise and when they reach high frequency, respectively. Values are means of 200 replicate simulation runs at each population recombination probability. Error bars are 95% CIs.

The number of genetic incompatibilities between populations is correlated with the amount of postzygotic reproductive isolation. In the presence of recombination, genetic incompatibilities between populations cause postzygotic reproductive isolation (RI) between them. Populations evolved at different population sizes under free-recombination (Extended Data Fig. 1a) or at a population size of $N = 100$ under different recombination probabilities (Extended Data Fig. 1b) evolve RI proportional to the number of IIs between them. For the data shown in Extended Data Fig. 1, adding an II between populations causes an increase in postzygotic RI of approximately 5% (one-crossover assay).

Discussion

Rosenzweig[42] noted that "no one denies that population size influences speciation rate. However, the direction of its effect is in doubt." In his classic neutral theory of biodiversity, Hubbell[43] assumed that all species have the same speciation rate per capita and that, therefore, more abundant species will have higher speciation rates. Similarly, Marzluff and Dial[44] proposed that a broad geographic distribution, a correlate of population size, promotes speciation. In contrast, Stanley[45] proposed that species whose populations contract and fragment due to, for example, predation or environmental deterioration, tend to experience high rates of speciation.

The few attempts to study this issue using explicit theoretical models have reached the same conclusion as ours, that low abundance promotes speciation. However, the models they considered differ from ours in important aspects. Mayr[46] proposed that speciation often results from "genetic

revolutions" precipitated by founder events. Others have developed similar models[47]. Most of them modeled ecological speciation and did not consider genetic incompatibilities as the basis for RI. Gavrilets and Hastings[48] did consider several DMI scenarios and showed that founder effect speciation is "plausible" under some of them, but did not consider the accumulation of multiple DMIs of arbitrary complexity. One question raised by founder effect speciation models is whether fluctuations of population size, such as those caused by founder events, may have a different effect on the accumulation of genetic incompatibilities when compared to constant (low) population size, as in our model.

Khatri and Goldstein showed that genetic incompatibilities accumulate more quickly in smaller populations in several models of gene regulation[14,15,49]. Like our model, their models included strong incompatibilities that create holes in the fitness landscape. However, the fitness of a viable genotype is a continuous function of the phenotype in their models. Khatri and Goldstein have argued that their result arises because smaller populations evolve a greater drift load and occupy regions of genotype space closer to the holes in the fitness landscape. We believe that the distance of a genotype to a hole plays the same role in our model because it determines the mutational robustness of the genotype. One limitation of the Khatri and Goldstein studies is that they considered only the SSWM regime and, therefore, were not able to evaluate the effect of recombination. Furthermore, under SSWM, neither genetic drift nor recombination would have any effect on the rate of accumulation of incompatibilities in our model. Therefore, our results must also involve evolutionary mechanisms not captured in the models studied by Khatri and Goldstein. We believe the ingredient they missed is segregating variation at loci involved in genetic incom-

patibilities. Such variation has been known to exist in natural populations of several species[50,51]. Corbett-Detig et al.[51] found evidence that multiple pairwise DMIs are currently segregating within natural populations of *D. melanogaster*. They surveyed a large panel of recombinant inbred lines (RILs) and found 22 incompatible pairs of alleles at unlinked loci in the RILs; of the 44 alleles, 27 were shared by two or more RILs, indicating that multiple DMIs are polymorphic within natural populations. They also found evidence for multiple DMIs in RIL panels in *Arabidopsis* and maize[51]. Corbett-Detig et al. did not attempt to identify DMIs among linked loci or higher-order DMIs and therefore are likely to have underestimated the actual number of segregating DMIs in the RILs. These observations suggest that the conditions for the effects reported here may occur in nature.

Maya-Lastra and Eaton[52] modelled genetically variable populations evolving in the presence of multiple pairwise DMIs like those envisioned by Orr and, like us, observed that smaller populations accumulated more DMIs. They incorporated recombination in their model but did not investigate its effect. They assumed that mutations involved in DMIs were beneficial and that the DMIs were weak. When DMIs were caused by neutral mutations, the effect of population size disappeared. Selection also played a decisive role in our model. Recombination generated selection for genetic robustness. Therefore, higher recombination probabilities and rates made selection stronger and larger population sizes made selection more efficient. If every genotype had the same genetic robustness in our model—i.e., if all mutations were truly neutral—, we would also expect the effect of population size on the accumulation of genetic incompatibilities to vanish.

