The Effects of Pollen and Seed Migration on Nuclear-Dicytoplasmic Systems. I. Nonrandom Associations and Equilibrium Structure With Both Maternal and Paternal Cytoplasmic Inheritance

Marjorie A. Asmussen* and Maria E. Orive*,†

**Department of Genetics, University of Georgia, Athens, Georgia 30602-7223 and †Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045-2106*

Manuscript received November 25, 1998
Accepted for publication February 14, 2000

ABSTRACT

We determine the nuclear-dicytoplasmic effects of unidirectional gene flow via pollen and seeds upon a mixed-mating plant population, focusing on nuclear-mitochondrial-chloroplast systems where mitochondria are inherited maternally and chloroplasts paternally, as in many conifers. After first delineating the general effects of admixture (via seeds or individuals) on the nonrandom associations in such systems, we derive the full dicytonuclear equilibrium structure, including when disequilibria may be indicators of gene flow. Substantial levels of permanent two- and three-locus disequilibria can be generated in adults by (i) nonzero disequilibria in the migrant pools or (ii) intermigrant admixture effects via different chloroplast frequencies in migrant pollen and seeds. Additionally, three-locus disequilibria can be generated by higher-order intermigrant effects such as different chloroplast frequencies in migrant pollen and seeds coupled with nuclear-mitochondrial disequilibria in migrant seeds, or different nuclear frequencies in migrant pollen and seeds coupled with mitochondrial-chloroplast disequilibria in migrant seeds. Further insight is provided by considering special cases with seed or pollen migration alone, complete random mating or selfing, or migrant pollen and seeds lacking disequilibria or intermigrant admixture effects. The results complete the theoretical foundation for a new method for estimating pollen and seed migration using joint cytonuclear or dicytonuclear data.

THERE are two different avenues for gene flow in plant populations: pollen migration and seed migration. Joint nuclear-cytoplasmic data can be especially useful for decomposing and estimating these two types of gene flow due to the asymmetrical inheritance of cytonuclear systems, where nuclear alleles are inherited through both parents while cytoplasmic alleles are usually inherited uniparentally (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992). This asymmetry and the differing effects of pollen and seed migration will be reflected in the cytonuclear structure of the plant population, through both the joint genotype frequencies and the pattern of statistical associations (disequilibria) generated between the cytoplasmic marker and the nuclear alleles and genotypes.

The existence of two uniparentally inherited organelles in plants greatly enhances the utility of cytonuclear data. Although cases of biparental inheritance of organelle genomes exist (Medgyesy *et al.* 1986; Wagner *et al.* 1991), the majority of plant species exhibit strict maternal inheritance of both mitochondria and chloroplasts (Kirk and Tilney-Bassett 1978; Birky 1988; Harrison and Doyle 1990). However, in two families of coniferous gymnosperms (including the largest family, the Pinaceae), chloroplast DNA (cpDNA) appears to be paternally inherited, while mitochondrial DNA (mtDNA) is maternally inherited (Neale *et al.* 1986; Wagner *et al.* 1987; Neale and Sederoff 1989; Mogensen 1996). Joint nuclear-mtDNA-cpDNA data should be uniquely valuable in estimating plant gene flow in such systems, since pollen will carry a haploid nuclear component and the chloroplast genome, while seeds will carry complete complements of all three genomes. Nuclear and chloroplast alleles will therefore have two opportunities for movement, while mitochondrial alleles can move only through seeds. Furthermore, although pollen migration can only carry nonrandom associations between nuclear and chloroplast alleles, seed migration can carry all possible pairwise associations between alleles and genotypes in the three genomes, as well as higher-order three-locus nuclear-mitochondrial-chloroplast associations.

Previous work (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992) has determined the effects of unidirectional pollen and seed migration on the standard, two-locus cytonuclear system with a nuclear locus and a single, uniparentally inherited cytoplasmic marker. With maternal cytoplasmic inheritance, Asmussen and Schnabel (1991) found that nonzero cytonuclear disequilibria are maintained only if migrant seeds carry nonrandom cytonuclear associations. These migrant associations would be found when selection or

Corresponding author: Maria E. Orive, Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045-2106. E-mail: orive@ukans.edu

other nonrandomizing factors such as assortative mating act on the source population (Arnold *et al.* 1988; Babcock and Asmussen 1996, 1998) as well as when seeds are contributed by multiple, genetically distinct sources, as might be expected in hybrid zones and other areas of admixture. Significant and long-lasting transient disequilibria can be generated, however, by migrant seeds lacking such associations. Pollen dispersal has only a small effect on the disequilibria caused by seed migration and cannot generate even transient disequilibria *de novo*.

In contrast, with paternal cytoplasmic inheritance (Schnabel and Asmussen 1992), pollen migration can significantly affect the cytonuclear structure of the recipient population through nonrandom cytonuclear associations in the migrant pollen, although pollen migration alone generally produces lower levels of disequilibria than does the equivalent amount of seed migration alone. Additionally, with paternal cytoplasmic inheritance, the presence of both types of gene flow can generate nonrandom cytonuclear associations via intermigrant admixture effects, such as differences in nuclear and chloroplast allele frequencies in migrant pollen and seeds. Such allele frequency differences in the two migrant pools could be caused by having separate sources for migrant pollen and seeds, common sources whose relative contributions depend on the type of gene flow, or selection or other evolutionary forces acting during the life cycle of the source population(s). Together, these results for cytonuclear disequilibria suggest that three-locus, nuclear-mtDNA-cpDNA data juxtaposing both forms of uniparental inheritance should be uniquely powerful for estimating pollen and seed migration rates, since they provide the greatest number of avenues for the accumulation of nonrandom associations.

