Associations Between Cytoplasmic and Nuclear Loci in Hybridizing Populations

Maria E. Orive*,†,[1] and Nicholas H. Barton†

**Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045 and †Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland, United Kingdom*

Manuscript received June 24, 2002
Accepted for publication August 9, 2002

ABSTRACT

We extend current multilocus models to describe the effects of migration, recombination, selection, and nonrandom mating on sets of genes in diploids with varied modes of inheritance, allowing us to consider the patterns of nuclear and cytonuclear associations (disequilibria) under various models of migration. We show the relationship between the multilocus notation recently presented by Kirkpatrick, Johnson, and Barton (developed from previous work by Barton and Turelli) and the cytonuclear parameterization of Asmussen, Arnold, and Avise and extend this notation to describe associations between cytoplasmic elements and multiple nuclear genes. Under models with sexual symmetry, both nuclear-nuclear and cytonuclear disequilibria are equivalent. They differ, however, in cases involving some type of sexual asymmetry, which is then reflected in the asymmetric inheritance of cytoplasmic markers. An example given is the case of different migration rates in males and females; simulations using 2, 3, 4, or 5 unlinked autosomal markers with a maternally inherited cytoplasmic marker illustrate how nuclear-nuclear and cytonuclear associations can be used to separately estimate female and male migration rates. The general framework developed here allows us to investigate conditions where associations between loci with different modes of inheritance are not equivalent and to use this nonequivalence to test for deviations from simple models of admixture.

THE mixing of genetically divergent populations, and assortative mating within them, can generate strong associations ("linkage disequilibria"), even between unlinked genes. Associations between alleles can therefore be used to give estimates of migration rates and mating patterns in hybrid zones. This approach has been applied to estimate dispersal rates, using autosomal markers sampled from linear transects (Barton 1983; Szymura and Barton 1986, 1991; Mallet *et al.* 1990); in these cases, random mating could be assumed. In parallel, associations between cytoplasmic and nuclear markers have been used to estimate both degree of assortment and migration rates, using samples from single populations subject to immigration (Asmussen *et al.* 1987, 1989; Arnold *et al.* 1988; Asmussen and Arnold 1991). Under many circumstances, pairwise associations between cytoplasmic and nuclear markers evolve in the same way as those between unlinked nuclear markers. They differ, however, when there is some type of sexual asymmetry; this asymmetry is then reflected in the asymmetric inheritance of cytoplasmic markers. In addition, higher-order associations (*e.g.*, between a cytoplasmic allele and heterozygosity at a neutral autosomal locus) can give information on asymmetries in mating preferences and differential dispersal between the sexes. For example, if females of one species hybridize more readily than females of the other, then the cytoplasm they pass on will tend to be found more often in individuals heterozygous for nuclear markers.

The array of genotypes in a hybrid zone contains a great deal of information; full analysis requires joint consideration of the deficit of heterozygotes; of discrepancies among associations involving autosomal, sex-linked, and cytoplasmic loci; of the relation between linkage disequilibrium and recombination rates; and of third- and higher-order associations. The large number of parameters involved for multilocus data, however, cannot generally be estimated without extremely large samples unless simplifying assumptions are made. Recent work by Pritchard *et al.* (2000) and Dawson and Belkhir (2001) has considered the problem of using multilocus genotype data from unlinked loci to infer population structure and assign individuals to (previously unknown) underlying populations. The general approach is to account for nonrandomness in genotype frequencies (departure from Hardy-Weinberg expectations and from linkage equilibrium) by partitioning the sampled individuals in some way. Both groups apply a Bayesian approach to this inference problem, with the evidence in favor of each possible partitioning or clustering of individuals represented by the posterior distribution of the populations and allele frequencies in those populations (Pritchard *et al.* 2000) or the partition of the sample (Dawson and Belkhir 2001). Other workers have also recently considered the use of multilocus genotypes to assign individuals to putative source popu-

[1]*Corresponding author:* Department of Ecology and Evolutionary Biology, University of Kansas, 1200 Sunnyside Ave., Lawrence, KS 66045. E-mail: morive@ku.edu

lations, where the source populations were known *a priori* (Rannala and Mountain 1997; Cornuet *et al.* 1999). As pointed out in Dawson and Belkhir (2001), the problem of assigning alleles within a genotype to source populations is closely related to the problem of inferring the distribution of a hybrid index, recently addressed in Barton (2000) using a maximum-likelihood approach and considering unlinked autosomal loci. All of these approaches have the related aim of estimating a limited number of parameters (different orders of multilocus associations in the case of Barton 2000) that approximate the full genotype frequencies. These parameters may be the migration rates or strengths of assortment themselves, or they may be measures of disequilibria, averaged over loci, which can later be used to estimate migration rates or other parameters.

The main motive of this article is to set methods for estimating and interpreting disequilibria (Szymura and Barton 1986, 1991; Asmussen *et al.* 1987, 1989; Arnold *et al.* 1988; Mallet *et al.* 1990) in the general multilocus context of Kirkpatrick *et al.* (2002), thus allowing genetic information from hybridizing populations to be used more fully. Kirkpatrick *et al.* (2002) extended the multilocus notation of Barton and Turelli (1991; Turelli and Barton 1994) to describe the effects of migration, recombination, selection, and nonrandom mating on sets of genes with arbitrary modes of inheritance. This is implemented in a set of packages, written in Mathematica 4.0, which generate recursions for allele frequencies and linkage disequilibria (available at http://helios.bto.ed.ac.uk/evolgen/index.html). We use this framework to derive patterns of nuclear and cytonuclear associations under various migration models; comparison among the various nuclear and cytonuclear associations gives additional information not available from consideration of each type alone. We use simulated data to illustrate how simple two-way nuclear and cytonuclear associations can be used to estimate male and female migration rates when these differ.

MODEL

Notation: The notation used here is based on that introduced by Kirkpatrick *et al.* (2002), which extended and generalized the approach of Barton and Turelli (1991) to allow for arbitrary modes of inheritance, including diploidy, polyploidy, sex linkage, cytoplasmic inheritance, and genomic imprinting. Here, we explicitly consider diploids. Kirkpatrick *et al.* (2002) define the word "gene" to mean a particular copy of a nonrecombining sequence at some locus in some individual. The "context" of a gene includes such information as the life-history stage it is found in, the deme it is found in, the sex of its carrier, and whether it was inherited from the individual's female or male parent (its "sex-of-origin"). A "position" is used to refer to a particular locus in a particular context. Consider a population of diploid individuals; a genome in such an individual is denoted by $\mathbb{G}$, which is the set of all positions in an individual. For a single diploid autosomal locus i, for example, the genome of a male might be written as $\mathbb{G} = \{i_{mm}, i_{mf}\}$, where the first subscript indicates the "sex-of-carrier" (in this case, male), and the second subscript indicates the sex of the parent from which the allele was inherited, the sex-of-origin. Similarly, in a female, $\mathbb{G} = \{i_{fm}, i_{ff}\}$. See Table 1 for frequently used notations.

The genotype of an individual at position $\mathbb{i}$ is represented by the indicator variable $X_{\mathbb{i}}$. Each $X_{\mathbb{i}}$ takes some value that describes the allelic state; for example, if we have only two alleles per locus, the alleles can be represented by $X_{\mathbb{i}} = 1$ or $X_{\mathbb{i}} = 0$; for this special case, the frequency of allele 1 at position $\mathbb{i}$ is written as $p_{\mathbb{i}}$ and the frequency of allele 0 as $q_{\mathbb{i}} = 1 - p_{\mathbb{i}}$. This representation generalizes to multiple alleles, where the indicator is taken to be a vector of length equal to the number of alleles; all entries are zero except for a 1 that indicates the allelic state.

The genetic state of the population is described by the set of statistical moments of $X_{\mathbb{i}}$ across loci, called "associations" by Kirkpatrick *et al.* (2002), which define the allele frequencies and linkage disequilibria. Following Barton and Turelli (1991), the association between the alleles at the positions in set $\mathbb{A}$ is defined as

$$D_{\mathbb{A}} = E_X[\zeta_{\mathbb{A}}], \tag{1}$$

where

$$\zeta_{\mathbb{A}} = \Pi_{\mathbb{i}\in\mathbb{A}}\zeta_{\mathbb{i}}, \tag{2}$$

$\zeta_{\mathbb{i}} = X_{\mathbb{i}} - p_{\mathbb{i}}$, and the expectation $E_X[\,]$ is taken over the distribution of genotype frequencies. Products over the empty set are defined to be 1, so that $D_{\varnothing} = 1$. The D's are the same as the C's in Barton and Turelli (1991). If these associations are defined relative to the actual allele frequency in the population (*i.e.*, as central moments), then the $D_{\mathbb{i}}$ are necessarily zero. However, it may be convenient to define these relative to allele frequencies at the *start* of the generation, in which case $D_{\mathbb{i}} = \Delta p_{\mathbb{i}}$, the difference between the current and the initial allele frequency. With two alleles per locus, moments involving repeated indices reduce to expressions involving the mean. For example, $D_{\mathbb{i}\mathbb{i}} = p_{\mathbb{i}}q_{\mathbb{i}}$ if moments are defined relative to the current allele frequency.

We consider here two different measures of cytonuclear disequilibrium. In our notation, $D_{\mathbb{i}\mathbb{c}}$ is equivalent to the allelic cytonuclear disequilibrium, D, in Asmussen *et al.* (1987) and measures the nonrandom association of a nuclear locus ($\mathbb{i}$) and a maternally inherited cytoplasmic marker ($\mathbb{c}$). The third-order association can be written as $D_{\mathbb{i}\mathbb{i}^*\mathbb{c}}$, where * is an operator that changes sex-of-origin (for example, if the sex-of-origin for $\mathbb{i}$ is f, that for $\mathbb{i}^*$ is m); it measures the association between a cytoplasmic marker and homozygosity at a nuclear locus. $D_{\mathbb{i}\mathbb{i}^*\mathbb{c}}$ is equivalent to Weir and Wilson's (1986) residual disequilibrium d and is a simple function of the genotypic cytonuclear disequilibria (D_1, D_2, and D_3) of Asmussen *et al.* (1987). With one nuclear and one

TABLE 1

Frequently used notation

Symbol	Usage
i, j	Loci
c	Label for cytoplasmic locus
m, f	Male, female; used to specify contexts
ω	Label for deme; used to specify context; for the model considered in the text, there are three demes ($\nu - 1$, ν, $\nu + 1$)
$\mathring{\imath} = i_{\{\omega,\mathrm{m,f}\}}$	The position corresponding to locus i in the context {ω, m, f}, that is, genes at the locus i in deme ω, in males (sex-of-carrier) inherited from females (sex-of-origin)
*	Operator that changes sex-of-origin (for example, if the sex-of-origin of $\mathring{\imath}$ is f, that for $\mathring{\imath}$* is m)
$\mathbb{U}$	A set of positions
$\mathbb{U}\backslash\mathbb{V}$	The set $\mathbb{U}$ with the elements of the set $\mathbb{V}$ removed, defined only when $\mathbb{V}$ is a subset of $\mathbb{U}$
$X_{\mathring{\imath}}$	Indicator variable that labels the allelic state of position $\mathring{\imath}$
$\zeta_{\mathring{\imath}} = X_{\mathring{\imath}} - p_{\mathring{\imath}}$	Deviation of an individual at position $\mathring{\imath}$ from the reference value (here, the allele frequency)
$\zeta_{\mathbb{U}} = \prod_{\mathring{\imath}\in\mathbb{U}} \zeta_{\mathring{\imath}}$	The product of the deviations over the set of positions $\mathbb{U}$
$D_{\mathbb{U}} = E_X[\zeta_{\mathbb{U}}]$	The association between the set of positions $\mathbb{U}$
$D'_{\mathbb{U}}$	Associations between positions in the set $\mathbb{U}$ after migration, but before recombination and random mating
$D''_{\mathbb{U}}$	Associations between positions in the set $\mathbb{U}$ after recombination and random mating
$D_{\mathbb{U}_\omega}$	Value of $D_{\mathbb{U}}$ in deme ω
$\delta D_{\mathbb{U}_\omega}$	The difference between $D_{\mathbb{U}}$ in deme ω and that in the deme under consideration (ν); $\delta D_{\mathbb{U}_\omega} = D_{\mathbb{U}_\omega} - D_{\mathbb{U}_\nu}$
$p_{\mathring{\imath}_\omega}$	Allele frequency at position $\mathring{\imath}$ in deme ω
$\Delta p_{\mathring{\imath}_\omega}$	Difference between the allele frequency at position $\mathring{\imath}$ in deme ω and that in the deme under consideration (ν) after migration; $\Delta p_{\mathring{\imath}_\omega} = p_{\mathring{\imath}_\omega} - p'_{\mathring{\imath}_\nu}$
$d_{\mathring{\imath}_\omega}$	Difference between the allele frequency at position $\mathring{\imath}$ in deme ω and that in the deme under consideration (ν); $d_{\mathring{\imath}_\omega} = p_{\mathring{\imath}_\omega} - p_{\mathring{\imath}_\nu}$
m_ω	Migration rate for deme ω; for model in text, $m_{\nu-1} = m_{\nu+1} = m/2$, $m_\nu = 1 - m$
$t_{\mathbb{A}\leftarrow\mathbb{B}}$	Proportion of genes at positions $\mathbb{A}$ that were transmitted from positions $\mathbb{B}$

cytoplasmic locus, both with two alleles, there are six distinct genotypes. Assuming that maternally and paternally derived alleles are equivalent, these genotype frequencies are determined by $p_{\mathring{\imath}}$, $p_{\mathring{c}}$, $D_{\mathring{\imath}\mathring{\imath}*}$, $D_{\mathring{\imath}\mathring{c}}$, and $D_{\mathring{\imath}\mathring{\imath}*\mathring{c}}$, where $D_{\mathring{\imath}\mathring{\imath}*}$ measures the heterozygote deficit at the nuclear locus.

Nonrandom mating can generate associations across diploid mating pairs (for example, $D_{\{i_{\mathrm{mm}},j_{\mathrm{mf}},c_{\mathrm{mf}},k_{\mathrm{fm}},l_{\mathrm{ff}}\}}$); these in turn produce associations across genomes in the diploid offspring. If there is random mating, there are no associations across mating pairs ($D_{\{i_{\mathrm{mm}},j_{\mathrm{mf}},c_{\mathrm{mf}},k_{\mathrm{fm}},l_{\mathrm{ff}}\}} = D_{\{i_{\mathrm{mm}},j_{\mathrm{mf}},c_{\mathrm{mf}}\}}D_{\{k_{\mathrm{fm}},l_{\mathrm{ff}}\}}$); if there is random union of gametes (no inbreeding), there are no associations across genomes ($D_{\{i_{\mathrm{mm}},j_{\mathrm{mf}},c_{\mathrm{mf}}\}} = D_{\{i_{\mathrm{mm}}\}}D_{\{j_{\mathrm{mf}},c_{\mathrm{mf}}\}}$), which allows us to follow haploid genotypes, as in Barton and Turelli (1991).

Kirkpatrick *et al.* (2002) extended the notation of Barton and Turelli (1991) to describe associations among arbitrary sets of loci, which may be in different individuals and which may have unusual modes of inheritance. This does not change the form of the equations for selection (Equations 10 and 12 of Barton and Turelli 1991 or Equation 19 of Turelli and Barton 1994), provided that relative fitness is defined as the relative contribution of each gene combination to the next generation (see also Price 1970; Frank and Slatkin 1990). Nonrandom mating between diploid individuals can be described in just the same way as selection, by defining fitness as the relative contribution of a diploid *pair*. This is analogous to Barton and Turelli's (1991) treatment of nonrandom mating between haploids.

Model of migration: We assume that migration of diploid individuals among demes occurs first, followed by random mating within demes (Figure 1). Initially, with only one migration rate for both males and females, we need follow only the sex-of-origin of each gene, denoted by the subsubscript f or m. Asmussen *et al.* (1989) calculate the allelic cytonuclear disequilibrium in diploid adults by considering the disequilibrium in the gametes that those adults would produce, given normal Mendelian segregation; in our notation, this is ($D_{\{i_{\mathrm{m}},c_{\mathrm{f}}\}} + D_{\{i_{\mathrm{f}},c_{\mathrm{f}}\}}$)/2, where the f or m subsubscript refers to a maternally or paternally inherited allele. (Note that the two components $D_{\{i_{\mathrm{m}},c_{\mathrm{f}}\}}$ and $D_{\{i_{\mathrm{f}},c_{\mathrm{f}}\}}$ cannot be measured separately without knowing the origin of the autosomal alleles.) Therefore, census scheme 2 from Asmussen *et al.* (1989) corresponds to first allowing individuals to migrate and then considering the haploid gametes produced by those individuals after meiosis.