These earlier studies, when combined with ours, lead to a clear prediction that postzygotic RI should build up more quickly in smaller populations. There have been multiple attempts to test this prediction experimentally by subjecting populations to founder events, sometimes repeatedly. In one of these studies, Ringo et al.[53] subjected 8 experimental populations of *D. simulans* to repeated bottlenecks of a single pair of individuals every 6 months. Between bottlenecks these "drift" lines were allowed to expand to large size. Another 8 populations were maintained at a larger size and subjected to a variety of artificial selection regimes. The drift lines experienced stronger inbreeding than the "selection" lines. An ancestral "base" population was kept at larger size than either treatment group throughout. Ringo et al. found that after 6 bottleneck/expansion cycles the drift lines evolved stronger postzygotic RI relative to the base population than did the selection lines (28% vs 4% reduction in hybrid fitness), in agreement with our prediction. However, the results from three similar experiments on other species of *Drosophila* were inconclusive[54–57].

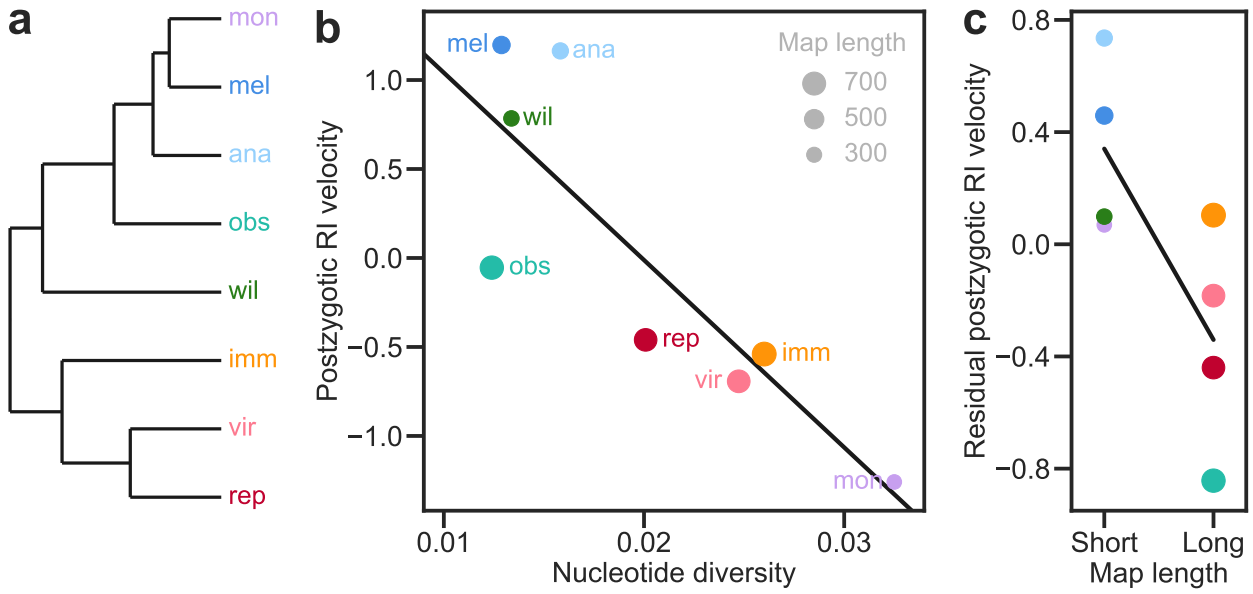


Fig. 6: ***Drosophila*** **species with higher levels of neutral polymorphism and higher recombination accumulate postzygotic RI more slowly.**

Fig. 6 (full caption): *Drosophila* species with higher levels of neutral polymorphism and higher recombination accumulate postzygotic RI more slowly. (a) Phylogenetic relationships between 8 *Drosophila* species groups[58] (*ananassae, immigrans, melanogaster, montium, obscura, repleta, virilis, willistoni*). **(b)** The rates with which postzygotic RI accumulates with genetic distance in the species groups shown in **(a)** were estimated by Rabosky & Matute[58] based on the proportion of sterile or inviable F_1 offspring in interspecific crosses[59,60]. Positive and negative values represent faster- and slower-than-average RI velocities. The levels of neutral polymorphism in individual species were measured by the synonymous nucleotide diversity (π_s) at the *Adh* locus (see Supplementary Data 1). Values are the mean π_s for the species or subspecies in each *Drosophila* species group. Multiple estimates for a single species were averaged before calculating the overall mean for the species group. The line shows a fit by linear ordinary least-squares regression of RI velocity against π_s ($\text{OLS}: \text{ slope and } 95\% \text{ CI} = -105 \pm 67,\ R^2 = 0.71,\ P = 0.008$). The results of a phylogenetic generalized least-squares regression analysis assuming Brownian motion were almost identical ($\text{PGLS}: R^2 = 0.82,\ P = 0.002$). **(c)** Residuals in RI velocity from the regression in **(b)** in the species groups with either shorter or longer total map lengths than average. The point areas in **(b)** and **(c)** are proportional to the mean total map length in cM in the species group (see Supplementary Data 2). Adding map length to the regression model did not improve fit in a statistically significant manner ($\text{multiple regression OLS}: \text{ coefficient} = -0.0016 \pm 0.0021,\ R^2 = 0.84,\ P = 0.11;\ \text{PGLS}: R^2 = 0.89,\ P = 0.13$). None of the traits showed a significant phylogenetic signal based on either the K or λ statistics[61,62] (see Supplementary Data 3).

