Sex-Specific Incompatibility Generates Locus-Specific Rates of Introgression between Species

Diana Fusco*
Marcy K. Uyenoyama§

*Computational Biology and Bioinformatics Program
Box 90090
Duke University
Durham, NC 27708-0090
USA
diana.fusco@duke.edu

§Department of Biology
Box 90338
Duke University
Durham, NC 27708-0338
USA
marcy@duke.edu
Running head: Locus-specific introgression

Keywords: population structure, locus-specific divergence, introgression, incompatibility selection

Address for correspondence:

Marcy K. Uyenoyama
Department of Biology
Box 90338
Duke University
Durham, NC 27708-0338
USA

Tel: 919-660-7350
Fax: 919-660-7293
e-mail: marcy@duke.edu
ABSTRACT

Disruption of interactions among ensembles of epistatic loci has been shown to contribute to reproductive isolation among various animal and plant species. Under the Bateson-Dobzhansky-Muller model, such interspecific incompatibility arises as a by-product of genetic divergence in each species, and the Orr-Turelli model indicates that the number of loci involved in incompatible interactions may “snowball” over time. We address the combined effect of multiple incompatibility loci on the rate of introgression at neutral marker loci across the genome. Our analysis extends previous work by accommodating sex-specificity: differences between the sexes in the expression of incompatibility, in rates of crossing-over between neutral markers and incompatibility loci, and in transmission of markers or incompatibility factors. We show that the evolutionary process at neutral markers in a genome subject to incompatibility selection is well-approximated by a purely neutral process with migration rates appropriately scaled to reflect the influence of selection targeted to incompatibility factors. We confirm that in the absence of sex-specificity and functional epistasis among incompatibility factors, the barrier to introgression induced by multiple incompatibility factors corresponds to the product of the barriers induced by the factors individually. A new finding is that barriers to introgression due to sex-specific incompatibility depart in general from multiplicativity. Our partitioning of variation in relative reproductive rate suggests that such departures derive from associations between sex and incompatibility and between sex and neutral markers. Concordant sex-specific incompatibility (for example, greater impairment of male hybrids or longer map lengths in females) induces lower barriers (higher rates of introgression) than expected under multiplicativity, and discordant sex-specific incompatibility induces higher barriers.
Gene flow may persist between incipient species for extended periods, even in the face of severe disruptions in the fertility, viability, or behavior of hybrids or their descendants. For example, introgression appears to persist between *Drosophila pseudoobscura* and *D. persimilis*, in spite of the rise, about 850,000 years ago (?), of a second-chromosome inversion associated with multiple mechanisms of reproductive isolation. In these species, at least three inverted chromosomal regions appear to contribute to hybrid incompatibility (?), and in general, the number of incompatible interactions among loci is expected to “snowball” as divergence time increases (?). Here we address the effects of multiple incompatibility loci on the rate of introgression at neutral marker loci throughout the genome. This study departs from previous work (e.g., ?) by addressing sex-specificity in expression of disruptions associated with hybridization, in rates of crossing-over between incompatibility factors and marker loci, and in linkage to sex chromosomes or other regions transmitted in a sex-specific manner.

**Nonuniform divergence among genomic regions**

Among the most robust patterns that have emerged from genome-scale surveys is the nonuniformity across genomic regions of divergence among populations. Geography constitutes a major explanatory variable for global patterns of variation in human genomes (?). Humans show striking locus-specific $F_{ST}$, the classical index of population structure (?), with many regions showing extreme values (beyond three empirical standard deviations from the mean for a chromosome; ?). This pattern is observed in a wide variety of non-human species as well (e.g., ??).

The widely-used IM program (?) uses observations of nucleotide variation at multiple, independent loci to infer introgression rates and other demographic parameters assumed to be common to all genomic regions. ? applied their method separately to each of 14 markers in *Drosophila pseudoobscura* and *D. persimilis*, obtaining estimates of introgression rates that varied over orders of magnitude: for example, a scaled rate of flow from *D. persimilis* to *D. pseudoobscura* of $M_2 = 0.09$ for *per* (X-linked) and $M_2 = 4.25$ for *Adh* (autosomal).
Earlier work (?) had also indicated substantial differences between loci in rates of gene exchange.

Genetic analyses across interspecific hybridization zones provided the first demonstrations of the differential permeability of reproductive barriers across the genome (?). Incompatibility is expected to impede introgression throughout the genome, but more intensely in regions harboring incompatibility factors (?). A number of studies have reported lower rates of introgression in regions shown to contain factors contributing to interspecific incompatibility. Using markers located on all 17 chromosomes of sunflowers *Helianthus annuus* and *H. petiolaris*, ? found significantly higher divergence between the species in 26 regions, of which 16 were associated with reduced pollen fertility in hybrids. A more recent study of these species found numerous amino acid differences fixed between the species near breakpoints of inverted regions, where quantitative trait loci for hybrid pollen sterility tend to cluster (?). Lower introgression rates in the vicinity of known incompatibility factors have been detected between house mice *Mus musculus musculus* and *M. m. domesticus* (?), among host races of pea aphids (?), and between *Drosophila pseudoobscura* and *D. persimilis* (?).

One interpretation of nonuniform divergence is that genomic regions that show extreme variation among demes have experienced recent selective sweeps restricted to certain geographical localities (reviewed by ??). This view holds that demographic structure should affect the pattern of variation in a uniform manner, with locus-specificity indicative of selection (e.g., ?). Beyond sweeps, distinct species or local subpopulations of the same species may undergo genetic divergence through various processes, including adaptation to local ecological conditions (e.g. ?).

Here, we use “incompatibility selection” to describe selective regimes characterized by neutrality within a specific demographic unit and negative selection elsewhere. Under this process, an incompatibility locus is generally monomorphic within populations for selectively equivalent forms of the wild-type allele, except for occasional introductions of deleterious foreign alleles, which are fully functional in their home deme. While incompatibility selection
may derive from disruption of deme-specific ensembles of interacting loci formed during past episodes of selective sweeps, the present genomic state is non-transient, maintained by a balance between negative selection and introgression.

**Barriers to neutral gene flow**

addressed the effect of interspecific incompatibility on introgression of regions linked to incompatibility loci, characterizing the “gene flow factor” as the probability that a newly-arrived gene at a neutral marker locus will become incorporated into the local gene pool. characterized reproductive barriers due to multiple, epistatic incompatibility factors. used this framework to study the rate of accumulation of incompatibility factors between species. In the absence of sex-specificity of any kind, loci unlinked to incompatibility factors face identical barriers to introgression and the joint barrier induced by multiple incompatibility factors corresponds to the product of the barriers generated by the factors individually.

Here, we address the implications for neutral introgression of sex-specificity. Sex-specificity may include differences between the sexes in fertility or viability of hybrids and their descendants, in rates of crossing-over between a neutral marker locus and incompatibility factors, or in transmission of the marker or incompatibility loci (including sex-linkage or sex-limited transmission). For clarity, we restrict consideration to a pair of hybridizing species.

Our results depart markedly from expectations in the absence of sex-specificity. Reproductive barriers generally show epistasis, even in the absence of functional interactions among incompatibility factors. This form of epistasis in introgression rate reflects associations between sex and incompatibility factors and between sex and neutral markers. Such associations tend to generate locus-specific introgression rates, even in the absence of physical linkage of marker loci to incompatibility factors. In general, concordance among incompatibility factors in their relative expression between the sexes (for example, more severe effects in hybrid males than females or greater crossover rates in females) generates lower reproductive barriers (greater introgression) than the multiplicative expectation and discordance generates higher barriers. Submultiplicativity of barriers arising under concordance appears
to reflect the greater efficiency of selection (???) in eliminating incompatibility factors that more severely impair the same sex.

THEORETICAL FRAMEWORK

Our major results fall into two classes. First, we describe the derivation of relative reproductive rate, the expected contribution to future generations of a foreign marker gene introduced by a migrant relative to a marker gene in a resident. Related to the “gene flow factor” of ?, relative reproductive rate is the central quantity in a coalescent-based approximation to the evolutionary process at neutral sites in a genome subject to incompatibility selection. Second, we address the barrier to introgression generated by multiple incompatibility factors that show no functional epistasis. This analysis presents a partitioning of variation in relative reproductive rate of neutral marker genes on all possible backgrounds.

For simplicity, we use terminology applicable to species with chromosomal sex determination in which males are heterogametic. In the context of Table ??, we indicate the modifications required to address cases with heterogametic females.

Genealogical migration rate

Effects of selection on gene flow: In the population genetics literature, the backward migration rate generally refers to the proportion of the local gene pool that derives from the gene pool of a different deme in the immediately preceding generation; the forward migration rate denotes the rate at which genes contribute in the immediately succeeding generation to a gene pool different from the one in which they presently reside. In the presence of selection, the forward and backward rates differ, even under a time-reversible migration process. Further, contribution of introgressed genes to future generations depends not only on the rate of mixing of gene pools but also on the reproductive success of their carriers (?).

Selection can change the fundamental structure of the coalescent process at a target of selection (??) and influences the coalescent process at neutral sites linked to targets of selection. Of central importance to the latter process is the waiting time between migration
events traced backward along the line of descent of a randomly sampled gene. Modeling the
distribution of waiting time as exponential, we refer to the parameter of that distribution as
the genealogical migration rate. Neutral substitution proceeds at rates expected in the ab-
sence of selection (?), with neutral divergence between demes dependent on the genealogical
migration rate. Our index of introgression differs from earlier definitions that refer to the
frequencies of species-specific marker alleles (e.g., ??).

Relative reproductive rate: We model the waiting time between migration events
traced backward along the line of descent of a neutral marker gene as exponential, assigning
its parameter as

\[ g = m \omega, \]  

for \( g \) the genealogical migration rate in the deme from which the gene currently resides, \( m \)
the backward migration rate, and \( \omega \) the relative reproductive rate, representing the expected
contribution to future generations of the marker gene. Relative reproductive rate reflects
the number and expression of foreign incompatibility alleles in the background of the focal
marker gene over many generations into the future.

At each of \( n \) incompatibility loci, let 1 denote the foreign incompatibility allele and 0
the local wild-type allele (the form that functions well in the local genome). A gene at the
neutral marker locus may reside on \( 2^{n+1} \) possible backgrounds, defined by sex (female or
male) and the states of the alleles at the \( n \) incompatibility loci (1 or 0).

Females bearing the 1 incompatibility allele in heterozygous form at locus \( i \) and no other
locus have fitness \( \sigma_{f,i} \) relative to females bearing only 0 alleles; \( \sigma_{m,i} \) denotes the relative
fitness of a male carrying the 1 allele only at locus \( i \). We assume no functional epistasis
among the incompatibility factors with respect to fitness: the fitness of carriers of multiple
foreign incompatibility alleles at the same or different loci corresponds to the product of the
\( \sigma \) values of those alleles.

As in earlier work (e.g., ?), we assume a monotonic decrease in the number of foreign
incompatibility alleles in the line of descent of a neutral marker gene traced forward in time.
This assumption reflects the virtual absence within the local population of those deleterious genes; we address the implications of within-population polymorphism at incompatibility loci in a separate work (??).

At the point of zygote formation, \( \mathbf{v} \) represents the vector of number of descendants of the focal marker gene on the \( 2^{n+1} \) backgrounds. After a single generation, the expected distribution of descendants becomes \( \mathbf{vC} \), for \( \mathbf{C} \) a product of matrices representing selection (\( \mathbf{S} \)) on zygotes followed by genetic transmission of the marker across the possible backgrounds (\( \mathbf{T} \)):

\[
\mathbf{C} = \mathbf{ST}.
\]  

To establish a common basis for transmission of the focal marker gene through male and female carriers, we scale the contributions of males by the reproductive value of males (??). Because males transmit gametes to the next generation only to the extent that they succeed in fertilizing eggs, males have reproductive value \( f/(1 - f) \), for \( f \) the proportion of females among reproductives, relative to unity for females.