As a first step toward formally testing this idea, we extend these previous pollen and seed dispersal models to the three-locus nuclear-dicytoplasmic (dicytonuclear) system parameterized by Schnabel and Asmussen (1989). We first review the dicytonuclear frequencies and associations for this system and then delimit the effects of population admixture on the three-locus disequilibria. We next determine the dynamics and full equilibrium structure when the two cytoplasmic markers are inherited through opposite parents, including when each type of two- and three-locus association is generated. Further insight into the distinctive dicytonuclear effects of pollen and seed flow is provided by numerical examples showing the magnitude and duration of the disequilibria they generate and by the equilibrium structure for a series of special cases subsumed within our general framework. Together, these new three-locus results complete the theoretical foundation for a new method for estimating pollen and seed flow in plant populations using joint cytonuclear or dicytonuclear data. This final critical step is presented in a companion article (Orive and Asmussen 2000).

TABLE 1

Joint nuclear-mitochondrial-chloroplast genotype frequencies

	Nuclear genotype			
Joint cytotype	*AA*	*Aa*	*aa*	Total
MC	U_{11}	V_{11}	W_{11}	X_{11}
Mc	U_{12}	V_{12}	W_{12}	X_{12}
mC	U_{21}	V_{21}	W_{21}	X_{21}
mc	U_{22}	V_{22}	W_{22}	X_{22}
Total	U	V	W	1

THE NUCLEAR DICYTOPLASMIC SYSTEM

We consider the nuclear-dicytoplasmic structure of a diploid plant population. The dicytonuclear system consists of an autosomal nuclear locus with two alleles, *A* and *a*, and two haploid cytoplasmic loci: a mitochondrial marker with two alleles (cytotypes) *M* and *m* and a chloroplast marker with two alleles (cytotypes) *C* and *c*. As is true in the Pinaceae (Mogensen 1996), we assume that the mitochondria are inherited maternally and chloroplasts paternally, with no cross-parental leakage of either organelle. The definitions of frequency and nonrandom association variables follow those of Schnabel and Asmussen (1989), and we use the notation of that article throughout, except where noted below.

Frequency variables: The frequencies of the 12 joint, three-locus genotypes are given in Table 1 as U_{ij}, V_{ij}, and W_{ij}, where the first index ($i = 1, 2$) indicates *M* or *m* alleles and the second index ($j = 1, 2$) indicates *C* or *c* alleles. The joint cytotype frequencies (X_{ij}) are obtained by summing across each row and the nuclear genotype frequencies (U, V, and W) by summing down each column. The nuclear allele frequencies are

$$P = \text{freq}(A) = U + \tfrac{1}{2}V \tag{1a}$$

and

$$Q = \text{freq}(a) = W + \tfrac{1}{2}V = 1 - P, \tag{1b}$$

where freq denotes "frequency of." The cytotype frequencies for the two cytoplasmic markers are

$$\begin{aligned} X_M &= \text{freq}(M) = X_{11} + X_{12} \\ Y_M &= \text{freq}(m) = X_{21} + X_{22} \\ X_C &= \text{freq}(C) = X_{11} + X_{21} \\ Y_C &= \text{freq}(c) = X_{12} + X_{22}, \end{aligned} \tag{2}$$

where $Y_M = 1 - X_M$ and $Y_C = 1 - X_C$. It is also useful to specify the frequencies of the three-locus triallelic combinations where, for each joint cytotype combination $i, j = 1, 2$,

TABLE 2

Nuclear-mitochondrial and nuclear-chloroplast genotype frequencies

	Nuclear genotype			
Cytotype	*AA*	*Aa*	*aa*	Total
M	$U_{1M} = U_{11} + U_{12}$	$V_{1M} = V_{11} + V_{12}$	$W_{1M} = W_{11} + W_{12}$	X_M
m	$U_{2M} = U_{21} + U_{22}$	$V_{2M} = V_{21} + V_{22}$	$W_{2M} = W_{21} + W_{22}$	Y_M
Total	U	V	W	1
C	$U_{1C} = U_{11} + U_{21}$	$V_{1C} = V_{11} + V_{21}$	$W_{1C} = W_{11} + W_{21}$	X_C
c	$U_{2C} = U_{12} + U_{22}$	$V_{2C} = V_{12} + V_{22}$	$W_{2C} = W_{12} + W_{22}$	Y_C
Total	U	V	W	1

$$P_{ij} = U_{ij} + \tfrac{1}{2}V_{ij} \tag{3a}$$

and

$$Q_{ij} = W_{ij} + \tfrac{1}{2}V_{ij}. \tag{3b}$$

For instance, $P_{11} = \text{freq}(A/M/C)$ is the probability that a randomly sampled individual has the joint cytotype *MC* and that a randomly selected nuclear allele from that individual is *A*. Note that although we normally use slashes (/) to separate different markers, we omit the slash between the two organellar loci when viewed as a joint cytotype.

The final frequency variables are those for the two-locus nuclear-cytoplasmic combinations within each of the two cytonuclear subsystems. The two sets of joint, two-locus genotypic frequencies are given in Table 2. The joint diallelic frequencies for the nuclear-mitochondrial subsystem represent the ovule frequencies, while those for the nuclear-chloroplast subsystem represent the pollen frequencies in the population (Table 3).