There is also comparative evidence supporting our prediction. Huang et al.[63] tested for correlations between a measure of the level of neutral polymorphism within species—a proxy for effective population size, N_e—and speciation rates estimated from phylogenies in three genera of lichen-forming fungi. They found a significantly negative correlation in one genus, in agreement with our prediction, but no significant correlations in the other two genera. A larger phylogenetic study found significant negative correlations between speciation rates and a different proxy for N_e—geographic range size—in both birds and mammals[52]. These results, while encouraging, have the limitation that they have used overall speciation rate, which may not be driven primarily by postzygotic RI. Furthermore, the estimates of speciation rate may themselves be problematic because many different temporal patterns of speciation and extinction rates can have equal probability for a given phylogeny[64,65]. To address these limitations we tested for a correlation between the level of neutral polymorphism within species and the rate with which postzygotic RI accumulates with genetic distance in *Drosophila*.[58] We found a significant negative correlation (Fig. 6b). Note, however, that these comparative correlations could have alternative explanations. For example, reductions in population size might be a *consequence* of speciation, rather than its cause[66].

Another prediction from our work is that differences in the size of diverging populations will cause asymmetric accumulation of genetic incompatibilities: the expected number of IIs will be greater in introgressions from a large population to a small population than in the opposite direction. Asymmetric postzygotic RI has been observed between reciprocal crosses of many closely related species—a phenomenon termed Darwin's corollary to Haldane's rule[67]. We propose that this kind of asymmetry might be explained by differences in genetic robustness between the recipient populations to asymmetric introgression of cytoplasmic organelles and of the X chromosome in male offspring. Yukilevich[60] surveyed RI asymmetry between many *Drosophila* species. We were able to determine the level of neutral polymorphism within both species of 8 species pairs in which he found asymmetric postzygotic RI. The pattern of asymmetry in RI between these species pairs showed a statistically significant trend in the direction we predicted (Extended Data Fig. 2).

Environmental stochasticity selects directly for environmental robustness and indirectly for genetic robustness[20,36,68]. Therefore, we predict that temporal variation in environmental conditions should hinder speciation. This prediction is consistent with the observation that speciation rates are higher in tropical lineages than in temperate lineages of amphibians[69], birds[70], and mammals[71]. Furthermore, Yukilevich[72] found that postzygotic RI (but not prezygotic RI) accumulated more than twice as quickly between tropical species pairs of *Drosophila* than between temperate and subtropical species pairs of *Drosophila*. However, many other hypotheses have been proposed to explain these patterns[73,74].

As with population size, no one denies that recombination influences speciation rate. But,

again, the direction of its effect is disputed. Butlin[75] argued that "speciation can be viewed as the evolution of restrictions on the freedom of recombination". In general, low recombination is found to promote speciation in the *presence* of gene flow between diverging populations because it restricts introgression. For example, chromosomal rearrangements often suppress recombination along a large region of the genome. Navarro and Barton[76] showed that such chromosomal rearrangements facilitate the accumulation of DMIs in the presence of gene flow. However, these results are not universal. Bank et al. studied the effect of recombination on the maintenance of a DMI involving two loci in a haploid with gene flow from a continental to an island population[77]. They showed that recombination can have opposing effects on the maintenance of the DMI depending on the fitnesses of the genotypes at the DMI loci. In some scenarios, the DMI is maintained more easily through selection against immigrants when recombination is low. But in other scenarios, a DMI is maintained more easily through selection against hybrids when recombination is high. Founder effect speciation models have also tended to conclude that high recombination should promote speciation[78].

There is some comparative evidence for the prediction that recombination slows down the accumulation of postzygotic RI. We found a negative correlation between the genome-wide recombination rate and the rate with which postzygotic RI accumulates with genetic distance in *Drosophila* species groups[58], after removing the effect of the level of neutral polymorphism (Fig. 6c). Including recombination rate in the ordinary least-squares regression model increased the proportion of variance in RI velocity explained by the predictors from 71% to 84%. However, this improvement was not statistically significant ($P = 0.11$).

In contrast to the results from earlier models, we found that recombination within populations can hinder speciation in the *absence* of gene flow. Our results indicate that recombination suppresses the accumulation of genetic incompatibilities within populations, which in turn slows down their accumulation between populations. In other words, we found that microevolution and macroevolution are *coupled* in our model. This coupling is also a feature of our result that genetic drift promotes speciation. In both cases, processes acting within populations to alter genetic robustness explain variation in speciation rates, contrary to Stanley's[4] conjecture.

