Evolution, 2023, **XX(XX)**, 1–14
https://doi.org/10.1093/evolut/qpad021
Advance access publication 13 February 2023
Original Article

Recombination and selection against introgressed DNA

Carl Veller[1,2], Nathaniel B. Edelman[3,4], Pavitra Muralidhar[1,2], Martin A. Nowak[5,6]

[1]Center for Population Biology, University of California, Davis, CA, United States
[2]Department of Evolution and Ecology, University of California, Davis, CA, United States
[3]Yale Institute for Biospheric Studies, Yale University, New Haven, CT, United States
[4]Yale School for the Environment, Yale University, New Haven, CT, United States
[5]Department of Mathematics, Harvard University, Cambridge, MA, United States
[6]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, United States

Corresponding author: Center for Population Biology, Department of Evolution and Ecology, University of California, Davis, CA, USA. Email: carl.veller@gmail.com

Abstract

Introgressed DNA is often deleterious at many loci in the recipient species' genome, and is therefore purged by selection. Here, we use mathematical modeling and whole-genome simulations to study the influence of recombination on this process. We find that aggregate recombination controls the genome-wide rate of purging in the early generations after admixture, when purging is most rapid. Aggregate recombination is influenced by the number of chromosomes and heterogeneity in their size, and by the number of crossovers and their locations along chromosomes. A comparative prediction is that species with fewer chromosomes should purge introgressed ancestry more profoundly, and should therefore exhibit weaker genomic signals of historical introgression. Turning to within-genome patterns, we show that, in species with autosomal recombination in both sexes, more purging is expected on sex chromosomes than autosomes, all else equal. The opposite prediction holds for species without autosomal recombination in the heterogametic sex. Finally, positive correlations between recombination rate and introgressed ancestry have recently been observed within the genomes of several species. We show that these correlations are likely driven not by recombination's effect in unlinking neutral from deleterious introgressed alleles, but by recombination's effect on the rate of purging of deleterious introgressed alleles themselves.

Keywords: recombination, introgression, hybridization, speciation

Introduction

It has become clear in recent years that hybridization and subsequent genetic introgression are common features of the evolutionary histories of many species, including our own (Edelman & Mallet, 2021; Taylor & Larson, 2019). It has therefore become a major focus of evolutionary genetics to understand the impact of introgression on the population genomics of species, and conversely to learn about historical admixture from genomic patterns of introgressed ancestry (Moran et al., 2021).

Introgressed DNA is typically deleterious in the recipient species. This can be for a number of reasons: donor-species alleles could be maladapted to the recipient species' ecology (Schluter, 2009) or genome (Dobzhansky, 1937; Muller, 1942; Orr, 1995), and the donor species could carry a higher genetic load than the recipient species (Harris & Nielsen, 2016; Juric et al., 2016). Importantly, the deleterious effect of introgressed ancestry is often spread across a large number of loci. For example, it has been estimated that Neanderthal alleles were deleterious in humans at ~1,000 loci (Harris & Nielsen, 2016; Juric et al., 2016), and comparable estimates have been obtained for other species (e.g., Aeschbacher et al., 2017; Schumer et al., 2014).

Since introgressed alleles initially appear in the recipient population in perfect linkage disequilibrium, recombination will clearly be influential in determining their fate. The influence of recombination on the purging of deleterious introgressed alleles can be interpreted in two complementary ways. The first is that recombination breaks up the initially very long blocks of introgressed DNA, which are strongly selected against, into smaller and smaller blocks, which are more weakly selected against (Barton, 1983). The second is that recombination, by distributing introgressed ancestry more and more evenly among more and more individuals, reduces ancestry variance over time and thus reduces the efficiency of selection against deleterious introgressed ancestry (Harris & Nielsen, 2016). Both interpretations indicate that recombination should, over time, reduce the rate at which deleterious introgressed ancestry is purged.

This process has been well studied in the context of stylized genetic maps (Bengtsson, 1985) and small genomic segments (e.g., Barton, 1983; Barton & Bengtsson, 1986). However, recent evidence suggests that introgressed alleles can be deleterious at many loci throughout the genome, and that many species—with a great diversity of recombination processes—have experienced introgression. Moreover, recently developed software enables evolutionary simulations of whole genomes (Haller & Messer, 2017, 2019; Messer, 2013), and such simulations have already been applied fruitfully to the study of introgression and speciation (e.g., Duranton & Pool, 2022; Harris & Nielsen, 2016; Yamaguchi et al., 2022). In light of these developments, it is worth revisiting the role that recombination plays in the purging of introgressed DNA, taking a whole-genome perspective.

Received July 25, 2022; revisions received January 19, 2023; accepted February 6, 2023

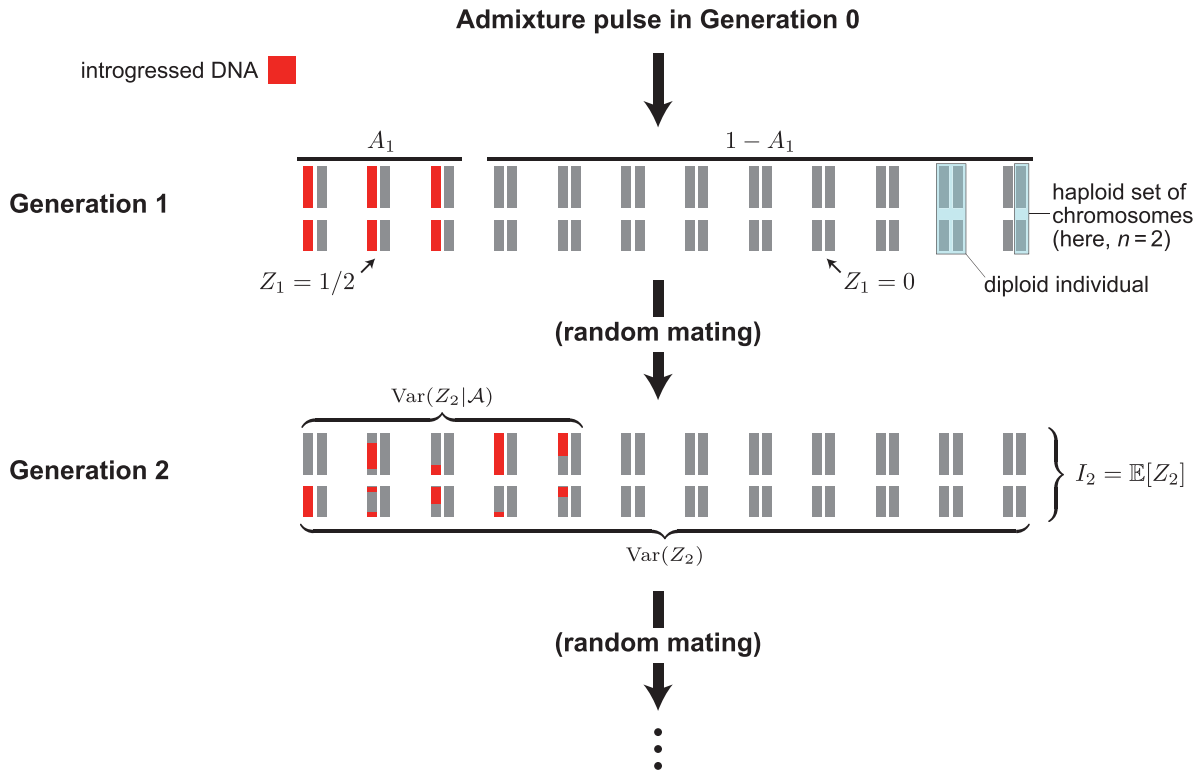

Figure 1. Schematic of the model that we study and the variables used in our calculations.

Here, we use mathematical models and computer simulations to study the role of recombination in selection against introgressed DNA. First, we ask what features of the recombination process affect the efficiency with which introgressed ancestry is purged genome-wide. Cross-species variation in these features would be expected to predict variation in the retention of introgressed ancestry, and thus the strength of the genomic signal of historical admixture. Second, inspired by recent empirical findings in several taxa that introgressed ancestry is preferentially retained in high-recombination regions of the genome (Brandvain et al., 2014; Calfee et al., 2021; Edelman et al., 2019; Martin et al., 2019; Schumer et al., 2018), we investigate the mechanisms that generate within-genome correlations between recombination rate and introgressed ancestry.

Model

We study a model in which selection acts additively against introgressed DNA: if a proportion p of an individual's genome is introgressed, the individual's relative fitness is $1 - pS$. This is the additive version of the model in Barton (1983), and corresponds to a situation where introgressed alleles are deleterious in the recipient species at a large number of loci, with fitness effects additive within and across loci. The loci at which introgressed alleles are deleterious are assumed to be spaced uniformly along the physical map of the genome, although this assumption can be relaxed with minimal change to the interpretation of our results.

This model, with additive fitness effects at many loci, applies best to situations where introgressed ancestry is deleterious because of higher load in the donor species or broad maladaptation of donor-species alleles to the recipient species' ecology. We later briefly consider the case where introgressed ancestry is deleterious because of negative epistatic interactions between donor- and recipient-species alleles (Dobzhansky-Muller incompatibilities).

Notice that, in assuming that an individual's fitness is determined by the fraction of its genome that is introgressed, we have ignored the distinction between deleterious and neutral (or beneficial) introgressed alleles. If introgressed alleles are deleterious at sufficiently many loci throughout the genome, this assumption will have little effect on the overall rate at which introgressed ancestry is purged (Supplementary Figure S1). To understand how and when recombination can cause differential retention of deleterious and neutral introgressed DNA, we later carry out simulations with neutral loci included amongst the deleterious loci. We also later discuss the case where introgressed alleles are strongly deleterious at only a few loci, which pertains in some important systems.

Similar to previous work (e.g., Harris & Nielsen, 2016; Juric et al., 2016; Steinrücken et al., 2018), we assume that hybridization occurs as a pulse in a single generation, such that a certain fraction of individuals in the next generation are F1 hybrids, each with half its genome introgressed. Mating is assumed to be random with respect to ancestry (the role of nonrandom mating in this model is studied in Muralidhar et al., 2022). For simplicity, we ignore sex chromosomes at first. Finally, we assume that the population is large enough that drift can be ignored (in our simulations, the population size $N = 100{,}000$).

All simulations were carried out in SLiM 3 (Haller & Messer, 2019). Code is available at github.com/cveller/RecombinationAndIntrogression.