TABLE 3

Joint diallelic frequencies for the nuclear-mitochondrial and nuclear-chloroplast subsystems

	Nuclear allele		
Cytotype	*A*	*a*	Total
M	$P_{1M} = U_{1M} + \tfrac{1}{2}V_{1M}$	$Q_{1M} = W_{1M} + \tfrac{1}{2}V_{1M}$	X_M
m	$P_{2M} = U_{2M} + \tfrac{1}{2}V_{2M}$	$Q_{2M} = W_{2M} + \tfrac{1}{2}V_{2M}$	Y_M
Total	P	Q	1
C	$P_{1C} = U_{1C} + \tfrac{1}{2}V_{1C}$	$Q_{1C} = W_{1C} + \tfrac{1}{2}V_{1C}$	X_C
c	$P_{2C} = U_{2C} + \tfrac{1}{2}V_{2C}$	$Q_{2C} = W_{2C} + \tfrac{1}{2}V_{2C}$	Y_C
Total	P	Q	1

Two-locus disequilibria: We consider three sets of pairwise disequilibria among the three loci. These disequilibria are calculated in the same general way as are linkage disequilibria among nuclear loci and equal the difference between the frequency of each joint combination and that expected under random association between the two components involved. The first two involve the two-locus associations in each cytonuclear subsystem as defined in Asmussen *et al.* (1987). To facilitate the development of concise general formulas in the three-locus system, they are denoted here by a more informative notation that explicitly includes the nuclear and cytoplasmic components involved. The two *allelic disequilibria*

$$D_{A/M} = \text{freq}(A/M) - \text{freq}(A)\text{freq}(M) = P_{1M} - PX_M$$
$$D_{A/C} = \text{freq}(A/C) - \text{freq}(A)\text{freq}(C) = P_{1C} - PX_C \tag{4}$$

(D_M and D_C in Schnabel and Asmussen 1989) measure the nonrandom association between the nuclear alleles and each cytoplasmic marker. The six *genotypic disequilibria*

$$\begin{aligned} D_{AA/M} &= U_{1M} - UX_M, & D_{AA/C} &= U_{1C} - UX_C \\ D_{Aa/M} &= V_{1M} - VX_M, & D_{Aa/C} &= V_{1C} - VX_C \\ D_{aa/M} &= W_{1M} - WX_M, & D_{aa/C} &= W_{1C} - WX_C \end{aligned} \tag{5}$$

similarly measure the nonrandom associations between each nuclear genotype and each cytoplasmic marker, where here we have substituted *AA*, *Aa*, *aa* for the usual genotype placeholders 1, 2, and 3. For example, $D_{AA/M} = \text{freq}(AA/M) - \text{freq}(AA)\text{freq}(M)$ measures the nonrandom association between *AA* nuclear homozygotes and the mitochondrial marker *M* (D_{1M} in Schnabel and Asmussen 1989). Within each of the two cytonuclear subsystems, each disequilibrium can be written as a linear combination of any two of the others via the relations

$$D_{AA/*} + D_{Aa/*} + D_{aa/*} = 0 \tag{6a}$$

and

$$D_{A/*} = D_{AA/*} + \tfrac{1}{2}D_{Aa/*}, \tag{6b}$$

where each * stands for *M* for the nuclear-mitochondrial associations and each * stands for *C* for the nuclear-chloroplast associations.

The third and final type of two-locus disequilibria is the *cytoplasmic disequilibrium*,

$$D_{M/C} = \text{freq}(MC) - \text{freq}(M)\text{freq}(C) = X_{11} - X_M X_C \tag{7}$$

(D_{MC} in Schnabel and Asmussen 1989), which is new to the three-locus system and measures the nonrandom association between alleles at the two cytoplasmic loci.

TABLE 4

Decomposition of two-locus cytonuclear genotype frequencies

	Nuclear genotype			
Cytotype	AA	Aa	aa	Total
M	$U_{1M} = UX_M + D_{AA/M}$	$V_{1M} = VX_M + D_{Aa/M}$	$W_{1M} = WX_M + D_{aa/M}$	X_M
m	$U_{2M} = UY_M - D_{AA/M}$	$V_{2M} = VY_M - D_{Aa/M}$	$W_{2M} = WY_M - D_{aa/M}$	Y_M
Total	U	V	W	1
C	$U_{1C} = UX_C + D_{AA/C}$	$V_{1C} = VX_C + D_{Aa/C}$	$W_{1C} = WX_C + D_{aa/C}$	X_C
c	$U_{2C} = UY_C - D_{AA/C}$	$V_{2C} = VY_C - D_{Aa/C}$	$W_{2C} = WY_C - D_{aa/C}$	Y_C
Total	U	V	W	1

The two-locus genotypic frequencies (U_{iM}, U_{iC}, etc.), joint diallelic frequencies (P_{iM}, P_{iC}, etc.), and joint cytotype frequencies (X_{ij}) can be written in terms of the marginal frequencies and the corresponding cytonuclear and cytoplasmic disequilibria, as shown in Tables 4–6.

Three-locus disequilibria: The three-locus system can also result in higher-order associations that involve all three loci, both pairwise three-locus associations and full three-way associations. These are specified fully in Schnabel and Asmussen (1989) and for convenience are briefly summarized here. Throughout, we focus on the associations involving the M and C cytotypes; the interrelations in the three-locus system allow us to derive all of the other associations from these (Schnabel and Asmussen 1989). The basic three-locus disequilibria are the three *joint genotypic disequilibria*

$$\begin{aligned} D_{AA/MC} &= \text{freq}(AA/MC) - \text{freq}(AA)\text{freq}(MC) = U_{11} - UX_{11} \\ D_{Aa/MC} &= \text{freq}(Aa/MC) - \text{freq}(Aa)\text{freq}(MC) = V_{11} - VX_{11} \\ D_{aa/MC} &= \text{freq}(aa/MC) - \text{freq}(aa)\text{freq}(MC) = W_{11} - WX_{11} \end{aligned} \tag{8}$$

and the *joint allelic disequilibrium*

$$\begin{aligned} D_{A/MC} &= \text{freq}(A/MC) - \text{freq}(A)\text{freq}(MC) \\ &= P_{11} - PX_{11}, \end{aligned} \tag{9}$$

which, respectively, measure the nonrandom associations between the MC joint cytotype and the three genotypes and two alleles at the nuclear locus.