The extent to which our results can be generalized to real organisms requires further investigation. Our model makes several strong assumptions, such as, that all viable genotypes have the same viability (although they can differ in overall fitness), all genetic incompatibilities result in complete inviability, strong epistatic interactions are common, haploidy, and populations of constant size diverge in allopatry under the same selection regime (i.e., mutation-order speciation[35]). The effect of each of these assumptions can be investigated within the framework of our model. Such an exploration has the potential to serve as the foundation for a new theory of the coupling between microevolution and macroevolution.

Methods

Genotype and phenotype: The genotype is an RNA sequence of length L. In all simulations summarized in this paper we used $L = 100$ nucleotides and assumed that genotypes are haploid. The phenotype associated with a particular genotype is its minimum free-energy secondary structure, as determined using the ViennaRNA package version 2.1.9 with default parameters[79].

Selection: The viability of sequence i is defined as in Kalirad and Azevedo[13]

$$w_i = \begin{cases} 1 & \text{if} \quad \beta_i > \alpha \quad \text{and} \quad \delta_i \leqslant \alpha \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where β_i is the number of base pairs in the secondary structure of sequence i, δ_i is the base-pair distance between the secondary structure of sequence i and a reference secondary structure, and α is an arbitrary threshold. In all simulations summarized in this paper we used $\alpha = 12$.

Recombination: From a population of size N, we randomly sample with replacement two sets of N viable genotypes, S_1 and S_2. Genotypes from S_1 and S_2 are paired, such that the ith genotypes from each set recombine.

Free recombination: In the free recombination regime, every offspring is recombinant. Crossovers occur independently with probability 0.5 at each of the $L-1$ internucleotide sites in the genome.

Limited recombination: In the limited (not free) recombination regime, offspring are recombinant with probability r resulting from a single crossover located at random at one of the $L-1$ internucleotide sites in the genome.

No recombination: If there is no recombination, the offspring is a copy of a randomly selected parent.

Mutation: Mutations arise according to a binomial process where each site mutates with probability equal to the per-site mutation rate μ in each generation. All types of point mutations occur with equal probability. Insertions and deletions are not considered. In all simulations summarized in this paper we used $\mu = 0.001$, resulting in a genome-wide mutation rate of $U = \mu L = 0.1$.

Evolution: *Ancestor:* We start by generating a random RNA sequence with $L = 100$ nucleotides and defining its minimum free-energy secondary structure as the reference provided it is viable (see Equation 3). This initial sequence is not representative of the holey fitness landscape it defines because it has the reference structure (i.e., $\delta = 0$, see Equation 3), whereas most viable sequences are $\delta \approx \alpha$ base pairs away from the reference structure[13]. To obtain a representative sequence, we subject the initial sequence to a random walk of 200 steps, allowing multiple hits at individual sites. We define the resulting sequence as the ancestor.

Initial evolution: The ancestor is cloned N times to create a population. We allow this population to evolve for t_1 generations to reach an approximate mutation-recombination-selection-drift equilibrium ($t_1 = 1000$ for all regimes except $N = 500$ with free-recombination, which required $t_1 = 4000$ to reach equilibrium). We made an exact copy of the resulting population to obtain two identical ancestral populations.

Divergence: The two ancestral populations are allowed to evolve for t_2 generations under a Wright-Fisher model following a selection-recombination-mutation life cycle ($t_2 = 500$ for all regimes except $N = 500$ with free-recombination, which required $t_2 = 2000$ to reach similar levels of mean genetic divergence).

Incompatible introgressions: Two viable genotypes, i and j, differ at k sites. The number of incompatible introgressions[13] (IIs) from genotype i to genotype j, Π_{ij}, was measured by introgressing each allele at a divergent site from genotype i to genotype j one at a time and counting the number of inviable introgressed genotypes ($w = 0$, see Equation 3). Π_{ij} is not necessarily equal to Π_{ji}. As populations diverged, we periodically estimated Π_{ij} in a particular direction (e.g., from

population 1 to population 2) from 10 randomly chosen pairs of individuals from the two populations, sampled with replacement (except when $N = 1$, where the assay was conducted between the single individuals).

Segregating IIs: We also counted IIs within populations as they diverged. We used the same approach as for IIs between populations, except that the 10 pairs of individuals were sampled randomly with replacement from the same population.

Mutational robustness: The mutational robustness of a genotype was measured as the proportion ν of its $3L$ single mutations that did not render the genotype inviable. The robustness of each replicate population was estimated as the average ν of 10 individuals sampled at random with replacement (except when $N = 1$, where it was simply the ν of the individual).

Reproductive isolation: The level of postzygotic reproductive isolation (RI) between two populations is defined as:

$$\mathrm{RI} = 1 - \frac{1}{N}\sum_{i=1}^{N} w_i \quad , \tag{4}$$

where w_i is the fitness of a genotype resulting the from recombination between the ith genotypes of sets S_1 and S_2, where set S_j contains N genotypes drawn randomly with replacement from diverging population j. Recombinant sequences were produced by either a single crossover or free recombination.