In a generation in the remote future, the expected total number of descendant genes is \( \lim_{t \to \infty} \mathbf{vC}^t \mathbf{e} \), for \( \mathbf{e} \) the vector with all elements equal to 1. Relative to a marker gene on a pure local background, the contribution of a marker gene introduced by a migrant corresponds to

\[
\omega = \lim_{t \to \infty} \frac{\mathbf{v}_m \mathbf{C}^t \mathbf{e}}{\mathbf{v}_r \mathbf{C}^t \mathbf{e}},
\]  

for \( \mathbf{v}_m \) denoting the distribution of backgrounds of the marker gene in the hybrid offspring of the migrant and \( \mathbf{v}_r \) the analogous vector for a marker gene in offspring of a resident in the same generation as the hybrid. ?? proposed a similar measure, which they called the “neutral effective migration rate.” Although they formulated their definition in terms of the frequency of an allele introduced by migrants, they noted its relationship to the history of a randomly sampled lineage.

As the number of generations into the future (\( t \)) grows large, the spectral radius \( \lambda \) of \( \mathbf{C} \) comes to dominate (??). For diagonalizable \( \mathbf{C} \), \( \lambda \) corresponds to a simple, positive root
of the characteristic equation and the set of eigenvectors span the space (Appendix ?? addresses more general cases). Transmission of the focal marker genes through $t$ generations is determined from

$$C^t = Q D^t Q^{-1},$$

for $D$ a diagonal matrix with the eigenvalues of $C$ along the diagonal, $Q$ a matrix of right eigenvectors written as columns, and $Q^{-1}$ a matrix of left eigenvectors written as rows. For sufficiently large $t$, $\lambda^t$ comes to dominate the non-zero entries of $D^t$, which implies

$$\omega = \frac{v_m q}{v_r q}, \tag{4}$$

for $q$ the right eigenvector associated with the dominant eigenvalue $\lambda$ of $C$. In keeping with the Fisherian notion, we term $q$ (??) the reproductive value vector, the $i^{th}$ element of which represents the expected ultimate contribution to future generations of a marker gene on the $i^{th}$ background.

In the Mathematica notebook (matrix_builder.nb) provided as Supplementary Online Material, the matrices reflect an ordering first of females bearing backgrounds containing the maximum number of foreign (1) incompatibility alleles through successively fewer down to zero foreign alleles (resident background) and then of males bearing these backgrounds. In the case of introgression at an autosomal marker induced by incompatibility factor 1 alone, for example, the reproductive value vector (??) corresponds to

$$q = (\eta_{f,1}, \eta_{f,0}, \eta_{m,1} f/(1 - f), \eta_{m,0} f/(1 - f))',$$

in which $\eta_{f,1}$ denotes the relative contribution to future generations of the marker gene borne by a female carrier of the foreign incompatibility allele relative to that of a non-carrier female or male ($\eta_{f,0} = \eta_{m,0} = 1$), $\eta_{m,1}$ the analogous quantity for a male carrier, and the prime the transpose. Relative reproductive rate (??) corresponds to

$$\omega = \frac{(f, 0, 1 - f, 0)q}{(0, f, 0, 1 - f)q} = (\eta_f + \eta_m)/2.$$
in which \( \mathbf{v}_m = (f, 0, 1-f, 0) \) and \( \mathbf{v}_r = (0, f, 0, 1-f) \) reflect that the focal migrant and the focal resident each produce daughters and sons in proportions \( f \) and \((1 - f)\). Appendix ?? explicitly presents the matrices and the reproductive value vector \( \mathbf{q} \) for the case of an autosomal marker gene linked to a single autosomal incompatibility locus.

We use \( \omega \), defined in terms of contribution to future generations, to approximate the backward-perspectived genealogical migration rate \( g \). Results of simulation studies indicate excellent agreement between the expected and observed distributions of number of migration events traced backward along a randomly sampled lineage (Fig. ??, Appendix ??). The supplementary Mathematica notebook (matrix_builder.nb) constructs transmission matrix \( C \) and produces the reproductive value vector \( \mathbf{q} \) under user-specified values for the number of incompatibility loci, genomic locations of the marker and incompatibility loci, and sex-specific selection coefficients and recombination parameters.

**Partitioning of variation in long-term contribution**

We assume the absence of functional epistasis among incompatibility loci in expression of the deleterious effects of incompatibility: the fitness of individuals bearing multiple foreign incompatibility alleles is equal to the product of the fitnesses induced by the alleles separately. Even so, our results indicate pervasive departures of the overall reproductive barrier from the product of the barriers induced by incompatibility factors individually. Here, we introduce a decomposition of variation in relative reproductive rate \( \omega \) in terms of interactions among incompatibility factors and sex.

**Epistasis:** Expression (??) defines relative reproductive rate as a linear combination of elements of the reproductive value vector \( \mathbf{q} \), which gives the contributions to future generations of a marker gene on all possible backgrounds.

We explore the basis of departures from multiplicativity of the overall relative reproductive rate \( \omega \) at a neutral marker locus induced by multiple incompatibility factors.

We view reproductive value vector \( \mathbf{q} \) as a multiway table of a Fisherian factorial design experiment. Binary factors representing sex and the allelic state of the incompatibility
loci affect the response variable of contribution to future generations of the focal marker gene.

Our index of epistasis corresponds to an analogue of multilocus disequilibrium as proposed by ?, in which, for example, the two-way association corresponds to a covariance:

$$E[(A_i - \overline{A}_i)(A_j - \overline{A}_j)].$$

In the case at hand, we interpret $E[A_i]$ as the expected relative contribution to future generations of the focal marker gene held by an individual bearing the foreign incompatibility allele at locus $i$ and no other locus:

$$E[A_i] = \overline{A}_i = \omega_i,$$

with

$$E[\prod_{i\in\Omega} A_i] = \omega_\Omega$$

the expected long-term contribution in the presence of foreign incompatibility alleles at all loci having an index in the set $\Omega$ and no other loci. Superscripts specify the context: for example, $\omega^{X,f}_\Omega$ denotes the relative reproductive rate at an X-linked marker gene introduced by a female migrant, and $\omega^{X,m}_\Omega$ the rate in the case of a male migrant.

Two-way epistasis reflects the departure from multiplicativity of barriers to introgression:

$$e_{ij} = E[(A_i - \overline{A}_i)(A_j - \overline{A}_j)] = \omega_{ij} - \omega_i \omega_j.$$

Appendix ?? shows that for an autosomal marker (superscript $A$), $k$-way epistasis among $k$ autosomal incompatibility loci corresponds to

$$e_{\Omega[k]}^{A,f} = e_{\Omega[k]}^{A,m} = \sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega[j] \subseteq \Omega[k]} \delta_{\Omega[j]} \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b,$$

(5)

for $\Omega[k]$ a set of $k$ indices representing incompatibility loci, $\Omega[j]$ a subset of $\Omega[k]$ comprising $j$ indices, $\Omega[k] \setminus \Omega[j]$ the set of indices in $\Omega[k]$ but not in $\Omega[j]$, and

$$\delta_{\Omega[j]} = \omega_{\Omega[j]} - \prod_{i \in \Omega[j]} \omega_i$$

(6)
the departure of relative reproductive rate from multiplicativity. This partitioning of \( k \)-way epistasis includes components for all subsets of \( 2, \ldots, k \) loci, giving a total of \( \sum_{j=2}^{k} \binom{k}{j} = 2^k - (k + 1) \) terms. It has the intuitively appealing property of zero epistasis for sets that include any neutral locus \( j \) \((E[A_j] = \omega_j = 1)\), reflecting cancellation between terms of the form \( E[\ldots A_j \ldots] \) and \( E[\ldots A^\prime_j \ldots] \).

We apply this partitioning of epistasis to relative reproductive rate (\(??\)) for an arbitrary number of incompatibility factors, assuming no linkage among the factors or the marker. Relative reproductive rate at an autosomal marker corresponds to

\[
\omega^{A,f}_{\Omega[k]} = \omega^{A,m}_{\Omega[k]} = \left( \eta_{f,\Omega[k]} + \eta_{m,\Omega[k]} \right)/2, \tag{7}
\]

reflecting the equal transmission of the marker through the female and male hybrid offspring of migrants of either sex. Expression (\(??\)) also applies to an X-linked marker introduced by a female migrant \((\omega^{X,f}_{\Omega[k]})\); for a male migrant, \(\omega^{X,m}_{\Omega[k]} = \eta_{f,\Omega[k]}\), reflecting that only female offspring carry the marker.

For cases in which relative reproductive rate corresponds to (\(??\)), Appendix ?? shows that the overall departure from multiplicativity (\(??\)) decomposes into indices of interaction between the foreign allele at incompatibility locus \( i \) and sex:

\[
\Delta_i = (\eta_{f,i} - \eta_{m,i})/2. \tag{8}
\]

Positive \(\Delta_i\) signifies a positive effect of femaleness: neutral marker genes borne by females contribute to future generations at higher rates than those borne by males. Negative \(\Delta_i\) signifies that a marker gene on background \( i \) carried by a male contributes to future generations at a higher rate than a gene on the same background in a female. Sex-specific departures from multiplicativity correspond to

\[
\delta_{f,\Omega[j]} = \eta_{f,\Omega[j]} - \prod_{i \in \Omega[j]} \eta_{f,i}, \tag{9a}
\]

\[
\delta_{m,\Omega[j]} = \eta_{m,\Omega[j]} - \prod_{i \in \Omega[j]} \eta_{m,i}. \tag{9b}
\]
Appendix ?? shows that
\[
e_{\Omega_{[k]}} = \begin{cases} 
\frac{1}{2} \sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega_{[j]}} \left( \delta f_{\Omega_{[j]}} + \delta m_{\Omega_{[j]}} \right) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_{b} k \text{ odd} \\
\frac{1}{2} \sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega_{[j]}} \left( \delta f_{\Omega_{[j]}} + \delta m_{\Omega_{[j]}} \right) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_{b} + \prod_{i=1}^{k} \Delta_{i} k \text{ even}.
\end{cases}
\] (10)

**RESULTS**

**Barriers due to single sex-specific incompatibility factors**

A major objective is to explore the effects of sex-specificity on barriers to neutral introgression. Appendix ?? presents in detail the case of X-linkage of both the marker locus and an incompatibility locus, illustrating the modifications to transmission matrix \( T \) to accommodate sex-specificity. Here, we address the effect on the relative reproductive rate \( \omega \) of the genomic locations of the marker locus and of a single incompatibility locus.

Table ?? presents reproductive values at a marker locus in a genome together with a single incompatibility locus, with the loci at various locations in the genome, including autosomal (A), X-linked (X), Y-linked (Y), and mitochondrial (mt). The Reproductive value columns list the relative long-term contribution to future generations of female \( \eta_{f} \) and male \( \eta_{m} \) carriers of the foreign incompatibility factor, and the Factor-sex association column gives the difference between them \( \eta_{f,1} - \eta_{m,1} \); compare (11)). The final columns show the overall relative reproductive rates (12) of marker alleles introduced by female \( \omega^{f} \) and male migrants \( \omega^{m} \). For cases in which offspring of both sexes can transmit the marker and carry the incompatibility factor (A or X), the relative reproductive rate (13) of a female migrant corresponds to the average of the reproductive values of its female and male offspring:

\[
(\eta_{f,1} + \eta_{m,1})/2
\]
(compare (11)). In the case of an autosomal marker and an X-linked incompatibility factor, a marker allele introduced by a male migrant has higher relative reproductive rate:

\[
(\eta_{f} + 1)/2,
\]
reflecting that sons of male migrants are free of the incompatibility factor. In the case of either an autosomal or an X-linked incompatibility factor, the relative reproductive rate at
an X-linked marker allele introduced by a male migrant corresponds to $\eta_f$ alone, reflecting that only daughters of male migrants carry the marker.

