

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

Diana Fusco* and Marcy K. Uyenoyama^{†,1}

*Computational Biology and Bioinformatics Program, Duke University, Durham, North Carolina 27708-0090 and [†]Department of Biology, Duke University, Durham, North Carolina 27708-0338

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, ~850,000 years ago (Leman *et al.* 2005), 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 (Noor *et al.* 2001), and in general, the number of incompatible interactions among loci is expected to “snowball” as divergence time increases (Orr 1995; Orr and

Turelli 2001). 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.*, Barton and Bengtsson 1986) 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 (Rosenberg *et al.* 2002; Coop *et al.* 2009). Humans show striking locus-specific F_{ST} , the classical index of population structure (Akey *et al.* 2002; Holsinger and Weir 2009), with many regions showing extreme values beyond three empirical standard deviations

Copyright © 2011 by the Genetics Society of America

doi: 10.1534/genetics.111.130732

Manuscript received May 14, 2011; accepted for publication June 19, 2011

Available freely online through the author-supported open access option.

Supporting information is available online at <http://www.genetics.org/content/suppl/2011/06/24/genetics.111.130732.DC1>.

¹Corresponding author: Department of Biology, Box 90338, Duke University, Durham, NC 27708-0338. E-mail: marcy@duke.edu

from the mean for a chromosome; (Weir *et al.* 2005). This pattern is observed in a wide variety of nonhuman species as well (*e.g.*, Rieseberg *et al.* 1999; Machado and Hey 2003; Teeter *et al.* 2009).

The widely used IM program (isolation with migration, Hey and Nielsen 2004, 2007) 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. Hey and Nielsen (2004) applied their method separately to each of 14 markers in *D. 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 (Wang *et al.* 1997) 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 (Barton and Hewitt 1985). Incompatibility is expected to impede introgression throughout the genome, but more intensely in regions harboring incompatibility factors (Bengtsson 1985; Barton and Bengtsson 1986; Navarro and Barton 2003). 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*, Rieseberg *et al.* (1999) 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 (Strasburg *et al.* 2009). Lower introgression rates in the vicinity of known incompatibility factors have been detected between house mice *Mus musculus musculus* and *M. m. domesticus* (Payseur *et al.* 2004; Payseur and Nachman 2005), among host races of pea aphids (Via 2009), and between *D. pseudoobscura* and *D. persimilis* (Kulathinal *et al.* 2009).

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 Schlötterer 2003; Beaumont and Balding 2004; Storz 2005). This view holds that demographic structure should affect the pattern of variation in a uniform manner, with locus specificity indicative of selection (*e.g.*, Akey *et al.* 2002). 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.*, Charlesworth *et al.* 1997).

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 nontransient, maintained by a balance between negative selection and introgression.

Barriers to neutral gene flow

Bengtsson (1985) 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. Barton and Bengtsson (1986) characterized reproductive barriers due to multiple, epistatic incompatibility factors. Navarro and Barton (2003) 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 in 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 (Haldane 1957; Hill and Robertson 1966; Barton 1995) 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 Bengtsson (1985), 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 1, 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 (Karlin 1982).

Selection can change the fundamental structure of the coalescent process at a target of selection (Krone and Neuhauser 1997; Neuhauser and Krone 1997) 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 absence of selection (Birky and Walsh 1988), 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.*, Barton 1979; Gavrillets 1997; Kobayashi *et al.* 2008).

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

$$g = m\omega, \quad (1)$$

for g the genealogical migration rate in the deme from which the gene currently resides, m the backward migration

rate, and ω 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 σ -values of those alleles.

As in earlier work (*e.g.*, Bengtsson 1985), 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 (Fusco and Uyenoyama 2011).

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}):

$$\mathbf{C} = \mathbf{ST}. \quad (2)$$

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 (Fisher 1930). 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 $\lim_{t \rightarrow \infty} \mathbf{vC}^t \mathbf{e}$, 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

$$\omega = \lim_{t \rightarrow \infty} \frac{\mathbf{v}_m \mathbf{C}^t \mathbf{e}}{\mathbf{v}_r \mathbf{C}^t \mathbf{e}}, \quad (3)$$

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. Kobayashi *et al.* (2008) 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 λ of \mathbf{C} comes to dominate (Equation 3). For diagonalizable \mathbf{C} , λ corresponds to a simple, positive root of the characteristic equation and the set of eigenvectors that span the space (*Appendix A* addresses more general cases). Transmission of the focal marker genes through t generations is determined from

$$\mathbf{C}^t = \mathbf{QD}^t\mathbf{Q}^{-1},$$

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

$$\omega = \frac{\mathbf{v}_m\mathbf{q}}{\mathbf{v}_r\mathbf{q}}, \quad (4)$$

for \mathbf{q} the right eigenvector associated with the dominant eigenvalue λ of \mathbf{C} . In keeping with the Fisherian notion, we term \mathbf{q} (Equation 4) 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 in *Supporting Information*, [File S1](#) and [File S2](#), the matrices reflect an ordering first of females bearing backgrounds containing the maximum number of foreign (Equation 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 (Equation 4) corresponds to

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

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 noncarrier female or male ($\eta_{f,0} = \eta_{m,0} = 1$), $\eta_{m,1}$ is the analogous quantity for a male carrier, and the prime is the transpose. Relative reproductive rate (Equation 4) corresponds to

$$\omega = \frac{(f, 0, 1-f, 0)\mathbf{q}}{(0, f, 0, 1-f)\mathbf{q}} = \frac{\eta_{f,1} + \eta_{m,1}}{2}$$

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 B* explicitly presents the matrices and the reproductive value vector \mathbf{q} (Equation 4) for the case of an autosomal marker gene linked to a single autosomal incompatibility locus.

We use ω (Equation 4), defined in terms of contribution to future generations, to approximate the backward-perspectived genealogical migration rate g (Equation 1). 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 (Figure 2 and *Appendix C*). The Mathematica notebook `matrix_builder.nb` ([File S1](#) and [File S2](#)) constructs transmission matrix \mathbf{C} and produces the reproductive value vector \mathbf{q} (Equation 4) 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 (Equation 4) in terms of interactions among incompatibility factors and sex.

Epistasis: Expression (4) 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 ω (Equation 3) at a neutral marker locus induced by multiple incompatibility factors.

We view reproductive value vector \mathbf{q} (Equation 4) 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 analog of multilocus disequilibrium as proposed by Bennett (1954), in which, for example, the two-way association corresponds to a covariance:

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

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,

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

with

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

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

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

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

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

$$e_{\Omega_{[k]}^{A,f}} = e_{\Omega_{[k]}^{A,m}} = \sum_{j=2}^k (-1)^{k-j} \sum_{\Omega_{[j]} \subseteq \Omega_{[k]}} \delta_{\Omega_{[j]}} \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b, \quad (5)$$

for $\Omega_{[k]}$ a set of k indexes representing incompatibility loci, $\Omega_{[j]}$ a subset of $\Omega_{[k]}$ comprising j indexes, $\Omega_{[k]} \setminus \Omega_{[j]}$ the set of indexes in $\Omega_{[k]}$ but not in $\Omega_{[j]}$, and

$$\delta_{\Omega_{[j]}} = \omega_{\Omega_{[j]}} - \prod_{i \in \Omega_{[j]}} \omega_i \quad (6)$$

the departure of relative reproductive rate from multiplicativity. This partitioning of k -way epistasis includes components for all subsets of $2, \dots, k$ 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 $E[\dots A_j \dots]$ and $E[\dots \bar{A}_j \dots]$.

We apply this partitioning of epistasis to relative reproductive rate (Equation 4) 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

$$\omega_{\Omega_{[k]}^{A,f}} = \omega_{\Omega_{[k]}^{A,m}} = \frac{\eta_{f,\Omega_{[k]}} + \eta_{m,\Omega_{[k]}}}{2}, \quad (7)$$

reflecting the equal transmission of the marker through the female and male hybrid offspring of migrants of either sex. Expression (7) also applies to an X-linked marker introduced by a female migrant ($\omega_{\Omega_{[k]}^{X,f}}$); for a male migrant, $\omega_{\Omega_{[k]}^{X,m}} = \eta_{f,\Omega_{[k]}}$, reflecting that only female offspring carry the marker.