Species differences in the genome-wide rate of purging

Figure 2 shows the genome-wide purging of introgressed DNA in simulations of the model described above, under the

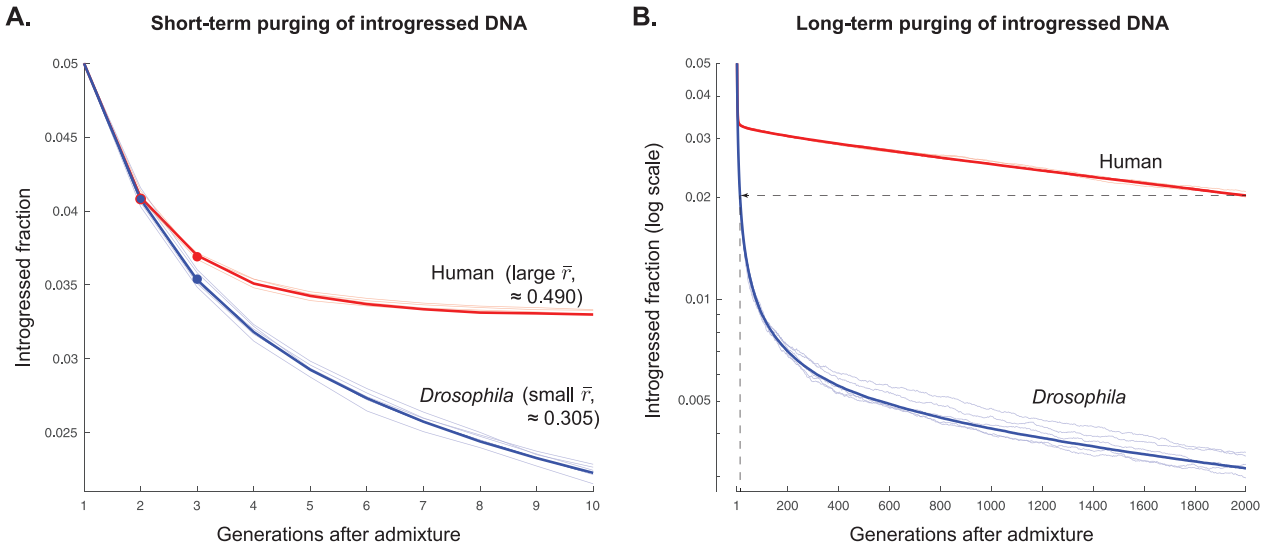

Figure 2. Genome-wide purging of introgressed DNA following an admixture pulse, under the recombination processes of humans and *Drosophila melanogaster*. The recombination process of *D. melanogaster* causes much more introgressed ancestry to be purged in the first few generations after admixture (A), and thus by later generations as well (B), because it is associated with a lower aggregate recombination rate, driven predominantly by the small karyotype of *D. melanogaster* (2 major autosomes) relative to humans (22 autosomes). The dots in A are analytical predictions from Equations (6) and (10). The bold lines are trajectories averaged across 100 replicate simulations; the faint lines are representative trajectories. The dotted line in B marks that *Drosophila* purges as much introgressed DNA in 12 generations as humans do in 2,000 generations.

recombination processes of humans and *Drosophila melanogaster* (using linkage maps produced by Kong et al. (2010) and Comeron et al. (2012), respectively). All other parameters are identical between the two cases, and were chosen to resemble parameters inferred for Neanderthal–human introgression (Harris & Nielsen, 2016; Juric et al., 2016): the initial introgression fraction is 5%, introgressed alleles are deleterious at 1,000 loci, and F1 hybrids suffer a 20% fitness reduction. Simulations were run for 2,000 generations, approximately the number of human generations since Neanderthal introgression. Several features of the trajectories in Figure 2 are noteworthy.

First, in both humans and *Drosophila*, introgressed DNA is purged very rapidly in the first few generations after admixture. The rate of purging subsequently decreases. This effect, expected from previous theory (Bengtsson, 1985) and observed in recent simulations (Harris & Nielsen, 2016; Petr et al., 2019), is so extreme that, for both species, most of the purging that has eventually occurred after 2,000 generations occurred in the first five generations.

Second, *Drosophila* purges introgressed DNA much more efficiently than humans. After 2,000 generations, *Drosophila* has purged 94% of the initial introgressed fraction, while humans have purged only 59%. Put differently, the amount of introgressed DNA that it takes humans 2,000 generations to purge, *Drosophila* has purged after just 12 generations. Therefore, the recombination process of *D. melanogaster* is a much more effective barrier to gene flow than the recombination process of humans.

The observations above are robust to alternative specifications of our model in which the fitness effects of introgressed alleles are variable across loci (Supplementary Figure S2) or driven by Dobzhansky-Muller incompatibilities (Supplementary Figure S3).

Our aim in the following subsections is to derive analytical expressions that help us understand these observations.

Preliminary calculations

Let Z_t be the fraction of a random generation-t zygote's genome that is introgressed, and let G_t be the introgressed fraction of a random gamete produced by generation-t adults. Let A_t be the proportion of generation-t zygotes that have hybrid ancestry (in the pedigree sense—i.e., some of them might nevertheless have inherited no introgressed DNA), and A'_t the corresponding proportion of adults after viability selection (and therefore also the proportion of successful gametes produced by generation t). Let $\mathcal{A}$ signify the property of having hybrid ancestry (which applies to a proportion A_t of zygotes and A'_t of adults and gametes). A graphical representation of the variables used in these calculations is given in Figure 1.

We are interested in how the overall fraction of introgressed DNA, $\mathbb{E}[Z_t]$, is reduced over time by selection—this is the process displayed in Figure 2. We show in SI Section S1 that, in general, the fraction of introgressed DNA purged in the t-th generation is

$$\Delta_t = \frac{\mathbb{E}[Z_t] - \mathbb{E}[Z_{t+1}]}{\mathbb{E}[Z_t]} = \frac{S\mathrm{Var}(Z_t)}{\mathbb{E}[Z_t](1 - S\mathbb{E}[Z_t])}. \quad (1)$$

Therefore, to understand what factors govern the rate of purging of introgressed DNA, we must understand what factors govern the trajectory of $\mathrm{Var}(Z_t)$, the variance across individuals in how much introgressed DNA they carry (Harris & Nielsen, 2016). $\mathrm{Var}(Z_t)$ can be decomposed into two components: a contribution from the fact that some individuals have introgressed ancestry and some do not, and a contribution from variance among those individuals who do have introgressed ancestry. We shall call these the "between group" and "within group" components of the overall variance, respectively. The precise decomposition is

$$\mathrm{Var}(Z_t) = \underbrace{A_t(1 - A_t)(\mathbb{E}[Z_t|\mathcal{A}])^2}_{\text{between}} + \underbrace{A_t\mathrm{Var}(Z_t|\mathcal{A})}_{\text{within}}. \quad (2)$$

An analogous decomposition holds for genetic ancestry variance across gametes:

$$\mathrm{Var}(G_t) = A'_t(1-A'_t)(\mathbb{E}[G_t|\mathcal{A}])^2 + A'_t\mathrm{Var}(G_t|\mathcal{A}). \quad (3)$$

We begin by calculating $\mathrm{Var}(Z_1)$ and $\mathrm{Var}(Z_2)$. For later t, we provide an approximate calculation of $\mathrm{Var}(Z_t)$ in the special case where selection is weak.

Short-term purging of introgressed DNA

Selection in the first generation.

A zygote in the first generation after admixture either carries no introgressed DNA (probability $1-A_1$) or is an F1 hybrid, with half its genome introgressed (probability A_1). Therefore,

$$\mathbb{E}[Z_1] = \frac{A_1}{2} \quad (4)$$

and

$$\mathrm{Var}(Z_1) = \frac{A_1(1-A_1)}{4}. \quad (5)$$

Notice that all of the ancestry variance in generation 1 is due to differences between hybrids and non-hybrids—there is no "within group" contribution, because all hybrids have the same introgressed fraction.

From Equation (1), the proportion of introgressed DNA removed by selection in the first generation is

$$\Delta_1 = \frac{\mathbb{E}[Z_1]-\mathbb{E}[Z_2]}{\mathbb{E}[Z_1]} = \frac{S\mathrm{Var}(Z_1)}{\mathbb{E}[Z_1](1-S\mathbb{E}[Z_1])} = \frac{S(1-A_1)}{2-SA_1}. \quad (6)$$

This expression is, of course, independent of the recombination process, explaining why the introgressed fraction in generation-2 zygotes in our simulations is the same for humans and *Drosophila* (Figure 2A).

Selection in the second generation.

Some generation-2 zygotes have F1 hybrid parents; recombination in these parents will generate ancestry variance among their offspring. From Equation (1) in Veller et al. (2020),

$$\mathrm{Var}(G_1|\mathcal{A}) = \frac{1}{2}\left(\frac{1}{2}-\bar{r}\right) \quad (7)$$

among those gametes produced by generation-1 hybrids, where $\bar{r}$ is the average recombination rate across all pairs of loci in the genome (Veller et al., 2019). When the recombination process differs between the sexes—as is usually the case (Lenormand & Dutheil, 2005; Sardell & Kirkpatrick, 2020)—the sex-averaged value of $\bar{r}$ applies in Equation (7).

Noting that $\mathbb{E}[G_1|\mathcal{A}]=1/2$ and applying Equation (3), we find that

$$\begin{aligned}\mathrm{Var}(G_1) &= A'_1(1-A'_1)(\mathbb{E}[G_1|\mathcal{A}])^2 + A'_1\mathrm{Var}(G_1|\mathcal{A}) \\ &= \frac{A'_1(1-A'_1)}{4} + \frac{A'_1}{2}\left(\frac{1}{2}-\bar{r}\right), \end{aligned} \quad (8)$$

where $A'_1 = A_1\frac{1-S/2}{1-SA_1/2}$ is the fraction of generation-1 gametes that derive from hybrid parents.

Since a zygote's introgressed fraction is the average of the introgressed fractions of the gametes that produced it, and since we have assumed that mating is random with respect to ancestry,

$$\mathrm{Var}(Z_2) = 2\mathrm{Var}\left(\frac{G_1}{2}\right) = \frac{A'_1(1-A'_1)}{8} + \frac{A'_1}{4}\left(\frac{1}{2}-\bar{r}\right), \quad (9)$$

where the first term is the "between group" contribution and the second term is the "within group" contribution.