We assume that males do not transmit mitochondria, with the implication that incompatibility due to foreign mitochondria (mt) presents no barrier to autosomal or X-linked markers borne by a male migrant ($\omega_{mt}^{A,m} = \omega_{mt}^{X,m} = 1$). Similarly, incompatibility factors on the X-chromosome or mitochondria present no barrier to Y-linked markers ($\omega_{X}^{Y} = \omega_{mt}^{Y} = 1$) because the marker and the factor are never transmitted to the same offspring.

Birds and various other organisms exhibit ZW sex determination, with male homogamety (ZZ) and female heterogamety (ZW). Reproductive values under this system correspond to those given in Table ??, with Z substituted for X and W for Y. Because complete cosegregation of mitochondria and the W chromosome constitutes in essence complete linkage between these regions, the reproductive values for cases involving a marker or a factor on mitochondria can be obtained from Table ?? by substituting mt for W.

**Barriers due to multiple incompatibility factors**

We describe conditions under which the relative reproductive rate $\omega$ (??) induced by multiple incompatibility factors corresponds to the product of the relative reproductive rates induced by the factors individually. Beyond this case (absence of sex-specificity in expression of incompatibility, crossover rates, or genetic transmission), the joint barrier to introgression engendered by multiple incompatibility loci departs in general from the product of the barriers induced by the factors in isolation. Epistasis in relative reproductive rate as defined in this section and in Appendix ?? reflects such departures from multiplicativity.

**Multiplicative barriers to introgression**: In the absence of sex-specificity, the Kronecker structure of the transition matrix $C$ (??) for an autosomal marker extends to multiple incompatibility loci with multiplicative effects on the survival or reproduction of carriers. Here, we show that the total barrier to introgression at a neutral autosomal marker locus corresponds to the product of the barriers generated by the incompatibility loci individually under (1) sex-independent incompatibility selection and recombination, (2) in-
dependent transmission of the factors conditional on transmission of the marker, and (3) sex-independent genetic transmission.

We consider \( l \) incompatibility loci, for which the foreign allele at locus \( i \) reduces the viability and fertility of its carriers by a factor of \( \sigma_i \) relative to non-carriers. Crossing-over occurs between the focal marker gene and incompatibility locus \( i \) at rate \( r_i \), independently of crossing-over in other regions. This assumption implies that at most two incompatibility loci (flanking the marker locus) can reside on the chromosome bearing the marker locus, with all other incompatibility loci unlinked (\( r_{f,j} = r_{m,j} = 1/2 \)).

Selection matrix (\( ?? \)) generalizes to

\[
\begin{pmatrix}
1 & 0 \\
0 & \frac{f}{1-f}
\end{pmatrix} \otimes \tilde{S},
\]

for \( \tilde{S} = S_1 \otimes \ldots S_l \) with

\[
S_i = \begin{pmatrix}
\sigma_i & 0 \\
0 & 1
\end{pmatrix}.
\]

Transmission of the autosomal marker and sex retains the form (\( ?? \)). Independent transmission of the factors conditional on transmission of the marker implies

\[
F = \begin{pmatrix}
1 & 1 \\
1 & 1
\end{pmatrix} \otimes \tilde{R},
\]

for \( \tilde{R} = R_1 \otimes \ldots R_l \) with

\[
R_i = \begin{pmatrix}
1 - r_i & r_i \\
0 & 1
\end{pmatrix}
\]

(compare (\( ?? \))). The full transition matrix corresponds to

\[
C = \left[ \begin{pmatrix}
1 & 0 \\
0 & \frac{f}{1-f}
\end{pmatrix} \otimes \tilde{S} \right] \left[ \frac{1}{2} (f, 1-f) \otimes \begin{pmatrix}
1 \\
1
\end{pmatrix} \otimes \tilde{R} \right] = \left[ \begin{pmatrix}
1 & 0 \\
0 & \frac{f}{1-f}
\end{pmatrix} \right] \left[ \frac{1}{2} (f, 1-f) \otimes \begin{pmatrix}
1 \\
1
\end{pmatrix} \right] \otimes \tilde{S} \tilde{R}
\]

(compare (\( ?? \))). The eigenvalues of \( \tilde{S} \tilde{R} = S_1 R_1 \otimes \ldots S_l R_l \) correspond to products of the eigenvalues of the component matrices and its eigenvectors to Kronecker products of the eigenvectors of those matrices. In particular, elements of the reproductive value \( q \) (\( ?? \)) have
the form

\[ \eta_{f,\Omega} = \prod_{i \in \Omega} \eta_{f,i} \quad \eta_{m,\Omega} = \prod_{i \in \Omega} \eta_{m,i}, \]  

for \( \Omega \) any subset of indices in \( \{1, \ldots, l\} \) and \( \eta_{f,i} = \eta_{m,i} \) the reproductive value of female and males carriers of the foreign incompatibility allele at locus \( i \). Consequently, the overall relative reproductive rate (??) corresponds to the product of the reproductive rates across the loci individually (??):

\[ \omega_{1,\ldots,l}^{A,f} = \omega_{1,\ldots,l}^{A,m} = \prod_{i=1}^{l} \frac{\sigma_i r_i}{1 - \sigma_i (1 - r_i)}. \]  

Multiplicative barriers also arise in the case of markers that are transmitted without recombination through only one sex: those on mitochondria or in the non-recombining male (female)-specific region of the Y (W) chromosome, for example. For markers transmitted only through females, the selection matrix (??) reduces to \( \tilde{S} = S_1 \otimes \ldots \otimes S_l \), with \( S_i \) containing only the female selection parameter \( \sigma_{f,i} \), and the recombination matrix \( \tilde{R} = R_1 \otimes \ldots \otimes R_l \) contains the crossover rates for females alone. As the transition matrix reduces to \( C = \tilde{S} \tilde{R} \), the reproductive rate vector \( q \) has elements \( \eta_{f,\Omega} = \prod_{i \in \Omega} \eta_{f,i} \) and the overall relative reproductive rate (??) is given by (??), with the selection and recombination parameters corresponding to the values for females. In the absence of linkage between the marker and the \( i^{th} \) factor \( (r_{f,i} = 1/2) \), the \( i^{th} \) multiplicand of (??) reduces to

\[ \frac{\sigma_{f,i}}{2 - \sigma_{f,i}} \]  

(compare (??)). Complete linkage \( (r_{f,i} = 0) \) implies a zero overall relative reproductive rate (no introgression at the marker).

In this section, we have assumed the complete absence of sex-specificity. The absence of interaction between any incompatibility locus and sex \( (\Delta_i = 0 \ \forall i) \) and of all sex-specific departures from multiplicativity \( (\delta_{f,\Omega_{[j]}} = \delta_{m,\Omega_{[j]}} = 0 \ \forall j) \) (??) as well as of the overall departure (??) implies the absence of epistasis of all orders.

**Nonmultiplicative barriers to introgression:** Violation of any of the postulates enumerated in the preceding section induces a departure from multiplicativity of the barriers to
introgression generated by more than one incompatibility locus. In this sense, the postulates represent minimal conditions for multiplicativity. To explore the nature and implications of nonmultiplicative barriers, we address the relative reproductive rate at a marker locus (autosomal or X-linked) subject to incompatibility generated by two loci: both autosomal or one autosomal and one X-linked.

**Epistasis between autosomal factors with sex-specific expression or recombination:** We explore the effects of sex-specificity in expression of incompatibility or in crossover rates on the rate of introgression at an autosomal marker flanked by autosomal incompatibility loci 1 and 2.

Conditional on transmission of the marker, transmission from female parents of the two factors occurs independently of each other:

\[
\begin{pmatrix}
1 - r_{f,1} & r_{f,1} \\
0 & 1
\end{pmatrix}
\otimes
\begin{pmatrix}
1 - r_{f,2} & r_{f,2} \\
0 & 1
\end{pmatrix},
\]

for \( r_{f,i} \) the rate of crossing-over in females between the marker and incompatibility locus \( i \). A similar transmission matrix holds for male parents, with \( r_{m,i} \), the crossover rate in males, substituted for \( r_{f,i} \). In the absence of physical linkage, \( r_{f,i} = r_{m,i} = 1/2 \). The foreign incompatibility allele at locus \( i \) reduces the viability of its female (male) carriers by a factor of \( \sigma_{f,i} (\sigma_{m,i}) \), with multiplicative effects in carriers bearing foreign alleles at both loci (e.g., \( \sigma_{f,12} = \sigma_{f,1}\sigma_{f,2} \)).

Migrants of either sex produce female and male hybrid offspring in frequencies \( f \) and \( 1 - f \). The relative reproductive rate at the marker locus (\( \omega \)) corresponds to

\[
\omega_{12}^{A,f} = \omega_{12}^{A,m} = \frac{(f, 0, 0, 0, 1 - f, 0, 0, 0)q}{(0, 0, 0, 0, 0, 0, 1 - f)}q = (\eta_{f,12} + \eta_{m,12})/2,
\]

for \( \eta_{f,12} \) and \( \eta_{m,12} \) the contribution to future generations of the marker gene in females and males that bear the foreign factor at both incompatibility loci.

Our index of epistasis (\( \delta \)) among incompatibility loci reduces in this case to

\[
\delta_{12} = (\delta_{f,12} + \delta_{m,12})/2 + \Delta_1 \Delta_2 = 2\Delta_1 \Delta_2 / \Gamma,
\]

for \( \Delta_1 \) and \( \Delta_2 \) the effects of incompatibility at each locus.
in which
\[
\Delta_i = \frac{(\eta_{f,i} - \eta_{m,i})/2 = \frac{\sigma_{f,i}r_{f,i} - \sigma_{m,i}r_{m,i} - \sigma_{f,i}\sigma_{m,i}(r_{f,i} - r_{m,i})}{2 - \sigma_{f,i}r_{f,i} - \sigma_{m,i}r_{m,i}}}{16}
\] (16)

(compare (??)) and
\[
\Gamma = 2 - \sigma_{f,1}\sigma_{f,2}(1 - r_{f,1})(1 - r_{f,2}) - \sigma_{m,1}\sigma_{m,2}(1 - r_{m,1})(1 - r_{m,2}).
\] (17)

In this case, the sex-specific measures of epistasis are proportional to each other:
\[
\delta_{f,12} = \eta_{f,12} - \eta_{f,1}\eta_{f,2} = 2\sigma_{f,1}\sigma_{f,2}(1 - r_{f,1})(1 - r_{f,2})\Delta_1\Delta_2/\Gamma
\]
\[
\delta_{m,12} = \eta_{m,12} - \eta_{m,1}\eta_{m,2} = 2\sigma_{m,1}\sigma_{m,2}(1 - r_{m,1})(1 - r_{m,2})\Delta_1\Delta_2/\Gamma.
\]

These autosomal factors undergo independent transmission (??) and cause no functional epistasis with respect to fitness of carriers. Furthermore, each in isolation induces the relative reproductive rate given by ?, with the average crossover and viability rates substituted for the non-sex-specific parameters (??). Even so, the joint barrier to introgression departs in general from the product of the barriers induced by each locus separately (??).

To explore the nature of epistasis due to sex-specificity in the impairment of carriers of foreign incompatibility loci, we determined the relative reproductive rate (??) at an autosomal marker induced by two autosomal incompatibility factors in the absence of physical linkage \((r_{f,i} = r_{m,i} = 1/2 \text{ for } i = 1, 2 \text{ in (??)})\). In Figure ??, the relatively flat plane corresponds to the expectation under multiplicative reproductive barriers, while the saddle-shaped surface shows the actual joint barrier, with heavy lines marking the only points of agreement between the actual and expected barriers. This example illustrates the general finding that incompatibility factors with concordant relative effects on the sexes (for example, both afflicting male carriers to a greater extent than female carriers) induce lower barriers to neutral introgression than the multiplicative expectation, and factors with discordant effects higher barriers.