For cases in which relative reproductive rate corresponds to Equation 7, *Appendix E* shows that the overall departure from multiplicativity (Equation 6) decomposes into indexes of interaction between the foreign allele at incompatibility locus i and sex:

$$\Delta_i = \frac{\eta_{f,i} - \eta_{m,i}}{2}. \quad (8)$$

Positive Δ_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 Δ_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

$$\delta_{f,\Omega_{[j]}} = \eta_{f,\Omega_{[j]}} - \prod_{i \in \Omega_{[j]}} \eta_{f,i} \quad (9a)$$

$$\delta_{m,\Omega_{[j]}} = \eta_{m,\Omega_{[j]}} - \prod_{i \in \Omega_{[j]}} \eta_{m,i}. \quad (9b)$$

Appendix E shows that

$$e_{\Omega_{[k]}} = \begin{cases} \frac{1}{2} \sum_{j=2}^k (-1)^{k-j} \sum_{\Omega_{[j]}} (\delta_{f,\Omega_{[j]}} + \delta_{m,\Omega_{[j]}}) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b & k \text{ odd} \\ \frac{1}{2} \sum_{j=2}^k (-1)^{k-j} \sum_{\Omega_{[j]}} (\delta_{f,\Omega_{[j]}} + \delta_{m,\Omega_{[j]}}) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b + \prod_{i=1}^k \Delta_i & k \text{ even.} \end{cases} \quad (10)$$

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 D* presents in detail the case of X-linkage of both the marker locus and an incompatibility locus, illustrating the modifications to transmission matrix \mathbf{T} (Equation 2) to accommodate sex specificity. Here, we address the effect on the relative reproductive rate ω (Equation 4) of the genomic locations of the marker locus and of a single incompatibility locus.

Table 1 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 (η_f) and male (η_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 Equation 8). The final columns show the overall relative reproductive rates (Equation 4) of marker alleles introduced by female (ω^f) and male migrants (ω^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 (Equation 4) of a female migrant corresponds to the average of the reproductive values of its female and male offspring:

$$\frac{\eta_{f,1} + \eta_{m,1}}{2}$$

(compare Equation 7). 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,

$$\frac{\eta_{f,1} + 1}{2},$$

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 η_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 1, 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 1 by substituting mt for W.

Barriers due to multiple incompatibility factors

We describe conditions under which the relative reproductive rate ω (Equation 4) induced by multiple incompatibility

Table 1 One-factor reproductive values

Position ^a		Relative contribution ^b		Factor–sex association ^c	Relative reproductive rate ^d	
Factor	Marker	η_f	η_m		Female migrant	Male migrant
A	A	$\frac{\sigma_f[r_f - \sigma_m(r_f - r_m)]/2^d}{1 - \bar{\sigma}(1 - \bar{r})}$	$\frac{\sigma_m[r_m + \sigma_f(r_f - r_m)]/2}{1 - \bar{\sigma}(1 - \bar{r})}$	$\frac{\sigma_f r_f - \sigma_m r_m - \sigma_f \sigma_m (r_f - r_m)}{1 - \bar{\sigma}(1 - \bar{r})}$	$\frac{\bar{\sigma}\bar{r}}{1 - \bar{\sigma}(1 - \bar{r})}$	$\frac{\bar{\sigma}\bar{r}}{1 - \bar{\sigma}(1 - \bar{r})}$
A	X	$\frac{\sigma_f(4 + \sigma_m)}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{\sigma_m(4 + \sigma_f)}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{4(\sigma_f - \sigma_m)}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{2(\sigma_f + \sigma_m) + \sigma_f \sigma_m}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{\sigma_f(4 + \sigma_m)}{8 - \sigma_f(2 + \sigma_m)}$
A	Y	—	$\frac{\sigma_m}{2 - \sigma_m}$	—	—	$\frac{\sigma_m}{2 - \sigma_m}$
A	mt	$\frac{\sigma_f}{2 - \sigma_f}$	—	—	$\frac{\sigma_f}{2 - \sigma_f}$	—
X	A	$\frac{\sigma_f(4 + \sigma_m)}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{\sigma_m(4 + \sigma_f)}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{4(\sigma_f - \sigma_m)}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{2(\sigma_f + \sigma_m) + \sigma_f \sigma_m}{8 - \sigma_f(2 + \sigma_m)}$	$\frac{4 + \sigma_f}{8 - \sigma_f(2 + \sigma_m)}$
X	X	$\frac{2\sigma_f r_f}{2 - \sigma_f(1 + \sigma_m)(1 - r_f)}$	$\frac{2\sigma_f \sigma_m r_f}{2 - \sigma_f(1 + \sigma_m)(1 - r_f)}$	$\frac{2\sigma_f(1 - \sigma_m)r_f}{2 - \sigma_f(1 + \sigma_m)(1 - r_f)}$	$\frac{\sigma_f(1 + \sigma_m)r_f}{2 - \sigma_f(1 + \sigma_m)(1 - r_f)}$	$\frac{2\sigma_f r_f}{2 - \sigma_f(1 + \sigma_m)(1 - r_f)}$
X	Y	—	σ_m	—	—	1
X	mt	$\frac{\sigma_f}{2 - \sigma_f}$	—	—	$\frac{\sigma_f}{2 - \sigma_f}$	—
mt	A	$\frac{\sigma_f \sigma_m}{2 - \sigma_f}$	σ_m	$-\frac{2\sigma_m(1 - \sigma_f)}{2 - \sigma_f}$	$\frac{\sigma_m}{2 - \sigma_f}$	1
mt	X	$\frac{\sigma_f \sigma_m}{2 - \sigma_f}$	σ_m	$-\frac{2\sigma_m(1 - \sigma_f)}{2 - \sigma_f}$	$\frac{\sigma_m}{2 - \sigma_f}$	1
mt	Y	—	—	—	—	1
Y	A	—	$\frac{\sigma_m}{2 - \sigma_m}$	$\frac{2(1 - \sigma_m)}{2 - \sigma_m}$	1	$\frac{1}{2 - \sigma_m}$
Y	X	—	—	—	1	1
Y	mt	—	—	—	1	—

^a Genomic location: autosomal (A), X-linked (X), Y-linked (Y), mitochondrial (mt).

^b Relative contribution to future generations of the marker held by a female (η_f) or a male (η_m) carrier of the factor.

^c $\eta_f - \eta_m$; compare Equation 8.

^d $\bar{\sigma}_F$ and $\bar{\sigma}$ are defined in Equation B8.

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 E* reflects such departures from multiplicativity.

Multiplicative barriers to introgression: In the absence of sex specificity, the Kronecker structure of the transition matrix \mathbf{C} (Equation B5) 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) independent 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 σ_i relative to noncarriers. 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 ($r_{f,j} = r_{m,j} = \frac{1}{2}$).

The selection matrix (Equation B1) generalizes to

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

for $\tilde{\mathbf{S}} = \mathbf{S}_1 \otimes \dots \otimes \mathbf{S}_l$ with

$$\mathbf{S}_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & 1 \end{pmatrix}.$$

Transmission of the autosomal marker and sex retains the form (Equation B2). Independent transmission of the factors conditional on transmission of the marker implies

$$\mathbf{F} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \tilde{\mathbf{R}},$$

for $\tilde{\mathbf{R}} = \mathbf{R}_1 \otimes \dots \otimes \mathbf{R}_l$ with

$$\mathbf{R}_i = \begin{pmatrix} 1-r_i & r_i \\ 0 & 1 \end{pmatrix}$$

(compare Equation B3). The full transition matrix corresponds to

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

(compare Equation B5). The eigenvalues of $\tilde{\mathbf{S}}\tilde{\mathbf{R}} = \mathbf{S}_1\mathbf{R}_1 \otimes \dots \otimes \mathbf{S}_l\mathbf{R}_l$ 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 \mathbf{q} (Equation 4) have the form

$$\eta_{f,\Omega} = \prod_{i \in \Omega} \eta_{f,i} \quad \eta_{m,\Omega} = \prod_{i \in \Omega} \eta_{m,i}, \quad (11)$$

for Ω any subset of indexes in $\{1, \dots, l\}$ 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 (Equation 4) corresponds to the product of the reproductive rates across the loci individually (Equation B6):

$$\omega_{1\dots l}^{A,f} = \omega_{1\dots l}^{A,m} = \prod_{i=1}^l \frac{\sigma_i r_i}{1 - \sigma_i(1 - r_i)}. \quad (12)$$