TABLE 5

Decomposition of joint cytonuclear diallelic frequencies

	Nuclear allele		
Cytotype	A	a	Total
M	$P_{1M} = PX_M + D_{A/M}$	$Q_{1M} = QX_M - D_{A/M}$	X_M
m	$P_{2M} = PY_M - D_{A/M}$	$Q_{2M} = QY_M + D_{A/M}$	Y_M
Total	P	Q	1
C	$P_{1C} = PX_C + D_{A/C}$	$Q_{1C} = QX_C - D_{A/C}$	X_C
c	$P_{2C} = PY_C - D_{A/C}$	$Q_{2C} = QY_C + D_{A/C}$	Y_C
Total	P	Q	1

Several true three-way measures of nonrandom associations can also be defined that measure associations among the three markers (nuclear, mitochondrial, and chloroplast) after taking into account all of the possible two-way associations (nuclear-mitochondrial, nuclear-chloroplast, and mitochondrial-chloroplast). For the M/C cytotype, we have three *three-way genotypic disequilibria*

$$\begin{aligned} D_{AA/M/C} &= U_{11} - UX_MX_C - UD_{M/C} - X_MD_{AA/C} - X_CD_{AA/M} \\ D_{Aa/M/C} &= V_{11} - VX_MX_C - VD_{M/C} - X_MD_{Aa/C} - X_CD_{Aa/M} \\ D_{aa/M/C} &= W_{11} - WX_MX_C - WD_{M/C} - X_MD_{aa/C} - X_CD_{aa/M} \end{aligned} \tag{10}$$

and the *three-way allelic disequilibrium*

$$D_{A/M/C} = P_{11} - PX_MX_C - PD_{M/C} - X_MD_{A/C} - X_CD_{A/M}, \tag{11}$$

which is analogous to the three-way gametic disequilibrium for three nuclear loci (Bennett 1954). The sets of four joint and four three-way associations each satisfy the relations given in (6a) and (6b), with $* = MC$ for the joint disequilibria and $* = M/C$ for the three-way disequilibria. We also note that each three-way disequilibrium can be written in terms of the corresponding joint dicytonuclear and two-locus disequilibria and the marginal cytotype frequencies, with

$$D_{N/M/C} = D_{N/MC} - X_MD_{N/C} - X_CD_{N/M} \tag{12}$$

for each nuclear component, $N = A$, AA, Aa, or aa.

Finally, only 11 independent variables are necessary to completely describe the full 12 joint genotype system. One such parameterization (Table 7) includes five two-locus disequilibria ($D_{A/M}$, $D_{A/C}$, $D_{Aa/M}$, $D_{Aa/C}$, and $D_{M/C}$), two three-locus disequilibria ($D_{A/MC}$ and $D_{Aa/MC}$), the nuclear, mitochondrial, and chloroplast allele frequencies (P, X_M, and X_C), and the nuclear heterozygote frequency (V). However, for completeness, our analysis below provides the results for an allelic, homozygote, and heterozygote association for each disequilibrium category: $D_{A/MC}$, $D_{AA/MC}$, and $D_{Aa/MC}$ for the three-locus joint

TABLE 6

Decomposition of joint cytotype frequencies

Chloroplast locus	Mitochondrial locus		Total
	M	m	
C	$X_{11} = X_M X_C + D_{M/C}$	$X_{21} = Y_M X_C - D_{M/C}$	X_C
c	$X_{12} = X_M Y_C - D_{M/C}$	$X_{22} = Y_M Y_C + D_{M/C}$	Y_C
Total	X_M	Y_M	1

disequilibria and their counterparts for the three-way disequilibria.

ADMIXTURE FORMULAS

After migration by seeds (or adults), the dicytonuclear makeup of the new population will be a mixture of that of residents and migrants. We consider here the admixture effects for the general case where n genetically distinct sources contribute to a single population. Define the expected value of a variable Z across all n sources as

$$E(Z) = \sum_{i=1}^{n} m_i Z^{(i)}, \tag{13}$$

where m_i is the fraction contributed by source i and $Z^{(i)}$ is the value of the variable Z in source i. Frequencies after admixture are simply their weighted average (expectation) across all sources,

$$Z^T = E(Z), \tag{14}$$

where ^{T} indicates values in the total population after admixture.

Admixture has a more complicated effect on disequilibria since these can be generated in the total population by a two-locus Wahlund effect if the sources are genetically distinct (Nei and Li 1973; Asmussen and Arnold 1991; Goodisman and Asmussen 1997). Following admixture, the two-way disequilibria will be the weighted average of the disequilibria across the n sources plus the covariance across the sources between the frequencies of the two genetic components being considered (Asmussen and Arnold 1991). For example, the allelic and genotypic cytonuclear disequilibria after admixture are

$$D^T_{N/M} = E(D_{N/M}) + \mathrm{Cov}(F_N, X_M) \tag{15a}$$

$$D^T_{N/C} = E(D_{N/C}) + \mathrm{Cov}(F_N, X_C), \tag{15b}$$

where $N = A$, AA, Aa, or aa, and F_N is the frequency of N. The cytoplasmic disequilibrium after admixture has the analogous form

$$D^T_{M/C} = E(D_{M/C}) + \mathrm{Cov}(X_M, X_C) \tag{16}$$

while the joint allelic and joint genotypic disequilibria become

$$D^T_{N/MC} = E(D_{N/MC}) + \mathrm{Cov}(F_N, X_{11}), \tag{17}$$

where, again, $N = A$, AA, Aa, or aa.