To verify the accuracy of our expression for the genealogical migration rate \(g\) (??), we conducted a numerical simulation study using SFS_CODE (??), modified to identify the lineage involved in each migration event. We specified an autosomal marker gene and two autosomal
incompatibility loci under multiplicative fitnesses and no linkage among any of the loci. We simulated 100,000 generations forward in time and then counted the number of migration events in randomly sampled line of descent traced backward in time. We compared the counts observed in 15,000 independent replicate runs to a Poisson distribution with parameter determined by (??) under the population parameters specified in the simulation (see caption of Fig. ??). Figure ?? indicates an excellent fit of the simulated data to our prediction. It also provides another illustration that incompatibility factors with concordant effects on the sexes induce lower barriers than expected under multiplicativity (left panel) and factors with discordant effects higher barriers (right panel).

Sex-specific recombination rates can generate epistasis even in the absence of sex-specificity in carrier fitness ($\sigma_{f,i} = \sigma_{m,i} = \sigma < 1$). In this case, (??) reduces to

$$\Delta_i = \frac{\sigma(1 - \sigma)(r_{f,i} - r_{m,i})}{2 - \sigma(r_{f,i} + r_{m,i})}.$$ 

Rates of introgression at the marker exceed those expected under multiplicative barriers ($e_{12} > 0$) if crossing-over between the marker and each of the flanking incompatibility factors occurs at higher rates in females than males ($\Delta_1, \Delta_2 > 0$) or if both rates are higher in males ($\Delta_1, \Delta_2 < 0$).

**X-linked marker with autosomal and X-linked incompatibility factors:** We address barriers generated by one autosomal (A) and one X-linked (X) incompatibility locus to introgression at an X-linked marker introduced by a female migrant (Appendix ?? gives the expressions for a male migrant). We assume no crossing-over between the X-linked factor and the marker in (hemizygous) males ($r_{m,X} = 0$).

Epistasis in relative reproductive rate (??) between the X-linked and autosomal incompatibility loci corresponds to

$$e_{AX}^X = \omega_{AX}^X - \omega_A^X \omega_X^X = (\delta_{f,AX}^X + \delta_{m,AX}^X)/2 + \Delta_A^X \Delta_X^X.$$ 

Table ?? gives the index of association between the autosomal factor and sex,

$$\Delta_X^X = (\eta_{f,A}^X - \eta_{m,A}^X)/2 = \frac{2(\sigma_{f,A} - \sigma_{m,A})}{8 - \sigma_{f,A}(2 + \sigma_{m,A})},$$ 

(19)
and between the X-linked factor and sex,

\[
\Delta_X = (\eta_{f,X}^X - \eta_{m,X}^X)/2 = \frac{\sigma_{f,X}r_{f,X}(1 - \sigma_{m,X})}{2 - \sigma_{f,X}(1 + \sigma_{m,X})(1 - r_{f,X})},
\]  

(20)

Within-sex interactions correspond to

\[
\delta_{f,AX}^X = \eta_{f,AX}^X - \eta_{f,A}^X\eta_{f,X}^X = 4\sigma_{f,A}\sigma_{f,X}(1 - r_{f,X})\Delta_A^X\Delta_X^X/\Gamma
\]  

(21a)

\[
\delta_{m,AX}^X = \eta_{m,AX}^X - \eta_{m,A}^X\eta_{m,X}^X = 8\sigma_{f,A}\sigma_{f,X}(1 - r_{f,X})\sigma_{m,A}\sigma_{m,X}\Delta_A^X\Delta_X^X/\Gamma,
\]  

(21b)

giving an overall two-way epistasis between the incompatibility factors of

\[
e_{AX}^X = 8\Delta_A^X\Delta_X^X/\Gamma,
\]

(22)

in which

\[\Gamma = 8 - \sigma_{f,A}\sigma_{f,X}(1 - r_{f,X})(2 + \sigma_{m,A}\sigma_{m,X}).\]

Because a male transmits its entire X chromosome (the marker together with any foreign incompatibility allele) to all daughters and no sons,

\[\eta_{m,X}^X = \eta_{f,X}^X\sigma_{m,X}\]

(Table ??). Accordingly, (??) indicates that the X-linked factor obstructs transmission of the X-linked marker through males more than through females ($\Delta_X^X > 0$) unless the X-linked factor causes complete sterility or inviability in females ($\sigma_{f,X} = 0$), has no deleterious effects in males ($\sigma_{m,X} = 1$), or shows absolute linkage to the marker ($r_{f,X} = 0$). The nature of epistasis, sex-specific (??) as well as overall (??), depends on the effect of the autosomal factor ($\Delta_A^X$). Expression (??) indicates positive epistasis (greater introgression than expected under multiplicativity) if the autosomal factor impairs male carriers to a greater extent ($\sigma_{f,A} > \sigma_{m,A}$).

**Autosomal marker with autosomal and X-linked incompatibility factors:** We now consider introgression at an autosomal marker in a genome containing one autosomal (A) incompatibility locus, possibly linked to the marker, and one X-linked (X) locus. We assume a female migrant (Appendix ?? provides the expressions for a male migrant).
The overall two-way epistasis between the incompatibility factors corresponds to

$$ e_{AX}^A = \frac{4\Delta_A^AD_f}{[8 - \sigma_{f,X}(2 + \sigma_{m,X})]\Gamma} \tag{23} $$

for

$$ \Delta_A^A = \frac{(\eta_{f,A}^A - \eta_{m,A}^A)/2 = \frac{\sigma_{f,A}r_f - \sigma_{m,A}r_m - \sigma_{f,A}\sigma_{m,A}(r_f - r_m)}{2 - \sigma_{f,A}r_f - \sigma_{m,A}r_m}}{\Gamma} \tag{24} $$

(compare to (??) and (??)), for $D_f$ an index of association between sex and the X-linked factor (??) and $\Gamma$ a positive quantity (Appendix ??). The joint barrier to introgression departs from multiplicativity only if the autosomal factor shows an interaction with sex ($\Delta_A^A \neq 0$), reflecting sex-specific expression of incompatibility ($\sigma_{f,A} \neq \sigma_{m,A}$) or sex-specific rates of crossing-over with the marker ($r_{f,A} \neq r_{m,A}$). Whether the overall epistasis $e_{AX}^A$ is positive or negative depends on the nature of the female-specific epistasis between the incompatibility factors ($\delta_{f,A}^A \propto D_f$ (??)), which in turn depends on parameters of both the autosomal and X-linked factors.

Indices $D_f$ and $D_m$ (??) function as indicators of the effect of sex on the barrier to transmission of the autosomal marker locus induced by the X-linked incompatibility factor: positive values indicate higher transmission of the marker through female than male carriers of the incompatibility factor. In general, $D_m \leq 0$, signifying that femaleness reduces transmission by carriers of the X-linked incompatibility factor, with no effect ($D_m = 0$) only if the factor does not in fact contribute to incompatibility ($\sigma_{f,X} = \sigma_{m,X} = 1$). This property may reflect that while female carriers pass the X-linked incompatibility factor to offspring of both sexes, male carriers pass it only to daughters.

Another indication of the intrinsic association of an X-linked incompatibility factor with sex is that it can affect the sign of the overall epistasis $e_{AX}^A$ even if it exhibits no sex-specificity in expression of incompatibility or crossing-over ($\sigma_{f,X} = \sigma_{m,X}$ and $r_{f,X} = r_{m,X}$). In this case, the autosomal marker has lower reproductive rates in female carriers of the X-linked incompatibility factor than in male carriers ($D_f, D_m < 0$).

DISCUSSION

22
We have explored the implications of sex-specific interspecific incompatibility for neutral introgression across the genome. We proposed a scaling of the backward migration rate (??) that accounts for sex-specific incompatibility selection and linkage and showed that the distribution of waiting times along a random lineage traced backward conforms well to the distribution determined by this scaling (Fig. ??). Among the key qualitative findings are that (1) neutral introgression generally occurs at locus-specific rates, even in the absence of linkage of marker loci to incompatibility factors and (2) the joint barrier generated by multiple incompatibility factors is generally nonmultiplicative, even in the absence of functional epistasis among the incompatibility factors.

Here, we provide a qualitative discussion of these results, illustrate some implications for the interpretation of patterns of genetic variation in closely related Drosophila species, and suggest that sex-specificity is a pervasive feature of interspecific hybridization in plants and animals.

**Locus-specificity in reproductive barriers reflects associations with sex**

Factors contributing to interspecific incompatibility impede introgression of neutral markers at a rate that depends on the level of linkage, selection intensity, and functional epistasis among incompatibility factors (???). In the absence of sex-specificity, the locus-specific nature of barriers to introgression extends only to regions immediately adjacent to targets of selection. For example, Table ?? indicates that in the absence of linkage with a neutral marker \( r_f = r_m = 1/2 \), a single factor transmitted through both sexes that induces non-sex-specific incompatibility \( \sigma_f = \sigma_m = \sigma \leq 1 \) either imposes no barrier to introgression at the marker or reduces it by a factor of \( \sigma/(2 - \sigma) \) (compare the “gene flow factor” (??) of ?). Factors borne on the mitochondria also have this effect, while Y-linked factors impede introgression on autosomes by a factor of \( 1/(2 - \sigma) \). In contrast, sex-specific incompatibility induces locus-specific barriers even at unlinked marker loci.

Sex-specificity in expression of incompatibility, transmission, or level of linkage to incompatibility factors impedes introgression at neutral markers to unequal extents across genomic
regions, even in the absence of functional epistasis. Figure ?? illustrates, for a female or male migrant, relative reproductive rates (??) at neutral marker loci at various locations in the genome (color-coded bars) with one or multiple incompatibility loci at locations indicated on the abscissa, assuming free crossing-over ($r = 1/2$) between the marker and any incompatibility factor. Because a male migrant never transmits both its X and Y chromosomes to the same offspring, an incompatibility factor on one chromosome poses no barrier to introgression of the other chromosome. Accordingly, Fig. ?? shows relative reproductive rates of 1 for $\omega_{X,m}^{Y,m}$ (magenta bar for an incompatibility factor only on the X) and $\omega_{Y,m}^{X,m}$ (cyan bar for an incompatibility factor only on the Y), regardless of sex-specificity. In all other cases, the differences in heights of the bars for a given set of locations for incompatibility factors arise as a consequence of sex-specificity in intensity of incompatibility selection or transmission of the marker or incompatibility loci.

Locus-specific reproductive barriers of this kind reflect associations between sex and incompatibility factors and between sex and marker loci (last section of RESULTS and Appendix ??). Because neutral markers borne on X chromosomes or on mitochondria, for example, descend preferentially or exclusively through females, they experience incompatibility primarily in a female context. If male hybrids experience greater impairment, female-associated markers face lower barriers to introgression than male-associated markers. A Y-linked incompatibility factor introduced by a male migrant imposes no barrier to introgression at an X-linked marker because only its daughters carry the marker, but it does impede introgression at an autosomal marker, which accompanies the Y-linked factor in half of the sons of a male carrier. In Table ??, the Factor-sex association column shows a measure of interaction (??) between a single incompatibility factor and sex in determining relative reproductive rate $\omega$ (??).

Differences between the sexes in levels of recombination have consequences similar to those of sex-specific expression of incompatibility. Greater rates of crossing-over in females than in males, for example, permit greater introgression of female-associated markers linked
to incompatibility factors.

**Epistasis in relative reproductive rate**

Our analysis of the joint barrier to neutral introgression induced by multiple incompatibility factors assumes the absence of functional epistasis: the fitness of individuals bearing multiple factors corresponds to the product of the fitnesses induced by the factors individually. Under equal impairment of male and female carriers of foreign incompatibility factors, transmission of incompatibility factors on autosomes, and equal rates of crossing-over between the sexes, we confirm the multiplicativity of the barriers (??), as commonly assumed. These conditions appear to be minimal, as violation of any one causes departures from multiplicativity (**Nonmultiplicative barriers to introgression** section).