Finally, we observe that $\mathbb{E}[Z_2] = A'_1/2$ and substitute Equation (9) into Equation (1) to find the proportion of introgressed DNA that is purged in the second generation after admixture:

$$\begin{aligned}\Delta_2 &= \frac{\mathbb{E}[Z_2]-\mathbb{E}[Z_3]}{\mathbb{E}[Z_2]} = \frac{S\mathrm{Var}(Z_2)}{\mathbb{E}[Z_2](1-S\mathbb{E}[Z_2])} \\ &= \frac{S\left(\frac{A'_1(1-A'_1)}{8}+\frac{A'_1}{4}\left(\frac{1}{2}-\bar{r}\right)\right)}{\frac{A'_1}{2}\left(1-S\frac{A'_1}{2}\right)} = \frac{S(1-A'_1+1-2\bar{r})}{4-2SA'_1}. \end{aligned} \quad (10)$$

Equation (10) reveals that the rate of purging of introgressed DNA in the second generation after admixture depends on the aggregate recombination rate, quantified by $\bar{r}$. $\bar{r}$ can be measured from various kinds of data, including cytological data of crossover positions at meiosis I, sequence data from gametes, and linkage maps (Veller et al., 2019). $\bar{r}$ is influenced by several features of the recombination process: the number of chromosomes and heterogeneity in their size, the number of crossovers and their locations along the chromosomes, and the spatial relationships among crossovers (i.e., crossover interference). In most species, the dominant contribution to $\bar{r}$ is from independent assortment of chromosomes at meiosis (Crow, 1988; Veller et al., 2019). Therefore, the primary cause of variation in $\bar{r}$ across taxa is variation in chromosome number, with crossovers playing a secondary role.

These considerations explain a key feature of the trajectories in Figure 2A—that in the first few generations after admixture, *D. melanogaster* purges introgressed DNA much more rapidly than humans do. *D. melanogaster* has only two major autosomes, the independent assortment of which at meiosis contributes relatively little genetic shuffling. Furthermore, crossing over is absent in males. Veller et al. (2020) estimated a sex-averaged autosomal value of $\bar{r} = 0.305$ for *D. melanogaster*, substantially less than the theoretical maximum of 1/2. In contrast, humans have 22 autosomes, the independent assortment of which generates much genetic shuffling. Veller et al. (2019) estimated a sex-averaged autosomal value of $\bar{r} = 0.490$ in humans, close to the theoretical maximum. Substituting these values into Equation (10), we obtain predictions of how much introgressed DNA is retained by humans and by *Drosophila* in the third generation after admixture. These analytical predictions agree well with our simulations (Figure 2A).

We conclude that, relative to humans, *Drosophila* purges introgressed DNA more efficiently in the early generations after admixture because its aggregate recombination process generates substantially less genetic shuffling, owing primarily to its small karyotype and the absence of crossing over in males. Generalizing, this suggests that karyotypic variation across species could be a primary driver of species differences in the retention of introgressed DNA and thus the strength of the genomic signal of historical introgression (Supplementary Figure S5; Discussion).

Long-term purging of introgressed DNA

After the third generation post-admixture, the complex interaction of recombination and selection prevents tractable analytical calculations in the general case. Nevertheless, we can make analytical progress in understanding the impact of recombination on the purging of introgressed DNA in these

later generations by considering the special case where selection against introgressed ancestry is weak ($S \ll 1$). In this case, selection does not appreciably alter the ancestry variance in the population; therefore, we need only calculate the ancestry variance that would obtain if introgressed ancestry were neutral.

We begin by assuming that the initial introgressed fraction is small ($A_0 \ll 1$). In this case, with selection not appreciably limiting the number of descendants of the initial hybrids, the fraction of the population with introgressed ancestry grows exponentially: $A_t = A_0 \cdot 2^t$. Concomitantly, because selection does not appreciably deplete the overall fraction of introgressed DNA ($\mathbb{E}[Z_t] = A_0$), the average fraction among those individuals with introgressed ancestry declines exponentially: $\mathbb{E}[Z_t|\mathcal{A}] = 1/2^t$.

Because the initial introgressed fraction is small, each generation-t individual with introgressed ancestry descends from a single introgressing ancestor in generation 0, and, because selection has not appreciably affected the distribution of genetic ancestry among these generation-t individuals, the ancestry variance among them is equal to the variance of a random individual's genetic relatedness to one of its t-th degree great-grandparents. This quantity is calculated in Veller et al. (2020):

$$\mathrm{Var}(Z_t|\mathcal{A}) = \frac{1}{2^{t+1}}\left(\overline{(1-r)^{t-1}} - \frac{1}{2^{t-1}}\right). \tag{11}$$

Here, $\overline{(1-r)^{t-1}}$ is the average value of $(1-r_{ij})^{t-1}$ taken over all locus pairs (i,j), with r_{ij} the sex-averaged recombination rate between loci i and j [notice that this expression is not the same as $(1-\bar{r})^{t-1}$]. Equation (11), and the quantities below that depend on it, can be interpreted in terms of average covariances, or linkage disequilibria, between introgressed alleles (Veller et al., 2020), recognizing that these linkage disequilibria are simply remnants of the originally perfect linkage disequilibrium in which introgressed alleles initially appeared. Substituting Equation (11) into Equation (2), we find that the overall variance of genetic ancestry in generation t is

$$\begin{aligned}\mathrm{Var}(Z_t) &= A_t(1-A_t)(\mathbb{E}[Z_t|\mathcal{A}])^2 + A_t\mathrm{Var}(Z_t|\mathcal{A}) \\ &= A_0 2^t(1-A_0 2^t)\left(\frac{1}{2^t}\right)^2 + A_0 2^t \frac{1}{2^{t+1}}\left(\overline{(1-r)^{t-1}} - \frac{1}{2^{t-1}}\right) \\ &= \frac{1}{2}A_0\overline{(1-r)^{t-1}} - A_0^2 \qquad (12)\\ &\approx \frac{1}{2}A_0\overline{(1-r)^{t-1}}. \qquad (13)\end{aligned}$$

Substituting this result into Equation (1), we find that the rate of purging of introgressed DNA in generation t is

$$\Delta_t = \frac{\mathbb{E}[Z_t] - \mathbb{E}[Z_{t+1}]}{\mathbb{E}[Z_t]} \approx \frac{1}{2}S\overline{(1-r)^{t-1}}. \tag{14}$$

We now relax the assumption that the initial introgressed fraction is small, still assuming that introgressed ancestry is only weakly selected against. Observe that, in the special case above where the initial introgressed fraction is small, if an individual does have an introgressing ancestor (property $\mathcal{A}$) and has inherited DNA from that ancestor (probability p_t, calculated in Donnelly, 1983), then they have almost certainly inherited only one contiguous block of introgressed DNA if the number of generations since admixture is large. Equation (11) is therefore informative of the variance of introgressed block length, an observation that carries over to the more general case where the initial introgressed fraction need not be small. In Supplementary Section S2.1, we combine this observation with the assumption that, many generations after admixture, distinct introgressed blocks are approximately independently inherited to derive an expression for the ancestry variance when t is large and introgressed ancestry is neutral or weakly deleterious, which turns out to be the same as in the case where the initial introgressed fraction is small[1]:

$$\mathrm{Var}(Z_t) = \frac{1}{2}A_0\overline{(1-r)^{t-1}}. \tag{15}$$

Therefore, the rate of purging in generation t is

$$\Delta_t = \frac{\mathbb{E}[Z_t] - \mathbb{E}[Z_{t+1}]}{\mathbb{E}[Z_t]} \approx \frac{1}{2}S\overline{(1-r)^{t-1}}. \tag{16}$$

We can make several observations from Equations (15) and (16). First, because the terms $(1-r_{ij})^{t-1}$ that contribute to the average $\overline{(1-r)^{t-1}}$ decline exponentially over time (since $1-r_{ij} < 1$), the variance of introgressed ancestry declines over time (Equation (15)), and therefore so does the rate of purging of introgressed DNA (Equation (16); Figure 2). Second, Equation (16) is informative of the genomic scales of recombination that most influence the rate of purging at different timepoints after admixture. When t is small, terms $(1-r_{ij})^{t-1}$ from unlinked loci ($r_{ij} = 1/2$) are not much smaller than terms $(1-r_{ij})^{t-1}$ from linked loci ($r_{ij} < 1/2$). Since, in most species, there are many more unlinked locus pairs than linked locus pairs (Crow, 1988; Veller et al., 2019), in the early generations after admixture, unlinked locus pairs will usually be the dominant contributors to the average $\overline{(1-r)^{t-1}}$ and thus to the rate of purging of introgressed DNA. In contrast, when t is large, only terms $(1-r_{ij})^{t-1}$ from tightly linked loci ($r_{ij} \approx 0$) are not negligibly small, and so, many generations after admixture, only tightly linked loci contribute meaningfully to the average $\overline{(1-r)^{t-1}}$ and thus to the rate of purging. A corollary of these observations is that, in the early generations after admixture, species differences in the rate of purging will be driven largely by species differences in karyotype, which determine the fraction of locus pairs that are unlinked, while in later generations, species differences in the rate of purging will be driven by species differences in fine-scale recombination rates.

Intuition

Our main result above, that species with lower aggregate recombination rates purge introgressed DNA more efficiently, has a simple intuition. Initially, introgressed DNA appears in the recipient population in long linkage blocks, each the size of a haploid genome. Recombination influences the subsequent rate of purging of introgressed DNA because it breaks up these initial long blocks into smaller, variably sized blocks. A species with a low aggregate recombination rate will maintain introgressed DNA in longer and/or more variably sized

1 Because this expression can be interpreted as the neutral ancestry variance expected many generations after admixture in a well-mixed population, it can be compared against the measured variance in particular cases to test for selection or population structure or other forces that might cause the variance to deviate from this expectation.

blocks, and will therefore purge introgressed DNA more efficiently (Barton, 1983).

The average size of the blocks into which recombination chops the initial long blocks is determined predominantly by the number of chromosomes and crossovers, while the variability of block size is determined largely by heterogeneity in the sizes of chromosomes and the spatial arrangement of crossovers along chromosomes. To see this, note that a crossover anywhere along a block of introgressed DNA will break that block up into two smaller blocks with an average size of half the parent block. The position of the crossover along the block, however, will determine how variably sized the two descendant blocks are.

Thus, all else equal, a species with more chromosomes and/or more crossovers will purge introgressed DNA less efficiently, because it more rapidly breaks up the initial introgressed blocks into smaller blocks, which are less deleterious. And, all else equal, a species that situates crossovers near the tips of chromosomes will purge introgressed DNA more efficiently than a species with a uniform distribution of crossovers, because, in the former species, the resulting blocks of introgressed DNA will vary greatly in size. (Similarly, crossover interference—which causes even spacing of crossovers along chromosomes and therefore more even sizing of introgressed blocks—will reduce the efficiency with which introgressed DNA is purged.)

In Supplementary Section S2.2, we analyze a simple model that distinguishes the contributions of these two factors—the mean and variance of introgressed block length—to ancestry variance across individuals and thus the rate at which introgressed DNA is purged. The model applies best when sufficiently many generations have elapsed since the initial admixture pulse that different blocks of introgressed DNA can be assumed to have been inherited independently. Applying this model to the recombination process of *D. melanogaster*, we find that (a) the model accurately predicts the rate of purging of introgressed DNA (Supplementary Figure S6A), (b) the contribution of block length variance is greatest in the early generations after admixture, when crossover locations and variation in chromosome size have greatest effect (Supplementary Figure S6B), and (c) block number variance across individuals—which in this model is proportional to the average block length—becomes the most important contributor to ancestry variance in the long term (Supplementary Figure S6B). These results are again consistent with the view that aggregate recombination—as quantified by $\bar{r}$ and analogs—determines the short-run rate of purging of introgressed DNA, while fine-scale recombination rates determine the long-run rate of purging.