Multiplicative barriers also arise in the case of markers that are transmitted without recombination through only one sex: those on mitochondria or in the nonrecombining male- (female-)specific region of the Y (W) chromosome, for example. For markers transmitted only through females, the selection matrix (Equation B1) reduces to $\tilde{\mathbf{S}} = \mathbf{S}_1 \otimes \dots \otimes \mathbf{S}_l$, with \mathbf{S}_i containing only the female selection parameter $\sigma_{f,i}$, and the recombination matrix $\tilde{\mathbf{R}} = \mathbf{R}_1 \otimes \dots \otimes \mathbf{R}_l$ contains the crossover rates for females alone. As the transition matrix reduces to $\mathbf{C} = \tilde{\mathbf{S}}\tilde{\mathbf{R}}$, the reproductive rate vector \mathbf{q} has elements $\eta_{f,\Omega} = \prod_{i \in \Omega} \eta_{f,i}$ and the overall relative reproductive rate (Equation 4) is given by Equation 13, 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 ($r_{f,i} = \frac{1}{2}$), the i th multiplicand of Equation 12 reduces to

$$\frac{\sigma_{f,i}}{2 - \sigma_{f,i}}$$

(compare Equation B6). 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 ($\delta_{f,\Omega[j]} = \delta_{m,\Omega[j]} = 0 \forall j$) (Equation 11) as well as of the overall departure (Equation 12) 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,

$$\begin{pmatrix} 1 - r_{f,1} & r_{f,1} \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 - r_{f,2} & r_{f,2} \\ 0 & 1 \end{pmatrix}, \quad (13)$$

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, $r_{f,i} = r_{m,i} = \frac{1}{2}$. 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 (Equation 4) corresponds to

$$\omega_{12}^{A,f} = \omega_{12}^{A,m} = \frac{(f, 0, 0, 0, 1 - f, 0, 0, 0)\mathbf{q}}{(0, 0, 0, f, 0, 0, 0, 1 - f)\mathbf{q}} = \frac{\eta_{f,12} + \eta_{m,12}}{2}, \quad (14)$$

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 (Equation 10) among incompatibility loci reduces in this case to

$$e_{12}^A = \frac{\delta_{f,12} + \delta_{m,12}}{2} + \Delta_1\Delta_2 = \frac{2\Delta_1\Delta_2}{\Gamma}, \quad (15)$$

in which

$$\Delta_i = \frac{\eta_{f,i} - \eta_{m,i}}{2} = \frac{\sigma_{f,i}r_{f,i} - \sigma_{m,i}r_{m,i} - \sigma_{f,i}\sigma_{m,i}(r_{f,i} - r_{m,i})}{2 - \sigma_{f,i}r_{f,i} - \sigma_{m,i}r_{m,i}} \quad (16)$$

(compare Equation B9) and

$$\Gamma = 2 - \sigma_{f,1}\sigma_{f,2}(1 - r_{f,1})(1 - r_{f,2}) - \sigma_{m,1}\sigma_{m,2}(1 - r_{m,1})(1 - r_{m,2}).$$

In this case, the sex-specific measures of epistasis are proportional to each other:

$$\begin{aligned} \delta_{f,12} &= \eta_{f,12} - \eta_{f,1}\eta_{f,2} = \frac{2\sigma_{f,1}\sigma_{f,2}(1 - r_{f,1})(1 - r_{f,2})\Delta_1\Delta_2}{\Gamma} \\ \delta_{m,12} &= \eta_{m,12} - \eta_{m,1}\eta_{m,2} = \frac{2\sigma_{m,1}\sigma_{m,2}(1 - r_{m,1})(1 - r_{m,2})\Delta_1\Delta_2}{\Gamma}. \end{aligned} \quad (17)$$

These autosomal factors undergo independent transmission (Equation 13) and cause no functional epistasis with respect to fitness of carriers. Furthermore, each in isolation induces the relative reproductive rate given by Bengtsson (1985), with the average crossover and viability rates substituted for the non-sex-specific parameters (Equation B7). Even so, the joint barrier to introgression departs in general from the product of the barriers induced by each locus separately (Equation 15).

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 (Equation 4) at an autosomal marker induced by two autosomal incompatibility factors in the absence of physical linkage ($r_{f,i} = r_{m,i} = \frac{1}{2}$ for $i = 1, 2$ in Equation 16). In Figure 1, 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 induce higher barriers.

To verify the accuracy of our expression for the genealogical migration rate g (Equation 1), we conducted a numerical simulation study using SFS_CODE (Hernandez 2008), 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 a 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 the parameter determined by Equation 1 under the population parameters specified in the simulation (Figure 2 legend). Figure 2 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 (Figure 2, left) and factors with discordant effects induce higher barriers (Figure 2, right).

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, Equation 16 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 ($e_{12}^A > 0$) if crossing over between the marker and each of the flanking incompatibility factors occurs at higher rates in females than in 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 A and one X incompatibility locus to introgression at an X-linked marker introduced by a female migrant (*Appendix F* gives the expressions for a male migrant). We assume no crossing over between the X-linked factor and the marker in (hemizygous) males ($r_{m,X} = 0$).

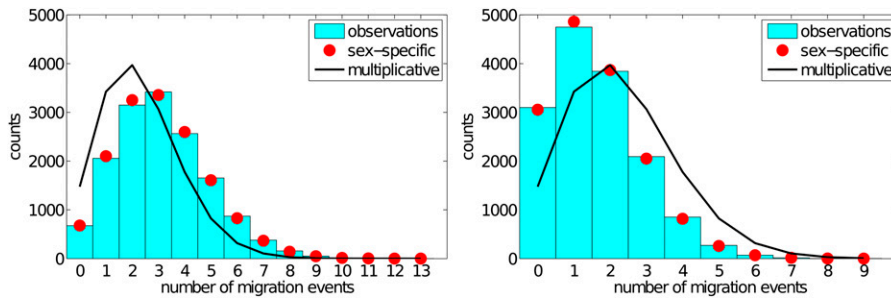
Epistasis in relative reproductive rate (Equation 10) between the X-linked and autosomal incompatibility loci corresponds to

$$e_{AX}^X = \omega_{AX}^X - \omega_{AX}^X \omega_X^X = \frac{\delta_{f,AX}^X + \delta_{m,AX}^X}{2} + \Delta_A^X \Delta_X^X. \quad (18)$$

Table 1 gives the index of association between the autosomal factor and sex,

$$\Delta_A^X = \frac{\eta_{f,A}^X - \eta_{m,A}^X}{2} = \frac{2(\sigma_{f,A} - \sigma_{m,A})}{8 - \sigma_{f,A}(2 + \sigma_{m,A})}, \quad (19)$$

and between the X-linked factor and sex,



backward in time. Each of the two populations comprised 1000 genes (500 diploid individuals), with the proportion of migrant genes each generation set to $m_{12} = m_{21} = 1 \times 10^{-4}$. Left, 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 ($\chi^2 = 12.1$, 11 d.f.) to the Poisson distribution determined from our approximation (Equation 1) of the parameter of the exponential waiting time between migration events ($g = 3.096$, red circles). In contrast, the distribution predicted by assuming multiplicative barriers (black line) showed a poor fit ($\chi^2 = 4924$, 9 d.f.), predicting too few migration events ($g = 2.318$). Right, 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 to 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 circles) again fitted the observations well ($\chi^2 = 6.3$, 7 d.f.), contrary to the multiplicative prediction ($g = 2.318$, $\chi^2 = 3758$, 7 d.f.). Discordance between the incompatibility loci on the relative impairment of males and females induces a greater barrier (fewer migration events) than expected under multiplicativity.