In general, discordance among incompatibility factors in the nature of association with sex tends to increase barriers to introgression. For example, Figure ?? depicts the relative reproductive rate of an autosomal marker in a genome containing two incompatibility loci with sex-specific expression ($\sigma_{f,i} \neq \sigma_{m,i}$) in the absence of linkage between any pair of loci. It illustrates that the highest barriers (lowest $\omega$ values, bluer regions) derive from a positive interaction (??) with femaleness in one factor ($\Delta_i > 0$) and a negative interaction in another factor ($\Delta_j < 0$). Positive epistasis arises under concordant associations ($\Delta_i \Delta_j > 0$), implying higher rates of gene flow (redder regions). For extreme manifestations of **HALDANE’s rule** ($\sigma_f = 1, \sigma_m = 0$ or $\sigma_m = 1, \sigma_f = 0$), the actual rate of introgression is nearly twofold greater than expected under multiplicative reproductive barriers (red corners of Fig. ??).

Rapid elimination of the deleterious foreign incompatibility factors permits greater neutral introgression. The selection process is more efficient if the factors tend to occur together (??): in the same sex, in the present context. Concordance of the sex-specific effects of multiple incompatibility factors (greater impairment of male carriers than female carriers, for example) implies their association with the same sex.

**Interspecific incompatibility in Drosophila**

?? studied patterns of nucleotide divergence and polymorphism in sister species **Drosophila**
yakuba and D. santomea in 29 genomic regions, including sites on the X chromosome, Y chromosome, autosomes, and mitochondria. Regions on the X and Y contribute to severe reductions in hybrid male fertility (??) and quantitative trait loci contributing to both prezygotic and postzygotic isolation occur on the X chromosome and the autosomes (??). ? found significantly lower ratios of shared to fixed polymorphisms between the species on the X-chromosome than on autosomes, which they considered consistent with reduced introgression of X-linked sites due to the disproportionate contribution of the X to hybrid male sterility.

To account for the difference in effective number of genes under X- and autosomal-linkage and its effects on levels of polymorphism, they also conducted coalescent simulations under IM-based (?) estimates of demographic parameters. However, this approach indicated no significant reduction in introgression rate at X-linked relative to autosomal sites.

In confirmation of the expectations of ?, our model indicates that incompatibility due to a single X-linked factor permits greater introgression in autosomal than X-linked regions:

\[
\omega_{X,f}^{A} > \omega_{X,f}^{X} \quad \text{and} \quad \omega_{X,m}^{A} > \omega_{X,m}^{X},
\]

regardless of the intensity of incompatibility selection \((\sigma_{f,X}, \sigma_{m,X})\), the rate of crossing-over on the X \((r_{f,X})\), or the sex of the initial migrant (Table ??). The blue (autosomal marker) and cyan (X-linked marker) bars in Fig. ?? for the case of a single X-linked factor (X on the abscissa) illustrates this effect.

In contrast, incompatibility due to a single autosomal factor introduced by a male migrant permits greater introgression in X-linked regions than in autosomal regions unlinked to the factor \((r_{f} = r_{m} = 1/2)\):

\[
\omega_{X,m}^{A} > \omega_{A,m}^{X}
\]

(for example, blue (A) and cyan (X) bars in Fig. ?? over A for a male migrant). For autosomal factors introduced by a female migrant, higher introgression in X-linked regions

\[
\omega_{X,f}^{A} > \omega_{A,f}^{X}
\]

holds for factors that impair male carriers more than female carriers \((\sigma_{f} > \sigma_{m})\).
Many aspects of incompatibility between *D. yakuba* and *D. santomea* show sex-specificity, with multiple factors on the X-chromosome, Y-chromosome, and autosomes contributing to hybrid male sterility (??). Accordingly, we expect the overall barrier to introgression to show sex-related epistasis. To illustrate the implications, we compare the relative reproductive rates at X-linked and autosomal markers induced by incompatibility factors both on the X-chromosome and on an autosome unlinked to the autosomal factor ($r_{f,A} = r_{m,A} = 1/2$). For the case in which the X-linked factor causes complete sterility in hemizygous form ($\sigma_{m,X} = 0$), with maximal crossing-over in females ($r_{f,X} = 1/2$; $r_{m,X} = 0$ in hemizygous males), our model indicates uniformly higher introgression rates in X-linked than autosomal regions for the case of a male migrant ($\omega_{A,X}^{m} > \omega_{A,X}^{A}$). For female migrants, with $\sigma_{m,X} = 0$,

$$\omega_{A,X}^{X,f} - \omega_{A,X}^{A,f} \propto \sigma_{f,A} \sigma_{f,X} \sigma_{m,A} (\sigma_{f,A} - \sigma_{m,A}).$$

This expression implies equal barriers to introgression on the autosome and the X-chromosome ($\omega_{A,X}^{X,f} = \omega_{A,X}^{A,f}$) if either factor expresses dominant lethality or sterility ($\sigma_{f,A} \sigma_{f,X} \sigma_{m,A} = 0$). Otherwise, greater introgression is expected at X-linked than autosomal sites ($\omega_{A,X}^{X,f} > \omega_{A,X}^{A,f}$) if the effects of the autosomal factor conform to Haldane’s rule ($\sigma_{f,A} > \sigma_{m,A}$).

Contrary to the expectation of ??, these results indicate that sex-specificity in crossing-over and in postzygotic incompatibility can induce a greater barrier to introgression in autosomal regions unlinked to any incompatibility factor than in regions on the X that are loosely linked ($r_{f} = 1/2$) to an incompatibility factor. The presence of an incompatibility factor on the Y, as reported by ??, would reinforce this trend by inhibiting introgression of autosomes but not the X-chromosome because foreign X-linked markers and the foreign Y never occur in the same genome.

Even under parameter values for which $\omega_{A,X}^{X,f} < \omega_{A,X}^{A,f}$ (e.g., Fig. ??), positive epistasis due to concordant sex-specific effects among factors (for example, greater impairment of male hybrids or higher rates of crossing-over in females) may cause the joint barrier to exceed the multiplicative expectation considerably (Fig. ??). In such cases, the introgression rate in unlinked autosomal regions may only slightly exceed the rate on the X, consistent with the
non-significant differences reported by ?.

Unlike the expectation in the absence of sex-specificity, Fig. ?? illustrates marked differences in introgression rates among markers freely recombining with incompatibility loci. Our analysis of locus-specific introgression rates may contribute toward a basis for inferring the existence and location of incompatibility factors from the pattern of neutral variation throughout the genome. In particular, the ensemble response of marker loci in the four genomic regions shown in Fig. ?? differ among models for the location of incompatibility loci and also between the sexes of the migrants within a given model.

**Sex-specific transmission or expression of incompatibility factors**

As sex-specificity in expression of interspecific incompatibility or in crossover rates has been widely observed, our findings raise the possibility of pervasive locus-specific neutral divergence among species.

? have reviewed studies documenting differences in map length between males and females in both plants and animals, including humans (?). In Drosophila, the premiere model organism for the experimental investigation of interspecific incompatibility, crossing-over is suppressed altogether in males (?).

Assessment of factors contributing to postzygotic interspecific incompatibility that have been identified to the level of operons confirm the traditional view that speciation arises as a by-product of divergence of all manner of genes (reviewed in ??). While the homeobox gene Ods induces sterility in male hybrids between *Drosophila simulans* and *D. mauritiana* (?), the oncogene *Xmrk* contributes to tumor formation in backcross Xiphophorus hybrids of both sexes (?), suggesting that sex-specific incompatibility is common but not universal.

Many of the iconic traits associated with intrinsic postzygotic isolation (?) affect reproduction, often showing sex-limited expression or different expression in male and female hybrids. Haldane’s (?) rule holds in a number of animals, including Drosophila, in which hybrid males tend to suffer much more severe post-zygotic incompatibility than do hybrid females (?, Chap. 8). Direct experiments have demonstrated that introgression of chromo-
osomal segments from *Drosophila mauritiana* into a *D. simulans* background (??) or into a *D. sechellia* background (?) induces male sterility at rates several-fold higher than female sterility or inviability in both sexes.

Numerous studies have described pervasive sex-specificity in gene regulation and expression (???). Table 1 of ? indicates that of the 4,776 coding sequences surveyed, over 61% showed sex-biased expression in *Drosophila melanogaster*. Further, they found that genes expressed at higher levels in males showed significantly greater divergence in expression levels between *D. melanogaster* and *D. simulans*.

Together with Haldane's (?) rule, the disproportionately large effect of the X-chromosome constitutes a major rule of speciation (?). Experiments introgressing small segments between Drosophila genomes indicate that a higher proportion of (hemizygous) introgressions on the X chromosome induce hybrid male sterility than do homozygous autosomal introgressions (???). A number of workers have provided discriminating assessments of leading proposals for evolutionary mechanisms that could contribute to the large-X effect (????). Whether this major trend reflects faster fixation of advantageous variants on hemizygous X chromosomes than on autosomes (??), accelerated divergence of genes affecting sex or reproduction (?), segregation distortion of sex chromosomes (?), or other processes, the large-X effect itself suggests that a substantial proportion of factors contributing to interspecific incompatibility may generate locus-specific barriers to neutral introgression of the kind studied here.

**Local adaptation within structured populations of conspecifics**

Incompatibility selection may arise not only as a consequence of interspecific hybridization, but also through divergence due to local adaptation of subpopulations of the same species. Table ?? indicates that autosomal incompatibility factors reduce introgression of mitochondrial markers by a factor of \( \sigma_f/(2 - \sigma_f) \) and of Y-linked markers by a factor of \( \sigma_m/(2 - \sigma_m) \). This finding suggests that sex-specific differences in expression of locally-adapted alleles might contribute to differences in divergence at mitochondrial and Y-linked markers (e.g., ???), even if propensity to migrate and mating success of migrants are inde-
Pendente of sex.

Polymorphisms for incompatibility factors within demes might be expected under incompatibility derived from local adaptation within species. In a separate work, we explore the effects of such characteristics on rates of introgression of neutral markers throughout the genome (?).

Acknowledgments

Public Health Service grant GM 37841 (MKU) provided partial funding for this research. We gratefully acknowledge support for the National Evolutionary Synthesis Center (NES-Cent) working group on Genomic Introgression, and the constructive comments of Associate Editor Dr. Naoyuki Takahata and the anonymous reviewers.

LITERATURE CITED


Coyne, J. A., and H. A. Orr, 1989 Two rules of speciation. In Speciation and its con-
sequences, edited by D. Otte and J. Endler. Sinauer Assoc., Inc., Sunderland, MA, 
180–270.


Oxford.

Fusco, D., and M. K. Uyenoyama, 2011 Effects of polymorphism in locally-adapted 
genes on rates of neutral introgression in structured populations. Theor. Pop. Biol. In 
press.

York.

Gavrilets, S., 1997 Hybrid zones with Dobzhansky-type epistatic selection. Evolution 51: 
1027–1035.

Hopkins Univ. Press, Baltimore, MD.

Haldane, J. B. S., 1922 Sex ratio and unisexual sterility in animal hybrids. J. Genetics 


Hernandez, R. D., 2008 A flexible forward simulator for populations subject to selection 

Hey, J., and R. Nielsen, 2004 Multilocus methods for estimating population sizes, mi-
gration rates and divergence time, with applications to the divergence of Drosophila pseud-


Schartl, M., 2008 Evolution of *Xmrk*: an oncogene, but also a speciation gene? BioEssays **30**: 822–832.


APPENDICES
Appendix 1: Relative reproductive rate

Because $C$, describing genetic transmission and selection (??), is non-negative, its spectral radius $\lambda$ corresponds to a non-negative characteristic value with modulus exceeded by no other eigenvalue and the eigenvector associated with $\lambda$ is non-negative (?, Chapter XIII, §3). It has a Schur decomposition of the form

$$C = QUQ^{-1},$$

for $U$ a triangular matrix with the eigenvalues of the transformation along the diagonal (see, for example, ?). For convenience, we choose $Q$ such that the eigenvalues appear along the diagonal in order of their moduli, with $\lambda$ the first element. For $t$ sufficiently large, $\lambda^t$ dominates

$$C^t = QU^tQ^{-1},$$

implying

$$\omega = \frac{v_m q_0}{v_r q_0}$$

(compare (??)), for $q_0$ the column or the sum of the columns of $Q$ that correspond to $\lambda$.