Migration is by diploid individuals; the location of an individual can be specified in the context in the same

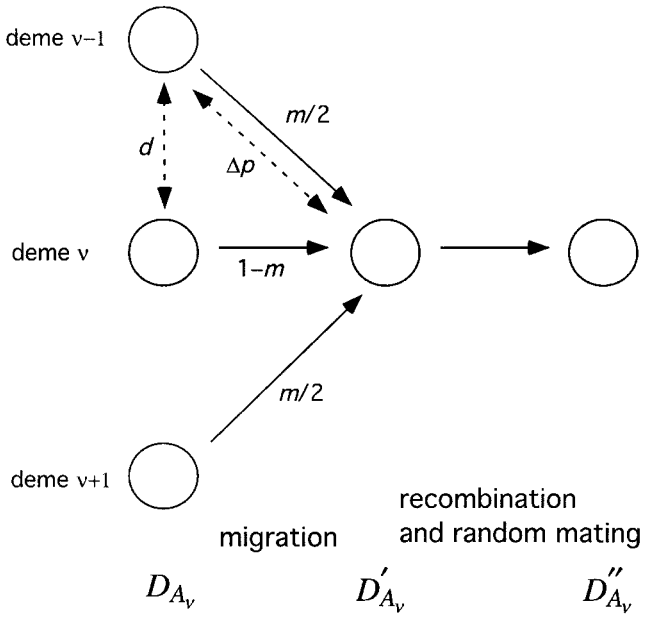

FIGURE 1.—Diagram of life cycle for migration model; diploid individuals migrate between demes and then mate at random within demes (census scheme 2). The measures of association between loci in the deme of interest ν are given by $D_{\mathbb{A}_\nu}$, $D'_{\mathbb{A}_\nu}$, and $D''_{\mathbb{A}_\nu}$ at various stages of the life cycle. Differences in allele frequencies between demes are d; differences in allele frequencies between demes before and after migration are given by Δp.

way that we specify sex-of-origin. The demes are labeled ω. Because we use central moments to describe disequilibria, which are measured relative to different reference points in each deme, we need an equation for admixture that takes into account the differences in allele frequency of each deme providing migrants (ω) and the deme under consideration (ν). Let $\Delta p_{\mathbb{i}_\omega} = (p_{\mathbb{i}_\omega} - p'_{\mathbb{i}_\nu})$ be the difference between the allele frequency in the contributing deme ω before migration and that in the deme under consideration after migration, where the shorthand notation $\mathbb{A}_\omega$ or $\mathbb{i}_\omega$ indicates the set of genes $\mathbb{A}$ or gene $\mathbb{i}$ located in deme ω. Note that this usage is inconsistent with the notation of KIRKPATRICK *et al.* (2002), but we adopt it here for convenience. We can also express the difference in allele frequencies between each deme providing migrants and the deme of interest prior to migration,

$$d_{\mathbb{i}_\omega} = p_{\mathbb{i}_\omega} - p_{\mathbb{i}_\nu} \quad \text{and} \quad d_{\mathbb{U}_\omega} = \prod_{\mathbb{i}\in\mathbb{U}}(p_{\mathbb{i}_\omega} - p_{\mathbb{i}_\nu}). \quad (3)$$

The allele frequency in deme ν after migration is given by

$$p'_{\mathbb{i}_\nu} = \sum_\omega m_\omega p_{\mathbb{i}_\omega}, \quad (4)$$

where the migration rate m_ω gives the proportion of individuals contributed by deme ω. This is equivalent to Equation 29 of KIRKPATRICK *et al.* (2002) if we allow $\Sigma_\omega m_\omega = m$ for $\omega \neq \nu$ and $m_\nu = 1 - m$. After migration, the associations in the deme ω relative to the allele frequencies in the deme of interest after migration are given by

$$\begin{aligned} D'_{\mathbb{A}_\omega} &= E_X[\textstyle\prod_{\mathbb{i}\in\mathbb{A}}(X_{\mathbb{i}_\omega} - p'_{\mathbb{i}_\nu})] \\ &= E_X[\textstyle\prod_{\mathbb{i}\in\mathbb{A}}(X_{\mathbb{i}_\omega} - p_{\mathbb{i}_\omega} + \Delta p_{\mathbb{i}_\omega})] \\ &= \textstyle\sum_{\mathbb{U}\subseteq\mathbb{A}} D_{\mathbb{A}\backslash\mathbb{U}_\omega} \prod_{\mathbb{i}\in\mathbb{U}} \Delta p_{\mathbb{i}_\omega}, \end{aligned} \quad (5)$$

where $\mathbb{A}\backslash\mathbb{U}$ stands for the positions in set $\mathbb{A}$ that are left after those in set $\mathbb{U}$ are taken away. Because genes change their contexts as they move from one deme to another, migration can be considered a form of transmission (an event that changes the contexts of genes). For this reason, (5) is simply (15) from KIRKPATRICK *et al.* (2002), giving the associations after a change in reference values from $p_{\mathbb{i}_\omega}$ to $p'_{\mathbb{i}_\nu}$.

The associations in the deme of interest (ν) after migration are then a linear mixture of the associations of different demes, measured relative to the same reference point:

$$D'_{\mathbb{A}_\nu} = \textstyle\sum_\omega m_\omega \sum_{\mathbb{U}\subseteq\mathbb{A}} D_{\mathbb{A}\backslash\mathbb{U}_\omega} \prod_{\mathbb{i}\in\mathbb{U}} \Delta p_{\mathbb{i}_\omega}. \quad (6)$$

If we instead use the allele frequencies in the deme of interest *prior* to migration as the reference point, so that we use the deme of interest before and after migration as the shift in reference, we require two changes of reference. When we again consider the deme of interest (ν), we find that (6) can be written as

$$\begin{aligned} D'_{\mathbb{A}_\nu} &= \textstyle\sum_\omega m_\omega D'_{\mathbb{A}_\omega} \\ &= \textstyle\sum_\omega m_\omega E_X[\prod_{\mathbb{i}\in\mathbb{A}}(X_{\mathbb{i}_\omega} - p'_{\mathbb{i}_\nu})] \\ &= \textstyle\sum_\omega m_\omega E_X[\prod_{\mathbb{i}\in\mathbb{A}}(X_{\mathbb{i}_\omega} - \sum_x m_x p_{\mathbb{i}_x})] \\ &= m_\nu \textstyle\sum_{\mathbb{U}\subseteq\mathbb{A}} D_{\mathbb{A}\backslash\mathbb{U}_\nu} \prod_{\mathbb{i}\in\mathbb{U}} (\sum_x (-m_x) d_{\mathbb{i}_x}) \\ &\quad + \textstyle\sum_{\omega\neq\nu} m_\omega \sum_{\mathbb{U}\subseteq\mathbb{A}} \prod_{\mathbb{i}\in\mathbb{U}} (\sum_x (-m_x) d_{\mathbb{i}_x}) \sum_{\mathbb{V}\subseteq\mathbb{A}\backslash\mathbb{U}} d_{\mathbb{V}_\omega} D_{(\mathbb{A}\backslash\mathbb{U})\backslash\mathbb{V}_\omega}. \end{aligned} \quad (7)$$

If we consider only one migrant deme, and let $m_\nu = 1 - m$, $m_\omega = m$, $D_{\mathbb{A}_\nu} = D^R_{\mathbb{A}}$, and $D_{\mathbb{A}_\omega} = D^M_{\mathbb{A}}$, (7) is equivalent to Equation (30) from KIRKPATRICK *et al.* (2002),

$$\begin{aligned} D'_{\mathbb{A}} &= \textstyle\sum_{\mathbb{U}\subseteq\mathbb{A}} (-m)^{|\mathbb{U}|} d_{\mathbb{U}} \{(1 - m) D^R_{\mathbb{A}\backslash\mathbb{U}} + m \sum_{\mathbb{V}\subseteq\mathbb{A}\backslash\mathbb{U}} d_{\mathbb{V}} D^{\mathrm{M}}_{(\mathbb{A}\backslash\mathbb{U})\backslash\mathbb{V}}\} \\ &= (1 - m) \textstyle\sum_{\mathbb{U}\subseteq\mathbb{A}} (-m)^{|\mathbb{U}|} d_{\mathbb{U}} D^R_{\mathbb{A}\backslash\mathbb{U}} \\ &\quad + m \textstyle\sum_{\mathbb{U}\subseteq\mathbb{A}} (-m)^{|\mathbb{U}|} d_{\mathbb{U}} \sum_{\mathbb{V}\subseteq\mathbb{A}\backslash\mathbb{U}} d_{\mathbb{V}} D^{\mathrm{M}}_{(\mathbb{A}\backslash\mathbb{U})\backslash\mathbb{V}} \end{aligned} \quad (8)$$

where $|\mathbb{U}|$ means the number of positions in set $\mathbb{U}$.

All of the above assumes that male and female migration rates are equal; we generalize this later to include different male and female migration rates and discuss differences that arise in males versus females.

Description of recombination and random mating: Under the assumption that there is Mendelian segregation, with no genetic variation for the rules of transmission (*e.g.*, no meiotic drive or modifiers of recombination), the effects of recombination and random mating on the associations and allele frequencies for a popula-

tion can be described by a simple equation (Equation 12, Kirkpatrick *et al.* 2002),

$$D''_{\mathbb{A}} = \sum_{\mathbb{U}:\, U=A} t_{\mathbb{A}\leftarrow\mathbb{U}} D'_{\mathbb{U}}, \tag{9}$$

when the reference values are chosen to be equal for all positions at each locus. The summation is over all sets of positions $\mathbb{U}$ that could become set $\mathbb{A}$ following transmission. The notation "$\mathbb{U}$: $U = A$" means that the genes in the sets $\mathbb{U}$ and $\mathbb{A}$ must be equal when the context information is stripped from them (*i.e.*, when $U = A$). The transmission coefficient $t_{\mathbb{A}\leftarrow\mathbb{U}}$ is defined as the probability that the positions in set $\mathbb{A}$ were inherited from positions in set $\mathbb{U}$. As an example, for autosomal loci in a haploid population with two sexes, the transmission coefficient $t_{\{i_{\mathrm{m}}, j_{\mathrm{m}}\}\leftarrow\{i_{\mathrm{f}}, j_{\mathrm{f}}\}}$ is the probability that the genes at loci i and j in a male were both inherited from a female (the mother), which is $(1 - r_{ij})/2$, where r_{ij} is the recombination rate between loci i and j.

For transmission in diploids, the sex-of-origin for each position in set $\mathbb{A}$ must equal the sex-of-carrier for the corresponding position in set $\mathbb{U}$, since that is the sex of the parent from which a gene in set $\mathbb{A}$ is descended. When A is a set of cytoplasmic loci or a mixture of cytoplasmic and nuclear genes, then certain transmissions do not occur; for example, under the assumption of a maternally inherited cytoplasmic locus (c) and an autosomal locus (i), $t_{\{i_{\mathrm{xm}}, c_{\mathrm{xm}}\}\leftarrow\{i_{\mathrm{mm}}, c_{\mathrm{mf}}\}}$ and $t_{\{i_{\mathrm{xm}}, c_{\mathrm{xm}}\}\leftarrow\{i_{\mathrm{mf}}, c_{\mathrm{mf}}\}}$ (where x is f or m) are both zero, because a male cannot transmit his maternally inherited cytoplasmic allele to his sons or daughters.

As an example, consider the genotypic cytonuclear disequilibrium, $D_{\mathrm{i\dot{i}}*\mathbb{c}}$, in female zygotes. Here, $\mathbb{A} = \{i_{\mathrm{fm}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}\}$. These might have come from four possible sets of positions $\mathbb{U}$ in the previous generation of diploids, $\{i_{\mathrm{mf}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}\}$, $\{i_{\mathrm{mm}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}\}$, $\{i_{\mathrm{mf}}, i_{\mathrm{fm}}, c_{\mathrm{ff}}\}$, and $\{i_{\mathrm{mm}}, i_{\mathrm{fm}}, c_{\mathrm{ff}}\}$. The transition coefficients for all of these equal ¼; to see this, consider $t_{\{i_{\mathrm{fm}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}\}\leftarrow\{i_{\mathrm{mf}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}\}}$. This is the probability that, in a female, the locus i inherited from her male parent was inherited from that parent's female parent, which is ½, times the probabilty that the locus i inherited from her female parent was inherited from that parent's female parent, which also equals ½. The cytoplasmic locus must pass through the female lineage, with probability 1. Putting this together gives

$$D''_{\mathrm{i\dot{i}}*\mathbb{c}} = \frac{1}{4}(D'_{i_{\mathrm{mf}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}} + D'_{i_{\mathrm{mm}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}} + D'_{i_{\mathrm{mf}}, i_{\mathrm{fm}}, c_{\mathrm{ff}}} + D'_{i_{\mathrm{mm}}, i_{\mathrm{fm}}, c_{\mathrm{ff}}}). \tag{10}$$

This shows that any associations between the maternally derived allele and the paternally derived allele at the same locus must have originated in associations between alleles from the different diploid parents (and, in each parent, could have involved the maternally derived or paternally derived gene). If we assume random mating, there are no associations between alleles in different diploid individuals (so that $D'_{i_{\mathrm{mf}}, i_{\mathrm{ff}}, c_{\mathrm{ff}}} = D'_{i_{\mathrm{ff}}, c_{\mathrm{ff}}} D'_{i_{\mathrm{mf}}}$, etc.), and $D'_{\mathbb{i}} = 0$ (we have defined the associations relative to the actual allele frequencies in the population, as discussed earlier, and thus the disequilibria are central moments), so that $D''_{\mathrm{i\dot{i}}*\mathbb{c}} = 0$.

We now consider the effects of this simple model of migration followed by meiosis and random mating on both the allelic cytonuclear disequilibrium ($D_{\mathrm{i}\mathbb{c}}$) and the genotypic cytonuclear disequilibrium ($D_{\mathrm{i\dot{i}}*\mathbb{c}}$).

Allelic cytonuclear disequilibrium, $D_{\mathrm{i}\mathbb{c}}$: The allelic cytonuclear disequilibrium is $D_{\mathrm{i}\mathbb{c}} = E[(X_{\mathrm{i}} - p_{\mathrm{i}})(X_{\mathrm{c}} - p_{\mathrm{c}})] = E[X_{\mathrm{i}}X_{\mathrm{c}}] - p_{\mathrm{i}}p_{\mathrm{c}}$. In Asmussen *et al.*'s (1989) notation, $D_{\mathrm{i}\mathbb{c}} = D = D_1 + \frac{1}{2}D_2$. We consider a specific model of continued immigration at rate $\frac{1}{2}m$ into a deme of interest (ν) from two source demes ($\nu - 1$, $\nu + 1$) and consider the equilibrium under migration. Again, when considering only one migration rate for both males and females, we need follow only the sex-of-origin of each gene.

We use Equation 6 and note that the possible sets of positions for $\mathbb{U}$ include $\{\mathbb{i}, \mathbb{c}\}$, $\{\mathbb{i}\}$, $\{\mathbb{c}\}$, and $\{\varnothing\}$. We define $D_{\varnothing} = 1$ and $\Delta p_{\varnothing} = 1$, and, since we are using central moments, defined relative to each deme, $D_{\mathbb{i}} = 0$ by definition, giving us

$$\begin{aligned} D'_{\mathrm{i}\mathbb{c}_\nu} = {} & (1 - m) D_{\mathrm{i}\mathbb{c}_\nu} + \frac{m}{2}(D_{\mathrm{i}\mathbb{c}_{\nu-1}} + D_{\mathrm{i}\mathbb{c}_{\nu+1}}) \\ & + (1 - m)\Delta p_{i_{\{\nu\}}} \Delta p_{c_{\{\nu\}}} + \frac{m}{2}(\Delta p_{i_{\{\nu-1\}}} \Delta p_{c_{\{\nu-1\}}} + \Delta p_{i_{\{\nu+1\}}} \Delta p_{c_{\{\nu+1\}}}). \end{aligned} \tag{11}$$

We can express this result in terms of deviations from moments and allele frequencies in the deme of interest (ν) by using (3) and allowing $\delta D_{\mathbb{A}_\omega} = D_{\mathbb{A}_\omega} - D_{\mathbb{A}_\nu}$. (Note that these deviations are across demes at the same time point, whereas $\Delta p_{i_{\{\omega\}}}$ involves deviations across space and time.) The context of a gene now includes both its deme-of-origin and its sex-of-origin (in that order). For an i allele with a female sex-of-origin, we then have

$$\begin{aligned} D'_{i_{\{\nu,\mathrm{f}\}} c_{\{\nu,\mathrm{f}\}}} = {} & D_{i_{\{\nu,\mathrm{f}\}} c_{\{\nu,\mathrm{f}\}}} + \frac{m}{2}(\delta D_{i_{\{\nu-1,\mathrm{f}\}} c_{\{\nu-1,\mathrm{f}\}}} + \delta D_{i_{\{\nu+1,\mathrm{f}\}} c_{\{\nu+1,\mathrm{f}\}}}) \\ & + \frac{m}{2}(d_{i_{\{\nu-1,\mathrm{f}\}}} d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{f}\}}} d_{c_{\{\nu+1,\mathrm{f}\}}}) \\ & - \frac{m^2}{4}(d_{i_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{f}\}}})(d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{c_{\{\nu+1,\mathrm{f}\}}}), \end{aligned} \tag{12}$$

where gene context is given as {deme, sex-of-origin}.