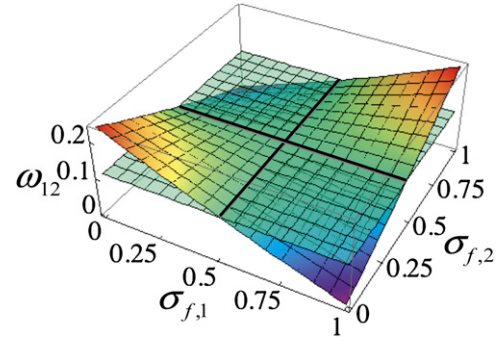


Figure 1 Joint relative reproductive rate at a neutral autosomal marker locus induced by two autosomal factors in the absence of physical linkage. Each of the horizontal axes represents the fitness (relative viability or fertility) of female carriers of an autosomal factor ($\sigma_{f,i}$, $i = 1, 2$), 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. Thick lines indicate the intersection between the actual and expected barriers.

$$\Delta_X^X = \frac{\eta_{f,X}^X - \eta_{m,X}^X}{2} = \frac{\sigma_{f,X} r_{f,X} (1 - \sigma_{m,X})}{2 - \sigma_{f,X} (1 + \sigma_{m,X}) (1 - r_{f,X})}. \quad (20)$$

Within-sex interactions correspond to

$$\delta_{f,AX}^X = \eta_{f,AX}^X - \eta_{f,A}^X \eta_{f,AX}^X = \frac{4\sigma_{f,A}\sigma_{f,X}(1 - r_{f,X})\Delta_A^X \Delta_X^X}{\Gamma} \quad (21a)$$

$$\delta_{m,AX}^X = \eta_{m,AX}^X - \eta_{m,A}^X \eta_{m,X}^X = \frac{8\sigma_{f,A}\sigma_{f,X}(1 - r_{f,X})\sigma_{m,A}\sigma_{m,X}\Delta_A^X \Delta_X^X}{\Gamma}, \quad (21b)$$

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

$$e_{AX}^X = \frac{8\Delta_A^X \Delta_X^X}{\Gamma}, \quad (22)$$

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 (Hernandez 2008), we simulated 100,000 generations forward in time and then counted the number of migration events in a randomly sampled line of descent traced

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 1). Accordingly, Equation 20 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 (Equation 21) as well as overall (Equation 22), depends on the effect of the autosomal factor (Δ_X^X). Expression (19) 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 A incompatibility locus, possibly linked to the marker, and one X locus. We assume a female migrant (*Appendix F* provides the expressions for a male migrant).

The overall two-way epistasis (Equation 10) between the incompatibility factors corresponds to

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

for

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

(compare to Equations 16 and 19), for D_f an index of association between sex and the X-linked factor (Equation F4) and Γ a positive quantity (*Appendix F*). 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 ($\delta_{f,AX}^A \propto D_f$ (Equation F4), which in turn depends on parameters of both the autosomal and X-linked factors.

Indexes D_f and D_m (Equation F4) 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 through 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

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 (Equation 3) 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 (Figure 2). 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 (Bengtsson 1985; Barton and Bengtsson 1986; Navarro and Barton 2003). 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 1 indicates that in the absence of linkage with a neutral marker ($r_f = r_m = \frac{1}{2}$), 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” (Equation F6) of Bengtsson 1985]. 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 3 illustrates, for a female or male migrant, relative reproductive rates (Equation 4) 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 ($r = \frac{1}{2}$) 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, Figure 3 shows relative reproductive rates of 1 for $\omega_X^{Y,m}$ (magenta bar for an incompatibility factor only on the X) and $\omega_Y^{X,m}$ (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 E*). 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 1, the factor–sex association column shows a measure of interaction (Equation 8) between a single incompatibility factor and sex in determining relative reproductive rate ω (Equation 4).

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 incompat-

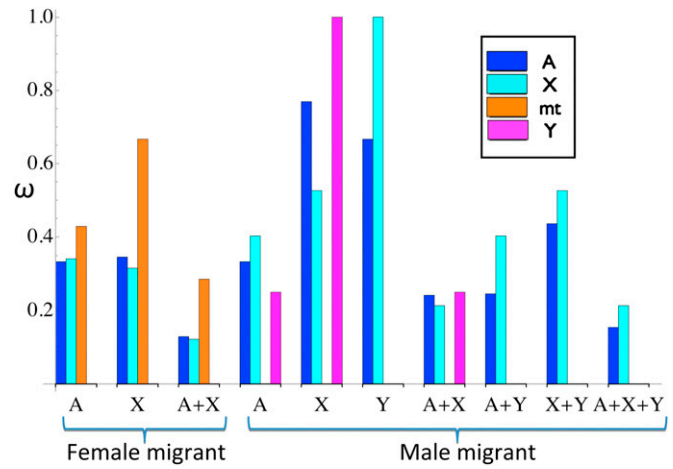


Figure 3 Relative reproductive rates at a neutral marker locus induced by incompatibility factors at various genomic locations for female or male migrants. Bars indicate the relative reproductive rate (Equation 4) of a neutral marker located on an autosome (A, blue) unlinked to any incompatibility locus ($r_{f,A} = r_{m,A} = \frac{1}{2}$), on the X chromosome (X, cyan), on the mitochondria (mt, orange), or on 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 ($r_{f,X} = \frac{1}{2}$) 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 $\sigma_{f,\cdot} + \sigma_{m,\cdot} = 1$. 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, $\sigma_{m,Y} = 0.5$.

ibility factors, transmission of incompatibility factors on autosomes, and equal rates of crossing over between the sexes, we confirm the multiplicativity of the barriers (Equation 12), 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 1 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 ω -values, bluer regions) derive from a positive interaction (Equation 8) 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 in Figure 1).

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 (Hill and Robertson 1966; Barton 1995): in the same sex, in the present context. Concordance of the

sex-specific effects of multiple incompatibility factors (greater impairment of male carriers than of female carriers, for example) implies their association with the same sex.

Interspecific incompatibility in *Drosophila*

Llopart *et al.* (2005) studied patterns of nucleotide divergence and polymorphism in sister species *D. yakuba* and *D. santomea* in 29 genomic regions, including sites on the X chromosome, the Y chromosome, autosomes, and mitochondria. Regions on the X and Y contribute to severe reductions in hybrid male fertility (Coyne *et al.* 2004) and quantitative trait loci contributing to both prezygotic and postzygotic isolation occur on the X chromosome and the autosomes (Moehring *et al.* 2006a,b). Llopart *et al.* (2005) 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 (Hey and Nielsen 2004) 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 Llopart *et al.* (2005), our model indicates that incompatibility due to a single X-linked factor permits greater introgression in autosomal than in X-linked regions,

$$\omega_X^{A,f} > \omega_X^{X,f} \quad \text{and} \quad \omega_X^{A,m} > \omega_X^{X,m},$$

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 1). The blue (autosomal marker) and cyan (X-linked marker) bars in Figure 3 for the case of a single X-linked factor (X on the abscissa) illustrate 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 ($r_f = r_m = \frac{1}{2}$):

$$\omega_A^{X,m} > \omega_A^{A,m}$$

[for example, blue (A) and cyan (X) bars in Figure 3 over A for a male migrant]. For autosomal factors introduced by a female migrant, higher introgression in X-linked regions

$$\omega_A^{X,f} > \omega_A^{A,f}$$

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, the Y chromosome, and autosomes contributing to hybrid male sterility (Coyne *et al.* 2004;

Moehring *et al.* 2006a). 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 ($r_{f,A} = r_{m,A} = \frac{1}{2}$). 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 ($r_{f,X} = \frac{1}{2}$; $r_{m,X} = 0$ in hemizygous males), our model indicates uniformly higher introgression rates in X-linked than in autosomal regions for the case of a male migrant ($\omega_{A,X}^{X,m} > \omega_{A,X}^{A,m}$). 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 at 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 Llopart *et al.* (2005), 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 ($r_f = \frac{1}{2}$) to an incompatibility factor. The presence of an incompatibility factor on the Y, as reported by Coyne *et al.* (2004), 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., Figure 3), 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 (Figure 1). 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 Llopart *et al.* (2005).

Unlike the expectation in the absence of sex specificity, Figure 3 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 Figure 3 differs 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.

Lenormand and Duthiel (2005) have reviewed studies documenting differences in map length between males and females in both plants and animals, including humans (Li *et al.* 1998). In *Drosophila*, the premiere model organism for the experimental investigation of interspecific incompatibility, crossing over is suppressed altogether in males (Morgan 1914).

Assessment of factors contributing to postzygotic interspecific incompatibility that have been identified to the level of operons confirms the traditional view that speciation arises as a by-product of divergence of all manner of genes (reviewed in Orr *et al.* 2006; Araripe *et al.* 2010). While the homeobox gene *Ods* induces sterility in male hybrids between *D. simulans* and *D. mauritiana* (Ting *et al.* 1998), the oncogene *Xmrk* contributes to tumor formation in backcross *Xiphophorus* hybrids of both sexes (Schartl 2008), suggesting that sex-specific incompatibility is common but not universal.