Variation in the rate of purging across the genome

The intuition above applies not only to different species but also to different regions within a given species' genome. In genomic regions with low aggregate rates of recombination, deleterious introgressed alleles will be maintained in larger and/or more variably sized blocks, and will therefore be purged more efficiently. Thus, for example, a chromosome that typically receives few crossovers per meiosis will, all else equal, purge introgressed DNA more rapidly than a chromosome that receives more crossovers (Supplementary Figure S7A). Similarly, if two chromosomes receive the same number of crossovers per meiosis, but crossovers tend to be terminally situated on the one chromosome and uniformly distributed along the other, then the chromosome with the more terminal distribution will purge introgressed DNA more rapidly, owing to its lower chromosome-specific value of $\bar{r}$ (Supplementary Figure S7B).

Below, we study two implications of recombination differences across the genome for the purging of introgressed DNA.

Sex chromosomes

Sex chromosomes often show unusual signatures of admixture. Most notably, X and Z chromosomes tend to retain less introgressed ancestry than autosomes (Martin & Jiggins, 2017). Several factors could account for the weaker signal of admixture on X/Z chromosomes, including their enrichment for alleles that reduce hybrid fitness (Charlesworth et al., 1987; Presgraves, 2008) and their hemizygous expression in the heterogametic sex (Turelli & Orr, 1995).

Here, we explore an additional factor that affects the retention of introgressed ancestry on sex chromosomes versus autosomes: recombination differences. In species with a degenerate sex-specific chromosome (the Y or W), recombination along most of the X/Z chromosome is typically restricted to the homogametic sex (XX females or ZZ males). This affects the average rate at which the X/Z chromosome recombines, relative to the autosomes.

Species with autosomal recombination in the heterogametic sex.

First, consider a species with autosomal recombination in both sexes (i.e., most heterogametic species). Unless the X/Z chromosome recombines at a substantially elevated rate in the homogametic sex, its lack of recombination in the heterogametic sex will ensure that, on average across generations, it experiences less recombination than the autosomes. Therefore, all else equal, we expect sex chromosomes to retain less introgressed DNA than autosomes in such species. This prediction was confirmed in simulations of our model augmented to include sex chromosomes (Figure 3A and C).

Species without autosomal recombination in the heterogametic sex.

In some species, the autosomes do not cross over in the heterogametic sex, with recombination along both the X/Z chromosome and the autosomes therefore limited to the homogametic sex (e.g., *Drosophila*, Lepidoptera). Assuming an even sex ratio, in a given generation of such a species, one half of the copies of each autosome are present in the homogametic sex and therefore have an opportunity to recombine; the other half are present in the heterogametic sex and do not recombine. In contrast, two thirds of the copies of the X/Z chromosome are present in the homogametic sex and can therefore recombine. So, in this case, the X/Z chromosome actually experiences more recombination than the autosomes. All else equal, we therefore expect the sex chromosomes to retain more introgressed DNA than the autosomes in such species. This prediction too was confirmed in simulations (Figure 3B and D).

Thus, recombination differences alone can generate differences in the retention of introgressed DNA between sex chromosomes and autosomes. Of course, recombination differences will not fully explain the ancestry differences actually observed between sex chromosomes and autosomes, since

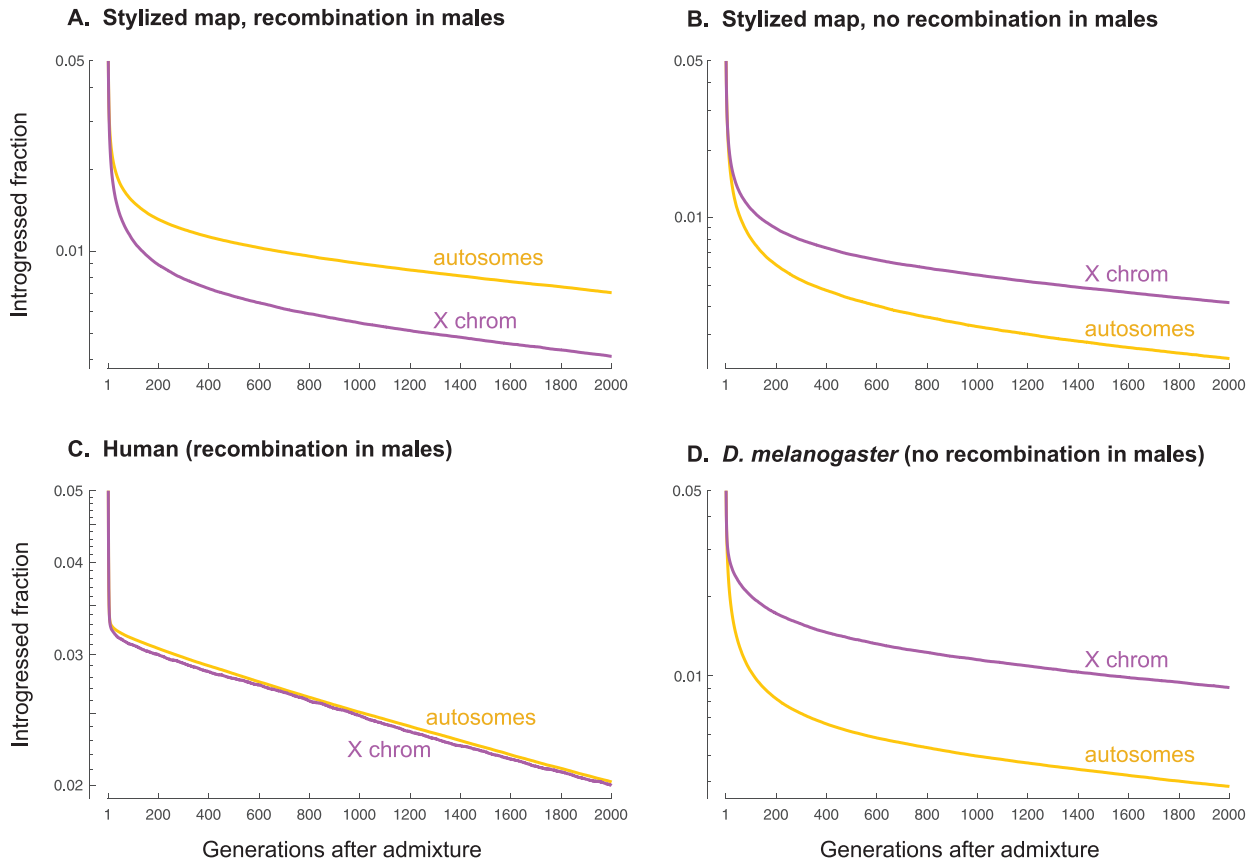

Figure 3. Recombination differences between sex chromosomes and autosomes generate differences in the rate of purging of introgressed DNA. (A,C) In species with autosomal recombination in both sexes, the lack of recombination along the sex chromosome in the heterogametic sex (here, the X chromosome in males) leads to a lower average rate of recombination than on autosomes. The sex chromosome therefore purges introgressed DNA more rapidly than the autosomes, all else equal. (B,D) In species without autosomal recombination in the heterogametic sex, the sex chromosome has a higher average rate of recombination than the autosomes, because 2/3 of its copies are in the (recombining) homogametic sex, compared to only 1/2 of autosomes. Therefore, in such cases, all else equal, the sex chromosome purges less introgressed DNA than the autosomes. The simulations here assume no dosage compensation in the heterogametic sex, such that, for a heterogametic individual with autosomal and sex-linked introgressed fractions p_A and p_X, the overall introgressed fraction is calculated as $(2L_A p_A + L_X p_X) / (2L_A + L_X)$, where L_A and L_X are the total haploid lengths of the autosomes and the X chromosome. Trajectories are averages across 100 replicate simulations. The stylized genomes in A and B involve one autosome, of equal size to the sex chromosome (500 deleterious loci each), with the sex chromosome and the autosome each receiving, on average, two crossovers in the homogametic sex, with the location of each crossover sampled independently and uniformly along its chromosome. In A, the autosomal recombination process is identical in the heterogametic and homogametic sexes.

additional features of sex chromosomes are known to be important in this regard (Coyne & Orr, 2004; Payseur et al., 2018). Indeed, in some cases, the predicted impact of recombination differences is opposite in direction to the observed ancestry disparity between sex chromosomes and autosomes. For example, in *Heliconius* butterflies, with female heterogamety and no autosomal recombination in females, Z chromosomes are especially depleted for introgressed ancestry (Van Belleghem et al., 2018; Martin et al., 2019; though see Zhang et al., 2016 for an interesting exception). This is in spite of the Z having a higher average recombination rate than the autosomes, per the argument above. In such cases, the factors that underlie the reduced retention of introgressed ancestry on the sex chromosome must be even stronger than previously thought, since they must work against the countervailing effect of recombination differences between the sex chromosomes and the autosomes.

A large X-contribution to hybrid fitness variance.

The "large X-effect" refers to the disproportionate contribution of the X (or Z) chromosome to the reduced fitness of early generation hybrids (Coyne, 1992), owing, among other possibilities, to its hemizygous expression in the heterogametic sex and its enrichment for alleles that are deleterious in hybrids (Presgraves, 2008). The results above suggest another way in which the X chromosome can play a disproportionate role in determining hybrid fitness: its outsized contribution to the variance of hybrid fitness, owing to its lower aggregate recombination rate than the autosomes (in species with autosomal recombination in the heterogametic sex). Here, we quantify this additional large X-effect in the case of F2 hybrids.

To do so, we compare two secenarios: one where the entire genome lies on a single autosome, and one where the entire genome lies on an X chromosome. The number of loci at which the introgressed allele is deleterious is large and equal among the two cases, as is the genomic spacing of these loci. The recombination process in females is the same in the two cases, and, in the autosomal case, is the same in males. Under these assumptions, we can calculate the variance in fitness among F2 hybrid offspring in the two cases. We focus on female F2s, since comparison of the fitness variance of male

F2s in the two cases is complicated by assumptions about X-chromosome dosage compensation.