If individuals within demes are formed by random union of gametes, the association between the paternal genotype and the maternally inherited cytotype must be zero. In that case, the association $D_{i_{\{\omega,\mathrm{m}\}} c_{\{\omega,\mathrm{f}\}}}$ (where m indicates a male sex-of-origin and f a female sex-of-origin) is initially zero before migration, leaving

$$\begin{aligned} D'_{i_{\{\nu,\mathrm{m}\}} c_{\{\nu,\mathrm{f}\}}} = {} & \frac{m}{2}(d_{i_{\{\nu-1,\mathrm{m}\}}} d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{m}\}}} d_{c_{\{\nu+1,\mathrm{f}\}}}) \\ & - \frac{m^2}{4}(d_{i_{\{\nu-1,\mathrm{m}\}}} + d_{i_{\{\nu+1,\mathrm{m}\}}})(d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{c_{\{\nu+1,\mathrm{f}\}}}), \end{aligned} \tag{13}$$

where, again, context is given as {deme, sex-of-origin}. We next allow meiosis by applying Equation 9 and find

$$D''_{i_{\{\nu,\mathrm{f}\}}c_{\{\nu,\mathrm{f}\}}} = D''_{i_{\{\nu,\mathrm{m}\}}c_{\{\nu,\mathrm{f}\}}} = \frac{1}{2}(D'_{i_{\{\nu,\mathrm{f}\}}c_{\{\nu,\mathrm{f}\}}} + D'_{i_{\{\nu,\mathrm{m}\}}c_{\{\nu,\mathrm{f}\}}}). \quad (14)$$

Using (12) and (13), we find the full recursion for the allelic cytonuclear disequilibrium

$$\begin{aligned} D''_{\mathrm{i}\mathbb{c}_\nu} = {} & \frac{1}{2}D_{i_{\{\nu,\mathrm{f}\}}c_{\{\nu,\mathrm{f}\}}} + \frac{m}{4}(\delta D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}} + \delta D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}}) \\ & + \frac{m}{4}(d_{i_{\{\nu-1,\mathrm{f}\}}}d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu-1,\mathrm{m}\}}}d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{f}\}}}d_{c_{\{\nu+1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{m}\}}}d_{c_{\{\nu+1,\mathrm{f}\}}}) \\ & - \frac{m^2}{8}(d_{i_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{f}\}}} + d_{i_{\{\nu-1,\mathrm{m}\}}} + d_{i_{\{\nu+1,\mathrm{m}\}}})(d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{c_{\{\nu+1,\mathrm{f}\}}}). \end{aligned} \quad (15)$$

At equilibrium under this symmetric model of migration, $d_{i_{\{\nu-1,\mathrm{f}\}}} = -d_{i_{\{\nu+1,\mathrm{f}\}}} = d_{i_\mathrm{f}}$, $d_{i_{\{\nu-1,\mathrm{m}\}}} = -d_{i_{\{\nu+1,\mathrm{m}\}}} = d_{i_\mathrm{m}}$, and $d_{c_{\{\nu-1,\mathrm{f}\}}} = -d_{c_{\{\nu+1,\mathrm{f}\}}} = d_{c_\mathrm{f}}$, so that we have

$$\hat{D}_{\mathrm{i}\mathbb{c}_\nu} = m\left[\frac{1}{2}(\delta D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}} + \delta D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}}) + (d_{i_\mathrm{f}} + d_{i_\mathrm{m}})d_{c_\mathrm{f}}\right]. \quad (16)$$

Note that we have calculated the change in a haploid association after migration and meiosis; this corresponds to census scheme 2 in Asmussen *et al.* (1989), for which they find

$$\hat{D} = \frac{m}{1+m}[D^{(1)} + D^{(2)} + \frac{1}{2}(p_1 - p_2)(x_1 - x_2)]. \quad (17)$$

Here, $D^{(i)}$ gives the disequilibrium in deme i, demes 1 and 2 correspond to demes $\nu - 1$ and $\nu + 1$, and $m_1 = m_2 = m/2$. This is the result given in (16) if we assume $p_{i_\mathrm{m}} = p_{i_\mathrm{f}}$, since $D^{(1)} = \frac{1}{2}(D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}} + D_{i_{\{\nu-1,\mathrm{m}\}}c_{\{\nu-1,\mathrm{f}\}}})$, $D^{(2)} = \frac{1}{2}(D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}} + D_{i_{\{\nu+1,\mathrm{m}\}}c_{\{\nu+1,\mathrm{f}\}}})$ and, initially, $D_{i_\mathrm{m}c_\mathrm{f}} = 0$. The $(1 + m)$ term arises when we replace the deviations of moments ($\delta D_{\mathrm{i}\mathbb{c}_\omega}$) with the values of the moments in the neighboring demes ($D_{\mathrm{i}\mathbb{c}_\omega}$). Note that the behavior of the disequilibrium between a nuclear allele and a cytoplasmic marker under random mating and migration is the same as that between two unlinked nuclear loci (D_{ij}); the pairwise disequilibrium caused by admixture is an average of the disequilibria in the source populations, plus an additional amount caused by the covariance between the allele frequencies in the source populations (Nei and Li 1973; Prout 1973).

Genotypic cytonuclear disequilibrium $D_{\mathrm{ii}^*\mathbb{c}}$: The genotypic cytonuclear disequilibrium measures nonrandom association between the same nuclear locus from two different genomes and a cytoplasmic locus. We use the definition $D_{\mathrm{ii}^*\mathbb{c}} = E[(X_{i_\mathrm{f}} - p_{i_\mathrm{f}})(X_{i_\mathrm{m}} - p_{i_\mathrm{m}})(X_{c_\mathrm{f}} - p_{c_\mathrm{f}})]$, where the f subsubscript indicates female sex-of-origin, for diallelic loci with indicator X_i. We let $X_\mathrm{i} = 1$ if P_i is present at locus i and $X_\mathrm{i} = 0$ if Q_i is present at locus i, and similarly for other loci. From the definitions of D_1, D_2, and D_3 (Asmussen *et al.* 1989),

$$\begin{aligned} D_1 &= \psi(P_{i_\mathrm{f}}P_{i_\mathrm{m}}P_{c_\mathrm{f}}) - \psi(P_{i_\mathrm{f}}P_{i_\mathrm{m}})\psi(P_{c_\mathrm{f}}) = \psi(P_{i_\mathrm{f}}P_{i_\mathrm{m}}P_{c_\mathrm{f}}) - \psi(P_{i_\mathrm{f}}P_{i_\mathrm{m}})p_{c_\mathrm{f}} \\ D_2 &= \psi(P_{i_\mathrm{f}}Q_{i_\mathrm{m}}P_{c_\mathrm{f}}) - \psi(P_{i_\mathrm{f}}Q_{i_\mathrm{m}})p_{c_\mathrm{f}} \\ D_3 &= \psi(Q_{i_\mathrm{f}}Q_{i_\mathrm{m}}P_{c_\mathrm{f}}) - \psi(Q_{i_\mathrm{f}}Q_{i_\mathrm{m}})p_{c_\mathrm{f}}, \end{aligned} \quad (18)$$

where, for example, $\psi(P_{i_\mathrm{f}}P_{i_\mathrm{m}}P_{c_\mathrm{f}})$ gives the frequency of $P_{i_\mathrm{f}}P_{i_\mathrm{m}}P_{c_\mathrm{f}}$. Using these definitions and allowing $p_{i_\mathrm{f}} = p_{i_\mathrm{m}}$ gives $D_{\mathrm{ii}^*\mathbb{c}} = D_1 - 2p_\mathrm{i}D = D_3p_\mathrm{i}^2 - D_2p_\mathrm{i}q_\mathrm{i} + D_1q_\mathrm{i}^2$, where $q_\mathrm{i} = 1 - p_\mathrm{i}$. Note that, necessarily, $D_1 + D_2 + D_3 = 0$, so that the genotypic cytonuclear disequilibria as defined by Asmussen *et al.* are not independent, as was originally noted by Asmussen *et al.* (1987).

To consider the case of admixture, we once again apply the migration model given earlier; (6) again gives the change in the disequilibrium due to admixture where, here, $\mathbb{A} = \{\mathrm{i}, \mathrm{i}^*, \mathbb{c}\}$. Since $D_\varnothing = 1$, and $\Delta p_\varnothing = 1$, we have

$$\begin{aligned} D'_{\mathrm{ii}^*\mathbb{c}_\nu} = {} & (1-m)D_{\mathrm{ii}^*\mathbb{c}_\nu} + \frac{m}{2}(D_{\mathrm{ii}^*\mathbb{c}_{\nu-1}} + D_{\mathrm{ii}^*\mathbb{c}_{\nu+1}}) \\ & + (1-m)(D_{i_{\{\nu,\mathrm{f}\}}i_{\{\nu,\mathrm{m}\}}}\Delta p_{c_{\{\nu,\mathrm{f}\}}} + D_{i_{\{\nu,\mathrm{f}\}}c_{\{\nu,\mathrm{f}\}}}\Delta p_{i_{\{\nu,\mathrm{m}\}}} + D_{i_{\{\nu,\mathrm{m}\}}c_{\{\nu,\mathrm{f}\}}}\Delta p_{i_{\{\nu,\mathrm{f}\}}}) \\ & + \frac{m}{2}(D_{i_{\{\nu-1,\mathrm{f}\}}i_{\{\nu-1,\mathrm{m}\}}}\Delta p_{c_{\{\nu-1,\mathrm{f}\}}} + D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}}\Delta p_{i_{\{\nu-1,\mathrm{m}\}}} + D_{i_{\{\nu-1,\mathrm{m}\}}c_{\{\nu-1,\mathrm{f}\}}}\Delta p_{i_{\{\nu-1,\mathrm{f}\}}}) \\ & + \frac{m}{2}(D_{i_{\{\nu+1,\mathrm{f}\}}i_{\{\nu+1,\mathrm{m}\}}}\Delta p_{c_{\{\nu+1,\mathrm{f}\}}} + D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}}\Delta p_{i_{\{\nu+1,\mathrm{m}\}}} + D_{i_{\{\nu+1,\mathrm{m}\}}c_{\{\nu+1,\mathrm{f}\}}}\Delta p_{i_{\{\nu+1,\mathrm{f}\}}}) \\ & + (1-m)(\Delta p_{i_{\{\nu,\mathrm{f}\}}}\Delta p_{i_{\{\nu,\mathrm{m}\}}}\Delta p_{c_{\{\nu,\mathrm{f}\}}}) \\ & + \frac{m}{2}(\Delta p_{i_{\{\nu-1,\mathrm{f}\}}}\Delta p_{i_{\{\nu-1,\mathrm{m}\}}}\Delta p_{c_{\{\nu-1,\mathrm{f}\}}} + \Delta p_{i_{\{\nu+1,\mathrm{f}\}}}\Delta p_{i_{\{\nu+1,\mathrm{m}\}}}\Delta p_{c_{\{\nu+1,\mathrm{f}\}}}), \end{aligned} \quad (19)$$

where, again, we indicate context by the subsubscript {deme, sex-of-origin}. If individuals within demes are formed by random union of gametes, there will necessarily be no association between maternal and paternal loci ($D_{\mathrm{ii}^*\mathbb{c}} = D_{\mathrm{i}\mathbb{c}}D_{\mathrm{i}^*} = 0$, etc.), and the equation above simplifies to

$$\begin{aligned} D'_{\mathrm{ii}^*\mathbb{c}_\nu} = {} & \frac{m}{2}(\delta D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}}d_{i_{\{\nu-1,\mathrm{m}\}}} + \delta D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}}d_{i_{\{\nu+1,\mathrm{m}\}}}) \\ & - \frac{m^2}{4}(\delta D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}} + \delta D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}})(d_{i_{\{\nu-1,\mathrm{m}\}}} + d_{i_{\{\nu+1,\mathrm{m}\}}}) \\ & - \frac{m(1-m)}{2}(d_{i_{\{\nu-1,\mathrm{f}\}}}d_{i_{\{\nu-1,\mathrm{m}\}}}d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{f}\}}}d_{i_{\{\nu+1,\mathrm{m}\}}}d_{c_{\{\nu+1,\mathrm{f}\}}}) \\ & - \frac{m^2(1-m)}{4}(d_{i_{\{\nu-1,\mathrm{f}\}}} + d_{i_{\{\nu+1,\mathrm{f}\}}})(d_{i_{\{\nu-1,\mathrm{m}\}}} + d_{i_{\{\nu+1,\mathrm{m}\}}})(d_{c_{\{\nu-1,\mathrm{f}\}}} + d_{c_{\{\nu+1,\mathrm{f}\}}}). \end{aligned} \quad (20)$$

Thus, the genotypic cytonuclear disequilibrium after migration is due to interactions between the allelic disequilibrium and the nuclear allele frequency in males plus additional disequilibrium caused by the differences in allele frequencies in the migrant demes relative to those in the deme under consideration. Once again, cytoplasmic inheritance is irrelevant at this stage; the disequilibrium after migration for this model would be the same for a nuclear genotype and an unlinked nuclear locus ($D_{i_\mathrm{f}i_\mathrm{m}j_\mathrm{f}}$) or (substituting k for i) for the three-locus association $D_{i_\mathrm{m}j_\mathrm{f}k_\mathrm{f}}$.

As shown earlier, after meiosis,

$$D''_{\mathbb{i}\mathbb{i}^*\mathbb{c}} = \frac{1}{4}(D'_{i_{\mathrm{mf}},i_{\mathrm{ff}},c_{\mathrm{ff}}} + D'_{i_{\mathrm{mm}},i_{\mathrm{ff}},c_{\mathrm{ff}}} + D'_{i_{\mathrm{mf}},i_{\mathrm{fm}},c_{\mathrm{ff}}} + D'_{i_{\mathrm{mm}},i_{\mathrm{fm}},c_{\mathrm{ff}}}), \quad (21)$$

since any association involving a maternally derived allele and a paternally derived allele must originate in alleles that were in different individuals in the previous generation. Once again, if we assume random mating, there are no associations between alleles in different diploid individuals and $D''_{\mathbb{i}\mathbb{i}^*\mathbb{c}} = 0$. Therefore, at equilibrium, the disequilibrium generated by migration is lost after random mating and meiosis ($\hat{D}_{\mathbb{i}\mathbb{i}^*\mathbb{c}} = 0$). Note that the same is true under random mating when considering the association between the same nuclear locus from the maternal and paternal genomes and another nuclear locus ($\hat{D}_{i_{\mathrm{f}}i_{\mathrm{m}}j_{\mathrm{f}}} = 0$), but it is not true for three different nuclear loci on the same genome ($\hat{D}_{i_{\mathrm{f}}j_{\mathrm{f}}k_{\mathrm{f}}}$), since then all three alleles are derived from the same parent.

We have allowed individuals to migrate and then mate and are now considering associations in the new diploid individuals formed after mating; this corresponds to census 2 from Asmussen *et al.* (1989). As shown above, $D_{\mathbb{i}\mathbb{i}^*\mathbb{c}} = D_{\mathbb{1}} - 2p_{\mathbb{i}}D$, and from Equations B12 and B13 of Asmussen *et al.* (1989), $\hat{D}_{\mathbb{1}} - 2p_{\mathbb{i}}\hat{D} = 0$, so the two methods give the same result.