Many of the iconic traits associated with intrinsic postzygotic isolation (Coyne and Orr 2004) affect reproduction, often showing sex-limited expression or different expression in male and female hybrids. Haldane's (1922) rule holds in a number of animals, including *Drosophila*, in which hybrid males tend to suffer much more severe postzygotic incompatibility than do hybrid females (Coyne and Orr 2004, Chap. 8). Direct experiments have demonstrated that introgression of chromosomal segments from *D. mauritiana* into a *D. simulans* background (True *et al.* 1996; Tao *et al.* 2003) or into a *D. sechellia* background (Masly and Presgraves 2007) induces male sterility at rates severalfold higher than those of female sterility or inviability in both sexes.

Numerous studies have described pervasive sex specificity in gene regulation and expression (Michalak and Noor 2003; Ranz *et al.* 2003; Civetta and Singh 2006). Table 1 of Ranz *et al.* (2003) indicates that of the 4776 coding sequences surveyed, >61% showed sex-biased expression in *D. 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 (1922) rule, the disproportionately large effect of the X chromosome constitutes a major rule of speciation (Coyne and Orr 1989). 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 (True *et al.* 1996; Tao *et al.* 2003; Masly and Presgraves 2007). A number of workers have provided discriminating assessments of leading proposals for evolutionary mechanisms that could contribute to the large-X effect (True *et al.* 1996; Tao *et al.* 2003; Coyne and Orr 2004; Presgraves 2008). Whether this major trend reflects faster fixation of advantageous variants on hemizygous X chromosomes than on autosomes (Charlesworth

et al. 1987; Kirkpatrick and Hall 2004), accelerated divergence of genes affecting sex or reproduction (Civetta and Singh 1998), segregation distortion of sex chromosomes (Tao *et al.* 2001), 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 1 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.*, Seielstad *et al.* 1998; Oota *et al.* 2001; Ségurel *et al.* 2008), even if the propensity to migrate and mating success of migrants are independent 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 (Fusco and Uyenoyama 2011).

Acknowledgments

We gratefully acknowledge the constructive comments of Associate Editor Naoyuki Takahata and the anonymous reviewers and support for the National Evolutionary Synthesis Center (NESCent) working group on genomic introgression. Public Health Service grant GM 37841 (to M.K.U.) provided partial funding for this research.

Literature Cited

- Akey, J. M., G. Zhang, K. Zhang, L. Jin, and M. D. Shriver, 2002 Interrogating a high-density SNP map for signatures of natural selection. *Genome Res.* 12: 1805–1814.
- Araripe, L. O., H. Montenegro, B. Lemos, and D. L. Hartl, 2010 Fine-scale genetic mapping of a hybrid sterility factor between *Drosophila simulans* and *D. mauritiana*: the varied and elusive functions of “speciation genes”. *BMC Evol. Biol.* 10: 385.
- Barton, N. H., 1979 Gene flow past a cline. *Heredity* 43: 333–339.
- Barton, N. H., 1995 Linkage and the limits to natural selection. *Genetics* 140: 821–841.
- Barton, N. H., and B. O. Bengtsson, 1986 The barrier to genetic exchange between hybridising populations. *Heredity* 57: 357–376.
- Barton, N. H., and G. M. Hewitt, 1985 Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* 16: 113–148.
- Beaumont, M. A., and D. J. Balding, 2004 Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13: 969–980.

- Bengtsson, B. O., 1985 The flow of genes through a genetic barrier, pp. 31–42 in *Evolution: Essays in Honor of John Maynard Smith*, edited by Greenwood, P. J., P. H. Harvey, and M. Slatkin. Cambridge University Press, New York.
- Bennett, J. H., 1954 On the theory of random mating. *Ann. Eugen.* 18: 311–317.
- Birky, C. W., and J. B. Walsh, 1988 Effects of linkage on rates of molecular evolution. *Proc. Natl. Acad. Sci. USA* 85: 6414–6418.
- Charlesworth, B., J. A. Coyne, and N. H. Barton, 1987 The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* 130: 113–146.
- Charlesworth, B., M. Nordborg, and D. Charlesworth, 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* 70: 155–174.
- Civetta, A., and R. S. Singh, 1998 Sex-related genes, directional sexual selection, and speciation. *Mol. Biol. Evol.* 15: 901–909.
- Civetta, A., and R. S. Singh, 2006 Gene regulation divergence is a major contributor to the evolution of Dobzhansky-Muller incompatibilities between species of *Drosophila*. *Mol. Biol. Evol.* 23: 1707–1714.
- Coop, G., J. K. Pickrell, J. Novembre, S. Kudaravalli, J. Li *et al.*, 2009 The role of geography in human adaptation. *PLoS Genet.* 5: e1000500.
- Coyne, J. A., and H. A. Orr, 1989 Two rules of speciation, pp. 180–270 in *Speciation and Its Consequences*, edited by Otte, D., and J. Endler. Sinauer Associates, Sunderland, MA.
- Coyne, J. A., and H. A. Orr, 2004 *Speciation*. Sinauer Associates, Sunderland, MA.
- Coyne, J. A., S. Elwyn, S. Y. Kim, and A. Llopart, 2004 Genetic studies of two sister species in the *Drosophila melanogaster* subgroup, *D. yakuba* and *D. santomea*. *Genet. Res.* 84: 11–26.
- Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*, Ed. 1. Oxford University Press, Oxford.
- Fusco, D., and M. K. Uyenoyama, 2011 Effects of polymorphism in locally-adapted genes on rates of neutral introgression in structured populations. *Theor. Popul. Biol.* 80: 121–131.
- Gantmacher, F. R., 1959 *The Theory of Matrices*, Vol. II. Chelsea Publishing, New York.
- Gavrilets, S., 1997 Hybrid zones with Dobzhansky-type epistatic selection. *Evolution* 51: 1027–1035.
- Golub, G. H., and C. F. Van Loan, 1996 *Matrix Computations*, Ed. 3. Johns Hopkins University Press, Baltimore.
- Haldane, J. B. S., 1922 Sex ratio and unisexual sterility in animal hybrids. *J. Genet.* 12: 101–109.
- Haldane, J. B. S., 1957 The cost of natural selection. *J. Genet.* 55: 511–524.
- Hernandez, R. D., 2008 A flexible forward simulator for populations subject to selection and demography. *Bioinformatics* 24: 2786–2787.
- Hey, J., and R. Nielsen, 2004 Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167: 747–760.
- Hey, J., and R. Nielsen, 2007 Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. USA* 104: 2785–2790.
- Hill, W. G., and A. Robertson, 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* 8: 269–294.
- Holsinger, K. E., and B. S. Weir, 2009 Genetics in geographically structured populations: defining, estimating and interpreting f_{ST} . *Nat. Rev. Genet.* 10: 639–650.
- Johnson, N. L., and S. Kotz, 1969 *Discrete Distributions*. John Wiley & Sons, New York.
- Karlin, S., 1982 Classifications of selection-migration structures and conditions for a protected polymorphism, pp. 61–204 in *Evolutionary Biology*, Vol. 14, edited by Hecht, M. K., B. Wallace, and G. T. Prance. Plenum Press, New York.
- Kirkpatrick, M., and D. W. Hall, 2004 Sexual selection and sex linkage. *Evolution* 58: 683–691.
- Kobayashi, Y., P. Hammerstein, and A. Telschow, 2008 The neutral effective migration rate in a mainland-island context. *Theor. Popul. Biol.* 74: 84–92.
- Krone, S. M., and C. Neuhauser, 1997 Ancestral processes with selection. *Theor. Popul. Biol.* 51: 210–237.
- Kulathinal, R. J., L. S. Stevison, and M. A. F. Noor, 2009 Genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet.* 5: e1000550.
- Leman, S. C., Y. Chen, J. E. Stajich, M. A. F. Noor, and M. K. Uyenoyama, 2005 Likelihoods from summary statistics: recent divergence between species. *Genetics* 171: 1419–1436.
- Lenormand, T., and J. Dutheil, 2005 Recombination difference between sexes: a role for haploid selection. *PLoS Biol.* 3: 396–403.
- Li, W., C. S. J. Fann, and J. Ott, 1998 Low-order polynomial trends of female-to-male map distance ratios along human chromosomes. *Hum. Hered.* 48: 226–270.
- Llopart, A., D. Lachaise, and J. A. Coyne, 2005 Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. *Genetics* 171: 197–210.
- Machado, C. A., and J. Hey, 2003 The causes of phylogenetic conflict in a classic *Drosophila* species group. *Proc. Biol. Sci.* 270: 1193–1202.
- Masly, J. P., and D. C. Presgraves, 2007 High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biol.* 5: 1890–1898.
- Michalak, P., and M. A. F. Noor, 2003 Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. *Mol. Biol. Evol.* 23: 1070–1076.
- Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. C. Mackay, 2006a The genetic basis of postzygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to hybrid male sterility. *Genetics* 173: 225–233.
- Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. C. Mackay, 2006b The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference. *Genetics* 173: 215–223.
- Morgan, T. H., 1914 No crossing over in the male of *Drosophila* of genes in the second and third pairs of chromosomes. *Biol. Bull.* 26: 195–204.
- Navarro, A., and N. H. Barton, 2003 Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* 57: 447–459.
- Neuhauser, C., and S. M. Krone, 1997 The genealogy of samples in models with selection. *Genetics* 145: 519–534.
- Noor, M. A. F., K. L. Grams, L. A. Bertucci, Y. Almendarez, J. Reiland *et al.*, 2001 The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. *Evolution* 55: 512–521.
- Oota, H., W. Settheetham-Ishida, D. Tiwawech, T. Ishida, and M. Stoneking, 2001 Human mtDNA and Y-chromosome variation is correlated with matrilineal vs. patrilineal residence. *Nat. Genet.* 29: 20–21.
- Orr, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* 139: 1805–1813.
- Orr, H. A., and M. Turelli, 2001 The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* 55: 1085–1094.
- Orr, H. A., J. P. Masly, and N. Phadnis, 2006 Speciation in *Drosophila*: from phenotypes to molecules. *J. Hered.* 98: 103–110.