The derivation of these general formulas depends on the fact that $\mathrm{Cov}(X,Y) = E(XY) - E(X)\,E(Y)$ for any two random variables X and Y. As an example, we derive the joint allelic disequilibrium after admixture, using (9) and (14), as

$$\begin{aligned} D^T_{A/MC} &= P^T_{11} - P^T X^T_{11} \\ &= E(P_{11}) - E(P)E(X_{11}) \\ &= E(P_{11}) - E(PX_{11}) + \mathrm{Cov}(P, X_{11}) \\ &= E(P_{11} - PX_{11}) + \mathrm{Cov}(P, X_{11}) \\ &= E(D_{A/MC}) + \mathrm{Cov}(P, X_{11}), \end{aligned}$$

which is (17), where $N = A$ and $F_N = P$.

For all of the two-way disequilibria, the covariance terms have a simple interpretation for the special case where there are only two source populations contributing to the total population. In this case the admixture formulas (15–17) become

$$\begin{aligned} D^T_{N/M} &= E(D_{N/M}) + m_1(1 - m_1)(F_N^{(1)} - F_N^{(2)})(X_M^{(1)} - X_M^{(2)}) \\ D^T_{N/C} &= E(D_{N/C}) + m_1(1 - m_1)(F_N^{(1)} - F_N^{(2)})(X_C^{(1)} - X_C^{(2)}) \\ D^T_{M/C} &= E(D_{M/C}) + m_1(1 - m_1)(X_M^{(1)} - X_M^{(2)})(X_C^{(1)} - X_C^{(2)}) \\ D^T_{N/MC} &= E(D_{N/MC}) + m_1(1 - m_1)(F_N^{(1)} - F_N^{(2)})(X_{11}^{(1)} - X_{11}^{(2)}), \end{aligned}$$

where $^{(i)}$ indicates source i and $N = A$, AA, Aa, or aa. One can see that the covariance term will be nonzero,

TABLE 7

Decomposition of three-locus genotype frequencies

Joint cytotype	Nuclear genotype		
	AA	Aa	aa
MC	$U_{11} = UX_{11} + D_{AA/MC}$	$V_{11} = VX_{11} + D_{Aa/MC}$	$W_{11} = WX_{11} + D_{aa/MC}$
Mc	$U_{12} = UX_{12} - D_{AA/MC} + D_{AA/M}$	$V_{12} = VX_{12} - D_{Aa/MC} + D_{Aa/M}$	$W_{12} = WX_{12} - D_{aa/MC} + D_{aa/M}$
mC	$U_{21} = UX_{21} - D_{AA/MC} + D_{AA/C}$	$V_{21} = VX_{21} - D_{Aa/MC} + D_{Aa/C}$	$W_{21} = WX_{21} - D_{aa/MC} + D_{aa/C}$
mc	$U_{22} = UX_{22} + D_{AA/MC} - D_{AA/M} - D_{AA/C}$	$V_{22} = VX_{22} + D_{Aa/MC} - D_{Aa/M} - D_{Aa/C}$	$W_{22} = WX_{22} + D_{aa/MC} - D_{aa/M} - D_{aa/C}$

To complete the decomposition, $U = P - \tfrac{1}{2}V$, $W = 1 - P - \tfrac{1}{2}V$, $D_{AA/*} = D_{A/*} - \tfrac{1}{2}D_{Aa/*}$, and $D_{aa/*} = -D_{A/*} - \tfrac{1}{2}D_{Aa/*}$, where * indicates M, C, or MC. The decomposition for the joint cytotype frequencies (X_{ij}) is given in Table 6.

causing an admixture effect, if and only if the two components both differ across the two source populations. In general, to have an admixture effect with more than two source populations, it is necessary but not sufficient that the frequencies of both components vary across the sources.

Three-way disequilibria after admixture are more complex than are two-way disequilibria, involving the weighted averages of the disequilibria and cytoplasmic frequencies, covariances between nuclear and cytoplasmic frequencies, and covariances between cytoplasmic frequencies and two-locus cytonuclear disequilibria. The general formula is

$$D^T_{N/M/C} = E(D_{N/M/C}) - E(X_M)\mathrm{Cov}(F_N,X_C) - E(X_C)\mathrm{Cov}(F_N,X_M)$$
$$+ \mathrm{Cov}(F_N,X_{11}) + \mathrm{Cov}(X_M,D_{N/C}) + \mathrm{Cov}(X_C,D_{N/M}) \quad (18)$$

for $N = A$, AA, Aa, or aa, where, once again, F_N is the frequency of N. An admixture effect will be found for three-way associations only if at least one of the contributing covariances is nonzero. As an example of the derivation, we consider the three-way allelic disequilibrium, $D_{A/M/C}$. From (11) and (14),

$$D^T_{A/M/C} = P^T_{11} - P^T X^T_M X^T_C - P^T D^T_{M/C} - X^T_M D^T_{A/C} - X^T_C D^T_{A/M},$$

which, using (7) and (15), can be rewritten as

$$D^T_{A/M/C} = P^T_{11} - P^T X^T_{11} - X^T_M D^T_{A/C} - X^T_C D^T_{A/M}$$
$$= E(P_{11}) - E(P)E(X_{11}) - E(X_M)[E(D_{A/C}) + \mathrm{Cov}(P,X_C)]$$
$$- E(X_C)[E(D_{A/M}) + \mathrm{Cov}(P,X_M)].$$