Recombination and random mating followed by migration: In the derivation above, we have assumed a model where individuals migrate and then mate; censusing the population after random mating means that associations between maternal and paternal genomes are lost and associations within genomes are reduced by random segregation and mating. If instead the population is censused after migration, we find

$$\hat{D}_{i_{\{\nu,\mathrm{f}\}}c_{\{\nu,\mathrm{f}\}}} = m[\delta D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}} + \delta D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}} + (2d_{i_{\mathrm{f}}} + d_{i_{\mathrm{m}}})d_{c_{\mathrm{f}}}] \quad (22)$$

$$\hat{D}_{i_{\{\nu,\mathrm{m}\}}c_{\{\nu,\mathrm{f}\}}} = md_{i_{\mathrm{m}}}d_{c_{\mathrm{f}}} \quad (23)$$

$$\hat{D}_{i_{\{\nu,\mathrm{f}\}}i_{\{\nu,\mathrm{m}\}}c_{\{\nu,\mathrm{f}\}}} = \frac{m}{2}(\delta D_{i_{\{\nu-1,\mathrm{f}\}}c_{\{\nu-1,\mathrm{f}\}}} - \delta D_{i_{\{\nu+1,\mathrm{f}\}}c_{\{\nu+1,\mathrm{f}\}}})d_{i_{\mathrm{m}}}, \quad (24)$$

where now the $\delta D_{\mathbb{A}_\omega}$ and $d_{\mathbb{i}}$ are deviations from moments and allele frequencies in the deme of interest after recombination, $d_{i_{\{\nu-1,x\}}} = -d_{i_{\{\nu+1,x\}}} = d_{i_x}$, where x is f or m, and we recall that, with random mating, $D_{i_{\{\omega,\mathrm{f}\}}i_{\{\omega,\mathrm{m}\}}}$, $D_{i_{\{\omega,\mathrm{m}\}}c_{\{\omega,\mathrm{f}\}}}$, and $D_{i_{\{\omega,\mathrm{f}\}}i_{\{\omega,\mathrm{m}\}}c_{\{\omega,\mathrm{f}\}}}$ are all zero after recombination. The overall allelic cytonuclear disequilibrium will again be $(\hat{D}_{i_{\{\nu,\mathrm{f}\}}c_{\{\nu,\mathrm{f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m}\}}c_{\{\nu,\mathrm{f}\}}})/2$, which corresponds to $\hat{D}$ for census 1 in Asmussen *et al.* (1989), assuming that $p_{i_{\mathrm{m}}} = p_{i_{\mathrm{f}}}$, and is twice the allelic cytonuclear disequilibrium found under census 2.

Nuclear-nuclear *vs.* cytonuclear disequilibria: In the case of differences between females and males, the asymmetric inheritance of cytoplasmic markers becomes important. For example, with two unlinked nuclear loci, the disequilibrium after meiosis is given by (9) as

$$D''_{\mathbb{i}\mathbb{j}} = (D'_{i_{\mathrm{f}}j_{\mathrm{f}}} + D'_{i_{\mathrm{m}}j_{\mathrm{m}}} + D'_{i_{\mathrm{f}}j_{\mathrm{m}}} + D'_{i_{\mathrm{m}}j_{\mathrm{f}}})/4, \quad (25)$$

where the f or m subsubscript indicates sex-of-origin. This will only be equivalent to that for the cytonuclear disequilibrium,

$$D''_{\mathbb{i}\mathbb{c}} = (D'_{i_{\mathrm{f}}c_{\mathrm{f}}} + D'_{i_{\mathrm{m}}c_{\mathrm{f}}})/2, \quad (26)$$

when there are no differences for within- and between-genome associations and allele frequencies between males and females ($D'_{i_{\mathrm{f}}j_{\mathrm{f}}} = D'_{i_{\mathrm{m}}j_{\mathrm{m}}}$ and $D'_{i_{\mathrm{f}}j_{\mathrm{m}}} = D'_{i_{\mathrm{m}}j_{\mathrm{f}}}$). Such differences could arise for a variety of reasons, such as imprinting, selection, or differences in allele frequencies or migration rates in males and females. We now consider the latter.

Unequal migration rates between the sexes: To incorporate different migration rates for males and females, we need to keep track of moments and allele frequencies separately in the two sexes. If we consider the allelic cytonuclear disequilibrium, in females after migration we have $D'_{i_{\{\mathrm{f,f}\}}c_{\{\mathrm{f,f}\}}}$ and $D'_{i_{\{\mathrm{f,m}\}}c_{\{\mathrm{f,f}\}}}$ as in (12) and (13) (where now the first subsubscript refers to sex-of-carrier and the second to sex-of-origin) and we replace the migration rate m with m_{f}, which refers only to migration of females. In males, $D'_{i_{\{\mathrm{m,f}\}}c_{\{\mathrm{m,f}\}}}$ and $D'_{i_{\{\mathrm{m,m}\}}c_{\{\mathrm{m,f}\}}}$ are given by equations identical to (12) and (13), with the exception that now all loci are in males and the migration rate is m_{m}. Because the mitochondrial locus cannot be paternally inherited, there are no associations such as $D_{i_{\{\mathrm{f,m}\}}c_{\{\mathrm{f,m}\}}}$, $D_{i_{\{\mathrm{m,m}\}}c_{\{\mathrm{m,m}\}}}$, $D_{i_{\{\mathrm{f,f}\}}c_{\{\mathrm{f,m}\}}}$, or $D_{i_{\{\mathrm{m,f}\}}c_{\{\mathrm{m,m}\}}}$.

We first consider migration followed by recombination and random mating, which corresponds to census scheme 2 of Asmussen *et al.* (1989). Recombination follows as before; any associations between the paternally inherited nuclear loci and the mitochondrial locus will be lost after the formation of gametes, so that $D''_{i_{\{\mathrm{f,m}\}}c_{\{\mathrm{f,f}\}}} = 0$ and $D''_{i_{\{\mathrm{m,m}\}}c_{\{\mathrm{m,f}\}}} = 0$. This leaves

$$D''_{i_{\{\mathrm{f,f}\}}c_{\{\mathrm{f,f}\}}} = (D'_{i_{\{\mathrm{f,f}\}}c_{\{\mathrm{f,f}\}}} + D'_{i_{\{\mathrm{f,m}\}}c_{\{\mathrm{f,f}\}}})/2 \quad (27)$$

and

$$D''_{i_{\{\mathrm{m,f}\}}c_{\{\mathrm{m,f}\}}} = (D'_{i_{\{\mathrm{f,f}\}}c_{\{\mathrm{f,f}\}}} + D'_{i_{\{\mathrm{f,m}\}}c_{\{\mathrm{f,f}\}}})/2 \quad (28)$$

(any association between the maternally derived nuclear allele and the mitochondrial locus in males must have originated in the females of the previous generation). This means that the cytonuclear disequilibrium will depend only on female migration rates and at equilibrium will be equal in males and females,

$$\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}c_{\{\nu,\mathrm{f,f}\}}} = \hat{D}_{i_{\{\nu,\mathrm{m,f}\}}c_{\{\nu,\mathrm{m,f}\}}} = \hat{D}_{i_{\{\nu,\mathrm{x,f}\}}c_{\{\nu,\mathrm{x,f}\}}} = m_{\mathrm{f}}\left[\frac{1}{2}(\delta D_{i_{\{\nu-1,\mathrm{x,f}\}}c_{\{\nu-1,\mathrm{x,f}\}}} + \delta D_{i_{\{\nu+1,\mathrm{x,f}\}}c_{\{\nu+1,\mathrm{x,f}\}}}) + (d_{i_{\{\mathrm{x,f}\}}} + d_{i_{\{\mathrm{x,m}\}}})d_{c_{\{\mathrm{x,f}\}}}\right], \quad (29)$$

where the gene context is now given as {deme, sex-of-carrier, sex-of-origin), $d_{i_{\{\nu-1,\mathrm{x,y}\}}} = -d_{i_{\{\nu+1,\mathrm{x,y}\}}} = d_{i_{\{\mathrm{x,y}\}}}$ un-

der a symmetric model of migration, m_f indicates the female migration rate, and x and y each are either f or m.

In contrast, for two unlinked nuclear loci, after meiosis the association between maternally derived alleles in both females and males is

$$D''_{i_{\{f,f\}}j_{\{f,f\}}} = D''_{i_{\{m,f\}}j_{\{m,f\}}} = (D'_{i_{\{f,f\}}j_{\{f,f\}}} + D'_{i_{\{f,m\}}j_{\{f,m\}}} + D'_{i_{\{f,f\}}j_{\{f,m\}}} + D'_{i_{\{f,m\}}j_{\{f,f\}}})/4 \tag{30}$$

and so depends only on the female migration rate (m_f). The association for the paternally derived alleles in both females and males is

$$D''_{i_{\{f,m\}}j_{\{f,m\}}} = D''_{i_{\{m,m\}}j_{\{m,m\}}} = (D'_{i_{\{m,f\}}j_{\{m,f\}}} + D'_{i_{\{m,m\}}j_{\{m,m\}}} + D'_{i_{\{m,f\}}j_{\{m,m\}}} + D'_{i_{\{m,m\}}j_{\{m,f\}}})/4 \tag{31}$$

and depends on the male migration rate (m_m). All of the cross-genome associations ($D''_{i_{\{f,m\}}j_{\{f,f\}}}$, $D''_{i_{\{f,f\}}j_{\{f,m\}}}$, $D''_{i_{\{m,f\}}j_{\{m,m\}}}$, $D''_{i_{\{m,m\}}j_{\{m,f\}}}$) would be zero under random mating. The two-locus disequilibrium measured in females is ($D_{i_{\{f,f\}}j_{\{f,f\}}} + D_{i_{\{f,m\}}j_{\{f,m\}}}$)/2, while that in males is ($D_{i_{\{m,f\}}j_{\{m,f\}}} + D_{i_{\{m,m\}}j_{\{m,m\}}}$)/2. In both cases, the disequilibrium would depend on both the female and male migration rates (m_f and m_m).

At equilibrium, the within-genome association in females equals the within-genome association in males,

$$\begin{aligned}(\hat{D}_{i_{\{f,f\}}j_{\{f,f\}}} + \hat{D}_{i_{\{f,m\}}j_{\{f,m\}}})/2 &= (\hat{D}_{i_{\{m,f\}}j_{\{m,f\}}} + \hat{D}_{i_{\{m,m\}}j_{\{m,m\}}})/2 \\ &= \frac{m_f + m_m}{4}\Big[\frac{1}{2}(\delta D_{i_{\{v-1,*,f\}}j_{\{v-1,*,f\}}} + \delta D_{i_{\{v+1,*,f\}}j_{\{v+1,*,f\}}} \\ &\quad + \delta D_{i_{\{v-1,*,m\}}j_{\{v-1,*,m\}}} + \delta D_{i_{\{v+1,*,m\}}j_{\{v+1,*,m\}}}) \\ &\quad + (d_{i_{\{*,f\}}} + d_{i_{\{*,m\}}})(d_{j_{\{*,f\}}} + d_{j_{\{*,m\}}})\Big],\end{aligned} \tag{32}$$

where, after random mating and assuming no selection, associations and allele frequencies must be equal in males and females, so that $D_{i_{\{\omega,f,f\}}j_{\{\omega,f,f\}}} = D_{i_{\{\omega,m,f\}}j_{\{\omega,m,f\}}} = D_{i_{\{\omega,*,f\}}j_{\{\omega,*,f\}}}$, $D_{i_{\{\omega,f,m\}}j_{\{\omega,f,m\}}} = D_{i_{\{\omega,m,m\}}j_{\{\omega,m,m\}}} = D_{i_{\{\omega,*,m\}}j_{\{\omega,*,m\}}}$, $p_{i_{\{f,f\}}} = p_{i_{\{m,f\}}} = p_{i_{\{*,f\}}}$, and $p_{i_{\{f,m\}}} = p_{i_{\{m,m\}}} = p_{i_{\{*,m\}}}$. Comparing (29) and (32), we see that the difference is due to inclusion of $j_{\{*,m\}}$ terms; since the mitochondrial locus cannot be paternally inherited, there are no similar $c_{\{*,m\}}$ terms in (29). Note that if we let $m_f = m_m$, once again $D_{i_{\{*,f\}}j_{\{*,f\}}} = D_{i_{\{*,m\}}j_{\{*,m\}}}$, and (32) takes the same form as (16).

If instead we consider recombination and random mating followed by migration (census scheme 1), the overall cytonuclear association will be made up of both within-genome and between-genome associations immediately following migration. At equilibrium, the between-genome association in females and males will depend on the female and male migration rates, respectively,

$$\hat{D}_{i_{\{v,f,m\}}c_{\{v,f,f\}}} = \frac{m_f}{2}(\delta D'_{i_{\{v-1,f,m\}}c_{\{v-1,f,f\}}} + \delta D'_{i_{\{v+1,f,m\}}c_{\{v+1,f,f\}}}) + m_f d_{i_{\{f,m\}}} d_{c_{\{f,f\}}} \tag{33}$$

$$\hat{D}_{i_{\{v,m,m\}}c_{\{v,m,f\}}} = \frac{m_m}{2}(\delta D'_{i_{\{v-1,m,m\}}c_{\{v-1,m,f\}}} + \delta D'_{i_{\{v+1,m,m\}}c_{\{v+1,m,f\}}}) + m_m d_{i_{\{m,m\}}} d_{c_{\{m,f\}}}. \tag{34}$$

The within-genome association in females will also depend on the female migration rate only,

$$\hat{D}_{i_{\{v,f,f\}}c_{\{v,f,f\}}} = m_f\Big[\frac{1}{2}(\delta D'_{i_{\{v-1,f,m\}}c_{\{v-1,f,f\}}} + \delta D'_{i_{\{v+1,f,m\}}c_{\{v+1,f,f\}}} + \delta D'_{i_{\{v-1,f,f\}}c_{\{v-1,f,f\}}} + \delta D'_{i_{\{v+1,f,f\}}c_{\{v+1,f,f\}}})\Big] + m_f(d_{i_{\{f,m\}}} + 2d_{i_{\{f,f\}}})\, d_{c_{\{f,f\}}} \tag{35}$$

but the within-genome association in males will depend both on the male and the female migration rates,

$$\begin{aligned}\hat{D}_{i_{\{v,m,f\}}c_{\{v,m,f\}}} &= \frac{m_f}{2}(\delta D'_{i_{\{v-1,f,m\}}c_{\{v-1,f,f\}}} + \delta D'_{i_{\{v+1,f,m\}}c_{\{v+1,f,f\}}} + \delta D'_{i_{\{v-1,f,f\}}c_{\{v-1,f,f\}}} + \delta D'_{i_{\{v+1,f,f\}}c_{\{v+1,f,f\}}}) \\ &\quad + \frac{m_m}{2}(\delta D'_{i_{\{v-1,m,f\}}c_{\{v-1,m,f\}}} + \delta D'_{i_{\{v+1,m,f\}}c_{\{v+1,m,f\}}}) \\ &\quad + m_f(d_{i_{\{f,m\}}} + d_{i_{\{f,f\}}})\, d_{c_{\{f,f\}}} + m_m d_{i_{\{m,f\}}} d_{c_{\{m,f\}}}.\end{aligned} \tag{36}$$

Derivations for Equations 33–36 are given in the APPENDIX. In contrast to this asymmetry for within-genome cytonuclear associations, within-genome associations for unlinked nuclear loci take the same general form in males and females (see APPENDIX).

Comparisons of nuclear-nuclear with cytonuclear disequilibria could therefore allow us to see if there are different migration rates for males and females and estimate those rates. However, differences between nuclear-nuclear and cytonuclear associations may arise due to selection on individual loci or assortative mating as opposed to migration. Consideration of a group of associations should average out selective effects on nuclear loci, but this would not rule out selection directly on the mitochondrial haplotype (*e.g.*, MACRAE and ANDERSON 1988; HUTTER and RAND 1995).

ESTIMATING MIGRATION

One approach to estimating migration rates from disequilibrium measurements is to assume that the population reaches a "quasi-equilibrium" (NAGYLAKI 1976; BARTON and GALE 1993) with the generation of associations due to migration or other forces such as selection balanced by their loss through recombination. After an initial period where large disequilibria break down quickly, linkage disequilibria change very slowly during the period when most allele frequency evolution takes place. In this case, the disequilibria in all of the demes converges to the same value, and the recursion in (14) gives

$$D_{ic} = 2md_i d_c, \tag{37}$$

where, under symmetric migration, $d_{i_{v-1}} = -d_{i_{v+1}} = d_i$ and $d_{c_{v-1}} = -d_{c_{v+1}} = d_c$, and assuming that there are no

differences in allele frequencies for maternally inherited and paternally inherited alleles ($d_{i_f} = d_{i_m} = d_i$). Thus, for a nuclear marker and a cytoplasmic marker, the migration rate could be estimated from the differences in allele frequencies and the disequilibrium (as is true for any two unlinked loci; *c.f.* Kruuk 1997). This corresponds to the results for the center of a continuous population in which migration into the center is approximated by diffusion (σ^2; Barton 1986); then, for two unlinked loci i and j,

$$D_{ij} = 2\sigma^2 \frac{\partial p_i}{\partial x} \frac{\partial p_j}{\partial x}. \quad (38)$$

This also corresponds to an approximation of a stepping-stone model, where migration is between neighbors that are in a similar state ("short-range" migration model). This will differ from the exact equilibrium result, given by (15), when there are differences in disequilibria between the demes ($\delta D_{i_{c_\omega}}$ nonzero).