- Payseur, B. A., and M. W. Nachman, 2005 The genomics of speciation: investigating the molecular correlates of X chromosome introgression across the hybrid zone between *Mus domesticus* and *Mus musculus*. *Biol. J. Linn. Soc. Lond.* 84: 523–534.
- Payseur, B. A., J. G. Krenz, and M. W. Nachman, 2004 Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* 58: 2064–2078.
- Presgraves, D. C., 2008 Sex chromosomes and speciation in *Drosophila*. *Trends Genet.* 24: 336–343.
- Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl, 2003 Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* 300: 1742–1745.
- Rieseberg, L. H., J. Whitton, and K. Gardner, 1999 Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* 152: 713–727.
- Rosenberg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd *et al.*, 2002 Genetic structure of human populations. *Science* 298: 2381–2385.
- Schartl, M., 2008 Evolution of *Xmrk*: an oncogene, but also a speciation gene? *BioEssays* 30: 822–832.
- Schlötterer, C., 2003 Hitchhiking mapping—functional genomics from the population genetics perspective. *Trends Genet.* 19: 32–38.
- Ségurel, L., B. na Martínez-Cruz, L. Quintana-Murci, P. Balaesque, M. Georges *et al.*, 2008 Sex-specific genetic structure and social organization in Central Asia: insights from a multi-locus study. *PLoS Genet.* 4: e1000200.
- Seielstad, M. T., E. Minch, and L. L. Cavalli-Sforza, 1998 Genetic evidence for a higher female migration rate in humans. *Nat. Genet.* 20: 279–280.
- Storz, J. F., 2005 Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.* 14: 671–688.
- Strasburg, J. L., C. Scotti-Saintagne, I. Scotti, Z. Lai, and L. H. Rieseberg, 2009 Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Mol. Biol. Evol.* 26: 1341–1355.
- Tao, Y., D. L. Hartl, and C. C. Laurie, 2001 Sex-ratio segregation distortion associated with reproductive isolation in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 98: 13183–13188.
- Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie, 2003 Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *Drosophila mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. *Genetics* 164: 1383–1397.
- Teeter, K. C., L. M. Thibodeau, Z. Gormpert, C. A. Buerkle, M. W. Nachman *et al.*, 2009 The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution* 64: 472–485.
- Ting, C.-T., S.-C. Tsaur, M.-L. Wu, and C.-I. Wu, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* 282: 1501–1504.
- True, J. R., B. S. Weir, and C. C. Laurie, 1996 A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* 142: 819–837.
- Via, S., 2009 Natural selection in action during speciation. *Proc. Natl. Acad. Sci. USA* 106: 9939–9946.
- Wang, R. L., J. Wakeley, and J. Hey, 1997 Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 147: 1091–1106.
- Weir, B. S., L. R. Cardon, A. D. Anderson, D. M. Nielsen, and W. G. Hill, 2005 Measures of human population structure show heterogeneity among genomic regions. *Genome Res.* 15: 1468–1476.

Communicating editor: N. Takahata

Appendixes

Appendix A: Relative Reproductive Rate

Because \mathbf{C} , describing genetic transmission and selection (Equation 2), is nonnegative, its spectral radius λ corresponds to a nonnegative characteristic value with modulus exceeded by no other eigenvalue and the eigenvector associated with λ is nonnegative (Gantmacher 1959, Chap. XIII, section 3). It has a Schur decomposition of the form

$$\mathbf{C} = \mathbf{Q}\mathbf{U}\mathbf{Q}^{-1},$$

for \mathbf{U} a triangular matrix with the eigenvalues of the transformation along the diagonal (see, for example, Golub and Van Loan 1996). For convenience, we choose \mathbf{Q} such that the eigenvalues appear along the diagonal in order of their moduli, with λ the first element. For t sufficiently large, λ^t dominates

$$\mathbf{C}^t = \mathbf{Q}\mathbf{U}^t\mathbf{Q}^{-1},$$

implying

$$\omega = \frac{\mathbf{v}_m \mathbf{q}_0}{\mathbf{v}_r \mathbf{q}_0}$$

(compare Equation 4), for \mathbf{q}_0 the column or the sum of the columns of \mathbf{Q} that correspond to λ .

Appendix B: 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 (\mathbf{v}_m or \mathbf{v}_r in Equation 4), selection first reduces the frequency of carriers of the foreign incompatibility factor, with female carriers surviving and reproducing at rate σ_f and males at rate σ_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

$$\mathbf{S} = \begin{pmatrix} \sigma_f & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \sigma_m \frac{f}{1-f} & 0 \\ 0 & 0 & 0 & \frac{f}{1-f} \end{pmatrix}, \quad (\text{B1})$$

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 $\frac{1}{2}$,

$$\frac{1}{2}(f, 1-f), \quad (\text{B2})$$

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

$$\mathbf{F} = \begin{pmatrix} 1-r_f & r_f \\ 0 & 1 \\ 1-r_m & r_m \\ 0 & 1 \end{pmatrix}, \quad (\text{B3})$$

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

$$\mathbf{T} = \frac{1}{2}(f, 1-f) \otimes \mathbf{F}, \quad (\text{B4})$$

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

$$\begin{aligned} \mathbf{C} = \mathbf{ST} &= \left[\begin{pmatrix} 1 & 0 \\ 0 & f \\ & 1-f \end{pmatrix} \otimes \begin{pmatrix} \sigma & 0 \\ 0 & 1 \end{pmatrix} \right] \left[\frac{1}{2}(f, 1-f) \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] \otimes \begin{pmatrix} 1-r & r \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & f \\ & 1-f \end{pmatrix} \left[\frac{1}{2}(f, 1-f) \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] \otimes \begin{pmatrix} \sigma(1-r) & \sigma r \\ 0 & 1 \end{pmatrix}. \end{aligned} \quad (\text{B5})$$

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, \mathbf{C} has dominant eigenvalue f (the product of the eigenvalues of the matrices) with right eigenvector

$$\mathbf{q} = \left(1, \frac{f}{1-f} \right)' \otimes \left(\frac{\sigma r}{1 - \sigma(1-r)}, 1 \right)'$$

Relative reproductive rate (Equation 4) corresponds to

$$\omega = \frac{(f, 0, 1-f, 0)\mathbf{q}}{(0, f, 0, 1-f)\mathbf{q}} = \frac{\sigma r}{1 - \sigma(1-r)}. \quad (\text{B6})$$

This expression, which has been derived numerous times (e.g., Equation A3 in Navarro and Barton 2003), corresponds to the “gene flow factor” of Bengtsson (1985) and “barrier strength” of Barton and Bengtsson (1986).