Statistical analysis: *Snowball exponent:* The model in Equation 2 was fit by least-squares regression. For a particular regime (e.g., SSWM in Fig. 3a) or combination of regimes (e.g., all populations summarized in Figs. 3 and 4) the snowball exponent c was estimated by maximum

likelihood (ML) while allowing each replicate simulation run to have different values of a and b. ML estimation was performed using the `mle2` function of the R package `bbmle` version 1.0.25 (`https://github.com/bbolker/bbmle`).

Rate of II accumulation: Once a value of c was estimated, we fit the model in Equation 2 for each replicate simulation run by linear regression ($\Pi = a + bx$, where $x = k^c$ for a fixed value of c). The rates of II accumulation summarized in Figs. 3c and 4c were calculated by averaging estimates of b from individual replicate runs.

Phylogenetic analyses: Phylogenetic generalized least-squares regression analyses were done using the R package `caper`[80] version 1.0.1. Phylogenetic signal was tested using the R packages `phytools`[81] version 1.0-3 and `geiger`[82] version 2.0.10.

Data and code availability: The software used to run all simulations was written in Python 2.7 with NumPy[83] version 1.14.5. The code used to analyze the resulting data was written in Python 3.7 with NumPy[83] version 1.21.0 and R[84] 4.1.0. All code and data are available at `https://github.com/Kalirad/RNA_pop_snowball`.

References

1. Cooney, C. R. & Thomas, G. H. Heterogeneous relationships between rates of speciation and body size evolution across vertebrate clades. *Nat. Ecol. Evol.* **5**, 101–110 (2021).

2. Jablonski, D. Species selection: Theory and data. *Annu. Rev. Ecol. Evol. Syst.* **39**, 501–524 (2008).

3. Harvey, M. G., Singhal, S. & Rabosky, D. L. Beyond reproductive isolation: Demographic controls on the speciation process. *Annu. Rev. Ecol. Evol. Syst.* **50**, 75–95 (2019).

4. Stanley, S. M. A theory of evolution above the species level. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 646–650 (1975).

5. Dobzhansky, T. *Genetics and the Origin of Species* (Columbia Univ. Press, New York, 1937).

6. Muller, H. J. Isolating mechanisms, evolution and temperature. *Biol. Symp.* **6**, 71–125 (1942).

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

8. Matute, D. R., Butler, I. A., Turissini, D. A. & Coyne, J. A. A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science* **329**, 1518–1521 (2010).

9. Matute, D. R. & Gavin-Smyth, J. Fine mapping of dominant X-linked incompatibility alleles in *Drosophila* hybrids. *PLoS Genet.* **10**, e1004270 (2014).

10. Wang, R. J., White, M. A. & Payseur, B. A. The pace of hybrid incompatibility evolution in house mice. *Genetics* **201**, 229–242 (2015).

11. Moyle, L. C. & Nakazato, T. Hybrid incompatibility "snowballs" between *Solanum* species. *Science* **329**, 1521–1523 (2010).

12. Sherman, N. A., Victorine, A., Wang, R. J. & Moyle, L. C. Interspecific tests of allelism reveal the evolutionary timing and pattern of accumulation of reproductive isolation mutations. *PLoS Genet.* **10**, e1004623 (2014).

13. Kalirad, A. & Azevedo, R. B. R. Spiraling complexity: A test of the snowball effect in a computational model of RNA folding. *Genetics* **206**, 377–388 (2017).

14. Khatri, B. S. & Goldstein, R. A. Simple biophysical model predicts faster accumulation of hybrid incompatibilities in small populations under stabilizing selection. *Genetics* **201**, 1525–37 (2015).

15. Khatri, B. S. & Goldstein, R. A. Biophysics and population size constrains speciation in an evolutionary model of developmental system drift. *PLoS Comput. Biol.* **15**, e1007177 (2019).

16. Kondrashov, A. S., Sunyaev, S. & Kondrashov, F. A. Dobzhansky–Muller incompatibilities in protein evolution. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14878–83 (2002).

17. Welch, J. J. Accumulating Dobzhansky-Muller incompatibilities: Reconciling theory and data. *Evolution* **58**, 1145–1156 (2004).

18. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940 (1999).

19. van Nimwegen, E., Crutchfield, J. P. & Huynen, M. Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9716–9720 (1999).

20. Ancel, L. W. & Fontana, W. Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool. (Mol. Dev. Evol.)* **288**, 242–283 (2000).

21. Wagner, A. Robustness and evolvability: a paradox resolved. *Proc. R. Soc. B* **275**, 91–100 (2008).

22. Draghi, J. A., Parsons, T. L., Wagner, G. P. & Plotkin, J. B. Mutational robustness can facilitate adaptation. *Nature* **463**, 353–355 (2010).