Applications of exact equilibrium results to estimate migration rates have used the assumption that migrants come from two genetically homogeneous parental species, so that disequilibria in the source demes are zero (Arnold *et al.* 1988; Asmussen *et al.* 1989). Under this "long-range" migration model, the equation corresponding to (37) is then

$$D_{ic} = \frac{2m}{(1 + m)} d_i d_c, \quad (39)$$

for the case of migration followed by meiosis and random mating. Again, the cytonuclear disequilibrium is a function only of the migration rate and of differences in allele frequencies between the parental species.

A simple method for estimating multilocus moments from genotype frequencies assumes that associations within and between genomes can be disentangled by estimating $D_{i_{[i^*]}}$ from the deficit of heterozygotes at individual loci (Barton 2000). Let $D_{0,2} = D_{i_{\{x,y\}} j_{\{x,y\}}}$, where x and y are f or m, stand for the usual measure of pairwise linkage disequilibrium between two genes derived from the same gamete, and let $D_{1,1} = D_{i_{\{x,y\}} i_{\{x,y^*\}}}$, where x and y are f or m and y^* indicates the sex opposite of y, be a measure of the deficit of heterozygotes (a measure of departure from Hardy-Weinberg equilibrium). We define an additive trait, $z = \sum_{i=1}^{n} (X_{i_f} + X_{i_m})$; the variance is then

$$\mathrm{Var}(z) = \sum_i \sum_j (D_{i_f j_f} + D_{i_f j_m} + D_{i_m j_f} + D_{i_m j_m}). \quad (40)$$

If we assume that between-genome associations are the same whether they involve the same or different loci [*i.e.*, $D_{i_{\{x,y\}} i_{\{x,y^*\}}} = D_{i_{\{x,y\}} j_{\{x,y^*\}}} = D_{1,1}$, where again x and y are either f or m, and y^* indicates the sex opposite of y], as is always true for admixture models (Barton 2000), the component of var(z) due to between-genome associations can be separated from that due to associations within genomes (Barton and Gale 1993; Kruuk 1997; Barton 2000), and the above becomes

$$\mathrm{Var}(z) = 2[\textstyle\sum_i p_i(1 - p_i) + n^2 D_{1,1} + n(n-1) D_{0,2}] \quad (41)$$

since $D_{i_i} = p_i(1 - p_i)$ and assuming $p_{i_f} = p_{i_m}$. In this way, genotype and allele frequencies allow us to estimate $D_{0,2}$ and $D_{1,1}$, which can be compared to the cytonuclear association. We give an example below, using simulated data.

This method can be extended to allow estimation of the whole matrix of moments, D_{ij} (where i is inherited from one gamete and j from the other) from a set of 3^n diploid genotype frequencies for all sets of loci, inherited maternally and paternally, assuming that the allele frequencies are known and are the same among male and female gametes (Barton 2000). The method can be straightforwardly extended to cytonuclear moments D_{ij_c}; it gives satisfactory estimates of individual moments, but it is not a good method for estimating the overall genotypic structure given by the full set of moments, since individual genotype frequencies may be negative (see Barton 2000 for discussion).

Estimation of female and male migration rates from cytonuclear and nuclear-nuclear associations: Under census 2, where we allow migration followed by recombination and random mating, cross-genome associations are lost and $D_{1,1} = 0$. Comparison of within-genome with cytonuclear associations will allow an estimation of m_f and m_m. Let $D_{1,c} = (\hat{D}_{i_{\{f,*\}} c_{\{f,*\}}} + \hat{D}_{i_{\{m,*\}} c_{\{m,*\}}})/2$ (so that we are averaging over the sex-of-carrier), averaged over all nuclear loci. From (22) and (23), under the assumption of short-range migration, the $\delta D = 0$, allowing an estimate of the female migration rate as

$$m_f = \frac{D_{1,c}}{(d_{i_{\{f,*\}}} + d_{i_{\{m,*\}}}) d_c} = \frac{D_{1,c}}{2 d_i d_c} \quad (42)$$

and the male migration rate as

$$m_m = \frac{4D_{0,2}}{(d_{i_{\{f,*\}}} + d_{i_{\{m,*\}}})(d_{j_{\{f,*\}}} + d_{j_{\{m,*\}}})} - \frac{D_{1,c}}{(d_{i_{\{f,*\}}} + d_{i_{\{m,*\}}}) d_c}$$
$$= \frac{D_{0,2}}{d_i d_j} - \frac{D_{1,c}}{2 d_i d_c}, \quad (43)$$

where, under the symmetric model of migration, $d_{i_{\{\nu-1,x,y\}}} = -d_{i_{\{\nu+1,x,y\}}} = d_{i_{\{x,y\}}}$, and we assume that $d_{i_{\{x,f\}}} = d_{i_{\{x,m\}}} = d_{i_{\{x,*\}}}$, where x is the sex-of-carrier and is either f or m, and we further allow the allele frequencies to be equal in the two sexes, so that $d_{i_{\{f,*\}}} = d_{i_{\{m,*\}}} = d_i$. We let $D_{0,2}$ be

$$(\hat{D}_{i_{\{f,f\}} j_{\{f,f\}}} + \hat{D}_{i_{\{f,m\}} j_{\{f,m\}}} + \hat{D}_{i_{\{m,f\}} j_{\{m,f\}}} + \hat{D}_{i_{\{m,m\}} j_{\{m,m\}}})/4, \quad (44)$$

averaged across all nuclear loci.

Assuming long-range migration, the disequilibria in the source demes are zero, and the equivalent equations to (42) and (43) are

$$m_{\mathrm{f}} = \frac{D_{1,c}}{(d_{i_{[\mathrm{f},*]}} + d_{i_{[\mathrm{m},*]}})\, d_{\mathrm{c}} - D_{1,c}} = \frac{D_{1,c}}{2d_i d_c - D_{1,c}} \qquad (45)$$

and

$$m_{\mathrm{m}} = \frac{2D_{0,2}}{2d_i d_j - D_{0,2}} - \frac{D_{1,c}}{2d_i d_c - D_{1,c}}. \qquad (46)$$

Under census scheme 1, we census after migration, so that the between-genome associations are not zero. Equations A26, A30, and A34 from the APPENDIX give the relationships between the measurable associations $D_{1,c}$, $D_{1,1}$, and $D_{0,2}$ and the two migration rates under short-range migration for census scheme 1, while Equations A38, A30, and A40 give these associations under long-range migration. When migration rates are equal in males and females, so that $m_{\mathrm{f}} = m_{\mathrm{m}} = m$, $D_{1,1}$ and $D_{0,2}$ under census scheme 1 are equivalent to $D_{[i,j^*]}(\kappa_{[i,j^*]})$ and $D_{[i,j]}(\kappa_{[i,j]})$, given in Equation 3 of BARTON (2000).

Example using simulated data: To illustrate these results, we simulate a population receiving one-way migration from two source demes, deme $\nu - 1$ and deme $\nu + 1$, as in Figure 1. Within each source deme, we have equal allele frequencies at $n = 2$, 3, 4, or 5 unlinked autosomal loci and also at a maternally inherited cytoplasmic locus; allele frequencies were 0 for deme $\nu - 1$ ($p_{i_{\nu-1}} = p_{c_{\nu-1}} = 0$) or 0.8 for deme $\nu + 1$ ($p_{i_{\nu+1}} = p_{c_{\nu+1}} = 0.8$). We assume that male and female migration rates differ, but that male and female allele frequencies are equal (which will be true as long as the ratio of female to male migration rates is equal in each source deme). After migration, individuals were allowed to mate randomly each generation (census scheme 2). After 100 generations had elapsed, allowing deme ν to reach approximate equilibrium, 1000 individuals were sampled from the population. Nuclear-nuclear and cytonuclear associations (D_{ij}, D_{ic}, D_{ii^*}, and D_{ii^*c}) were estimated using an extension of the moment-based method described in BARTON (2000) that allows for cytoplasmic inheritance. Note that, under this census scheme, between-genome associations are zero at equilibrium ($D_{ii^*} = D_{ii^*c} = 0$). Migration rates were estimated from these associations under the assumption of long-range migration (see above) using (45) for female migration rate (m_{f}) and (46) for male migration rate (m_{m}). The results are shown in Table 2 as the mean and standard deviation (SD) for 100 replicate samples for various values of m_{f} and m_{m}.

As one would expect, there is greater variance in the estimates of male migration rate (m_{m}) than in estimates of female migration rate (m_{f}; Table 2 and Figures 2 and 3), since both nuclear-nuclear and cytonuclear associations depend on the female migration rate, but the male migration rate influences only nuclear-nuclear associations. This is similar to the results found by ORIVE and ASMUSSEN (2000) in considering pollen- or seed-mediated gene flow in plants with dicytonuclear inheritance of chloroplast and mitochondrial markers; in such a system, seed migration rates affect all types of associations, but pollen migration rates affect only nuclear and nuclear DNA-chloroplast DNA associations, making pollen migration rates more difficult to estimate. From (43) and (46), we see that m_{m} is estimated as the difference between two estimated disequilibrium values and so has an inherently larger error associated with it.

Increasing the number of autosomal loci decreases the standard deviation for male migration rate (shaded circles, Figure 3); the greatest decrease is seen in going from two unlinked autosomal loci to three. For the top of Figure 3, the increase in SD for female migration rate (solid circles) from four nuclear loci to five nuclear loci is not significant ($F = 1.121$, $P > 0.25$, one-tailed f-test of variances). Note that the change in SD for female migration rate is nearly flat for increasing numbers of unlinked autosomal loci analyzed (Figure 3, solid circles); the cytonuclear associations give the greatest amount of information for this migration estimate, and these are relatively unaffected by adding greater numbers of autosomal markers. These results imply diminishing returns on analysis of large numbers of unlinked autosomal markers for estimating migration rates in this manner. However, a large number of nuclear markers may allow one to detect the possible effects of selection on the nuclear markers or on loci linked to the nuclear markers, increasing our confidence that the disequilibria measured are the results of admixture (migration) and not underlying selection.

DISCUSSION

Recently, KIRKPATRICK *et al.* (2002) presented a general framework for describing associations among multiple genes and their response to selection, nonrandom mating, and admixture. Multilocus associations are defined as moments of the distribution of allelic states. Here, we have used this framework to investigate the effects of migration on association among nuclear and cytoplasmic genes. In a simple model of admixture, where the two sexes are equivalent, the cytonuclear disequilibria (D_{ic} and $D_{\mathrm{ii^*c}}$) behave in the same way as the corresponding second- and third-order associations between unlinked nuclear alleles (D_{ij} and $D_{\mathrm{ii^*j}}$). This is not the case for more complex models involving some type of sexual asymmetry.

As a simple illustration, when migration rates differ between the sexes, the cytonuclear disequilibrium depends only on the female migration rate, whereas the association between two unlinked nuclear loci depends on the mean migration in males and females. This gives a simple way of separately estimating rates of migration in the two sexes. An example is given using simulated data where migration rates differ between the sexes but allele frequencies are equal; for this case, the moments-based method of BARTON (2000) can be used to estimate

TABLE 2

Simulation results from 100 replicate samples of 1000 individuals

	$n = 2$		$n = 3$		$n = 4$		$n = 5$	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
$m_f = 0.1$	0.1012	(0.0318)	0.1026	(0.0271)	0.1006	(0.0216)	0.1033	(0.0229)
$m_m = 0.15$	0.1471	(0.0709)	0.1497	(0.0489)	0.1553	(0.0380)	0.1458	(0.0344)
$D_{ij} = 0.0356$	0.0351	(0.0081)	0.0357	(0.0057)	0.0362	(0.0043)	0.0354	(0.0039)
$D_{ic} = 0.0145$	0.0146	(0.0042)	0.0148	(0.0036)	0.0146	(0.0029)	0.0149	(0.0030)
$D_{ii^*} = 0$	−0.0002	(0.0049)	0.0002	(0.0045)	−0.0003	(0.0036)	0.0000	(0.0037)
$D_{ii^*c} = 0$	0.0000	(0.0029)	0.0002	(0.0021)	−0.0001	(0.0018)	−0.0001	(0.0017)
$m_f = 0.15$	0.1446	(0.0336)	0.1504	(0.0270)	0.1486	(0.0252)	0.1537	(0.0227)
$m_m = 0.1$	0.1034	(0.0680)	0.1024	(0.0555)	0.1040	(0.0475)	0.0960	(0.0405)
$D_{ij} = 0.0356$	0.0351	(0.0081)	0.0358	(0.0060)	0.0358	(0.0051)	0.0355	(0.0043)
$D_{ic} = 0.0209$	0.0201	(0.0041)	0.0208	(0.0033)	0.0206	(0.0031)	0.0213	(0.0027)
$D_{ii^*} = 0$	0.0004	(0.0054)	−0.0003	(0.0049)	−0.0004	(0.0044)	0.0002	(0.0036)
$D_{ii^*c} = 0$	0.0000	(0.0028)	−0.0001	(0.0022)	0.0000	(0.0020)	−0.0003	(0.0017)
$m_f = 0.2$	0.2022	(0.0382)	0.1998	(0.0345)	0.2018	(0.0323)	0.2024	(0.0256)
$m_m = 0.25$	0.2465	(0.0867)	0.2504	(0.0515)	0.2534	(0.0490)	0.2420	(0.0450)
$D_{ij} = 0.0588$	0.0584	(0.0086)	0.0587	(0.0050)	0.0592	(0.0047)	0.0581	(0.0040)
$D_{ic} = 0.0267$	0.0268	(0.0043)	0.0265	(0.0038)	0.0268	(0.0036)	0.0269	(0.0028)
$D_{ii^*} = 0$	−0.0007	(0.0057)	−0.0008	(0.0042)	−0.0006	(0.0039)	0.0002	(0.0038)
$D_{ii^*c} = 0$	0.0000	(0.0029)	−0.0003	(0.0023)	0.0001	(0.0019)	0.0002	(0.0017)
$m_f = 0.25$	0.2478	(0.0430)	0.2512	(0.0340)	0.2483	(0.0362)	0.2462	(0.0328)
$m_m = 0.2$	0.2110	(0.0824)	0.1950	(0.0599)	0.2047	(0.0536)	0.2095	(0.0489)
$D_{ij} = 0.0588$	0.0595	(0.0075)	0.0583	(0.0056)	0.0590	(0.0041)	0.0593	(0.0042)
$D_{ic} = 0.0320$	0.0316	(0.0045)	0.0320	(0.0035)	0.0317	(0.0037)	0.0315	(0.0034)
$D_{ii^*} = 0$	−0.0007	(0.0053)	0.0003	(0.0044)	−0.0001	(0.0045)	−0.0005	(0.0041)
$D_{ii^*c} = 0$	−0.0002	(0.0028)	−0.0001	(0.0022)	−0.0003	(0.0019)	0.0000	(0.0017)
$m_f = 0.05$	0.0527	(0.0266)	0.0521	(0.0282)	0.0513	(0.0220)	0.0512	(0.0216)
$m_m = 0.2$	0.2090	(0.0672)	0.2010	(0.0469)	0.1971	(0.0380)	0.2032	(0.0398)
$D_{ij} = 0.0356$	0.0368	(0.0077)	0.0359	(0.0053)	0.0353	(0.0044)	0.0361	(0.0042)
$D_{ic} = 0.0076$	0.0079	(0.0039)	0.0078	(0.0041)	0.0077	(0.0032)	0.0077	(0.0031)
$D_{ii^*} = 0$	−0.0007	(0.0046)	−0.0005	(0.0042)	0.0000	(0.0038)	0.0000	(0.0039)
$D_{ii^*c} = 0$	0.0001	(0.0026)	0.0000	(0.0021)	0.0001	(0.0018)	−0.0015	(0.0019)
$m_f = 0.2$	0.2033	(0.0305)	0.1993	(0.0282)	0.2018	(0.0259)	0.1994	(0.0260)
$m_m = 0.05$	0.0513	(0.0670)	0.0518	(0.0535)	0.0554	(0.0436)	0.0478	(0.0361)
$D_{ij} = 0.0356$	0.0359	(0.0075)	0.0356	(0.0061)	0.0364	(0.0049)	0.0351	(0.0039)
$D_{ic} = 0.0267$	0.0270	(0.0034)	0.0265	(0.0031)	0.0268	(0.0029)	0.0265	(0.0029)
$D_{ii^*} = 0$	−0.0002	(0.0054)	−0.0001	(0.0044)	−0.0003	(0.0041)	0.0003	(0.0040)
$D_{ii^*c} = 0$	0.0000	(0.0027)	0.0002	(0.0023)	0.0002	(0.0021)	−0.0001	(0.0018)

Individuals migrate from populations with the same allele frequencies at $n = 2$, 3, 4, or 5 unlinked autosomal loci and a maternally inherited cytoplasmic locus; allele frequencies were 0 for deme $v - 1$ ($p_{i_{v-1}} = p_{c_{v-1}} = 0$) or 0.8 for deme $v + 1$ ($p_{i_{v+1}} = p_{c_{v+1}} = 0.8$). After migration, individuals were allowed to mate randomly (census scheme 2). Migration rates were estimated from calculated nuclear-nuclear and cytonuclear associations under the assumption of long-range migration (see text for details) using (45) for female migration rate (m_f) and (46) for male migration rate (m_m). The mean and standard deviation (SD) are given for each case. Note that the slight increases seen in SD of m_f with increasing n are not significant (P value > 0.25 in all cases, one-tailed f-test of variances).

nuclear-nuclear associations. These can then be used to estimate the migration rates under simplifying assumptions about migration (long-range or short-range migration).