In the autosomal case, the variance in introgressed ancestry among F2s is given by Equation (5) in Veller et al. (2020): $\mathrm{Var}(Z) = \frac{1}{4}\left(\frac{1}{2} - \bar{r}\right)$, where $\bar{r}$ is the average recombination fraction across pairs of loci on the autosome. In the case of the X chromosome, a female F2 receives a potentially recombinant X from her F1 mother and an unrecombined X from her (haploid) F1 father. With equal probability, the paternal X carries no introgressed alleles or only introgressed alleles; its ancestry variance is therefore 1/4. From Equation (1) in Veller et al. (2020), the ancestry variance of the maternal X is $\frac{1}{2}\left(\frac{1}{2} - \bar{r}\right)$. The overall ancestry variance among F2 females is therefore $\mathrm{Var}(Z) = \frac{1}{4}\left[\frac{1}{4} + \frac{1}{2}\left(\frac{1}{2} - \bar{r}\right)\right]$.

If the fitness reduction of a female F2 is proportional to her introgressed fraction, as in the primary model we have studied, then the fitness variance among F2 females is proportional to their ancestry variance. Therefore, the ratio of fitness variances in the X chromosome vs. the autosomal case is

$$\frac{\frac{1}{4}\left[\frac{1}{4} + \frac{1}{2}\left(\frac{1}{2} - \bar{r}\right)\right]}{\frac{1}{4}\left(\frac{1}{2} - \bar{r}\right)} = \frac{1 - \bar{r}}{1 - 2\bar{r}},$$

an increasing function of $\bar{r}$. Single-chromosome values of $\bar{r}$ typically lie between 1/5 and 1/3 (e.g., 1 and 2 Morgan chromosomes with uniform recombination and no crossover interference show values of 0.22 and 0.31 respectively; the mean sex-averaged value in humans is 0.27). By these values, and holding all else equal, the X overcontributes to fitness variance among F2s by ~30–100%, relative to its genomic length.

Note that, if there is no autosomal recombination in males, then the contributions to fitness variance among F2 females would be the same for the X chromosome and the autosome, since, in each case, an F2 female inherits a potentially recombined chromosome from her mother and an unrecombined chromosome from her father.

Of course, as for the long-term retention of introgressed ancestry on the X chromosome versus the autosomes, recombination differences alone will not fully explain the difference between the X's and the autosomes' contributions to hybrid fitness variance. If the X is enriched for alleles involved in DMIs, for example, then its contribution to hybrid fitness will be larger than expected based on its genomic length.

Correlations between regional recombination rate and introgressed ancestry

Several recent studies, encompassing a broad diversity of taxa, have identified positive within-genome correlations between local recombination rate and the amount of introgressed ancestry (Brandvain et al., 2014; Calfee et al., 2021; Edelman et al., 2019; Martin et al., 2019; Schumer et al., 2018). That is, regions of the genome that experience less recombination tend to retain less introgressed DNA. In trying to understand these correlations, it is important to note that, even if introgressed alleles are deleterious at many loci throughout the genome, the overwhelming majority of introgressed DNA is expected to be neutral. Therefore, the positive correlations that have been reported are driven by heterogeneity across the genome in the retention of neutral introgressed alleles (Schumer et al., 2018).

For a neutral introgressed allele ultimately to survive in the recipient population, it must recombine away from its flanking deleterious introgressed alleles before those deleterious alleles are eliminated by selection (Barton & Bengtsson, 1986; Bengtsson, 1985). Recombination affects this process in two ways (Schumer et al., 2018): (i) it affects the rate at which deleterious introgressed alleles are purged, and (ii) it affects the rate at which neutral introgressed alleles dissociate from their flanking deleterious alleles. These two effects of recombination constitute two distinct mechanisms by which local recombination rates across the genome come to be positively correlated with local ancestry.[2] The mechanisms are concordant: in regions of low recombination, deleterious alleles are purged more rapidly, and neutral alleles remain linked to their flanking deleterious alleles for a longer time.

The relative contributions of these two mechanisms to the overall correlation between recombination rate and introgressed ancestry can be quantitatively distinguished using the following decomposition, a graphical representation of which is given in Supplementary Figure S8:

$$\mathrm{Cov}(I, r) \approx \mathrm{Cov}(I_N, r) = \mathrm{Cov}(I_D, r) + \mathrm{Cov}(I_N - I_D, r). \quad (17)$$

Here, I is the overall introgressed fraction in a given genomic window, I_N and I_D are the introgressed fractions at neutral and deleterious loci respectively, and r is some measure of the recombination rate within the window (below, we use the average per-bp rate). The first term on the right hand side of Equation (17), $\mathrm{Cov}\left(I_D, r\right)$, captures the direct effect of recombination on the rate of purging of deleterious introgressed alleles—effect (i) above. The second term, $\mathrm{Cov}\left(I_N - I_D, r\right)$, captures the effect of recombination in unlinking neutral from deleterious introgressed alleles, thus decoupling their frequency trajectories—effect (ii) above. The relative importance of the "unlinking effect" is given by $\mathrm{Cov}\left(I_N - I_D, r\right)/\mathrm{Cov}\left(I, r\right)$.

To analyze the genomic correlation between recombination rate and introgressed ancestry in our simulations, we augmented our model to include neutral loci between the loci at which introgressed alleles are deleterious. We first employed a similar parameter configuration to our simulations above: introgressed alleles are deleterious at 1,000 loci and cause a 20% fitness reduction in F1 hybrids. Neutral allele frequencies were tracked at 10,000 evenly-spaced loci. Under this configuration, for the recombination processes of humans and *D. melanogaster*, and for various genomic window sizes, we made the following observations (Figure 4). (a) The positive correlation between recombination rate and introgressed ancestry builds up quickly, and once it has reached a maximum value, remains approximately stable for many generations thereafter (Figure 4A). (b) Under the *Drosophila* recombination process, the correlation is stronger when taken across larger windows of the genome (Figure 4A), while under the human recombination process, the strength of the correlation does not monotonically increase with window size. A positive effect of window size on correlation strength

2 To see that the two mechanisms are conceptually distinct, consider the case where every neutral introgressed allele is completely linked to a deleterious introgressed allele. Then no neutral allele ever dissociates from its flanking deleterious alleles, so that, eventually, all neutral alleles will be purged. Despite this lack of unlinking, in the long period before purging is completed, there will be a positive correlation between recombination rate and neutral introgressed ancestry, because a greater fraction of deleterious ancestry (and linked neutral ancestry) will have been purged in regions of low recombination.

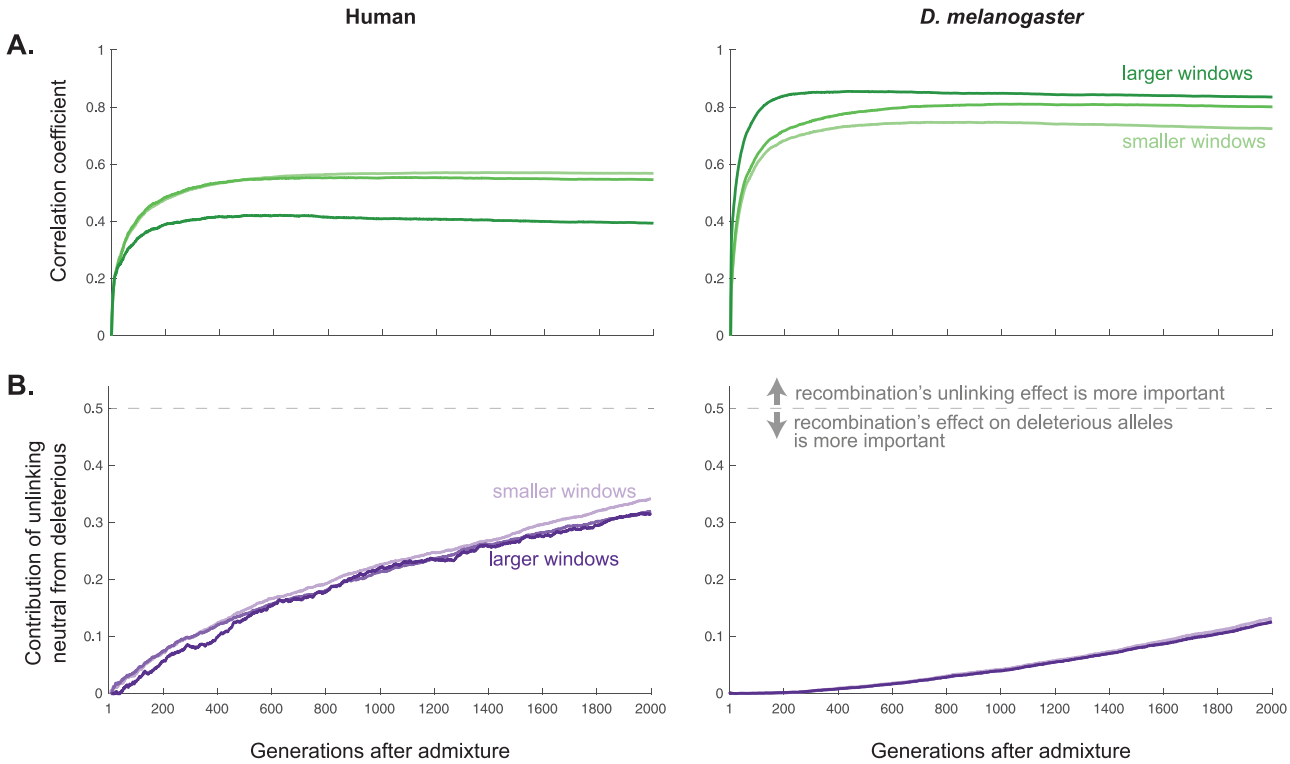

Figure 4. Genomic correlations between local recombination rate and introgressed ancestry. These simulations involve 10,000 evenly spaced loci at which introgressed alleles are neutral, and 1,000 loci at which introgressed alleles are deleterious. The genome is divided into windows of size 50, 100, or 150 neutral loci, with windows constrained to lie on the same chromosome. Each generation, we calculate, across all windows, the correlation coefficient between the average introgressed fraction per window (calculated at neutral loci) and the average recombination rate between adjacent neutral loci in the window. We then use Equation (17) to calculate the proportion of the correlation that is due to recombination's effect in unlinking neutral from deleterious introgressed alleles. Profiles are averaged across 100 replicate simulations. (A) The correlation between recombination rate and introgressed ancestry is set up quickly, and is stronger under the recombination process of *D. melanogaster* (right) than of humans (left). (B) For both recombination processes, recombination's effect in unlinking neutral from deleterious introgressed alleles (thus decoupling their frequency trajectories) contributes little to the correlation between recombination rate and introgressed ancestry. Instead, for many generations after the admixture pulse, the correlation is driven by differences across the genome in the efficiency with which deleterious introgressed alleles are purged. This conclusion is sensitive to the number of loci at which introgressed alleles are deleterious—see Supplementary Figure S9 and the Discussion.

has been observed previously (e.g., Schumer et al., 2018), and is presumably due to the per-window introgressed fraction changing in a less stochastic fashion over time in larger windows (and, in empirical settings, also because estimates of the introgressed fraction in larger windows are less noisy). Note that the window sizes we used were chosen such that each window included at least one deleterious locus, and are therefore substantially larger than those for which empirical recombination–introgression correlations have been measured. (c) The direct effect of recombination on the rate of purging of deleterious introgressed alleles is, at first, by far the more important mechanism in setting up the positive correlation between recombination rate and introgressed ancestry (Figure 4B). Recombination's effect in unlinking neutral from deleterious introgressed alleles gradually becomes more important, but for both humans and *Drosophila*, even 2,000 generations after admixture, it remains the minor mechanism (especially so for *Drosophila*).