23. Draghi, J. A., Parsons, T. L. & Plotkin, J. B. Epistasis increases the rate of conditionally neutral substitution in an adapting population. *Genetics* **187**, 1139–52 (2011).

24. Gavrilets, S. *Fitness Landscapes and the Origin of Species* (Princeton Univ. Press, Princeton, N.J., 2004).

25. Lipman, D. J. & Wilbur, W. J. Modelling neutral and selective evolution of protein folding. *Proc. R. Soc. B* **245**, 7–11 (1991).

26. Schuster, P., Fontana, W., Stadler, P. F. & Hofacker, I. L. From sequences to shapes and back: a case study in RNA secondary structures. *Proc. R. Soc. B* **255**, 279–284 (1994).

27. Azevedo, R. B. R., Lohaus, R., Srinivasan, S., Dang, K. K. & Burch, C. L. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature* **440**, 87–90 (2006).

28. Szöllősi, G. J. & Derényi, I. The effect of recombination on the neutral evolution of genetic robustness. *Math. Biosci.* **214**, 58–62 (2008).

29. Klug, A., Park, S.-C. & Krug, J. Recombination and mutational robustness in neutral fitness landscapes. *PLoS Comput. Biol.* **15**, e1006884 (2019).

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

31. Livingstone, K. *et al.* A stochastic model for the development of Bateson–Dobzhansky–Muller incompatibilities that incorporates protein interaction networks. *Math. Biosci.* **238**, 49–53 (2012).

32. Gillespie, J. H. A simple stochastic gene substitution model. *Theor. Popul. Biol.* **23**, 202–215 (1983).

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

34. Fraïsse, C., Elderfield, J. A. D. & Welch, J. J. The genetics of speciation: are complex incompatibilities easier to evolve? *J. Evol. Biol.* **27**, 688–699 (2014).

35. Schluter, D. Evidence for ecological speciation and its alternative. *Science* **323**, 737–741 (2009).

36. de Visser, J. A. G. M. *et al.* Perspective: Evolution and detection of genetic robustness. *Evolution* **57**, 1959–1972 (2003).

37. Jost, L. G_{ST} and its relatives do not measure differentiation. *Mol. Ecol.* **17**, 4015–4026 (2008).

38. Gardner, A. & Kalinka, A. T. Recombination and the evolution of mutational robustness. *J. Theor. Biol.* **241**, 707–715 (2006).

39. Misevic, D., Ofria, C. & Lenski, R. E. Sexual reproduction reshapes the genetic architecture of digital organisms. *Proc. R. Soc. B* **273**, 457–464 (2006).

40. Martin, O. C. & Wagner, A. Effects of recombination on complex regulatory circuits. *Genetics* **183**, 673–684 (2009).

41. Lohaus, R., Burch, C. L. & Azevedo, R. B. R. Genetic architecture and the evolution of sex. *J. Hered.* **101**, S142–S157 (2010).

42. Rosenzweig, M. L. Loss of speciation rate will impoverish future diversity. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 5404–5410 (2001).

43. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, N.J., 2001).

44. Marzluff, J. M. & Dial, K. P. Life history correlates of taxonomic diversity. *Ecology* **72**, 428–439 (1991).

45. Stanley, S. M. Population size, extinction, and speciation: The fission effect in Neogene Bivalvia. *Paleobiology* **12**, 89–110 (1986).

46. Mayr, E. Change of genetic environment and evolution. In Huxley, J., Hardy, A. C. & Ford, E. B. (eds.) *Evolution as a Process*, 157–180 (Allen & Unwin, 1954).

47. Carson, H. L. & Templeton, A. R. Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annu. Rev. Ecol. Syst.* **15**, 97–131 (1984).

48. Gavrilets, S. & Hastings, A. Founder effect speciation: A theoretical reassessment. *Am. Nat.* **147**, 466–491 (1996).

49. Khatri, B. S. & Goldstein, R. A. A coarse-grained biophysical model of sequence evolution and the population size dependence of the speciation rate. *J. Theor. Biol.* **378**, 56–64 (2015).

50. Cutter, A. D. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. *Tr. Ecol. Evol.* **27**, 209–218 (2012).

51. Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L. & Ayroles, J. F. Genetic incompatibilities are widespread within species. *Nature* **504**, 135–7 (2013).

52. Maya-Lastra, C. A. & Eaton, D. A. R. Genetic incompatibilities do not snowball in a demographic model of speciation. *bioRxiv* (2021).

53. Ringo, J., Wood, D., Rockwell, R. & Dowse, H. An experiment testing two hypotheses of speciation. *Am. Nat.* **126**, 642–661 (1985).

54. Powell, J. R. The founder-flush speciation theory: an experimental approach. *Evolution* **32**, 465–474 (1978).

55. Kilias, G., Alahiotis, S. N. & Pelecanos, M. A multifactorial genetic investigation of speciation theory using *Drosophila melanogaster*. *Evolution* **34**, 730–737 (1980).