Once again using the definition of the covariance given above, we have

$$D^T_{A/M/C} = E(P_{11}) - E(PX_{11}) + \mathrm{Cov}(P,X_{11}) - E(X_M D_{A/C})$$
$$+ \mathrm{Cov}(X_M,D_{A/C})$$
$$- E(X_M)\mathrm{Cov}(P,X_C) - E(X_C D_{A/M}) + \mathrm{Cov}(X_C,D_{A/M})$$
$$- E(X_C)\mathrm{Cov}(P,X_M).$$

Finally, using the definition of the three-way allelic disequilibrium ($D_{A/M/C}$) in (11) gives us the formula corresponding to (18), with $N = A$ and $F_N = P$,

$$D^T_{A/M/C} = E(D_{A/M/C}) - E(X_M)\mathrm{Cov}(P,X_C) - E(X_C)\mathrm{Cov}(P,X_M)$$
$$+ \mathrm{Cov}(P,X_{11}) + \mathrm{Cov}(X_M,D_{A/C}) + \mathrm{Cov}(X_C,D_{A/M}).$$

MODEL OF POLLEN AND SEED MIGRATION

The model of pollen and seed migration (summarized in Figure 1) represents a three-locus extension of the two-locus cytonuclear migration models considering only maternal or only paternal cytoplasmic inheritance in Asmussen and Schnabel (1991) and Schnabel and Asmussen (1992). We consider a population with nonoverlapping generations and no seed dormancy, from which we census adults. Migration is modeled via a continent-island framework with unidirectional migration, from the source population(s) to the population under consideration (Figure 2). We assume that the migration rates and the genetic composition(s) of the source population(s) are constant over time, and we ignore the effects of selection, mutation, and genetic drift. Mating is a mixture of selfing, which occurs at rate s, and outcrossing, which occurs at rate $1 - s$ in accordance with the mixed-mating model (Clegg 1980). During mating, pollen migration occurs at rate M, such that the fraction of outcrossed pollen that comes from the migrant pollen pool is M, and the fraction of outcrossed pollen that comes from within the population is $1 - M$. The total fraction of migrant pollen per generation is thus $M(1 - s)$. After pollination and seed maturation, seed migration occurs at a rate m, which means that each generation the seed population is a mixture of migrant seeds (a fraction m) and resident seeds (a fraction $1 - m$). Following germination and growth, a new adult population is formed, completing the generation cycle.

We distinguish variables in the various life stages by letting uppercase letters represent variables in adults (*e.g.*, P, U, X_M, X_C, $D_{M/C}$) and lowercase letters represent the corresponding variables in the interim seed population (*e.g.*, p, u, x_M, x_C, $d_{M/C}$). Variables in the two migrant pools are distinguished by overbars, with lowercase letters again indicating seeds and uppercase letters now indicating pollen. Since migrant pollen carries only a haploid nuclear component and the chloroplast genome, it is characterized by its nuclear-chloroplast diallelic frequencies ($\overline{P}_{1C}$, $\overline{P}_{2C}$, $\overline{Q}_{1C}$, $\overline{Q}_{2C}$) together with its nuclear and chloroplast allele frequencies ($\overline{P}$, $\overline{X}_C$) and the nuclear-chloroplast allelic disequilibrium ($\overline{D}_{A/C}$). Migrant seeds, on the other hand, carry a full complement of all three genomes, and therefore have analogs of all the three-locus frequency and disequilibrium variables defined above (*e.g.*, $\overline{u}_{ij}$, $\overline{p}$, $\overline{x}_M$, $\overline{x}_C$, $\overline{d}_{A/M}$, $\overline{d}_{A/C}$, $\overline{d}_{M/C}$, $\overline{d}_{A/MC}$, $\overline{d}_{A/M/C}$). Disequilibria may occur in migrant seed and pollen if, for example, the source population itself is a mixture of genetically distinct populations or experiences appropriate forms of nonrandom mating (Arnold

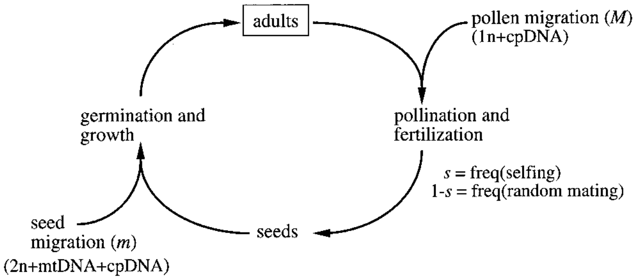

Figure 1.—Adult census model showing the two types of gene flow within a generation cycle. Pollen migration occurs first, at rate M, followed by seed migration at rate m. Mating is a mixture of selfing, which occurs at rate s, and outcrossing, which occurs at rate $1 - s$.

et al. 1988) or selection (Babcock and Asmussen 1996, 1998).

The interim seed variables, recursion equations, and equilibrium values for the nuclear-mitochondrial and nuclear-chloroplast subsystems are equivalent to those for the cases of strictly maternal and strictly paternal cytoplasmic inheritance, respectively (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992). To convert those results to the notation used here, we replace each cytoplasmic and cytonuclear variable z for maternal cytoplasmic inheritance with z_M (*e.g.*, x becomes x_M,) and those variables for paternal cytoplasmic inheritance with z_C (*e.g.*, x becomes x_C), with the nuclear component (A, AA, Aa, or aa) explicitly indicated in the cytonuclear disequilibria as above.