Appendix 2: Single autosomal incompatibility factor with an autosomal marker

To clarify the connection to earlier work, we address the case of an autosomal marker gene linked to a single autosomal incompatibility locus, for which a marker allele can reside on $4 = 2^2$ possible backgrounds.

Beginning with the distribution of backgrounds at the point of formation of the hybrid or resident offspring ($v_m$ or $v_r$ in (??)), selection first reduces the frequency of carriers of the foreign incompatibility factor, with female carriers surviving and reproducing at rate $\sigma_f$ and males at rate $\sigma_m$ relative to individuals bearing no foreign incompatibility allele. Further, we scale the contributions of males by the Fisherian male reproductive value $f/(1 - f)$. The selection matrix corresponds to

$$S = \begin{pmatrix} \sigma_f & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \sigma_m & 0 \\ 0 & 0 & 0 & \frac{f}{1-f} \end{pmatrix},$$

(A??1)
in which the ordering of the elements corresponds to females bearing background 1 and 0 at the incompatibility locus and males bearing these backgrounds. Carriers of the focal marker gene of either sex transmit it to female and male offspring at rate 1/2:

\[ \frac{1}{2}(f, 1 - f), \]

and transmission at the incompatibility locus given transmission of the marker gene is represented by

\[ F = \begin{pmatrix} 1 - r_f & r_f \\ 0 & 1 \\ 1 - r_m & r_m \\ 0 & 1 \end{pmatrix}, \]

in which \( (1 - r_f) \) represents the probability that a female bearing the foreign incompatibility factor transmits it together with the focal marker gene and \( (1 - r_m) \) the analogous quantity for male carriers. Joint transmission at the marker and incompatibility loci follows

\[ T = \frac{1}{2}(f, 1 - f) \otimes F, \]

for \( \otimes \) the Kronecker product.

In the absence of sex-specificity in rates of crossing-over (\( r_f = r_m = r \)) and in expression of incompatibility (\( \sigma_f = \sigma_m = \sigma \)), the transition matrix corresponds to

\[ C = ST = \left[ \begin{pmatrix} 1 \\ 0 \\ f \\ 1/f \end{pmatrix} \otimes \begin{pmatrix} \sigma & 0 \\ 0 & 1 \end{pmatrix} \right] \left[ \frac{1}{2}(f, 1 - f) \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 - r & r \\ 0 & 1 \end{pmatrix} \right] \]

\[ = \begin{pmatrix} 1 \\ 0 \\ f \\ 1/f \end{pmatrix} \left[ \frac{1}{2}(f, 1 - f) \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \sigma(1 - r) & \sigma r \\ 0 & 1 \end{pmatrix} \right]. \]

The rightmost matrix, representing incompatibility selection and the conditional transmission of the incompatibility factor, has a dominant eigenvalue of unity, the right eigenvector corresponding to which is

\[ \left( \frac{\sigma r}{1 - \sigma(1 - r)}, 1 \right). \]

The remaining matrix product, representing scaling of male contributions and transmission of the focal marker gene, has dominant eigenvalue \( f \), with right eigenvector

\[ \left( 1, \frac{f}{1 - f} \right). \]
As the Kronecker product of these matrices, \( C \) has dominant eigenvalue \( f \) (the product of the eigenvalues of the matrices) with right eigenvector
\[
q = \left( 1, \frac{f}{1-f} \right) \otimes \left( \frac{\sigma r}{1-\sigma(1-r)}, 1 \right).
\]
Relative reproductive rate (??) corresponds to
\[
\omega = \frac{(f, 0, 1-f, 0)q}{(0, f, 0, 1-f)q} = \frac{\sigma r}{1-\sigma(1-r)}. \tag{A??6}
\]
This expression, which has been derived numerous times (e.g., (A3) in ?), corresponds to the “gene flow factor” of ? and “barrier strength” of ?.

Under sex-specificity in selection (??) or transmission (??), the relative reproductive rates of female and male carriers correspond to
\[
\eta_f = \frac{\sigma_f [2r_f - \sigma_m (r_f - r_m)]}{2 - \sigma_f (1-r_f) - \sigma_m (1-r_m)},
\]
\[
\eta_m = \frac{\sigma_m [2r_m + \sigma_f (r_f - r_m)]}{2 - \sigma_f (1-r_f) - \sigma_m (1-r_m)},
\]
implying a relative reproductive rate at the marker locus of
\[
\omega = \frac{\eta_f + \eta_m}{2} = \frac{\bar{\sigma} \bar{r}}{1-\bar{\sigma}(1-\bar{r})}. \tag{A??7}
\]
for
\[
\bar{\sigma} = \frac{\sigma_f + \sigma_m}{2},
\]
\[
\bar{r} = \frac{\sigma_f r_f + \sigma_m r_m}{\sigma_f + \sigma_m}. \tag{A??8}
\]
(compare (??)).

In general, sex of the initial migrant and the incompatibility factor show epistasis with respect to the long-term contribution of the focal marker gene to future generations (compare (??)):
\[
\eta_f - \eta_m = \frac{2[\sigma_f r_f (1-\sigma_m) - \sigma_m r_m (1-\sigma_f)]}{2 - \sigma_f (1-r_f) - \sigma_m (1-r_m)}. \tag{A??9}
\]
No epistasis (\( \eta_f = \eta_m \)) arises in the absence of sex-specificity either in expression of incompatibility (\( \sigma_f = \sigma_m \)) or in recombination (\( r_f = r_m \)).
\((\sigma_f + \sigma_m), (\eta_f - \eta_m)\) increases with the excess viability or fertility of female over male carriers \((\sigma_f - \sigma_m)\). Similarly, for a given total magnitude of recombination \((r_f + r_m), (\eta_f - \eta_m)\) increases with the excess crossing-over in female over male carriers \((r_f - r_m)\).

**Appendix 3: Validation of the neutral approximation**

Among our main theses is that the evolutionary process at neutral marker loci in a genome containing incompatibility factors is well-approximated by a purely neutral process with a forward migration rate equal to our backward migration rate \((??)\). This section presents confirmation of this proposal, based on results of numerical simulations generated by SFS_CODE \((?)\), modified to track migration events along lineages at a neutral marker locus.

Our neutral approximation holds that the number of migration events along a lineage follows a Poisson distribution with parameter equal to the product of the number of generations observed and the per-generation backward migration rate \((??)\). This construction implies that the waiting time between migration events along a lineage has an exponential distribution with parameter equal to the backward migration rate. We assumed multiplicative effects of foreign incompatibility alleles, whether segregating at the same locus or at different loci.

We conducted forward simulations over \(t = 50,000\) generations, specifying a two-deme model with equal population sizes \((N_1 = N_2 = 1,000\) genes, or 500 diploid individuals) and symmetric forward migration rates \((N_1m_{12} = N_2m_{21} = 0.1)\). We assumed that the foreign incompatibility allele at each of two autosomal incompatibility loci reduces the viability of its carrier by a factor of \(\sigma = 0.65\) and that migrants compete for mates on an equal basis with residents. In the absence of linkage among any of the two incompatibility loci and an autosomal marker locus, these assumptions imply a relative reproductive rate \((??)\) of \(\omega = 0.4815\) and a genealogical migration rate \((??)\) of \(g = 4.8148 \times 10^{-5}\) at the neutral marker. A total of 36,233 migration events over 15,000 independent replicates showed an excellent fit \((X^2 = 3.946, 9\) df) to expectations under a Poisson distribution with parameter \(gt = 2.4074\).
Because relative reproductive rate \( \frac{w_1}{w_2} \) represents a limit for large numbers of generations and negligible frequencies of foreign incompatibility alleles, one might expect discrepancies to arise in cases with forward migration rates sufficiently high to permit the migration of an individual carrying a foreign incompatibility allele back into the other species. To explore the limits of our model, we repeated the simulation experiment under forward migration rates an order of magnitude higher \( (N_1m_{12} = N_2m_{21} = 1.0) \). We found that 7.5% of migration events traced backward along a lineage involved backgrounds that differed from the pure resident background, compared to 0.087% for the lower forward migration rate. The distribution of the number of migration events along a random lineage departs significantly from a Poisson distribution \( (X^2 = 90.4898, 32 \text{ df}) \): the observed mean \( (24.451) \) exceeds the expectation \( (b_t = 24.074) \) and the kurtosis value indicates excess mass in the tails. To provide a quantitative summary, we fit the observations to a Neyman’s Type A distribution, a compound distribution which assumes that the Poisson parameter is itself a Poisson-distributed random variable:

\[
\lambda / \phi \sim \text{Poisson}(\mu),
\]

for \( \phi \) an index of clumping (see ?, Chapter 9). The observations show an excellent fit \( (X^2 = 19.898, 24 \text{ df}) \) to a Neyman’s Type A distribution with mean set equal to the observed mean and a least-squares estimate of the clumping parameter \( \phi \) of 0.0141, obtained by minimizing the squared deviation between the observed variance up to the ninth moment and their expectations, using the probability generating function given in ?.

**Appendix 4: X-linkage**

We address the case of X-linkage of both the marker locus and a single incompatibility locus in an organism exhibiting X-Y sex determination with no crossing-over on the X chromosome in hemizygous males.

As in the autosomal case, female carriers transmit the focal marker gene symmetrically to female and male offspring:

\[
\frac{1}{2} (f, 1-f) \otimes \begin{pmatrix} 1 - r_f & r_f \\ 0 & 1 \end{pmatrix}
\]
In contrast, male carriers transmit the marker gene to all daughters, without recombination, and to no sons, with the last two rows of $T$ corresponding to

$$
\begin{pmatrix}
  f & 0 & 0 & 0 \\
  0 & f & 0 & 0
\end{pmatrix}
$$

(compare (??)).

For a female migrant, the relative reproductive rate (??) becomes

$$
\omega_X^{f} = \frac{(f, 0, 1 - f, 0)q}{(0, f, 0, 1 - f)q} = \frac{(\eta_f + \eta_m)/2 = \frac{2\sigma_f r_f (1 + \sigma_m)}{2 - \sigma_f (1 + \sigma_m)(1 - r_f)},}
$$

in which

$$
\eta_f = \frac{2\sigma_f r_f}{2 - \sigma_f (1 + \sigma_m)(1 - r_f)},
$$

$$
\eta_m = \frac{2\sigma_f \sigma_m r_f}{2 - \sigma_f (1 + \sigma_m)(1 - r_f)}.
$$

These expressions imply epistasis (??) between sex and the X-linked incompatibility locus proportional to

$$
\eta_f - \eta_m = \frac{2\sigma_f r_f (1 - \sigma_m)}{2 - \sigma_f (1 + \sigma_m)(1 - r_f)}
$$

(compare (??)). This expression indicates a positive interaction between femaleness and the foreign incompatibility factor under incomplete linkage ($r_f > 0$), regardless of whether male or female carriers suffer greater effects of incompatibility (sign of ($\sigma_f - \sigma_m$)). This association may reflect higher recombination rates in females ($r_f > r_m = 0$).

For a male migrant, only female descendants receive the focal marker allele, implying a relative reproductive rate of

$$
\omega_X^{m} = \frac{(f, 0, 0, 0)q}{(0, f, 0, 0)q} = \frac{2\sigma_f r_f}{2 - \sigma_f (1 + \sigma_m)(1 - r_f)}.
$$

Appendix 5: Epistasis among incompatibility factors

Here we describe a measure of epistasis among incompatibility factors (compare ?), applied to the relative reproductive rate (??) for an arbitrary number of incompatibility factors.
To establish (??), we expand
\[
e_{\Omega[k]} = E\left[ \prod_{i \in \Omega[k]} (A_i - \omega_i) \right]
\]
\[
= \sum_{j=0}^{k} (-1)^{k-j} \sum_{\Omega[j]} E\left[ \prod_{i \in \Omega[j]} A_i \right] \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b
\]
\[
= \sum_{j=1}^{k} (-1)^{k-j} \sum_{\Omega[j]} \left( E\left[ \prod_{i \in \Omega[j]} A_i \right] - \prod_{i \in \Omega[j]} \omega_i \right) \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b + \prod_{i \in \Omega[k]} \omega_b \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j}
\]
\[
= \sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega[j]} \delta_{\Omega[j]} \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b,
\]
(A??.1)
for \( \Omega[j] \) \((k \geq 2)\) a set of \( k \) indices representing \( k \) incompatibility loci, \( \Omega[j] \) a set of \( j \) indices in \( \Omega[k] \) \((j \leq k)\), \( \Omega[k] \setminus \Omega[j] \) denoting the set of indices in \( \Omega[k] \) but not in \( \Omega[j] \), and \( \delta_{\Omega[j]} \) representing the departure of relative reproductive rate from multiplicativity (??).