The reason that recombination's unlinking effect is relatively unimportant under this configuration of parameters—with introgressed alleles deleterious at 1,000 loci—is straightforward. Take a high-recombination case like humans, with ~20 chromosomes and ~1 crossover per chromosome per gamete. In the model as configured above, there would be ~50 deleterious loci per chromosome. Consider a neutral introgressed allele situated halfway between its two flanking deleterious alleles. For the neutral allele to dissociate from both its flanking deleterious alleles requires, first, a recombination event anywhere between the two deleterious alleles (rate ~1/50 → waiting time ~50 generations), which unlinks the neutral allele from one of the deleterious alleles; and then, subsequently, a recombination event between the neutral allele and the remaining linked deleterious allele (rate ~1/100 → additional waiting time ~100 generations). It therefore takes, on average, about 150 generations following the admixture pulse for the neutral allele to dissociate from both its flanking deleterious alleles (and even longer for neutral alleles that are closer to one flanking deleterious allele than the other). As we have seen, by the time 150 generations have elapsed since admixture, most of the purging of introgressed ancestry has already occurred, and substantial ancestry differences have been set up between low-recombination and high-recombination regions. These differences across the genome, the imprint of which will persist for many generations, are therefore driven predominantly by the direct effect of recombination on the rate of purging of deleterious introgressed alleles.

The logic above is even more forceful in the case of low-recombination species like *D. melanogaster*: neutral alleles take even longer to dissociate from their flanking deleterious alleles, allowing even greater ancestry differences across the genome to be set up in the mean time by the direct effect of recombination on the rate of purging of deleterious alleles.

This explains why, in our simulations of the *D. melanogaster* recombination process, 2,000 generations after the initial admixture pulse, the unlinking effect still accounts for less than 20% of the overall correlation between recombination rate and introgressed ancestry.

These arguments clearly depend on the number of loci at which introgressed alleles are deleterious—if this number is smaller, so that deleterious loci are more diffusely scattered through the genome, then neutral introgressed alleles will recombine more rapidly away from their flanking deleterious alleles, on average, and so the unlinking effect of recombination will be more important. Supplementary Figure S9 displays the recombination–introgression correlation over time, and the fraction of this correlation due to recombination's effect in unlinking neutral from deleterious introgressed alleles, in the case where introgressed alleles are deleterious at only 100 loci in the genome. As in the case of 1,000 deleterious loci, the positive correlation between recombination rate and introgressed ancestry again builds up quickly and then slowly dissipates, under both the human and *D. melanogaster* recombination processes (Supplementary Figure S9A). However, the effect of recombination in unlinking neutral from deleterious introgressed alleles is much more important than in the case of 1,000 deleterious alleles. Under the human recombination process, it grows to become more important than recombination's effect on the rate of purging of deleterious alleles themselves after about 75 generations; under the *D. melanogaster* recombination process, this takes about 500 generations (Supplementary Figure S9B). Thus, although the unlinking effect of recombination is more important in this case than in the case with more deleterious loci, the unlinking effect is still less important than recombination's effect on the rate of purging of deleterious alleles for many generations after admixture, during which time most of the purging of introgressed DNA occurs.

A corollary of the results above is that, when introgressed alleles are deleterious at many loci throughout the genome, recombination's role in determining the genome-wide rate of purging of introgressed DNA—and thus its role in generating species differences in the retention of introgressed DNA—is driven largely by recombination's impact on the rate of purging of deleterious introgressed alleles, rather than its effect in unlinking neutral from deleterious introgressed alleles. Even though most introgressed DNA is expected to be neutral, the persistence of the neutral alleles' initial linkage to deleterious alleles causes them to be purged genome-wide at a rate that is, for many generations, almost identical to the rate of purging of the deleterious alleles themselves (Supplementary Figure S1).

When introgressed alleles are deleterious at only a handful of loci across the genome, however, the genome-wide rates of purging of neutral and deleterious introgressed ancestry diverge sooner after admixture (Supplementary Figure S4), with the consequence that a greater fraction of introgressed ancestry is retained overall (Moran et al., 2021). This observation underscores the importance of discerning the nature of the fitness effects underlying the purging of introgressed ancestry (see Discussion).

Discussion

Recent genomic evidence indicates that, following admixture, introgressed alleles are often deleterious at many loci throughout the recipient species' genome. Here, we have studied the influence of the aggregate recombination process on the efficiency with which selection purges introgressed DNA genome-wide. We have shown that species (and genomic regions) with low aggregate recombination rates—as quantified by $\bar{r}$ and analogous metrics—purge introgressed DNA more rapidly and more profoundly than species (and genomic regions) with high aggregate recombination rates. These effects are driven predominantly by recombination's effect on the rate of purging of deleterious introgressed alleles, rather than its effect in unlinking neutral introgressed alleles from their deleterious counterparts.

Empirical predictions

The simplest prediction emerging from our analysis is that species with fewer chromosomes should exhibit a weaker genomic signal of historic introgression. This is because (a) species differences in the retention of introgressed DNA are typically set up in the first few generations after hybridization, when most purging of introgressed DNA occurs (Figure 2); and (b) the rate of purging in these first few generations is governed by the aggregate recombination rate, which is dominated by the effect of independent assortment of chromosomes and thus by the number of chromosomes (Veller et al., 2019).

We expect that it will soon be possible to test this prediction, as quantitative estimates of genome-wide introgressed fractions become available for many taxa owing to rapid accumulation of sequence data and the recent development of multiple complementary methods to identify introgressed tracts within sequence data (Dagilis et al., 2022). A particularly promising clade is *Drosophila*, because, with its small baseline karyotype, the variation in chromosome number observed in the genus (Bracewell et al., 2019) corresponds to substantial variation in aggregate recombination rate. For example, *D. melanogaster* has only two major autosomes, the independent assortment of which contributes ~0.25 to $\bar{r}$. *D. subobscura*, in contrast, has four major autosomes, the independent assortment of which contributes ~0.38 to $\bar{r}$ (this calculation makes use of chromosome lengths reported by Bracewell et al., 2019). This karyotypic advantage, together with substantial progress in the analysis of introgression across the genus (e.g., Suvorov et al., 2022), suggests that *Drosophila* will likely be among the most informative clades in which to test the prediction that chromosome number correlates positively with introgressed ancestry.

Introgression selects for lower recombination

We have shown that aggregate recombination affects the rate at which deleterious introgressed DNA is purged. Since we know hybridization and subsequent genetic introgression to be common, this raises the converse question: does introgression select for modification of the recombination process?

Introgression's effect on modifiers of local recombination rates is straightforward. A modifier allele in the recipient species that reduces its local recombination rate prevents deleterious introgressed alleles from recombining onto its background, and is thus favored by selection. For example, a segregating inversion keeps together a haplotype of non-introgressed alleles, and is therefore favored over the alternative haplotype whose orientation is the same as that in the donor species and which therefore admits deleterious introgressed alleles by recombination (Kirkpatrick & Barton, 2006).

Our results also point to how selection acts on global modifiers of the recombination process in the face of deleterious introgression. A modifier allele that reduces the aggregate recombination rate (i.e., $\bar{r}$ and analogous metrics) increases the variance among its descendants in how much introgressed DNA they carry (Veller et al., 2020). This allows selection to purge introgressed DNA more efficiently among descendants of the modifier allele, causing the modifier allele to end up in fitter genotypes and thus to be positively selected. It is interesting to contrast this logic with that underlying the standard result that, in a finite, well-mixed population with multiplicative selection across loci, a global modifier that increases the recombination rate is favored by selection because it increases fitness variance among its descendants (Barton, 1995; Barton & Otto, 2005; Burt, 2000). The reason is that, in this case, the interaction of selection and random drift tends to generate *negative* linkage disequilibra between deleterious alleles (Barton & Otto, 2005). These negative LDs reduce fitness variance; their destruction by recombination therefore increases fitness variance, and so selection favors higher recombination rates. In our case, in contrast, the deleterious alleles are introgressed into the recipient population in perfect *positive* LD. These positive LDs increase fitness variance; their destruction by recombination therefore decreases fitness variance, and so higher recombination rates are disfavored by selection.

Therefore, selection against introgressed ancestry can generate selection on both local and global modifiers to reduce the recombination rate. Local modifiers of recombination include structural rearrangements (Kirkpatrick, 2010), alterations to the binding sites of recombination-specifying proteins (Grey et al., 2018; Paigen & Petkov, 2018), and mutations that affect local chromatin structure in meiotic prophase (e.g., Stack et al., 2017). The simplest global modification of the aggregate recombination rate is a change in chromosome number (Veller et al., 2019), but introgression is not generally expected to select for reduced chromosome number owing to fertility problems in karyotype-heterozygous hybrids (White, 1978). Therefore, introgression is expected to select for global modification of the recombination process predominantly via modifiers of the number and spatial arrangement of crossovers. Our expanding knowledge of the molecular biology of meiosis and recombination (reviewed in Hunter, 2015; Zickler & Kleckner, 2015) suggests global modifiers of this form to be very common: they include mutations to key meiosis proteins, such as those that determine the lengths of chromosome axes in meiotic prophase (e.g., Hong et al., 2019; Novak et al., 2008), those that control the interference process along chromosome axes (e.g., Zhang et al., 2014), and those that globally specify recombination hotspots (Grey et al., 2018; Paigen & Petkov, 2018).

Frequent introgression is therefore expected to shape genetic variation in these factors toward reducing both local and global recombination. In this way, selection for reduced recombination acts as an indirect form of reinforcement, causing post-zygotic selection against introgressed DNA to be more efficient, and thus strengthening the barrier to gene flow between species.

Note that introgression can also select for reductions in recombination among loci underlying traits relevant to hybrid fitness without alteration of the recombination process itself, via turnover of and changes in allelic frequencies at these loci (Yeaman & Whitlock, 2011).

The influence of demography on recombination's role in selection against introgressed DNA

We have studied a very simple demographic scenario, where there is a single, immediate pulse of admixture, and mating is thereafter random in the admixed population.