56. Dodd, D. M. B. & Powell, J. R. Founder-flush speciation: an update of experimental results with *Drosophila*. *Evolution* **39**, 1388–1392 (1985).

57. Galiana, A., González-Candelas, F. & Moya, A. Postmating isolation analysis in founder-flush experimental populations of *Drosophila pseudoobscura*. *Evolution* **50**, 941–944 (1996).

58. Rabosky, D. L. & Matute, D. R. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15354–9 (2013).

59. Coyne, J. A. & Orr, H. A. Patterns of speciation in *Drosophila*. *Evolution* **43**, 362–381 (1989).

60. Yukilevich, R. Asymmetrical patterns of speciation uniquely support reinforcement in *Drosophila*. *Evolution* **66**, 1430–46 (2012).

61. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–84 (1999).

62. Blomberg, S. P., Garland, T., Jr & Ives, A. R. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**, 717–45 (2003).

63. Huang, J.-P., Leavitt, S. D. & Lumbsch, H. T. Testing the impact of effective population size on speciation rates—a negative correlation or lack thereof in lichenized fungi. *Sci. Rep.* **8**, 5729 (2018).

64. Louca, S. & Pennell, M. W. Extant timetrees are consistent with a myriad of diversification histories. *Nature* **580**, 502–505 (2020).

65. Louca, S. & Pennell, M. W. Why extinction estimates from extant phylogenies are so often zero. *Curr. Biol.* **31**, 3168–3173 (2021).

66. Glazier, D. S. Toward a predictive theory of speciation: The ecology of isolate selection. *J. Theor. Biol.* **126**, 323–333 (1987).

67. Turelli, M. & Moyle, L. C. Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics* **176**, 1059–1088 (2007).

68. Meir, E., von Dassow, G., Munro, E. & Odell, G. M. Robustness, flexibility, and the role of lateral inhibition in the neurogenic network. *Curr. Biol.* **12**, 778–786 (2002).

69. Pyron, R. A. & Wiens, J. J. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proc. R. Soc. B* **280**, 20131622 (2013).

70. Ricklefs, R. E. Estimating diversification rates from phylogenetic information. *Tr. Ecol. Evol.* **22**, 601–610 (2007).

71. Rolland, J., Condamine, F. L., Jiguet, F. & Morlon, H. Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.* **12**, e1001775 (2014).

72. Yukilevich, R. Tropics accelerate the evolution of hybrid male sterility in *Drosophila*. *Evolution* **67**, 1805–1814 (2013).

73. Mittelbach, G. G. *et al.* Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10**, 315–331 (2007).

74. Schluter, D. Speciation, ecological opportunity, and latitude. *Am. Nat.* **187**, 1–18 (2016).

75. Butlin, R. K. Recombination and speciation. *Mol. Ecol.* **14**, 2621–2635 (2005).

76. Navarro, A. & Barton, N. H. Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation. *Evolution* **57**, 447–459 (2003).

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

78. Templeton, A. R. The reality and importance of founder speciation in evolution. *BioEssays* **30**, 470–479 (2008).

79. Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms for Molecular Biology* **6**, 26 (2011).

80. Orme, D. *et al. caper: Comparative Analyses of Phylogenetics and Evolution in R.* R package version 1.0.1 (2018).

81. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Meth. Ecol. Evol.* **3**, 217–223 (2012).

82. Pennell, M. W. *et al.* geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* **30**, 2216–2218 (2014).

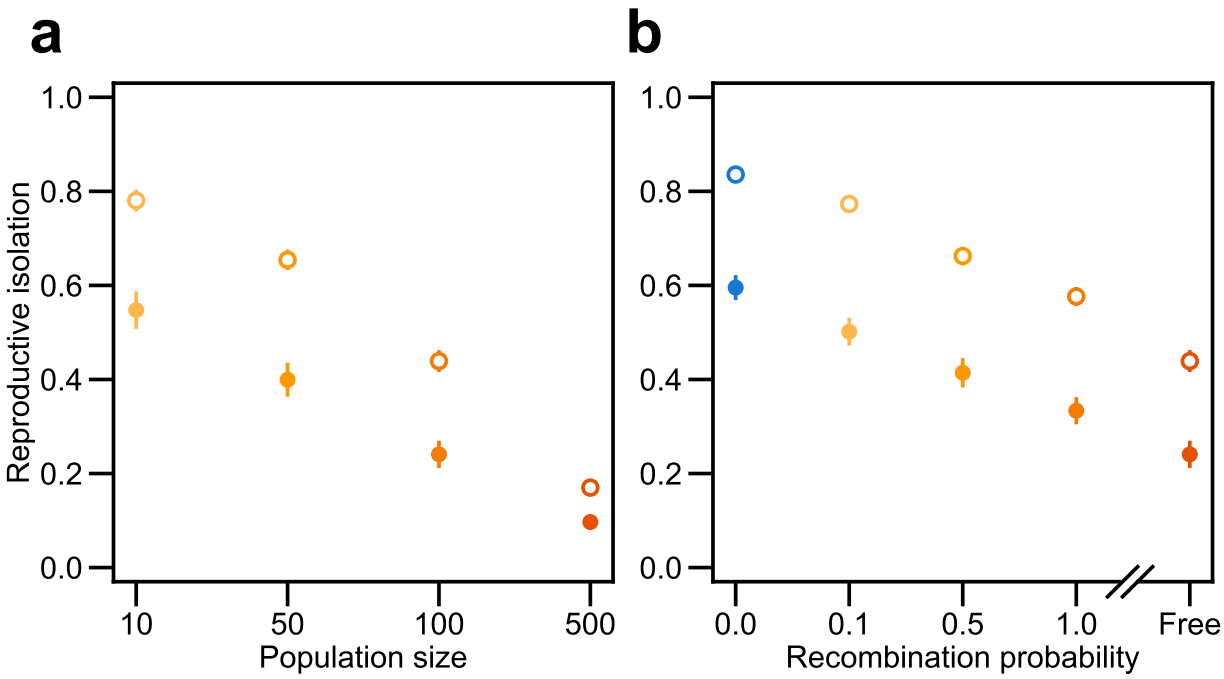
83. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