For relative reproductive rate corresponding to (??), we address the relationship between the overall departure from multiplicativity (??) and the sex-specific measures (??). Using that \( \eta_{f,i} = \omega_i + \Delta_i \) and \( \eta_{m,i} = \omega_i - \Delta_i \) for \( \Delta_i = (\eta_{f,i} - \eta_{m,i})/2 \) (??), we have
\[
\frac{1}{2} \left( \prod_{i \in \Omega[j]} \eta_{f,i} + \prod_{i \in \Omega[j]} \eta_{m,i} \right) = \frac{1}{2} \left[ \prod_{i \in \Omega[j]} (\omega_i + \Delta_i) + \prod_{i \in \Omega[j]} (\omega_i - \Delta_i) \right]
\]
(A??.2)
\[
= \prod_{i \in \Omega[j]} \omega_i + \xi_{\Omega[j]},
\]
for \( \xi_{\Omega[j]} \) representing terms of the form \( \Delta_i \). Note that
\[
\xi_{\Omega[j]} = \delta_{\Omega[j]} - (\delta_{f,\Omega[j]} + \delta_{m,\Omega[j]})/2.
\]
In the expansion of (??), terms involving an odd number of the \( \Delta_i \)s cancel out, while all the terms with an even number appear twice, implying
\[
\xi_{\Omega[k]} = \sum_{i=1}^{[k/2]} \sum_{\Omega[2i] \subseteq \Omega[k]} (\prod_{a \in \Omega[2i]} \Delta_a \prod_{c \in \Omega[k] \setminus \Omega[2i]} \omega_c),
\]
for \( \Omega[2i] \) representing all subsets of \( \Omega[j] \) that contain \( 2i \) elements. We rewrite (??) as
\[
e_{\Omega[k]} = \sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega[j]} \left( \frac{\delta_{f,\Omega[j]} + \delta_{m,\Omega[j]}}{2} + \xi_{\Omega[j]} \right) \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b,
\]
(A??.3)
We find that
\[
\sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega[j]} \xi_{\Omega[j]} \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b = \begin{cases} 
0 & \text{for } k \text{ odd} \\
\prod_{i=1}^{k} \Delta_i & \text{for } k \text{ even}
\end{cases}
\]
For \(k\) odd, this sum corresponds to
\[
\sum_{j=2}^{k} (-1)^{k-j} \sum_{\Omega[j]} \sum_{i=1}^{\lfloor j/2 \rfloor} \sum_{\Omega[2i] \subset \Omega[j]} \left( \prod_{a \in \Omega[2i]} \Delta_a \prod_{c \in \Omega[j] \setminus \Omega[2i]} \omega_c \right) \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b
\]
\[
= \sum_{j=2}^{k} (-1)^{k-j} \sum_{i=1}^{\lfloor j/2 \rfloor} \left( \frac{k}{j-2i} \right) \sum_{\Omega[2i]} \prod_{a \in \Omega[2i]} \Delta_a \prod_{b \in (\Omega[k] \setminus \Omega[2i])} \omega_b
\]
\[
= \sum_{i=1}^{\lfloor k/2 \rfloor} \prod_{a \in \Omega[2i]} \Delta_a \prod_{b \in (\Omega[k] \setminus \Omega[2i])} \omega_b \sum_{j=2i}^{k} \left( \frac{k}{j-2i} \right) (-1)^{k-j}
\]
\[
= 0,
\]
and for \(k\) even,
\[
\prod_{i=1}^{k} \Delta_i + \sum_{j=2}^{k-1} (-1)^{k-j} \sum_{\Omega[j]} \xi_{\Omega[j]} \prod_{b \in (\Omega[k] \setminus \Omega[j])} \omega_b = \prod_{i=1}^{k} \Delta_i.
\]
Substitution of these expressions into (??) produces (??).

Appendix 6: Epistasis between an X-linked and an autosomal factor

**X-linked marker (male migrant):** Because only female offspring of a male migrant receive the X-linked marker,
\[
\omega_{AX}^{X,m} = \eta_{f,AX}^X,
\]
with overall epistasis equal to the female-specific epistasis:
\[
e_{AX}^{X,m} = \eta_{f,AX}^X - \eta_{f,A}^X \eta_{f,x}^X = \delta_{f,AX}^X,
\]
for \(\delta_{f,AX}^X\) given in (??a). As in the case of the female migrant, the sign of epistasis depends on \(\Delta_{A}^X\) (??).

**Autosomal marker:** For a female migrant, epistasis (??) between the incompatibility loci corresponds to
\[
e_{AX}^{A,m} = \omega_{AX}^A - \omega_A^A \omega_X^A = (\delta_{f,AX}^A + \delta_{m,AX}^A) / 2 + \Delta_A^A \Delta_X^A,
\]
(A??.1)
for $\Delta_A^X$ given by (??) and
\[
\Delta_A^X = (\eta_{f,X}^A - \eta_{m,X}^A)/2 = \frac{2(\sigma_{f,X} - \sigma_{m,X})}{8 - \sigma_{f,X}(2 + \sigma_{m,X})}
\]  \hspace{1cm} (A??.2)

(Table ??). Within-sex interactions correspond to
\[
\delta_{f,AX}^A = \eta_{f,AX}^A - \eta_{f,A}^A \eta_{f,X}^A = \frac{2\sigma_{f,AX}(1 - r_{f,A})D_f\Delta_A^X}{[8 - \sigma_{f,X}(2 + \sigma_{m,X})]\Gamma}, \hspace{1cm} (A??.3a)
\]
\[
\delta_{m,AX}^A = \eta_{m,AX}^A - \eta_{m,A}^A \eta_{m,X}^A = \frac{2\sigma_{m,AX}(1 - r_{m,A})D_m\Delta_A^X}{[8 - \sigma_{f,X}(2 + \sigma_{m,X})]\Gamma}, \hspace{1cm} (A??.3b)
\]
in which
\[
D_f = 4(\sigma_{f,X} - \sigma_{m,X}) - \sigma_{m,A}\sigma_{m,X}(1 - r_{m,A})[4 - \sigma_{f,X}(3 + \sigma_{m,X})]
\]
\[
D_m = 2\sigma_{f,AX}(1 - r_{f,A})(\sigma_{f,X} - \sigma_{m,X}) - [4 - \sigma_{f,X}(3 + \sigma_{m,X})][4 - \sigma_{f,AX}(1 - r_{f,A})]
\]
\[
\Gamma = 8 - \sigma_{f,AX}(1 - r_{f,A})[2 + \sigma_{m,A}\sigma_{m,X}(1 - r_{m,A})], \hspace{1cm} (A??.4)
\]
giving an overall two-way epistasis between the incompatibility factors of (??). Comparison of (??) and (??) suggests that $D_f$ and $D_m$ assume the role of $\Delta_X$, an indicator of the nature of the interaction between sex and the X-linked incompatibility factor.

As indicated in the text following (??), $D_m$ is always nonpositive, and zero only if the X-linked factor does not contribute to incompatibility selection ($\sigma_{f,X} = \sigma_{m,X} = 1$). In contrast, $D_f$ is positive for
\[
\sigma_{f,X} > \sigma_{m,X}\{1 + \sigma_{m,A}(1 - r_{m,A})[1 - \sigma_{f,X}(3 + \sigma_{m,X})/4]\} \geq \sigma_{m,X}. \hspace{1cm} (A??.5)
\]

Sufficient conditions for positive epistasis $e_{AX}^A$ (higher introgression) include that both factors cause greater impairment to female carriers ($\sigma_{f,A} < \sigma_{m,A}$ and $\sigma_{f,X} < \sigma_{m,X}$) or that both cause sufficiently greater impairment to males, with positive $\Delta_A$ and (??) satisfied. In the absence of sex-specificity in expression of incompatibility due to the X-linked factor ($\sigma_{f,X} = \sigma_{m,X}$), $D_f$ is nonpositive, implying either no epistasis ($e_{AX}^A = 0$) or
\[
e_{AX}^A \propto -\Delta_A = -(r_{f,A} - r_{m,A}),
\]
indicating negative epistasis (less introgression) if the map distance between the autosomal marker and the autosomal factor is greater in females.
For a male migrant, all female offspring and none of the male offspring of the migrant carry the X-linked factor, while offspring of both sexes carry the autosomal factor:

\[
\omega_{AX}^{A,m} = \frac{(f,0,0,1-f,0,0,0,0)q}{(0,0,0,0,0,0,f,1-f)q} = (\eta_{f,AX}^{A} + \eta_{m,A}^{A})/2.
\]

Epistasis between the loci with respect to relative reproductive rate is given by

\[
e_{AX}^{A,m} = \delta_{f,AX}^{A}/2 + \Delta_{A}^{A}(\eta_{f,X}^{A} - 1)/2,
\]

with \(\Delta_{A}^{A}\) given by (??), the within-female departure from multiplicativity \(\delta_{f,AX}^{A}\) by (??a), and

\[
\eta_{f,X}^{A} - 1 = \frac{-2[4 - \sigma_{f,X}(3 + \sigma_{m,X})]}{8 - \sigma_{f,X}(2 + \sigma_{m,X})}.
\]

As in the case of a female migrant (??), epistasis arises only if the autosomal factor is sex-specific (\(\Delta_{A}^{A} \neq 0\)). Non-zero \((\eta_{f,X} - 1)\) implies that the X-linked factor contributes to incompatibility \((\sigma_{f,X} < 1\) or \(\sigma_{m,X} < 1\)). For a male migrant, the overall value for epistasis between the autosomal and X-linked incompatibility factors is

\[
e_{AX}^{A,m} = \frac{2D_{m}\Delta_{A}^{A}}{[8 - \sigma_{f,X}(2 + \sigma_{m,X})]\Gamma},
\]

for \(D_{m}\) and \(\Gamma\) given in (??). As \(D_{m}\) is negative if the X-linked factor contributes to incompatibility \((\sigma_{f,X} < 1\) or \(\sigma_{m,X} < 1\)), this expression indicates that the autosomal and X-linked factors show positive epistasis (higher introgression rates than expected under multiplicativity) if the autosomal factor permits greater reproductive rates through male than female carriers \((\Delta_{A}^{A} < 0)\) and negative epistasis if female carriers show higher reproductive rates.
Table ??: One-factor reproductive values