In reality, mating is unlikely to be random in early admixed populations. Instead, one might expect mates often to assort based partially on ancestry. Elsewhere, we have studied the consequences of ancestry-based assortative mating for selection against introgressed DNA, showing that assortative mating—by bundling introgressed ancestry together—increases ancestry variance in the population and thus allows selection to purge introgressed DNA more efficiently (Muralidhar et al., 2022). There, we discussed recombination's role in accelerating this "bundling effect" of assortative mating, which partially counteracts the effect of recombination that we have studied here, to reduce the rate at which introgressed ancestry is purged. Additionally, assortative mating might mute recombination's role in mediating the rate at which introgressed ancestry is purged, since, by bringing together like-with-like ancestry in individuals and thus decreasing the number of locus pairs at which individuals are doubly ancestry-heterozygous, assortative mating reduces "effective" recombination in these individuals (cf. Nordborg, 2000).

The "instant pulse" model of admixture that we have studied is likely also unrealistic. Even in cases where admixture is not perpetual as in stable clines, pulses of admixture will usually persist across multiple generations—or occur at separate points in time (e.g., Vilgalys et al., 2022). In such cases, early- and later-generation hybrids will interbreed, creating complex patterns of admixture that, in the absence of further simplifying assumptions such as that admixture pulses are very small (Barton, 1983), are not amenable to the kinds of simple calculations we have developed in this paper.

Finally, we have chosen the overall deleterious effect of introgressed ancestry to match that estimated for Neanderthal–human introgression (with F1 hybrids suffering a ~20% fitness reduction; Harris & Nielsen, 2016; Juric et al., 2016). While the rate at which introgressed ancestry is purged obviously depends on how deleterious it is, we do not expect recombination's role in this process to be particularly sensitive to this parameter.

The number of loci at which introgressed alleles are deleterious, and the nature of selection against introgressed ancestry

Recent genomic work has shown that, in many cases, introgressed alleles are deleterious at a great number of loci scattered throughout the genome, with alleles at these loci each having a small individual effect on fitness but together having a large combined effect such that early-generation hybrids—who carry many introgressed alleles—suffer substantial fitness reductions (e.g., Aeschbacher et al., 2017; Juric et al., 2016). Our calculations have focused on the role of recombination in mediating selection against introgressed ancestry in this "polygenic" case. However, in some systems, the reduced fitness of hybrids is predominantly due to very strong fitness effects at only a handful of loci (e.g., Powell et al., 2020; Presgraves, 2003).

A conclusion that should be particularly sensitive to the number of deleterious loci is the unimportance of recombination's effect in unlinking neutral from deleterious introgressed alleles, relative to its effect on the rate of purging of the

deleterious alleles themselves, in explaining both the overall rate of purging of introgressed alleles (Supplementary Figure S4) and genome-wide correlations between recombination rate and introgressed ancestry (Figure 4, Supplementary Figure S9). With fewer deleterious loci throughout the genome, the average genetic distance of a neutral locus to its nearest flanking deleterious loci is larger, and so the unlinking of neutral from deleterious introgressed alleles will be more rapid.

To illustrate the importance of the number of deleterious loci for genome-wide correlations between recombination rate and introgressed ancestry, we have compared cases with 1,000 loci (Figure 4) and 100 loci (Supplementary Figure S9) at which the introgressed alleles are deleterious. A technical difficulty in reducing the number of deleterious loci further than this in our simulations is that Equation (17), which decomposes the genome-wide correlation between recombination rate and introgressed ancestry into contributions from recombination's unlinking effect and its effect on the rate of purging of deleterious alleles, requires that there be at least one deleterious locus in each of the genomic windows across which the correlation is measured. With very few deleterious loci, these genomic windows would need to be unreasonably large—e.g., spanning chromosomes—for the decomposition to apply. Nonetheless, the increased importance of recombination's unlinking effect when there are fewer deleterious loci can be observed in the greater and more rapid divergence of the overall introgressed fractions at neutral and deleterious loci (Supplementary Figure S4). A corollary is that, holding the overall fitness reduction of early-generation hybrids constant, a greater fraction of introgressed ancestry persists when this overall deleterious effect is spread across fewer loci (Supplementary Figure S4; see also Figure 4 in Moran et al., 2021).

The number of loci at which introgressed alleles are deleterious seems often to depend on the nature of selection against introgressed ancestry (Moran et al., 2021). If this selection is due to higher genetic load in the donor species (owing to its lower effective poulation sizes), then the deleterious effect of introgressed ancestry is expected to be spread across a large number of loci at which weakly deleterious mutations accumulated in the donor species. Such a scenario has been argued to be relevant for Neanderthal–human introgression, and genomic patterns consistent with the same scenario have been found in other cases as well (e.g., Aeschbacher et al., 2017). In contrast, when introgressed ancestry is deleterious because of negative epistatic interactions between donor- and recipient-species alleles (DMIs), then individual selective effects are often strong, with individual incompatibilities substantially reducing hybrid fitness (e.g., Powell et al., 2020; Presgraves, 2003).

We have focused primarily on the role of recombination in selection against introgressed ancestry under a load-like model, where introgressed alleles are additively deleterious at many loci throughout the genome. This situation might also be expected to pertain if introgressed alleles are broadly maladapted to the recipient species' ecology. The role of recombination in the alternative scenario, where introgressed alleles are deleterious owing to their participation in DMIs, is complicated by two factors. First, two-locus DMIs fall into three epistatic dominance classes, depending on whether an individual must be homozygous for the incompatible allele at both, one, or neither of the two loci to suffer the deleterious effect of the DMI (type 2, 1, and 0 DMIs, respectively; Turelli & Orr, 2000). The effect of recombination in "uncovering" these three kinds of DMI varies: two recombination events are required for a type-2 DMI to be expressed, while type-0 DMIs are expressed already in F1 hybrids, with no recombination between the two loci required. Second, the loci involved in each pairwise DMI might not be independently chosen from across the genome, as we have assumed for simplicity in our DMI simulations (Supplementary Figure S3). For example, if ancestry mismatches between genes and their promoters lead to suboptimal gene expression in hybrids, then DMIs might disproportionately involve tightly linked locus pairs (Moran et al., 2021). In such cases, the effect of recombination on selection against incompatible alleles at these loci, and against the minor-parent ancestry more generally, will be influenced by a combination of local/fine-scale recombination rates, which govern how rapidly incompatible alleles recombine into the same genome, and aggregate recombination, which governs how efficiently selection purges collections of incompatible alleles that find themselves in these configurations.

Supplementary material

Supplementary material is available online at *Evolution* (https://academic.oup.com/evolut/qpad021)

Data availability

Simulation code is available at https://github.com/cveller/RecombinationAndIntrogression

Author contributions

C.V. conceived study. C.V., N.B.E., P.M., and M.A.N. performed modelling and simulations, and wrote the paper.

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

Acknowledgments

We are grateful to Graham Coop for many helpful discussions and suggestions on the content and framing of this paper, and to Yaniv Brandvain, Erin Calfee, Jeff Groh, Jim Mallet, Molly Schumer, and Sivan Yair for helpful comments. C.V. is supported by a Branco Weiss fellowship. NBE is supported by a G. Evelyn Hutchinson Environmental Postdoctoral Fellowship. P.M. is supported by a Center for Population Biology postdoctoral fellowship and an NSF postdoctoral fellowship. This work was supported in part by the National Institute of General Medical Sciences of the National Institutes of Health (grants NIH R01 GM108779 and R35 GM136290 to G. Coop).

References

Aeschbacher, S., Selby, J. P., Willis, J. H., & Coop, G. (2017). Population-genomic inference of the strength and timing of selection against gene flow. *Proceedings of the National Academy of Sciences*, 114(27), 7061–7066. https://doi.org/10.1073/pnas.1616755114

Barton, N. H. (1983). Multilocus clines. *Evolution*, 37(3), 454–471. https://doi.org/10.1111/j.1558-5646.1983.tb05563.x

Barton, N. H. (1995). A general model for the evolution of recombination. *Genetics Research*, 65(2), 123–144. https://doi.org/10.1017/S0016672300033140

Barton, N. H., & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, 57(3), 357–376. https://doi.org/10.1038/hdy.1986.135

Barton, N. H., & Otto, S. P. (2005). Evolution of recombination due to random drift. *Genetics*, 169(4): 2353–2370. https://doi.org/10.1534/genetics.104.032821

Bengtsson, B. O. (1985). The flow of genes through a genetic barrier. In Greenwood, P. J., Harvey, P. H., & Slatkin, M. (Eds.), *Evolution: Essays in honor of John Maynard Smith* (pp. 31–42). Cambridge University Press.

Bracewell, R., Chatla, K., Nalley, M. J., & Bachtrog, D. (2019). Dynamic turnover of centromeres drives karyotype evolution in *Drosophila*. *Elife*, 8, e49002. https://doi.org/10.7554/eLife.49002

Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., & Sweigart, A. L. (2014). Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genetics*, 10(6), e1004410. https://doi.org/10.1371/journal.pgen.1004410

Burt, A. (2000). Perspective: Sex, recombination, and the efficacy of selection—was Weismann right? *Evolution*, 54(2), 337–351. https://doi.org/10.1111/j.0014-3820.2000.tb00038.x

Calfee, E., Gates, D., Lorant, A., Perkins, M. T., Coop, G., & Ross-Ibarra, J. (2021). Selective sorting of ancestral introgression in maize and teosinte along an elevational cline. *PLoS Genetics*, 17(10), e1009810. https://doi.org/10.1371/journal.pgen.1009810

Charlesworth, B., Coyne, J. A., & Barton, N. H. (1987). The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist*, 130(1), 113–146. https://doi.org/10.1086/284701

Comeron, J. M., Ratnappan, R., & Bailin, S. (2012). The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genetics*, 8(10), e1002905. https://doi.org/10.1371/journal.pgen.1002905

Coyne, J. A. (1992). Genetics and speciation. *Nature*, 355(6360), 511–515. https://doi.org/10.1038/355511a0

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer.

Crow, J. F. (1988). The importance of recombination. In Michod, R. E. & Levin, B. R. (Eds.), *The evolution of sex: An examination of current ideas* (pp. 56–73). Sinauer.