Under sex specificity in selection (Equation B1) or transmission (Equation B3), the relative reproductive rates of female and male carriers correspond to

$$\begin{aligned} \eta_f &= \frac{\sigma_f[2r_f - \sigma_m(r_f - r_m)]}{2 - \sigma_f(1-r_f) - \sigma_m(1-r_m)} \\ \eta_m &= \frac{\sigma_m[2r_m + \sigma_f(r_f - r_m)]}{2 - \sigma_f(1-r_f) - \sigma_m(1-r_m)}, \end{aligned}$$

implying a relative reproductive rate at the marker locus of

$$\omega = \frac{\eta_f + \eta_m}{2} = \frac{\bar{\sigma} \bar{r}}{1 - \bar{\sigma}(1 - \bar{r})}, \quad (\text{B7})$$

for

$$\begin{aligned} \bar{\sigma} &= \frac{\sigma_f + \sigma_m}{2} \\ \bar{r} &= \frac{\sigma_f r_f + \sigma_m r_m}{\sigma_f + \sigma_m} \end{aligned} \quad (\text{B8})$$

(compare Equation B6).

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 Equation 8):

$$\eta_f - \eta_m = \frac{2[\sigma_f r_f (1 - \sigma_m) - \sigma_m r_m (1 - \sigma_f)]}{2 - \sigma_f (1 - r_f) - \sigma_m (1 - r_m)} \quad (\text{B9})$$

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 ($r_f = r_m$). For a given total magnitude of selection ($\sigma_f + \sigma_m$), ($\eta_f - \eta_m$) 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 C: 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 (Equation 1). This section presents confirmation of this proposal, on the basis of results of numerical simulations generated by SFS_CODE (Hernandez 2008), 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 the parameter equal to the product of the number of generations observed and the per-generation backward migration rate (Equation 1). This construction implies that the waiting time between migration events along a lineage has an exponential distribution with the 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 ($N_1 = N_2 = 1000$ genes or 500 diploid individuals) and symmetric forward migration rates ($N_1 m_{12} = N_2 m_{21} = 0.1$). 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 (Equation 4) of $\omega = 0.4815$ and a genealogical migration rate (Equation 1) 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 ($\chi^2 = 3.946$, 9 d.f.) to expectations under a Poisson distribution with parameter $gt = 2.4074$.

Because relative reproductive rate (Equation 3) 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 ($N_1 m_{12} = N_2 m_{21} = 1.0$). 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 ($\chi^2 = 90.4898$, 32 d.f.): the observed mean (24.451) exceeds the expectation ($b_i t = 24.074$) 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 that assumes that the Poisson parameter is itself a Poisson-distributed random variable,

$$\frac{\lambda}{\varphi} \sim \text{Poisson}(\mu),$$

for φ an index of clumping (see Johnson and Kotz 1969, Chap. 9). The observations show an excellent fit ($\chi^2 = 19.898$, 24 d. f.) to a Neyman's type A distribution with mean set equal to the observed mean and a least-squares estimate of the clumping parameter φ 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 Johnson and Kotz (1969).

Appendix D: 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 \mathbf{T} corresponding to

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

(compare Equation B4).

For a female migrant, the relative reproductive rate (Equation 4) becomes

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

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 (Equation 8) 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 Equation B9). 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

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

Appendix E: Epistasis Among Incompatibility Factors

Here we describe a measure of epistasis among incompatibility factors (compare Bennett 1954), applied to the relative reproductive rate (Equation 4) for an arbitrary number of incompatibility factors.

To establish Equation 5, we expand

$$\begin{aligned} e_{\Omega_{[k]}} &= E \left[\prod_{i \in \Omega_{[k]}} (A_i - \omega_i) \right] \\ &= \sum_{j=0}^k (-1)^{k-j} \sum_{\Omega_{[j]}} E \left[\prod_{i \in \Omega_{[j]}} A_i \right] \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b \\ &= \sum_{j=1}^k (-1)^{k-j} \sum_{\Omega_{[j]}} \left(E \left[\prod_{i \in \Omega_{[j]}} A_i \right] - \prod_{i \in \Omega_{[j]}} \omega_i \right) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b + \prod_{b \in \Omega_{[k]}} \omega_b \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} \\ &= \sum_{j=2}^k (-1)^{k-j} \sum_{\Omega_{[j]}} \delta_{\Omega_{[j]}} \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b, \end{aligned} \tag{E1}$$

for $\Omega_{[k]}$ ($k \geq 2$) a set of k indexes representing k incompatibility loci, $\Omega_{[j]}$ a set of j indexes in $\Omega_{[k]}$ ($j \leq k$), $\Omega_{[k]} \setminus \Omega_{[j]}$ denoting the set of indexes in $\Omega_{[k]}$ but not in $\Omega_{[j]}$, and $\delta_{\Omega_{[j]}}$ representing the departure of relative reproductive rate from multiplicativity (Equation 6).

For relative reproductive rate corresponding to Equation 7, we address the relationship between the overall departure from multiplicativity (Equation 6) and the sex-specific measures (Equation 9). Using that $\eta_{f,i} = \omega_i + \Delta_i$ and $\eta_{m,i} = \omega_i - \Delta_i$ for $\Delta_i = (\eta_{f,i} - \eta_{m,i})/2$ (Equation 8), we have

$$\begin{aligned} \frac{1}{2} \left(\prod_{i \in \Omega_{[j]}} \eta_{f,i} + \prod_{i \in \Omega_{[j]}} \eta_{m,i} \right) &= \frac{1}{2} \left[\prod_{i \in \Omega_{[j]}} (\omega_i + \Delta_i) + \prod_{i \in \Omega_{[j]}} (\omega_i - \Delta_i) \right] \\ &= \prod_{i \in \Omega_{[j]}} \omega_i + \xi_{\Omega_{[j]}}, \end{aligned} \quad (\text{E2})$$

for $\xi_{\Omega_{[j]}}$ representing terms of the form Δ_i . Note that

$$\xi_{\Omega_{[j]}} = \delta_{\Omega_{[j]}} - \frac{\delta_{f,\Omega_{[j]}} + \delta_{m,\Omega_{[j]}}}{2}.$$

In the expansion of Equation E2, terms involving an odd number of the Δ_i 's cancel out, while all the terms with an even number appear twice, implying

$$\xi_{\Omega_{[k]}} = \sum_{i=1}^{\lfloor k/2 \rfloor} \sum_{\Omega_{[2i]} \subset \Omega_{[k]}} \left(\prod_{a \in \Omega_{[2i]}} \Delta_a \prod_{c \in \Omega_{[k]} \setminus \Omega_{[2i]}} \omega_c \right),$$

for $\Omega_{[2i]}$ representing all subsets of $\Omega_{[j]}$ that contain $2i$ elements. We rewrite Equation E1 as

$$e_{\Omega_{[k]}} = \sum_{j=2}^k (-1)^{k-j} \sum_{\Omega_{[j]}} \left(\frac{\delta_{f,\Omega_{[j]}} + \delta_{m,\Omega_{[j]}}}{2} + \xi_{\Omega_{[j]}} \right) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b. \quad (\text{E3})$$

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

$$\begin{aligned} &\sum_{j=2}^k (-1)^{k-j} \sum_{\Omega_{[j]}} \sum_{i=1}^{\lfloor j/2 \rfloor} \sum_{\Omega_{[2i]} \subset \Omega_{[j]}} \left(\prod_{a \in \Omega_{[2i]}} \Delta_a \prod_{c \in \Omega_{[j]} \setminus \Omega_{[2i]}} \omega_c \right) \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b \\ &= \sum_{j=2}^k (-1)^{k-j} \sum_{i=1}^{\lfloor j/2 \rfloor} \binom{k-2i}{j-2i} \sum_{\Omega_{[2i]}} \prod_{a \in \Omega_{[2i]}} \Delta_a \prod_{b \in \Omega_{[k]} \setminus \Omega_{[2i]}} \omega_b \\ &= \sum_{i=1}^{\lfloor k/2 \rfloor} \sum_{\Omega_{[2i]}} \prod_{a \in \Omega_{[2i]}} \Delta_a \prod_{b \in \Omega_{[k]} \setminus \Omega_{[2i]}} \omega_b \sum_{j=2i}^k \binom{k-2i}{j-2i} (-1)^{k-j} \\ &= 0, \end{aligned}$$

and for k even,

$$\prod_{i=1}^k \Delta_i + \sum_{j=2}^{k-1} (-1)^{k-j} \sum_{\Omega_{[j]} \subset \Omega_{[k]}} \xi_{\Omega_{[j]}} \prod_{b \in (\Omega_{[k]} \setminus \Omega_{[j]})} \omega_b = \prod_{i=1}^k \Delta_i.$$

Substitution of these expressions into Equation E3 produces Equation 10.