We develop the recursion equations for the frequencies and representative disequilibria new to the dicytonuclear system for the complete model with mixed mating and both pollen and seed migration in two steps, first finding the interim values in the seeds following pollen migration and fertilization, and then calculating the new adult values following seed migration. We also consider eight important cases subsumed within this general framework.

Interim seed values: The joint genotype frequencies at the interim seed stage can be calculated by considering the contribution to each joint genotype due to self-fertilization of resident ovules and that due to random outcrossing. The contribution due to self-fertilization depends only on the joint genotype frequencies of the resident adults (Table 1), while the contribution due to outcrossing is determined by the joint nuclear-mitochondrial allele frequencies in resident ovules and the joint nuclear-chloroplast allele frequencies in the pollen pool (Table 3), which contains both resident and migrant pollen. For example, consider the frequency of $AA/M/C$ progeny. Such progeny result from the fertilization of a resident A/M ovule by A/C pollen. Under selfing (which occurs with probability s), the only individuals who can produce the right type of ovules and pollen are $AA/M/C$ (frequency U_{11}, who always do) and $Aa/M/C$ (frequency V_{11}, who produce each correct gamete type half of the time and the appropriate combination of gametes one-fourth of the time). The contribution due to outcrossing (probability $1 - s$) is similarly straightforward to derive. In this case an A/M ovule (frequency P_{1M}) is fertilized by either migrant A/C pollen (probability $M\ \bar{P}_{1C}$) or by resident A/C pollen [probability $(1 - M)\ P_{1C}$]. The frequency of $AA/M/C$ progeny is then

$$u_{11} = s(U_{11} + {}^1\!/\!_4V_{11}) + (1 - s)P_{1M}[M\bar{P}_{1C} + (1 - M)P_{1C}].$$

Repeating the same reasoning, we find that the frequencies for all of the interim seed genotypes are

$$u_{ij} = s(U_{ij} + {}^1\!/\!_4V_{ij}) + (1 - s)P_{iM}[M\bar{P}_{jC} + (1 - M)P_{jC}] \tag{19a}$$

$$\begin{aligned} v_{ij} = \frac{s}{2}V_{ij} &+ (1 - s)P_{iM}[M\bar{Q}_{jC} + (1 - M)Q_{jC}] \\ &+ (1 - s)Q_{iM}[M\bar{P}_{jC} + (1 - M)P_{jC}] \end{aligned} \tag{19b}$$

$$w_{ij} = s(W_{ij} + {}^1\!/\!_4V_{ij}) + (1 - s)Q_{iM}[M\bar{Q}_{jC} + (1 - M)Q_{jC}], \tag{19c}$$

where $i = 1$ or 2 indicates M or m at the mitochondrial locus and $j = 1$ or 2 indicates C or c at the chloroplast locus.

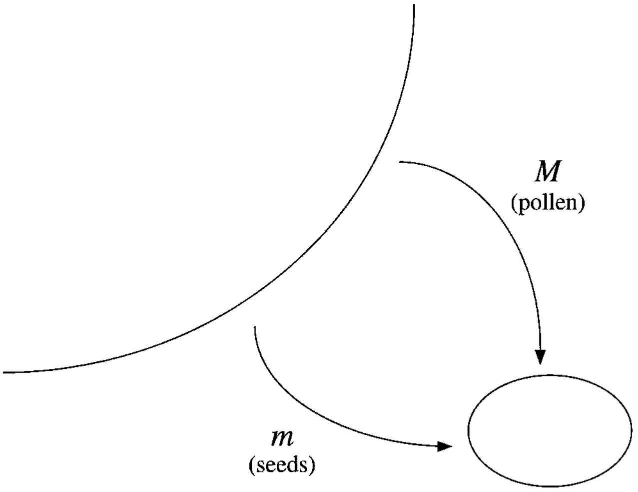

Figure 2.—Continent-island migration model with unidirectional pollen (M) and seed (m) migration.

From these interim dicytonuclear genotype frequencies, the marginal frequencies at the three loci can be found along with the interim two-locus cytonuclear disequilibria using the definitions in (4) and (5). Interim values for each of the two-locus subsystems that will be used in further derivations are given in appendix a, Equations A1–A8. The interim seed values for the other new three-locus variables can then be derived from the joint genotype frequencies using Table 1 and (7–11). The joint cytotype frequencies are

$$x_{ij} = sX_{ij} + (1 - s)X_{iM}[M\bar{X}_{jC} + (1 - M)X_{jC}], \tag{20}$$

where $X_{1M} = X_M$, $X_{2M} = Y_M$, $X_{1C} = X_C$, and $X_{2C} = Y_C$, and the two-locus cytoplasmic disequilibrium is

$$d_{M/C} = sD_{M/C}. \tag{21}$$

The formulas for the interim three-locus joint and three-way disequilibria in seeds are more complex and are given in appendix a, Equations A9–A12.

Recursion equations: Seed migration completes the life cycle. Using (13) and (14), each frequency variable in the new generation of adults is simply the weighted average of the corresponding value in migrant (m) and resident seeds ($1 - m$). The new frequencies of the joint genotypes and the joint cytotypes are then

$$U'_{ij} = m\bar{u}_{ij} + (1 - m)u_{ij}$$

$$V'_{ij} = m\bar{v}_{ij} + (1 - m)v_{ij}$$

$$W'_{ij} = m\bar{w}_{ij} + (1 - m)w_{ij} \tag{22}$$

and

$$X'_{ij} = m\bar{x}_{ij} + (1 - m)x_{ij}, \tag{23}$$

where the corresponding interim seed frequencies are given by (19) and (20).