Dagilis, A. J., Peede, D., Coughlan, J. M., Jofre, G. I., D'Agostino, E. R., Mavengere, H., Tate, A. D., & Matute, D. R. (2022). A need for standardized reporting of introgression: Insights from studies across eukaryotes. *Evolution Letters*, 6(5), 344–357. https://doi.org/10.1002/evl3.294

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

Donnelly, K. P. (1983). The probability that related individuals share some section of genome identical by descent. *Theoretical Population Biology*, 23(1), 34–63. https://doi.org/10.1016/0040-5809(83)90004-7

Duranton, M., & Pool, J. E. (2022). Interactions between natural selection and recombination shape the genomic landscape of introgression. *Molecular Biology and Evolution*, 39(7), msac122. https://doi.org/10.1093/molbev/msac122

Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R. B., García-Accinelli, G., Van Belleghem, S. M., Patterson, N., Neafsey, D. E., Challis, R., Kumar, S., Moreira, G. R. P., Salazar, C., Chouteau, M., Counterman, B. A., Papa, R., Blaxter, M., Reed, R. D., Dasmahapatra, K. K., Kronforst, M., Joron, M., Jiggins, C. D., McMillan, W. O., Palma, F. D., Blumberg, A. J., Wakeley, J., Jaffe, D., Mallet J. (2019). Genomic architecture and introgression shape a butterfly radiation. *Science*, 366, 594–599. https://doi.org/10.1126/science.aaw2090

Edelman, N. B., & Mallet, J. (2021). The prevalence and adaptive impact of introgression. *Annual Review of Genetics*, 55, 265–283. https://doi.org/10.1146/annurev-genet-021821-020805

Grey, C., Baudat, F., & de Massy, B. (2018). PRDM9, a driver of the genetic map. *PLoS Genetics*, 14(8), e1007479. https://doi.org/10.1371/journal.pgen.1007479

Haller, B. C., & Messer, P. W. (2017). SLiM 2: Flexible, interactive forward genetic simulations. *Molecular Biology and Evolution*, 34(1), 230–240. https://doi.org/10.1093/molbev/msw211

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. https://doi.org/10.1093/molbev/msy228

Harris, K., & Nielsen, R. (2016). The genetic cost of Neanderthal introgression. *Genetics*, 203(2), 881–891. https://doi.org/10.1534/genetics.116.186890

Hong, S., Joo, J. H., Yun, H., Kleckner, N., & Kim, K. P. (2019). Recruitment of Rec8, Pds5 and Rad61/Wapl to meiotic homolog pairing, recombination, axis formation and S-phase. *Nucleic Acids Research*, 47(22), 11691–11708. https://doi.org/10.1093/nar/gkz903

Hunter, N. (2015). Meiotic recombination: The essence of heredity. *Cold Spring Harbor Perspectives in Biology*, 7(12), a016618. https://doi.org/10.1101/cshperspect.a016618

Juric, I., Aeschbacher, S., & Coop, G. (2016). The strength of selection against Neanderthal introgression. *PLoS Genetics*, 12(11), e1006340. https://doi.org/10.1371/journal.pgen.1006340

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biology*, 8(9), e1000501. https://doi.org/10.1371/journal.pbio.1000501

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, 173(1), 419–434. https://doi.org/10.1534/genetics.105.047985

Kong, A., Thorleifsson, G., Gudbjartsson, D. F., Masson, G., Sigurdsson, A., Jonasdottir, A., Walters, G. B., Jonasdottir, A., Gylfason, A., Kristinsson, K. T., Gudjonsson, S. A., Frigge, M. L., Helgason, A., Thorsteinsdottir, U., & Stefansson, K. (2010). Fine-scale recombination rate differences between sexes, populations and individuals. *Nature*, 467(7319), 1099–1103. https://doi.org/10.1038/nature09525

Lenormand, T., & Dutheil, J. (2005). Recombination difference between sexes: a role for haploid selection. *PLoS Biology*, 3(3), e63. https://doi.org/10.1371/journal.pbio.0030063

Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biology*, 17(2), e2006288. https://doi.org/10.1371/journal.pbio.2006288

Martin, S. H., & Jiggins, C. D. (2017). Interpreting the genomic landscape of introgression. *Current Opinion in Genetics & Development*, 47, 69–74. https://doi.org/10.1016/j.gde.2017.08.007

Messer, P. W. (2013). SLiM: Simulating evolution with selection and linkage. *Genetics*, 194(4), 1037–1039. https://doi.org/10.1534/genetics.113.152181

Moran, B. M., Payne, C., Langdon, Q., Powell, D. L., Brandvain, Y., & Schumer, M. (2021). The genomic consequences of hybridization. *Elife*, 10, e69016. https://doi.org/10.7554/eLife.69016

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

Muralidhar, P., Coop, G., & Veller, C. (2022). Assortative mating enhances postzygotic barriers to gene flow via ancestry bundling. *Proceedings of the National Academy of Sciences*, 119(30), e2122179119. https://doi.org/10.1073/pnas.2122179119

Nordborg, M. (2000). Linkage disequilibrium, gene trees and selfing: An ancestral recombination graph with partial self-fertilization. *Genetics*, 154(2), 923–929. https://doi.org/10.1093/genetics/154.2.923

Novak, I., Wang, H., Revenkova, E., Jessberger, R., Scherthan, H., & Höög, C. (2008). Cohesin Smc1β determines meiotic chromatin axis loop organization. *Journal of Cell Biology*, 180(1), 83–90. https://doi.org/ 10.1083/jcb.200706136

Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, 139(4), 1805–1813. https://doi.org/10.1093/genetics/139.4.1805

Paigen, K., & Petkov, P. M. (2018). PRDM9 and its role in genetic recombination. *Trends in Genetics*, 34(4), 291–300. https://doi.org/10.1016/j.tig.2017.12.017

Payseur, B. A., Presgraves, D. C., & Filatov, D. A. (2018). Introduction: Sex chromosomes and speciation. *Molecular Ecology*, 27(19), 3745–3748. https://doi.org/10.1111/mec.14828

Petr, M., Pääbo, S., Kelso, J., & Vernot, B. (2019). Limits of long-term selection against Neandertal introgression. *Proceedings of the*

National Academy of Sciences, 116(5), 1639–1644. https://doi.org/10.1073/pnas.1814338116

Powell, D. L., García-Olazábal, M., Keegan, M., Reilly, P., Du, K., Díaz-Loyo, A. P., Banerjee, S., Blakkan, D., Reich, D., Andolfatto, P., Rosenthal, G. G., Schartl, M., & Schumer, M. (2020). Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. *Science*, 368(6492), 731–736. https://doi.org/10.1126/science.aba5216

Presgraves, D. C. (2003). A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. *Genetics*, 163(3), 955–972. https://doi.org/10.1093/genetics/163.3.955

Presgraves, D. C. (2008). Sex chromosomes and speciation in *Drosophila*. *Trends in Genetics*, 24(7), 336–343. https://doi.org/10.1016/j.tig.2008.04.007

Sardell, J. M., & Kirkpatrick, M. (2020). Sex differences in the recombination landscape. *American Naturalist*, 195(2): 361–379. https://doi.org/10.1086/704943

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, 323(5915), 737–741. https://doi.org/10.1126/science.1160006

Schumer, M., Cui, R., Powell, D. L., Dresner, R., Rosenthal, G. G., & Andolfatto, P. (2014). High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. *Elife*, 3, e02535. https://doi.org/10.7554/eLife.02535

Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Rosenthal, G. G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, 360(6389), 656–660. https://doi.org/10.1126/science.aar3684

Stack, S. M., Shearer, L. A., Lohmiller, L., & Anderson, L. K. (2017). Meiotic crossing over in maize knob heterochromatin. *Genetics*, 205(3), 1101–1112. https://doi.org/10.1534/genetics.116.196089

Steinrücken, M., Spence, J. P., Kamm, J. A., Wieczorek, E., & Song, Y. S. (2018). Model-based detection and analysis of introgressed Neanderthal ancestry in modern humans. *Molecular Ecology*, 27(19), 3873–3888. https://doi.org/10.1111/mec.14565

Suvorov, A., Kim, B. Y., Wang, J., Armstrong, E. E., Peede, D., D'agostino, E. R., Price, D. K., Waddell, P. J., Lang, M., Courtier-Orgogozo, V., David, J. R., Petrov, D., Matute, D. R., Schrider, D. R., & Comeault, A. A. (2022). Widespread introgression across a phylogeny of 155 *Drosophila* genomes. *Current Biology*, 32(1), 111–123.e5. https://doi.org/10.1016/j.cub.2021.10.052

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, 3(2), 170–177. https://doi.org/10.1038/s41559-018-0777-y

Turelli, M., & Orr, H. A. (1995). The dominance theory of Haldane's rule. *Genetics*, 140(1), 389–402. https://doi.org/10.1093/genetics/140.1.389

Turelli, M., & Orr, H. A. (2000). Dominance, epistasis and the genetics of postzygotic isolation. *Genetics*, 154(4), 1663–1679. https://doi.org/10.1093/genetics/154.4.1663

Van Belleghem, S. M., Baquero, M., Papa, R., Salazar, C., McMillan, W. O., Counterman, B. A., Jiggins, C. D., & Martin, S. H. (2018). Patterns of Z chromosome divergence among *Heliconius* species highlight the importance of historical demography. *Molecular Ecology*, 27(19), 3852–3872. https://doi.org/10.1111/mec.14560

Veller, C., Edelman, N. B., Muralidhar, P., & Nowak, M. A. (2020). Variation in genetic relatedness is determined by the aggregate recombination process. *Genetics*, 216(4), 985–994. https://doi.org/10.1534/genetics.120.303680

Veller, C., Kleckner, N., & Nowak, M. A. (2019). A rigorous measure of genome-wide genetic shuffling that takes into account crossover positions and Mendel's second law. *Proceedings of the National Academy of Sciences*, 116(5), 1659–1668. https://doi.org/10.1073/pnas.1817482116

Vilgalys, T. P., Fogel, A. S., Anderson, J. A., Mututua, R. S., Warutere, J. K., Siodi, I. L., Kim, S. Y., Voyles, T. N., Robinson, J. A., Wall, J. D., Archie, E. A., Alberts, S. C., & Tung, J. (2022). Selection against admixture and gene regulatory divergence in a long-term primate field study. *Science*, 377(6606), 635–641. https://doi.org/10.1126/science.abm4917

White, M. J. D. (1978). *Modes of speciation*. W. H. Freeman & Co.

Yamaguchi, R., Wiley, B., & Otto, S. P. (2022). The phoenix hypothesis of speciation. *Proceedings of the Royal Society B*, 289(1987), 20221186.

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration–selection balance. *Evolution; International Journal of Organic Evolution*, 65(7), 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

Zhang, L., Wang, S., Yin, S., Hong, S., Kim, K. P., & Kleckner, N. (2014). Topoisomerase II mediates meiotic crossover interference. *Nature*, 511(7511), 551–556. https://doi.org/10.1038/nature13442

Zhang, W., Dasmahapatra, K. K., Mallet, J., Moreira, G. R., & Kronforst, M. R. (2016). Genome-wide introgression among distantly related *Heliconius* butterfly species. *Genome Biology*, 17(1), 1–15.

Zickler, D., & Kleckner, N. (2015). Recombination, pairing, and synapsis of homologs during meiosis. *Cold Spring Harbor Perspectives in Biology*, 7(6), a016626. https://doi.org/10.1101/cshperspect.a016626