84. Team, R. C. *R: A Language and Environment for Statistical Computing*. Vienna, Austria (2021). URL `https://www.R-project.org/`.

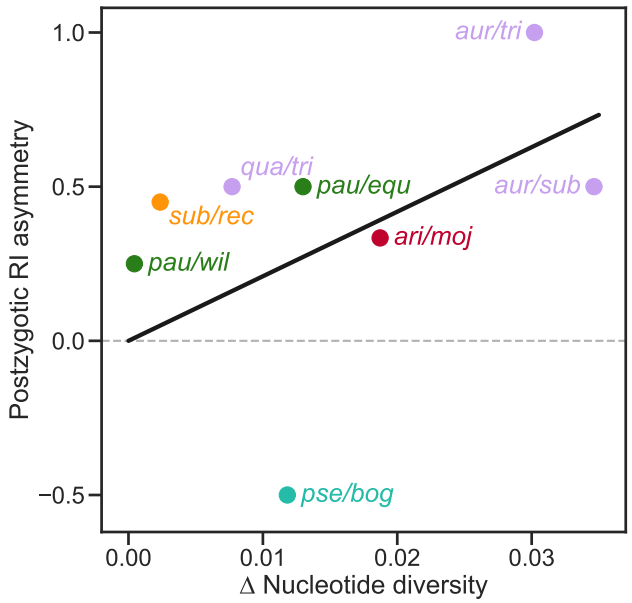
Acknowledgements This work was funded by grants from the National Science Foundation DEB-2014566 to RBRA and DEB-2014943 to CB. We used the Opuntia cluster from the Hewlett Packard Enterprise Data Science Institute at the University of Houston. D. Wiernasz and M. Servedio provided helpful comments on the manuscript.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to R.B.R.A. (email: razevedo@uh.edu).



Extended Data Fig. 1: Genetic drift promotes and recombination hinders speciation. (a) Postzygotic reproductive isolation (RI) in populations of different sizes evolving under free-recombination (summarized in Fig. 3b). **(b)** Postzygotic RI in populations of $N = 100$ individuals evolving under different recombination probabilities (summarized in Fig. 4a). RI assays were conducted at the end of the simulations with recombination probability $r = 1$ and either one crossover (filled circles) or free-recombination (open circles). Values are means of 200 replicate simulation runs per regime. Error bars are 95% CIs.



Extended Data Fig. 2: Differences in the level of neutral polymorphism of two species explain asymmetry in postzygotic RI between them. Postzygotic RI asymmetry for the *sp*1/*sp*2 pair was estimated[72] by taking the RI score from a ♀ *sp*1 × ♂ *sp*2 cross and subtracting the RI score from the reciprocal cross. The *x*-axis shows the difference in the synonymous nucleotide diversity $(\pi_{s1} - \pi_{s2})$ at the *Adh* locus (except the *rec/sub* pair for which we used the *Adhr* locus; see Supplementary Data 4). The colors of the points match the species groups shown in Fig. 6a except that the *quinaria* group (sister group to *immigrans*) is represented by orange: *montium* group: *auraria, subauraria, triauraria*; *obscura* group: *bogotana, pseudoobscura*; *quinaria* group: *recens, subquinaria*; *repleta* group: *arizonae, mojavensis*; *willistoni* group: *equinoxialis, paulistorum, willistoni*. The line shows a fit by linear ordinary least-squares regression through the origin (slope $= 20.9 \pm 17.8$, $R^2 = 0.52$, $P = 0.028$).