<table>
<thead>
<tr>
<th>Position&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Factor</th>
<th>Marker</th>
<th>Relative contribution&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Factor-sex association&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Relative reproductive rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\eta_f$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\eta_m$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td></td>
<td>$\frac{\sigma_f(r_f - \sigma_m(r_f - r_m))/2}{1-\sigma(r_f - r_m)}$</td>
<td>$\frac{\sigma_m(r_m + \sigma_f(r_f - r_m))/2}{1-\sigma(r_f - r_m)}$</td>
<td>$\frac{\sigma_f(r_f - \sigma_m(r_f - r_m))}{1-\sigma(r_f - r_m)}$</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td></td>
<td>$\frac{\sigma_f(4+\sigma_m)}{8-\sigma_f(2+\sigma_m)}$</td>
<td>$\frac{\sigma_m(4+\sigma_f)}{8-\sigma_f(2+\sigma_m)}$</td>
<td>$\frac{4(\sigma_f - \sigma_m)}{8-\sigma_f(2+\sigma_m)}$</td>
</tr>
<tr>
<td>A</td>
<td>Y</td>
<td></td>
<td>$\sigma_m$</td>
<td></td>
<td>$\sigma_m$</td>
</tr>
<tr>
<td>A</td>
<td>mt</td>
<td></td>
<td>$\frac{\sigma_f}{2-\sigma_f}$</td>
<td></td>
<td>$\frac{\sigma_f}{2-\sigma_f}$</td>
</tr>
<tr>
<td>X</td>
<td>A</td>
<td></td>
<td>$\frac{\sigma_f(4+\sigma_m)}{8-\sigma_f(2+\sigma_m)}$</td>
<td>$\frac{\sigma_m(4+\sigma_f)}{8-\sigma_f(2+\sigma_m)}$</td>
<td>$\frac{2(\sigma_f + \sigma_m) + \sigma_f \sigma_m}{8-\sigma_f(2+\sigma_m)}$</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td></td>
<td>$\frac{2\sigma_f r_f}{2-\sigma_f(1+\sigma_m)(1-r_f)}$</td>
<td>$\frac{2\sigma_f \sigma_m r_f}{2-\sigma_f(1+\sigma_m)(1-r_f)}$</td>
<td>$\frac{\sigma_f(1+\sigma_m)r_f}{2-\sigma_f(1+\sigma_m)(1-r_f)}$</td>
</tr>
<tr>
<td>X</td>
<td>Y</td>
<td></td>
<td>$\sigma_m$</td>
<td></td>
<td>$\sigma_m$</td>
</tr>
<tr>
<td>X</td>
<td>mt</td>
<td></td>
<td>$\frac{\sigma_f}{2-\sigma_f}$</td>
<td></td>
<td>$\frac{\sigma_f}{2-\sigma_f}$</td>
</tr>
<tr>
<td>mt</td>
<td>A</td>
<td></td>
<td>$\frac{\sigma_f \sigma_m}{2-\sigma_f}$</td>
<td>$\sigma_m$</td>
<td>$\frac{-2\sigma_m(1-\sigma_f)}{2-\sigma_f}$</td>
</tr>
<tr>
<td>mt</td>
<td>X</td>
<td></td>
<td>$\frac{\sigma_f \sigma_m}{2-\sigma_f}$</td>
<td>$\sigma_m$</td>
<td>$\frac{-2\sigma_m(1-\sigma_f)}{2-\sigma_f}$</td>
</tr>
<tr>
<td>mt</td>
<td>Y</td>
<td></td>
<td>$\sigma_m$</td>
<td></td>
<td>$\sigma_m$</td>
</tr>
<tr>
<td>Y</td>
<td>A</td>
<td></td>
<td>$\sigma_m$</td>
<td></td>
<td>$\frac{-2(1-\sigma_m)}{2-\sigma_m}$</td>
</tr>
<tr>
<td>Y</td>
<td>X</td>
<td></td>
<td>$\sigma_m$</td>
<td></td>
<td>$\sigma_m$</td>
</tr>
<tr>
<td>Y</td>
<td>mt</td>
<td></td>
<td>$\sigma_m$</td>
<td></td>
<td>$\sigma_m$</td>
</tr>
</tbody>
</table>

<sup>a</sup> Genomic location: autosomal (A), X-linked (X), Y-linked (Y), mitochondrial (mt)

<sup>b</sup> Relative contribution to future generations of the marker held by a female ($\eta_f$) or male ($\eta_m$) carrier of the factor

<sup>c</sup> $\eta_f - \eta_m$; compare (??)

<sup>d</sup> $\bar{r}$ and $\bar{\sigma}$ defined in (??)
FIGURE CAPTIONS

Figure ??.— Joint relative reproductive rate at a neutral autosomal marker locus induced by two autosomal factors in the absence of physical linkage. Each of the horizontal axes represents the fitness (relative viability or fertility) of female carriers of an autosomal factor ($σ_{f,i}, i = 1, 2$), with the fitness of a male carrier given by $σ_{m,i} = 1 - σ_{f,i}$. The saddle-shaped surface represents the relative reproductive rate at the marker locus induced by the two factors jointly and the relatively flat surface the rate under the multiplicative expectation. Heavy lines indicate the intersection between the actual and expected barriers.

Figure ??.— Simulated distribution of the number of migration events traced back along the line of descent of a randomly selected gene at the autosomal neutral marker locus in a genome containing two autosomal incompatibility loci, under multiplicative fitnesses and no linkage among any of the loci. For each of 15,000 independent replicates using a modified version of SFS_CODE (?), we simulated 100,000 generations forward in time and then counted the number of migration events in a randomly sampled line of descent traced backward in time. Each of the two populations comprised 1,000 genes (500 diploid individuals), with the proportion of migrant genes each generation set to $m_{12} = m_{21} = 1 \times 10^{-4}$. In the left panel, the foreign allele at the incompatibility loci reduced the fitness of their heterozygous male carriers ($σ_{m,1} = σ_{m,2} = 0.3$), with no effect on female carriers ($σ_{f,1} = σ_{f,2} = 1$). The observed counts (histogram) showed an excellent match ($X^2 = 12.1, 11$ df) to the Poisson distribution determined from our approximation (??) of the parameter of the exponential waiting time between migration events ($g = 3.096$, red dots). In contrast, the distribution predicted by assuming multiplicative barriers (black line) showed a poor fit ($X^2 = 4924, 9$ df), predicting too few migration events ($g = 2.318$). In the right panel, the histogram shows the counts under incompatibility due to loci with discordant relative effects on the sexes: the foreign allele at one locus caused greater detriment to males than females ($σ_{m,1} = 0.3, σ_{f,1} = 1$), with opposite effects at the other locus ($σ_{m,2} = 1, σ_{f,2} = 0.3$). Our approximation ($g = 1.5916$, red dots) again fit the observations well ($X^2 = 6.3, 7$ df), contrary to the multiplicative
prediction \( (g = 2.318, X^2 = 3758, 7 \text{ df}) \). Discordance between the incompatibility loci on the relative impairment of males and females induces a greater barrier (fewer migration events) than expected under multiplicativity.

**Figure 2:—** Relative reproductive rates at a neutral marker locus induced by incompatibility factors at various genomic locations for female or male migrants. Bars indicate the relative reproductive rate \( (??) \) of a neutral marker located on an autosome \( (A, \text{ blue}) \) unlinked to any incompatibility locus \( (r_{f,A} = r_{m,A} = 1/2) \), on the X-chromosome \( (X, \text{ cyan}) \), the mitochondria \( (\text{mt, orange}) \), or the Y-chromosome \( (Y, \text{ magenta}) \). Locations of one or more incompatibility factors are indicated on the abscissa (autosome only \( (A) \), autosome and X-chromosome \( (A+X) \), etc.). For X-linkage of both factor and marker, we assume free crossing-over in females \( (r_{f,X} = 1/2) \) and none in the hemizygous males \( (r_{m,X} = 0) \). To provide a basis for comparison, we assumed that the factors induce a common total intensity of incompatibility selection, constraining \( \sigma_{f} + \sigma_{m} = 1 \). Female carriers of the autosomal factor have viability \( \sigma_{f,A} = 0.6 \) and male carriers \( \sigma_{m,A} = 0.4 \), and for the X-linked factor, \( \sigma_{f,X} = 0.8 \) and \( \sigma_{m,X} = 0.2 \). For the Y-linked incompatibility locus, \( \sigma_{m,Y} = 0.5 \).
\[ P_s = (n - 1) \theta + M(n - 1) \theta^2 + \theta(n - 1 + Mn) + m \]

\[ P_d = M(n - 1) \theta^2 + \theta(n - 1 + Mn) + m \]

\[ M = 4 Nm_b \]

\[ \theta = 4 ... = \sigma r \]

\[ P(k) = e^{-m_b T(m_b T)^k} k! \]

\[ \sigma f, 1 + \sigma m, 1 = 1 \]

\[ \sigma f, 2 + \sigma m, 2 = 1 \]

\[ \omega_{12} = \omega_1 \omega_2 \]

**Figure 1**: Joint relative reproductive rate at a neutral autosomal marker locus induced by two autosomal factors in the absence of physical linkage. Each of the horizontal axes represents the fitness (relative viability or fertility) of female carriers of an autosomal factor \( (\sigma_{f,i}, i = 1, 2) \), with the fitness of a male carrier given by \( \sigma_{m,i} = 1 - \sigma_{f,i} \). The saddle-shaped surface represents the relative reproductive rate at the marker locus induced by the two factors jointly and the relatively flat surface the rate under the multiplicative expectation, with the heavy lines indicating the intersection between the actual and expected barriers.
Figure 2: Simulated distribution of the number of migration events traced back along the line of descent of a randomly selected gene at the autosomal neutral marker locus in a genome containing two autosomal incompatibility loci, under multiplicative fitnesses and no linkage among any of the loci. For each of 15,000 independent replicates using a modified version of SFS_CODE (?), we simulated 100,000 generations forward in time and then counted the number of migration events in a randomly sampled line of descent traced backward in time. Each of the two populations comprised 1,000 genes (500 diploid individuals), with the proportion of migrant genes each generation set to \( m_{12} = m_{21} = 1 \times 10^{-4} \). In the left panel, the foreign allele at the incompatibility loci reduced the fitness of their heterozygous male carriers (\( \sigma_{m,1} = \sigma_{m,2} = 0.3 \)), with no effect on female carriers (\( \sigma_{f,1} = \sigma_{f,2} = 1 \)). The observed counts (histogram) showed an excellent match (\( X^2 = 12.1, 11 \text{ df} \)) to the Poisson distribution determined from our approximation (?) of the parameter of the exponential waiting time between migration events (\( g = 3.096, \) red dots). In contrast, the distribution predicted by assuming multiplicative barriers (black line) showed a poor fit (\( X^2 = 4924, 9 \text{ df} \)), predicting too few migration events (\( g = 2.318 \)). In the right panel, the histogram shows the counts under incompatibility due to loci with discordant relative effects on the sexes: the foreign allele at one locus caused greater detriment to males than females (\( \sigma_{m,1} = 0.3, \sigma_{f,1} = 1 \)), with opposite effects at the other locus (\( \sigma_{m,2} = 1, \sigma_{f,2} = 0.3 \)). Our approximation (\( g = 1.5916, \) red dots) again fit the observations well (\( X^2 = 6.3, 7 \text{ df} \)), contrary to the multiplicative prediction (\( g = 2.318, X^2 = 3758, 7 \text{ df} \)). Discordance between the incompatibility loci on the relative impairment of males and females induces a greater barrier (fewer migration events) than expected under multiplicativity.
Figure 3: Relative reproductive rates at a neutral marker locus induced by incompatibility factors at various genomic locations for female or male migrants. Bars indicate the relative reproductive rate ($\omega$) of a neutral marker located on an autosome ($A$, blue) unlinked to any incompatibility locus ($r_{f,A} = r_{m,A} = 1/2$), on the X-chromosome ($X$, cyan), the mitochondria ($mt$, orange), or the Y-chromosome ($Y$, magenta). Locations of one or more incompatibility factors are indicated on the abscissa (autosome only ($A$), autosome and X-chromosome ($A+X$), etc.). For X-linkage of both factor and marker, we assume free crossing-over in females ($r_{f,X} = 1/2$) and none in the hemizygous males ($r_{m,X} = 0$). To provide a basis for comparison, we assumed that the factors induce a common total intensity of incompatibility selection, constraining $\sigma_{f} + \sigma_{m} = 1$. Female carriers of the autosomal factor have viability $\sigma_{f,A} = 0.6$ and male carriers $\sigma_{m,A} = 0.4$, and for the X-linked factor, $\sigma_{f,X} = 0.8$ and $\sigma_{m,X} = 0.2$. For the Y-linked incompatibility locus, $\sigma_{m,Y} = 0.5$. 