This is analogous to the use of the standardized variance in allele frequency (F_{ST}) of cytoplasmic and nuclear genes to estimate gene flow in the two sexes (or in plants, via seed or pollen; Dong and Wagner 1994; Ennos 1994; Latta and Mitton 1997). However, one should bear in mind that F_{ST} is proportional to the effective number of migrants (Nm), whereas linkage disequilibria are proportional to the proportion of migrants (m); since effective population size depends on the mode of transmission, Nm and m will vary between different modes in different ways. Different rates of migration between the sexes could also be inferred from differ-

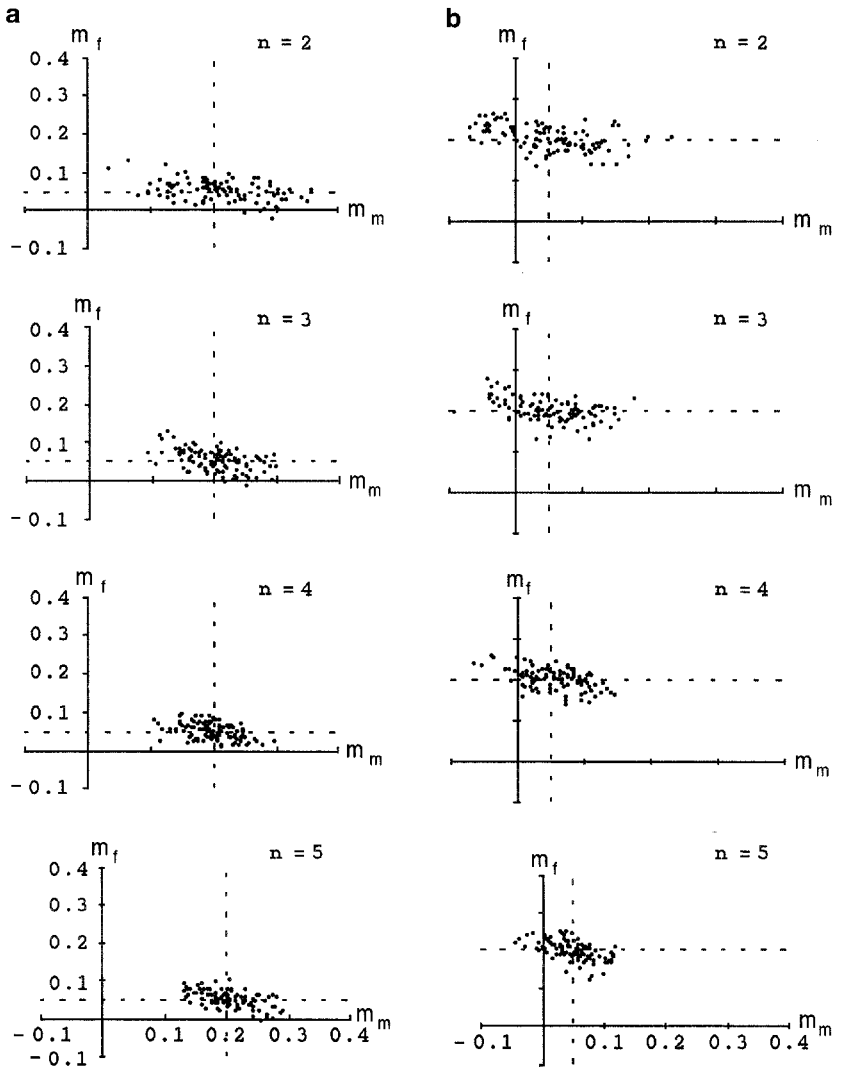

FIGURE 2.—Mean estimates of female (m_f) and male (m_m) migration rates; n gives the number of unlinked autosomal loci. True values are (a) $m_f = 0.05$, $m_m = 0.2$ and (b) $m_f = 0.2$, $m_m = 0.05$ and are indicated by dashed lines in each graph.

ences in cline position and width between nuclear or phenotypic markers and cytoplasmic markers (FERRIS *et al.* 1983; HARRISON 1989; YOUNG 1996; ROHWER *et al.* 2001) or from differences in allele frequency or in linkage disequilibrium measures in males and females after migration (*e.g.*, KRUUK 1997).

The general framework developed here allows us to investigate the conditions under which various measures of nuclear-nuclear and cytonuclear disequilibria would not be equivalent and use this nonequivalence to test for deviations from the simple model of admixture presented above. Such deviations might be due (for example) to nonrandom mating or differences in fitness among first-generation reciprocal hybrids, as well as to differential dispersal. Other kinds of linkage disequilibria may also be used to distinguish among different evolutionary processes. For example, if a hybrid population is maintained by immigration from two distinct sources (perhaps via some network of intermediate demes), then all orders of linkage disequilibria are proportional to the product of the allele frequency differences at the loci involved (BARTON 2000). This gives a testable null hypothesis that may be used to detect other processes such as selection that act on specific sets of loci. It also implies that long-range migration from highly divergent demes will maintain larger higher-order disequilibria than will short-range migration from genetically similar sources. Such genetically differentiated long-range migrants create leptokurtotic distributions resulting in larger higher-order moments. Hence, the relative magnitude of higher-order associations gives an estimate of the degree of divergence of the immigrants. Information may also come from comparison between heterozygote deficit D_{ii*} and within-genome associations D_{ij}; in models of admixture, there is a definite relation between these.

Statistical analysis of multilocus data raises interesting (and difficult) issues. We require methods of testing for differences between different kinds of association and for making composite estimates of parameters such as migration rate on the basis of multilocus genotypes. One attraction of the cytonuclear associations defined

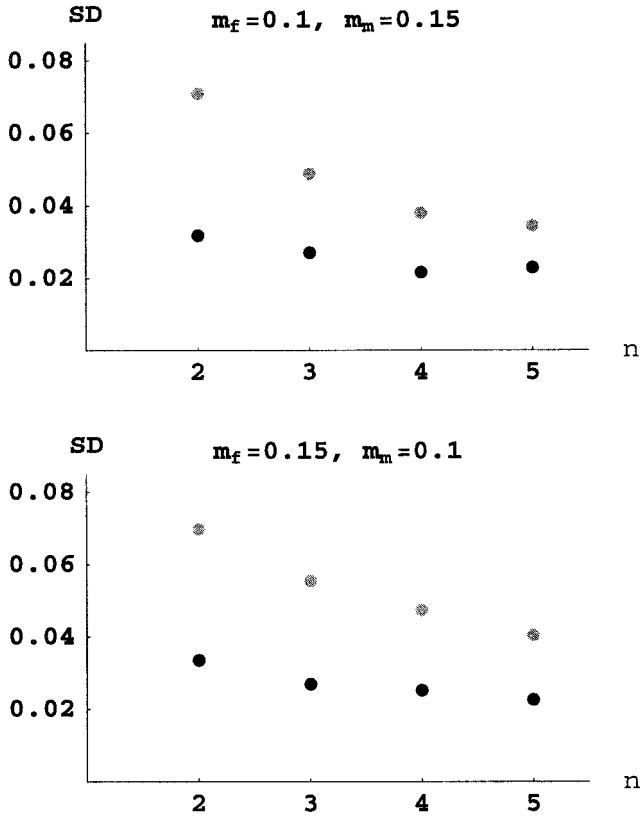

FIGURE 3.—Standard deviations (SD) of estimates of female (m_f) and male (m_m) migration rates as a function of the number of unlinked autosomal loci (n). Solid circles give values for m_f; shaded circles give values for m_m. Each point is calculated from 100 replicate samples of 1000 individuals (Table 2).

by ASMUSSEN *et al.* (1987) is that they can be estimated directly from the data. However, in general one must deal with the problem that the parental origin of alleles in a diploid is usually unknown and that *cis* and *trans* combinations cannot be distinguished. The likelihood-based methods developed in BARTON (2000) may be appropriate to such problems.

Another general difficulty is that, even with the large sample sizes used in these simulation examples (n = 1000 individuals), it is very difficult to accurately estimate disequilibria. The standard deviations for the migration estimates will scale approximately with $1/\sqrt{n}$, where n is the number of individuals sampled. Generally, estimation of disequilibria requires large samples unless allele frequencies are intermediate or the disequilibria to be estimated is maximal for the set of allele frequencies (BROWN 1975). To test for significance of migration rates estimated from multilocus disequilibria, one can either use simulations or find the sampling errors of the estimates of disequilibria used.

The main aims of this article have been to show the relation between different measures of multilocus association and to show how comparison among the various nuclear-nuclear and cytonuclear associations can be used in estimation of important population parameters. Combining nuclear-nuclear and cytonuclear associations should give more information than consideration of each alone. However, we have shown that there may be little gain with a small number of loci, since the majority of the associations are then with the cytoplasmic marker. Sampling of several nuclear genes and complete analysis of multilocus genotype data, including higher-order disequilibria and heterozygote deficit, may be especially useful in teasing apart the effects of selection and/or assortative mating from those of simple admixture.

The authors thank Toby Johnson for his helpful comments on this manuscript. This work was supported by a National Science Foundation NATO postdoctoral fellowship and National Science Foundation grants DEB-9813335 and DEB-0108242 to M.E.O.; N.H.B. gratefully acknowledges the support of the Darwin Trust of Edinburgh and the National Environmental Research Council.

LITERATURE CITED

ARNOLD, J., M. A. ASMUSSEN and J. C. AVISE, 1988 An epistatic mating system model can produce permanent cytonuclear disequilibria in a hybrid zone. Proc. Natl. Acad. Sci. USA **85:** 1893–1896.

ASMUSSEN, M. A., and J. ARNOLD, 1991 The effects of admixture and population subdivision on cytonuclear disequilibria. Theor. Popul. Biol. **39:** 273–300.

ASMUSSEN, M. A., J. ARNOLD and J. C. AVISE, 1987 Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. Genetics **115:** 755–768.

ASMUSSEN, M. A., J. ARNOLD and J. C. AVISE, 1989 The effects of assortative mating and migration on cytonuclear associations in hybrid zones. Genetics **122:** 923–934.

BARTON, N. H., 1983 Multilocus clines. Evolution **37:** 454–471.

BARTON, N. H., 1986 The effects of linkage and density-dependent regulation on gene flow. Heredity **57:** 415–426.

BARTON, N. H., 2000 Estimating multilocus linkage disequilibria. Heredity **84:** 373–389.

BARTON, N. H., and K. S. GALE, 1993 Genetic analysis of hybrid zones, pp. 13–45 in *Hybrid Zones and the Evolutionary Process*, edited by R. G. HARRISON. Oxford University Press, Oxford.

BARTON, N. H., and M. TURELLI, 1991 Natural and sexual selection on many loci. Genetics **127:** 229–255.

BROWN, A. H. D., 1975 Sample sizes required to detect linkage disequilibrium between two or three loci. Theor. Popul. Biol. **8:** 184–201.

CORNUET, J.-M., S. PIRY, G. LUIKART, A. ESTOUP and M. SOLIGNAC, 1999 New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics **153:** 1989–2000.

DAWSON, K. J., and K. BELKHIR, 2001 A Bayesian approach to the identification of panmictic populations and the assignment of individuals. Genet. Res. **78:** 59–77.

DONG, J., and D. B. WAGNER, 1994 Paternally inherited chloroplast polymorphism in Pinus: estimation of diversity and population subdivision, and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. Genetics **136:** 1187–1194.

ENNOS, R. A., 1994 Estimating the relative rates of pollen and seed migration among plant populations. Heredity **72:** 250–259.

FERRIS, S. D., R. D. SAGE, C.-M. HUANG, J. T. NIELSEN, U. RITTE *et al.*, 1983 Flow of mitochondrial DNA across a species boundary. Proc. Natl. Acad. Sci. USA **80:** 2290–2294.

FRANK, S. A., and M. SLATKIN, 1990 The distribution of allelic effects under mutation and selection. Genet. Res. **55:** 111–117.

HARRISON, R. G., 1989 Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. TREE **4:** 6–11.

Hutter, C. M., and D. M. Rand, 1995 Competition between mitochondrial haplotypes in distinct nuclear genetic environments: *Drosophila pseudoobscura vs. D. persimilis*. Genetics **140:** 537–548.
Kirkpatrick, M., T. Johnson and N. Barton, 2002 General models of multilocus evolution. Genetics **161:** 1727–1750.
Kruuk, L., 1997 Barriers to gene flow: a *Bombina* (fire-bellied toad) hybrid zone and multilocus theory. Ph.D. Thesis, University of Edinburgh, Edinburgh, UK.
Latta, R. G., and J. B. Mitton, 1997 A comparison of population differentiation across four classes of gene marker in limber pine (*Pinus flexilis* James). Genetics **146:** 1153–1163.
MacRae, A., and W. W. Anderson, 1988 Evidence for non-neutrality of mitochondrial DNA haplotypes in *Drosophila pseudoobscura*. Genetics **120:** 485–494.
Mallet, J. L. B., N. Barton, G. M. Lamas, J. C. Santisteban, M. M. Muedas *et al.*, 1990 Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. Genetics **124:** 921–936.
Nagylaki, T., 1976 The evolution of one- and two-locus systems. Genetics **83:** 583–600.
Nei, M., and W.-H. Li, 1973 Linkage disequilibrium in subdivided populations. Genetics **75:** 213–219.
Orive, M. E., and M. A. Asmussen, 2000 The effects of pollen and seed migration on nuclear-dicytoplasmic systems. II. A new method for estimating plant gene flow from joint nuclear-cytoplasmic data. Genetics **155:** 833–854.
Price, G. R., 1970 Selection and covariance. Nature **227:** 520–521.
Pritchard, J. K., M. Stephens and P. Donnelly, 2000 Inference of population structure using multilocus genotype data. Genetics **155:** 945–959.
Prout, T., 1973 Appendix to Mitton, J. B., and R. K. Koehn, 1973, Population genetics of marine pelecypods III. Epistasis between functionally related isoenzymes of *Mytilus edulis*. Genetics **73:** 493–496.
Rannala, B., and J. L. Mountain, 1997 Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA **94:** 9197–9221.
Rohwer, S., E. Bermingham and C. Wood, 2001 Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. Evolution **55:** 405–422.
Szymura, J. M., and N. H. Barton, 1986 Genetic analysis of a hybrid zone between fire-bellied toads *Bombina bombina* and *B. variegata*, near Cracow in Southern Poland. Evolution **40:** 1141–1159.
Szymura, J. M., and N. H. Barton, 1991 The genetic structure of the hybrid zone between fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. Evolution **45:** 237–261.
Turelli, M., and N. H. Barton, 1994 Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? Genetics **138:** 913–941.
Weir, B. S., and S. R. Wilson, 1986 Log-linear models for linked loci. Biometrics **42:** 665–670.
Young, N. D., 1996 Concordance and discordance: a tale of two hybrid zones in the Pacific Coast irises (Iridaceae). Am. J. Bot. **83:** 1623–1629.

Communicating editor: J. B. Walsh

APPENDIX: CYTONUCLEAR AND NUCLEAR ASSOCIATIONS UNDER RECOMBINATION FOLLOWED BY MIGRATION (CENSUS 1) WITH UNEQUAL MALE AND FEMALE MIGRATION RATES

Cytonuclear association: After recombination, the between-genome associations in both females and males are zero ($D'_{i_{\{f,m\}}c_{\{f,f\}}} = D'_{i_{\{m,m\}}c_{\{m,f\}}} = 0$), the within-genome associations are $D'_{i_{\{f,f\}}c_{\{f,f\}}} = D'_{i_{\{m,f\}}c_{\{m,f\}}} = (D_{i_{\{f,f\}}c_{\{f,f\}}} + D_{i_{\{f,m\}}c_{\{f,f\}}})/2$, and there are no associations such as $D'_{i_{\{f,m\}}c_{\{f,m\}}}$ and $D'_{i_{\{m,f\}}c_{\{m,m\}}}$.