Appendix F: 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 Equation 21a. As in the case of the female migrant, the sign of epistasis depends on Δ_A^X (Equation 19).

Autosomal marker

For a *female migrant*, epistasis (Equation 10) between the incompatibility loci corresponds to

$$e_{A,X}^A = \omega_{AX}^A - \omega_A^A \omega_X^A = \frac{\delta_{f,AX}^X + \delta_{m,AX}^X}{2} + \Delta_A^A \Delta_X^A, \quad (F1)$$

for Δ_X^A given by Equation 26 and

$$\Delta_X^A = \frac{\eta_{f,X}^A - \eta_{m,X}^A}{2} = \frac{2(\sigma_{f,X} - \sigma_{m,X})}{8 - \sigma_{f,X}(2 + \sigma_{m,X})} \quad (F2)$$

(Table 1). Within-sex interactions correspond to

$$\delta_{f,AX}^A = \eta_{f,AX}^A - \eta_{f,A}^A \eta_{f,X}^A = \frac{2\sigma_{f,A}\sigma_{f,X}(1 - r_{f,A})D_f\Delta_A^A}{[8 - \sigma_{f,X}(2 + \sigma_{m,X})]\Gamma} \quad (F3a)$$

$$\delta_{m,AX}^A = \eta_{m,AX}^A - \eta_{m,A}^A \eta_{m,X}^A = \frac{2\sigma_{m,A}\sigma_{m,X}(1 - r_{m,A})D_m\Delta_A^A}{[8 - \sigma_{f,X}(2 + \sigma_{m,X})]\Gamma}, \quad (F3b)$$

in which

$$\begin{aligned} D_f &= 4(\sigma_{f,X} - \sigma_{m,X}) - \sigma_{m,A}\sigma_{m,X}(1 - r_{m,A})[4 - \sigma_{f,X}(3 + \sigma_{m,X})] \\ D_m &= 2\sigma_{f,A}\sigma_{f,X}(1 - r_{f,A})(\sigma_{f,A} - \sigma_{m,X}) - [4 - \sigma_{f,X}(3 + \sigma_{m,X})][4 - \sigma_{f,X}\sigma_{f,A}(1 - r_{f,A})] \\ \Gamma &= 8 - \sigma_{f,A}\sigma_{f,X}(1 - r_{f,A})[2 + \sigma_{m,A}\sigma_{m,X}(1 - r_{m,A})], \end{aligned} \quad (F4)$$

giving an overall two-way epistasis between the incompatibility factors of Equation 23. Comparison of Equations 17 and F3 suggests that D_f and D_m assume the role of Δ_X , an indicator of the nature of the interaction between sex and the X-linked incompatibility factor.

As indicated in the text following Equation 24, 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

$$\sigma_{f,X} > \sigma_{m,X} \left\{ 1 + \sigma_{m,A}(1 - r_{m,A}) \left[1 - \frac{\sigma_{f,X}(3 + \sigma_{m,X})}{4} \right] \right\} \geq \sigma_{m,X}. \quad (F5)$$

Sufficient conditions for positive epistasis e_{AX}^X (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 Δ_A and Equation F5 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:

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

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

$$e_{AX}^{A,m} = \frac{\delta_{AX}^A}{2} + \frac{\Delta_A^A (\eta_{f,X}^A - 1)}{2},$$

with Δ_A^A given by Equation 24, the within-female departure from multiplicativity $\delta_{f,AX}^A$ by Equation F3a, 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 (Equation 23), epistasis arises only if the autosomal factor is sex specific ($\Delta_A^A \neq 0$). Nonzero $(\eta_{f,X}^A - 1)$ 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 Γ given in Equation F4. 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 through female carriers ($\Delta_A^A < 0$) and negative epistasis if female carriers show higher reproductive rates.

GENETICS

Supporting Information

<http://www.genetics.org/content/suppl/2011/06/24/genetics.111.130732.DC1>

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

Diana Fusco and Marcy K. Uyenoyama

FILE S1

Instructions for matrix_builder.nb

For arbitrary numbers of incompatibility factors, Mathematica notebook matrix_builder.nb computes the transmission matrix C for a single neutral marker and the right eigenvector q associated with the leading eigenvalue. It accommodates autosomal- and X-linkage of the neutral marker and of the incompatibility factors. It assumes no functional epistasis among the incompatibility factors in the determination of fitness. No linkage occurs among the factors, but the marker may be linked to one of the factors.

Specification of parameters is done in the section “Variables Settings”:

- `nloci`: number of incompatibility factors

- `locpos`: location of each incompatibility factor

Use `a` to specify the autosomal factors, followed by `x` for the X-linked factors.

- `markerpos`: location of the marker

Use `a` for an autosomal and `x` for an X-linked marker.

- `deltaf`: fitness parameters for female carriers of the factors

Ordering of the factors must adhere to `locpos`.

- `deltaf`: fitness parameters for male carriers of the factors

- `recf`: rate of crossing-over between the marker and each factor in females

The marker can be linked to at most one factor. For all unlinked factors, specify 1/2.

- `recfm`: rate of crossing-over between the marker and each factor in males

As males are assumed to be hemizygous for the X chromosome, an absence of crossing-over is assumed for X-linked factors and crossover rates are not specified. For other chromosomes, enter crossover rates as for `recf`.

Select “Evaluate Notebook” under the “Evaluate” menu to execute the script. Variable `c` contains the transmission matrix and variable `q` the right eigenvector q . The backgrounds on which the neutral marker can occur are indexed by $i = 1, \dots, 2 \cdot 2^{nloci}$, with the first 2^{nloci} backgrounds corresponding to female carriers and the remaining to male carriers. If $i \leq 2^{nloci}$ (female carrier), background i is given by the binary representation of $2^{nloci} - i$; if $i > 2^{nloci}$ (male carrier), it is given the binary representation of $2 \cdot 2^{nloci} - i$.

Example Consider an X-linked marker together with incompatibility loci 1 (autosomal) and 2 (X-linked). Crossing-over between the marker and locus 1 occurs at rate $1/2$ in both sexes and between the marker and locus 2 at rate r_f in females and r_m in males. Each foreign allele at locus i ($i = 1, 2$) reduces the fitness of a female carrier by a factor of $\sigma_{f,i}$ and the fitness of a male carrier by a factor of $\sigma_{m,i}$.

In the “Variables Settings” section, we set:

- `nloci=2`
- `locpos={a,X}`
- `markerpos=X`
- `deltaf={ $\sigma_{f,1},\sigma_{f,2}$ }`
- `deltam={ $\sigma_{m,1},\sigma_{m,2}$ }`
- `recf={ $1/2,r_f$ }`
- `recm={ $1/2$ }`

The order of the states in the matrix c and in the vector q follows the index given by the following table (1 means present, 0 means absent):

Index	Carrier	Factor 1	Factor 2
1	Female	1	1
2	Female	1	0
3	Female	0	1
4	Female	0	0
5	Male	1	1
6	Male	1	0
7	Male	0	1
8	Male	0	0

Eigenvector q corresponds to

$$\mathbf{q} = \left(\frac{1-f}{f} \eta_{f,12}, \frac{1-f}{f} \eta_{f,1}, \frac{1-f}{f} \eta_{f,2}, \frac{1-f}{f}, \eta_{m,12}, \eta_{m,1}, \eta_{m,2}, 1 \right).$$

For a female migrant, the reproductive value at the marker is given by

$$\omega = \frac{(f, 0, 0, 0, 1-f, 0, 0, 0) \mathbf{q}}{(0, 0, 0, f, 0, 0, 0, 1-f) \mathbf{q}} = (\eta_{f,12} + \eta_{m,12})/2,$$

in which f is the proportion of females among offspring. For a male migrant, the reproductive value at the marker corresponds to

$$\omega = \frac{(f, 0, 0, 0, 0, 0, 0, 0)\mathbf{q}}{(0, f, 0, 0, 0, 0, 0, 0)\mathbf{q}} = \frac{f\mathbf{q}(1)}{f\mathbf{q}(4)} = \eta_{f,12},$$

reflecting that the male migrant transmit its X-linked marker only to its daughters.

File S2

Matrix_builder.nb

File S2 is available for download as a Mathematica notebook file at
<http://www.genetics.org/content/suppl/2011/06/24/genetics.111.130732.DC1>.