After migration, the cross-genome cytonuclear association in females is

$$D''_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} = D'_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} + \frac{m_f}{2}(\delta D'_{i_{\{\nu-1,f,m\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,m\}}c_{\{\nu+1,f,f\}}}) + \frac{m_f}{2}(d_{i_{\{\nu-1,f,m\}}}d_{c_{\{\nu-1,f,f\}}} + d_{i_{\{\nu+1,f,m\}}}d_{c_{\{\nu+1,f,m\}}}) - \frac{m_f^2}{4}(d_{i_{\{\nu-1,f,m\}}} + d_{i_{\{\nu+1,f,m\}}})(d_{c_{\{\nu-1,f,f\}}} + d_{c_{\{\nu+1,f,f\}}}), \quad \text{(A1)}$$

where the full context for a gene is given as {deme, sex-of-carrier, sex-of-origin}. Since $D'_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} = 0$, $d_{i_{\{\nu-1,f,m\}}} = -d_{i_{\{\nu+1,f,m\}}} = d_{i_{\{f,m\}}}$, and $d_{c_{\{\nu-1,f,f\}}} = -d_{c_{\{\nu+1,f,f\}}} = d_{c_{\{f,f\}}}$ (under the symmetric model of migration), this becomes

$$D''_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} = \frac{m_f}{2}(\delta D'_{i_{\{\nu-1,f,m\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,m\}}c_{\{\nu+1,f,f\}}}) + m_f d_{i_{\{f,m\}}}d_{c_{\{f,f\}}}. \quad \text{(A2)}$$

At equilibrium, this is

$$\hat{D}_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} = \frac{m_f}{2}(\delta D'_{i_{\{\nu-1,f,m\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,m\}}c_{\{\nu+1,f,f\}}}) + m_f d_{i_{\{f,m\}}}d_{c_{\{f,f\}}}. \quad \text{(A3)}$$

Similarly, the cross-genome association in males will be

$$\hat{D}_{i_{\{\nu,m,m\}}c_{\{\nu,m,f\}}} = \frac{m_m}{2}(\delta D'_{i_{\{\nu-1,m,m\}}c_{\{\nu-1,m,f\}}} + \delta D'_{i_{\{\nu+1,m,m\}}c_{\{\nu+1,m,f\}}}) + m_m d_{i_{\{m,m\}}}d_{c_{\{m,f\}}}. \quad \text{(A4)}$$

The within-genome cytonuclear association in females after migration is

$$D''_{i_{\{\nu,f,f\}}c_{\{\nu,f,f\}}} = D'_{i_{\{\nu,f,f\}}c_{\{\nu,f,f\}}} + \frac{m_f}{2}(\delta D'_{i_{\{\nu-1,f,f\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,f\}}c_{\{\nu+1,f,f\}}}) + \frac{m_f}{2}(d_{i_{\{\nu-1,f,f\}}}d_{c_{\{\nu-1,f,f\}}} + d_{i_{\{\nu+1,f,f\}}}d_{c_{\{\nu+1,f,f\}}}) - \frac{m_f^2}{4}(d_{i_{\{\nu-1,f,f\}}} + d_{i_{\{\nu+1,f,f\}}})(d_{c_{\{\nu-1,f,f\}}} + d_{c_{\{\nu+1,f,f\}}}) = \frac{1}{2}(D_{i_{\{\nu,f,f\}}c_{\{\nu,f,f\}}} + D_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}}) + \frac{m_f}{2}(\delta D'_{i_{\{\nu-1,f,f\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,f\}}c_{\{\nu+1,f,f\}}}) + m_f d_{i_{\{f,f\}}}d_{c_{\{f,f\}}}. \quad \text{(A5)}$$

Note that this depends on the female migration rate only. At equilibrium, this becomes

$$\hat{D}_{i_{\{\nu,f,f\}}c_{\{\nu,f,f\}}} = m_f\left[\frac{1}{2}(\delta D'_{i_{\{\nu-1,f,m\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,m\}}c_{\{\nu+1,f,f\}}}) + \delta D'_{i_{\{\nu-1,f,f\}}c_{\{\nu-1,f,f\}}} + \delta D'_{i_{\{\nu+1,f,f\}}c_{\{\nu+1,f,f\}}}\right] + m_f(2d_{i_{\{f,f\}}} + d_{i_{\{f,m\}}})d_{c_{\{f,f\}}}. \quad \text{(A6)}$$

The within-genome association in males after migration is where the asymmetry inherent to this cytonuclear system comes into play,

$$D''_{i_{\{\nu,m,f\}}c_{\{\nu,m,f\}}} = D'_{i_{\{\nu,m,f\}}c_{\{\nu,m,f\}}} + \frac{m_m}{2}(\delta D'_{i_{\{\nu-1,m,f\}}c_{\{\nu-1,m,f\}}} + \delta D'_{i_{\{\nu+1,m,f\}}c_{\{\nu+1,m,f\}}}) + \frac{m_m}{2}(d_{i_{\{\nu-1,m,f\}}}d_{c_{\{\nu-1,m,f\}}} + d_{i_{\{\nu+1,m,f\}}}d_{c_{\{\nu+1,m,f\}}})$$

$$- \frac{m_m^2}{4}(d_{i\{\nu-1,m,f\}} + d_{i\{\nu+1,m,f\}})(d_{c\{\nu-1,m,f\}} + d_{c\{\nu+1,m,f\}})$$

$$= \frac{1}{2}(D_{i\{\nu,f,f\}c\{\nu,f,f\}} + D_{i\{\nu,f,m\}c\{\nu,f,f\}})$$

$$+ \frac{m_m}{2}(\delta D'_{i\{\nu-1,m,f\}c\{\nu-1,m,f\}} + \delta D'_{i\{\nu+1,m,f\}c\{\nu+1,m,f\}})$$

$$+ m_m d_{i\{m,f\}} d_{c\{m,f\}}. \quad \text{(A7)}$$

Note that associations found in *females* (the first term of the last line above) contribute after migration; thus at equilibrium, the within-genome association in males (unlike that in females) depends on both the female and male migration rates,

$$\hat{D}_{i\{\nu,m,f\}c\{\nu,m,f\}} = \frac{1}{2}(\hat{D}_{i\{\nu,f,f\}c\{\nu,f,f\}} + \hat{D}_{i\{\nu,f,m\}c\{\nu,f,f\}})$$

$$+ \frac{m_m}{2}(\delta D'_{i\{\nu-1,m,f\}c\{\nu-1,m,f\}} + \delta D'_{i\{\nu+1,m,f\}c\{\nu+1,m,f\}})$$

$$+ m_m d_{i\{m,f\}} d_{c\{m,f\}}$$

$$= \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,m\}c\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,m\}c\{\nu+1,f,f\}} + \delta D'_{i\{\nu-1,f,f\}c\{\nu+1,f,f\}}$$

$$+ \delta D'_{i\{\nu+1,f,f\}c\{\nu+1,f,f\}})$$

$$+ \frac{m_m}{2}(\delta D'_{i\{\nu-1,m,f\}c\{\nu-1,m,f\}} + \delta D'_{i\{\nu+1,m,f\}c\{\nu+1,m,f\}})$$

$$+ m_f(d_{i\{f,f\}} + d_{i\{f,m\}}) d_{c\{f,f\}} + m_m d_{i\{m,f\}} d_{c\{m,f\}}. \quad \text{(A8)}$$

Between-genome nuclear associations for unlinked loci: After recombination, all of the between-genome associations are zero, $D'_{i\{f,m\}j\{f,f\}} = D'_{i\{f,f\}j\{f,m\}} = D'_{i\{m,m\}j\{m,f\}} = D'_{i\{m,f\}j\{m,m\}} = 0$. After migration between demes, the between-genome association in females is

$$D''_{i\{\nu,f,m\}j\{\nu,f,f\}} = D'_{i\{\nu,f,m\}j\{\nu,f,f\}} + \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,m\}j\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,m\}j\{\nu+1,f,f\}})$$

$$+ \frac{m_f}{2}(d_{i\{\nu-1,f,m\}} d_{j\{\nu-1,f,f\}} + d_{i\{\nu+1,f,m\}} d_{j\{\nu+1,f,m\}})$$

$$- \frac{m_f^2}{4}(d_{i\{\nu-1,f,m\}} + d_{i\{\nu+1,f,m\}})(d_{j\{\nu-1,f,f\}} + d_{j\{\nu+1,f,f\}}). \quad \text{(A9)}$$

Since $D'_{i\{\nu,f,m\}j\{\nu,f,f\}} = 0$ and $d_{i\{\nu-1,f,m\}} = -d_{i\{\nu+1,f,m\}} = d_{i\{f,m\}}$, etc. (under the symmetric model of migration), this becomes

$$D''_{i\{\nu,f,m\}j\{\nu,f,f\}} = \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,m\}j\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,m\}j\{\nu+1,f,f\}}) + m_f d_{i\{f,m\}} d_{j\{f,f\}}. \quad \text{(A10)}$$

At equilibrium,

$$\hat{D}_{i\{\nu,f,m\}j\{\nu,f,f\}} = \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,m\}j\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,m\}j\{\nu+1,f,f\}}) + m_f d_{i\{f,m\}} d_{j\{f,f\}}. \quad \text{(A11)}$$

Similarly,

$$\hat{D}_{i\{\nu,f,f\}j\{\nu,f,m\}} = \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,f\}j\{\nu-1,f,m\}} + \delta D'_{i\{\nu+1,f,f\}j\{\nu+1,f,m\}}) + m_f d_{i\{f,f\}} d_{j\{f,m\}}. \quad \text{(A12)}$$

The equivalent equilibrium associations in males are

$$\hat{D}_{i\{\nu,m,m\}j\{\nu,m,f\}} = \frac{m_m}{2}(\delta D'_{i\{\nu-1,m,m\}j\{\nu-1,m,f\}} + \delta D'_{i\{\nu+1,m,m\}j\{\nu+1,m,f\}}) + m_m d_{i\{m,m\}} d_{j\{m,f\}} \quad \text{(A13)}$$

and

$$\hat{D}_{i\{\nu,m,f\}j\{\nu,m,m\}} = \frac{m_m}{2}(\delta D'_{i\{\nu-1,m,f\}j\{\nu-1,m,m\}} + \delta D'_{i\{\nu+1,m,f\}j\{\nu+1,m,m\}}) + m_m d_{i\{m,f\}} d_{j\{m,m\}}. \quad \text{(A14)}$$

Within-genome nuclear associations for unlinked loci: After recombination, the within-genome associations for maternally derived loci in females and in males depend on associations found in females in the previous generation,

$$D'_{i\{f,f\}j\{f,f\}} = D'_{i\{m,f\}j\{m,f\}}$$

$$= (D_{i\{f,f\}j\{f,f\}} + D_{i\{f,m\}j\{f,m\}} + D_{i\{f,f\}j\{f,m\}} + D_{i\{f,m\}j\{f,f\}})/4, \quad \text{(A15)}$$

while the within-genome associations for paternally derived loci in females and in males depend on associations found in males in the previous generation,

$$D'_{i\{f,m\}j\{f,m\}} = D'_{i\{m,m\}j\{m,m\}}$$

$$= (D_{i\{m,f\}j\{m,f\}} + D_{i\{m,m\}j\{m,m\}} + D_{i\{m,f\}j\{m,m\}} + D_{i\{m,m\}j\{m,f\}})/4. \quad \text{(A16)}$$

After migration, the within-genome nuclear association in females is

$$D''_{i\{\nu,f,f\}j\{\nu,f,f\}} = D'_{i\{\nu,f,f\}j\{\nu,f,f\}} + \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,f\}j\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,f\}j\{\nu+1,f,f\}})$$

$$+ \frac{m_f}{2}(d_{i\{\nu-1,f,f\}} d_{j\{\nu-1,f,f\}} + d_{i\{\nu+1,f,f\}} d_{j\{\nu+1,f,f\}})$$

$$- \frac{m_f^2}{4}(d_{i\{\nu-1,f,f\}} + d_{i\{\nu+1,f,f\}})(d_{j\{\nu-1,f,f\}} + d_{j\{\nu+1,f,f\}}). \quad \text{(A17)}$$

Since $d_{i\{\nu-1,f,f\}} = -d_{i\{\nu+1,f,f\}} = d_{i\{f,f\}}$, etc. (under the symmetric model of migration), this becomes

$$D''_{i\{\nu,f,f\}j\{\nu,f,f\}} = \frac{1}{4}(D_{i\{\nu,f,f\}j\{\nu,f,f\}} + D_{i\{\nu,f,m\}j\{\nu,f,m\}} + D_{i\{\nu,f,f\}j\{\nu,f,m\}} + D_{i\{\nu,f,m\}j\{\nu,f,f\}})$$

$$+ \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,f\}j\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,f\}j\{\nu+1,f,f\}}) + m_f d_{i\{f,f\}} d_{j\{f,f\}}. \quad \text{(A18)}$$

At equilibrium, the within-genome association for maternally derived alleles in females is

$$\hat{D}_{i\{\nu,f,f\}j\{\nu,f,f\}} = \frac{1}{4}(\hat{D}_{i\{\nu,f,f\}j\{\nu,f,f\}} + \hat{D}_{i\{\nu,f,m\}j\{\nu,f,m\}} + \hat{D}_{i\{\nu,f,f\}j\{\nu,f,m\}} + \hat{D}_{i\{\nu,f,m\}j\{\nu,f,f\}})$$

$$+ \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,f\}j\{\nu-1,f,f\}} + \delta D'_{i\{\nu+1,f,f\}j\{\nu+1,f,f\}}) + m_f d_{i\{f,f\}} d_{j\{f,f\}}. \quad \text{(A19)}$$

Similarly, for paternally derived alleles, the within-genome association in females is

$$\hat{D}_{i\{\nu,f,m\}j\{\nu,f,m\}} = \frac{1}{4}(\hat{D}_{i\{\nu,m,f\}j\{\nu,m,f\}} + \hat{D}_{i\{\nu,m,m\}j\{\nu,m,m\}} + \hat{D}_{i\{\nu,m,f\}j\{\nu,m,m\}} + \hat{D}_{i\{\nu,m,m\}j\{\nu,m,f\}})$$

$$+ \frac{m_f}{2}(\delta D'_{i\{\nu-1,f,m\}j\{\nu-1,f,m\}} + \delta D'_{i\{\nu+1,f,m\}j\{\nu+1,f,m\}}) + m_f d_{i\{f,m\}} d_{j\{f,m\}}. \quad \text{(A20)}$$

The equivalent equilibrium associations in males are

$$\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,f}\}}} = \frac{1}{4}(\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,f}\}}}) + \frac{m_\mathrm{m}}{2}(\delta D'_{i_{\{\nu-1,\mathrm{m,f}\}}j_{\{\nu-1,\mathrm{m,f}\}}} + \delta D'_{i_{\{\nu+1,\mathrm{m,f}\}}j_{\{\nu+1,\mathrm{m,f}\}}}) + m_\mathrm{m} d_{i_{\{\mathrm{m,f}\}}} d_{j_{\{\mathrm{m,f}\}}} \quad \text{(A21)}$$

for maternally derived alleles and

$$\hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,m}\}}} = \frac{1}{4}(\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,f}\}}}) + \frac{m_\mathrm{m}}{2}(\delta D'_{i_{\{\nu-1,\mathrm{m,m}\}}j_{\{\nu-1,\mathrm{m,m}\}}} + \delta D'_{i_{\{\nu+1,\mathrm{m,m}\}}j_{\{\nu+1,\mathrm{m,m}\}}}) + m_\mathrm{m} d_{i_{\{\mathrm{m,m}\}}} d_{j_{\{\mathrm{m,m}\}}} \quad \text{(A22)}$$

for paternally derived alleles.

Short-range migration model: Under the short-range migration model, we assume that the generation of associations due to migration or other forces such as selection is balanced by their loss due to recombination. After an initial period, linkage disequilibria change very slowly compared to allele frequency changes, and the disequilibria in all of the demes converge to the same value. Under this "quasi-equilibrium" model of migration, the $\delta D = 0$, and (A3), (A4), (A6), and (A8) become

$$\begin{aligned}
\hat{D}_{i_{\{\nu,\mathrm{f,m}\}}c_{\{\nu,\mathrm{f,f}\}}} &= m_\mathrm{f} d_{i_{\{\mathrm{f,m}\}}} d_{c_{\{\mathrm{f,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{m,m}\}}c_{\{\nu,\mathrm{m,f}\}}} &= m_\mathrm{m} d_{i_{\{\mathrm{m,m}\}}} d_{c_{\{\mathrm{m,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}c_{\{\nu,\mathrm{f,f}\}}} &= m_\mathrm{f}(2d_{i_{\{\mathrm{f,f}\}}} + d_{i_{\{\mathrm{f,m}\}}}) d_{c_{\{\mathrm{f,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}c_{\{\nu,\mathrm{m,f}\}}} &= m_\mathrm{f}(d_{i_{\{\mathrm{f,f}\}}} + d_{i_{\{\mathrm{f,m}\}}}) d_{c_{\{\mathrm{f,f}\}}} + m_\mathrm{m} d_{i_{\{\mathrm{m,f}\}}} d_{c_{\{\mathrm{m,f}\}}}. \quad \text{(A23)}
\end{aligned}$$

Since, when we measure the cytonuclear association ($D_{1,c}$), we are averaging over both maternally inherited and paternally inherited loci and also over males and females, our overall measure of cytonuclear association gives us

$$\begin{aligned}
D_{1,c} &= (\hat{D}_{i_{\{\nu,\mathrm{f,m}\}}c_{\{\nu,\mathrm{f,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}c_{\{\nu,\mathrm{m,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,f}\}}c_{\{\nu,\mathrm{f,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,f}\}}c_{\{\nu,\mathrm{m,f}\}}})/4 \\
&= \frac{3m_\mathrm{f}}{4}(d_{i_{\{\mathrm{f,f}\}}} + d_{i_{\{\mathrm{f,m}\}}}) d_{c_{\{\mathrm{f,f}\}}} + \frac{m_\mathrm{m}}{4}(d_{i_{\{\mathrm{m,f}\}}} + d_{i_{\{\mathrm{m,m}\}}}) d_{c_{\{\mathrm{m,f}\}}}. \quad \text{(A24)}
\end{aligned}$$

If $d_{i_{\{\mathrm{x,f}\}}} = d_{i_{\{\mathrm{x,m}\}}} = d_{i_{\{\mathrm{x,*}\}}}$, where x is either f or m, this simplifies to

$$D_{1,c} = \frac{3m_\mathrm{f}}{2} d_{i_{\{\mathrm{f,*}\}}} d_{c_{\{\mathrm{f,*}\}}} + \frac{m_\mathrm{m}}{2} d_{i_{\{\mathrm{m,*}\}}} d_{c_{\{\mathrm{m,*}\}}}, \quad \text{(A25)}$$

where the sole subsubscript now indicates sex-of-carrier. And if we further assume that allele frequencies are equal in the two sexes, then $d_{i_{\{\mathrm{f,*}\}}} = d_{i_{\{\mathrm{m,*}\}}} = d_i$, and we get

$$D_{1,c} = \left(\frac{3m_\mathrm{f} + m_\mathrm{m}}{2}\right) d_i d_c. \quad \text{(A26)}$$

For between-genome nuclear associations, Equations A11–A14 become

$$\begin{aligned}
\hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,f}\}}} &= m_\mathrm{f} d_{i_{\{\mathrm{f,m}\}}} d_{j_{\{\mathrm{f,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,m}\}}} &= m_\mathrm{f} d_{i_{\{\mathrm{f,f}\}}} d_{j_{\{\mathrm{f,m}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,f}\}}} &= m_\mathrm{m} d_{i_{\{\mathrm{m,m}\}}} d_{j_{\{\mathrm{m,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,m}\}}} &= m_\mathrm{m} d_{i_{\{\mathrm{m,f}\}}} d_{j_{\{\mathrm{m,m}\}}}. \quad \text{(A27)}
\end{aligned}$$

If we take $D_{1,1}$ as the average of the associations $\hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,f}\}}}$, $\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,m}\}}}$, $\hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,f}\}}}$, and $\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,m}\}}}$ for each locus, we find

$$D_{1,1} = \frac{m_\mathrm{f}}{4}(d_{i_{\{\mathrm{f,m}\}}} d_{j_{\{\mathrm{f,f}\}}} + d_{i_{\{\mathrm{f,f}\}}} d_{j_{\{\mathrm{f,m}\}}}) + \frac{m_\mathrm{m}}{4}(d_{i_{\{\mathrm{m,m}\}}} d_{j_{\{\mathrm{m,f}\}}} + d_{i_{\{\mathrm{m,f}\}}} d_{j_{\{\mathrm{m,m}\}}}). \quad \text{(A28)}$$

If $d_{i_{\{\mathrm{x,f}\}}} = d_{i_{\{\mathrm{x,m}\}}} = d_{i_{\{\mathrm{x,*}\}}}$, this further simplifies to

$$D_{1,1} = \frac{m_\mathrm{f}}{2} d_{i_{\{\mathrm{f,*}\}}} d_{j_{\{\mathrm{f,*}\}}} + \frac{m_\mathrm{m}}{2} d_{i_{\{\mathrm{m,*}\}}} d_{j_{\{\mathrm{m,*}\}}}, \quad \text{(A29)}$$

where, once again, the sole subsubscript now refers to sex-of-carrier. And if we further assume that allele frequencies are equal in the two sexes, we get

$$D_{1,1} = \left(\frac{m_\mathrm{f} + m_\mathrm{m}}{2}\right) d_i d_j. \quad \text{(A30)}$$

Note that, with an appropriate change of notation and letting $m_\mathrm{f} = m_\mathrm{m}$, this reduces to the result given in Equation 3 of Barton (2000) for $D^{\mathrm{short}}_{\{i,j^*\}}(\kappa^{\mathrm{short}}_{\{i,j^*\}})$. Finally, for within-genome nuclear associations, Equations A19–A22 become

$$\begin{aligned}
\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,f}\}}} &= \frac{1}{4}(\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,f}\}}}) + m_\mathrm{f} d_{i_{\{\mathrm{f,f}\}}} d_{j_{\{\mathrm{f,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,m}\}}} &= \frac{1}{4}(\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,f}\}}}) + m_\mathrm{f} d_{i_{\{\mathrm{f,m}\}}} d_{j_{\{\mathrm{f,m}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,f}\}}} &= \frac{1}{4}(\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,f}\}}}) + m_\mathrm{m} d_{i_{\{\mathrm{m,f}\}}} d_{j_{\{\mathrm{m,f}\}}} \\
\hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,m}\}}} &= \frac{1}{4}(\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,f}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,m}\}}} + \hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,f}\}}}) + m_\mathrm{m} d_{i_{\{\mathrm{m,m}\}}} d_{j_{\{\mathrm{m,m}\}}} \quad \text{(A31)}
\end{aligned}$$

If we take $D_{0,2}$ as the average of the associations $\hat{D}_{i_{\{\nu,\mathrm{f,f}\}}j_{\{\nu,\mathrm{f,f}\}}}$, $\hat{D}_{i_{\{\nu,\mathrm{f,m}\}}j_{\{\nu,\mathrm{f,m}\}}}$, $\hat{D}_{i_{\{\nu,\mathrm{m,f}\}}j_{\{\nu,\mathrm{m,f}\}}}$, and $\hat{D}_{i_{\{\nu,\mathrm{m,m}\}}j_{\{\nu,\mathrm{m,m}\}}}$ for each locus, using (A28), we find

$$D_{0,2} = \frac{m_\mathrm{f}}{4}(2d_{i_{\{\mathrm{f,f}\}}} d_{j_{\{\mathrm{f,f}\}}} + d_{i_{\{\mathrm{f,m}\}}} d_{j_{\{\mathrm{f,f}\}}} + d_{i_{\{\mathrm{f,f}\}}} d_{j_{\{\mathrm{f,m}\}}} + 2d_{i_{\{\mathrm{f,m}\}}} d_{j_{\{\mathrm{f,m}\}}}) + \frac{m_\mathrm{m}}{4}(2d_{i_{\{\mathrm{m,f}\}}} d_{j_{\{\mathrm{m,f}\}}} + d_{i_{\{\mathrm{m,m}\}}} d_{j_{\{\mathrm{m,f}\}}} + d_{i_{\{\mathrm{m,f}\}}} d_{j_{\{\mathrm{m,m}\}}} + 2d_{i_{\{\mathrm{m,m}\}}} d_{j_{\{\mathrm{m,m}\}}}). \quad \text{(A32)}$$

If $d_{i_{\{x,f\}}} = d_{i_{\{x,m\}}} = d_{i_{\{x,*\}}}$, where x is f or m, this further simplifies to

$$D_{0,2} = \frac{3m_f}{2} d_{i_{\{f,*\}}} d_{j_{\{f,*\}}} + \frac{3m_m}{2} d_{i_{\{m,*\}}} d_{j_{\{m,*\}}}, \qquad \text{(A33)}$$

where, again, the remaining subsubscript refers to sex-of-carrier. And if we further assume that allele frequencies are equal in the two sexes, we get

$$D_{0,2} = \frac{3(m_f + m_m)}{2} d_i d_j. \qquad \text{(A34)}$$

Again, allowing equal migration rates in males and females, this result reduces to that given in Barton (2000), Equation 3, for $D^{\text{short}}_{\{i,j\}}(\kappa^{\text{short}}_{\{i,j\}})$. If we compare (A26), (A30), and (A34), we see that both the within- and between-genome nuclear associations should be proportional to the quantity $m_f + m_m$, while the overall cytonuclear association is proportional to $3m_f + m_m$.

Long-range migration model: Under the long-range migration model, we assume that migrants come from source populations that are in linkage equilibrium, as would be the case if the sources are fixed for alternative alleles. Under this model of migration, the $\delta D'_{\mathbb{A}_{\nu-1}} = \delta D'_{\mathbb{A}_{\nu+1}} = -D'_{\mathbb{A}_\nu}$, and (A3), (A4), (A6), and (A8) become

$$\begin{aligned}
\hat{D}_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} &= m_f d_{i_{\{f,m\}}} d_{c_{\{f,f\}}} \\
\hat{D}_{i_{\{\nu,m,m\}}c_{\{\nu,m,f\}}} &= m_m d_{i_{\{m,m\}}} d_{c_{\{m,f\}}} \\
\hat{D}_{i_{\{\nu,f,f\}}c_{\{\nu,f,f\}}} &= \frac{m_f}{1 + m_f}[(1 - m_f) d_{i_{\{f,m\}}} + 2d_{i_{\{f,f\}}}] d_{c_{\{f,f\}}} \\
\hat{D}_{i_{\{\nu,m,f\}}c_{\{\nu,m,f\}}} &= \frac{m_f(1 - m_m)}{1 + m_f}(d_{i_{\{f,f\}}} + d_{i_{\{f,m\}}}) d_{c_{\{f,f\}}} \\
&\quad + m_m d_{i_{\{m,f\}}} d_{c_{\{m,f\}}}.
\end{aligned} \qquad \text{(A35)}$$

Since, when we measure the cytonuclear association, $D_{1,c}$, we are averaging over both maternally inherited and paternally inherited loci and also over males and females, our overall measure of cytonuclear association gives us

$$\begin{aligned}
D_{1,c} &= (\hat{D}_{i_{\{\nu,f,m\}}c_{\{\nu,f,f\}}} + \hat{D}_{i_{\{\nu,m,m\}}c_{\{\nu,m,f\}}} + \hat{D}_{i_{\{\nu,f,f\}}c_{\{\nu,f,f\}}} + \hat{D}_{i_{\{\nu,m,f\}}c_{\{\nu,m,f\}}})/4 \\
&= \frac{(3 - m_m) m_f}{4(1 + m_f)}(d_{i_{\{f,f\}}} + d_{i_{\{f,m\}}}) d_{c_{\{f,f\}}} + \frac{m_m}{4}(d_{i_{\{m,f\}}} + d_{i_{\{m,m\}}}) d_{c_{\{m,f\}}}.
\end{aligned} \qquad \text{(A36)}$$

If $d_{i_{\{x,f\}}} = d_{i_{\{x,m\}}} = d_{i_{\{x,*\}}}$, etc., this simplifies to

$$D_{1,c} = \frac{(3 - m_m) m_f}{2(1 + m_f)} d_{i_{\{f,*\}}} d_{c_{\{f,*\}}} + \frac{m_m}{2} d_{i_{\{m,*\}}} d_{c_{\{m,*\}}}, \qquad \text{(A37)}$$

where the remaining subsubscript indicates the sex-of-carrier. And if we further assume that allele frequencies are equal in the two sexes, we get

$$D_{1,c} = \frac{(3 + m_m)}{2(1 + m_f)} d_i d_c. \qquad \text{(A38)}$$

Between-genome nuclear associations are the same as for short-range migration, and the overall association is again given by (A28). For within-genome nuclear associations, Equations A19–A22 become

$$\begin{aligned}
\hat{D}_{i_{\{\nu,f,f\}}j_{\{\nu,f,f\}}} &= \frac{(1 - m_f)}{4}(\hat{D}_{i_{\{\nu,f,f\}}j_{\{\nu,f,f\}}} + \hat{D}_{i_{\{\nu,f,m\}}j_{\{\nu,f,m\}}} + \hat{D}_{i_{\{\nu,f,f\}}j_{\{\nu,f,m\}}} + \hat{D}_{i_{\{\nu,f,m\}}j_{\{\nu,f,f\}}}) \\
&\quad + m_f d_{i_{\{f,f\}}} d_{j_{\{f,f\}}} \\
\hat{D}_{i_{\{\nu,f,m\}}j_{\{\nu,f,m\}}} &= \frac{(1 - m_f)}{4}(\hat{D}_{i_{\{\nu,m,f\}}j_{\{\nu,m,f\}}} + \hat{D}_{i_{\{\nu,m,m\}}j_{\{\nu,m,m\}}} + \hat{D}_{i_{\{\nu,m,f\}}j_{\{\nu,m,m\}}} + \hat{D}_{i_{\{\nu,m,m\}}j_{\{\nu,m,f\}}}) \\
&\quad + m_f d_{i_{\{f,m\}}} d_{j_{\{f,m\}}} \\
\hat{D}_{i_{\{\nu,m,f\}}j_{\{\nu,m,f\}}} &= \frac{(1 - m_m)}{4}(\hat{D}_{i_{\{\nu,f,f\}}j_{\{\nu,f,f\}}} + \hat{D}_{i_{\{\nu,f,m\}}j_{\{\nu,f,m\}}} + \hat{D}_{i_{\{\nu,f,f\}}j_{\{\nu,f,m\}}} + \hat{D}_{i_{\{\nu,f,m\}}j_{\{\nu,f,f\}}}) \\
&\quad + m_m d_{i_{\{m,f\}}} d_{j_{\{m,f\}}} \\
\hat{D}_{i_{\{\nu,m,m\}}j_{\{\nu,m,m\}}} &= \frac{(1 - m_m)}{4}(\hat{D}_{i_{\{\nu,m,f\}}j_{\{\nu,m,f\}}} + \hat{D}_{i_{\{\nu,m,m\}}j_{\{\nu,m,m\}}} + \hat{D}_{i_{\{\nu,m,f\}}j_{\{\nu,m,m\}}} + \hat{D}_{i_{\{\nu,m,m\}}j_{\{\nu,m,f\}}}) \\
&\quad + m_m d_{i_{\{m,m\}}} d_{j_{\{m,m\}}}.
\end{aligned} \qquad \text{(A39)}$$

Once again, we take $D_{0,2}$ as the average of the associations $\hat{D}_{i_{\{\nu,f,f\}}j_{\{\nu,f,f\}}}$, $\hat{D}_{i_{\{\nu,f,m\}}j_{\{\nu,f,m\}}}$, $\hat{D}_{i_{\{\nu,m,f\}}j_{\{\nu,m,f\}}}$, and $\hat{D}_{i_{\{\nu,m,m\}}j_{\{\nu,m,m\}}}$ for each locus. If we make the simplifying assumptions that $d_{i_{\{x,f\}}} = d_{i_{\{x,m\}}} = d_{i_{\{x,*\}}}$ and $d_{i_{\{f,*\}}} = d_{i_{\{m,*\}}} = d_i$, this becomes

$$D_{0,2} = \frac{[6 - (m_f + m_m)](m_f + m_m)}{2(2 + m_f + m_m)} d_i d_j. \qquad \text{(A40)}$$

If migration rates are equal for the two sexes, so that $m_f = m_m = m$, this reduces to the familiar result

$$D_{0,2} = \frac{(3 - m) m}{(1 + m)} d_i d_j, \qquad \text{(A41)}$$

which, after a change in notation, is given as $D^{\text{long}}_{\{i,j\}}(\kappa^{\text{long}}_{\{i,j\}})$ in Equation 3 of Barton (2000).