The disequilibria in the new adults are the result of admixture between resident and migrant seeds, which can be calculated using (15–18). The cytoplasmic and joint three-locus disequilibria in new adults, for example, will be the weighted average of the corresponding disequilibria in migrant and resident seeds plus the covariance across these two seed populations between the frequencies of the two genetic components being considered. From (16), the cytoplasmic disequilibrium after admixture is

$$\begin{aligned} D'_{M/C} = {} & m\bar{d}_{M/C} + (1 - m)d_{M/C} \\ & + m(1 - m)(x_M - \bar{x}_M)(x_C - \bar{x}_C), \end{aligned} \tag{24}$$

while, using (17), the new joint three-locus disequilibria are found to be

$$\begin{aligned} D'_{N/MC} = {} & m\bar{d}_{N/MC} + (1 - m)d_{N/MC} \\ & + m(1 - m)(f_N - \bar{f}_N)(x_{11} - \bar{x}_{11}), \end{aligned} \tag{25}$$

for $N = A$, AA, Aa, or aa, where f_N is the frequency of the nuclear component N in seeds. The corresponding three-way disequilibria are similarly obtained from (18) as

$$\begin{aligned} D'_{N/M/C} = {} & m\bar{d}_{N/M/C} + (1 - m)d_{N/M/C} + m(1 - m)(f_N - \bar{f}_N) \\ & [d_{M/C} - \bar{d}_{M/C} + (2m - 1)(x_M - \bar{x}_M)(x_C - \bar{x}_C)] \\ & + m(1 - m)[(x_M - \bar{x}_M)(d_{N/C} - \bar{d}_{N/C}) \\ & + (x_C - \bar{x}_C)(d_{N/M} - \bar{d}_{N/M})]. \end{aligned} \tag{26}$$

Multiple alleles: Extension of this general framework to include multiple alleles and cytotypes is straightforward for the joint genotype and marginal frequencies. If we have n_1 nuclear alleles ($A_1, \ldots, A_{n1}$), n_2 mitochondrial cytotypes ($M_1, \ldots, M_{n2}$), and n_3 chloroplast cytotypes ($C_1, \ldots, C_{n3}$), for a total of $n_1(n_1 + 1)n_2n_3/2$ joint genotypes, we can let $F_{ij,k,l}$ indicate the frequency of adults with the A_iA_j nuclear genotype and the M_k mitochondrial and C_l chloroplast types. As an example of such a generalization, consider Equations 19a–19c giving interim joint genotype frequencies in seeds ($f_{ij,k,l}$). These would be replaced by two general equations, one for homozygotes,

$$f_{ii,k,l} = s\left(F_{ii,k,l} + {}^1\!/_4\sum_{j\neq i}F_{ij,k,l}\right) + (1 - s)P_{i,k}[M\bar{P}_{i,(l)} + (1 - M)P_{i,(l)}],$$

and one for heterozygotes,

$$\begin{aligned} f_{ij,k,l} = {} & \frac{s}{2}F_{ij,k,l} + (1 - s)P_{i,k}[M\bar{P}_{j,(l)} + (1 - M)P_{j,(l)}] \\ & + (1 - s)P_{j,k}[M\bar{P}_{i,(l)} + (1 - M)P_{i,(l)}], \end{aligned}$$

where $P_{i,k}$ is the frequency of resident A_i/M_k ovules, $P_{i,(l)}$ is the frequency of resident A_i/C_l pollen, $\bar{P}_{i,(l)}$ is the frequency of migrant A_i/C_l pollen, and $i, j = 1, 2, \ldots, n_1$, $k = 1, 2, \ldots, n_2$, and $l = 1, 2, \ldots, n_3$. Note that we do not distinguish between the nuclear genotypes A_iA_j and A_jA_i, so that, for $j > i$, $f_{ij,k,l} = f_{ji,k,l}$ and $F_{ij,k,l} = F_{ji,k,l}$. Substituting these interim seed values into the analogs of (22) then gives the adult genotypic recursions for the case of multiple alleles.

The corresponding increase in the number of disequilibria, however, is not as straightforward to analyze. Disequilibria could be defined, in the manner of Equations 4–12, for each possible two-way or three-way association between the various alleles, genotypes, and cytotypes following the multiallelic approach for cytonuclear systems in Asmussen and Basten (1996). Although there is no difficulty in setting up such definitions, interpretation of such a complex system is left for further work. For the remainder of this article, we focus on the basic case of diallelic loci.

DICYTONUCLEAR EQUILIBRIUM STRUCTURE

The utility of cytonuclear and dicytonuclear data for decomposing and estimating gene flow in plant populations depends on the extent to which such data reflect the differential effects of pollen and seed migration. To address this issue, we now turn to the equilibrium state for the dicytonuclear system, which is determined by the equilibrium values of the frequency and disequilibrium variables within the parameterization shown in Tables 6 and 7. These are calculated by setting each value after one generation of migration and mating equal to its previous value (for example, $P' = P$) and solving. Although not shown here, the stability of the three-locus equilibrium and the full dynamical behavior of the dicytonuclear system are determined by the explicit time-dependent solutions for the values of the independent populational variables in each generation t. The dynamical solutions for each of the two-locus cytonuclear systems are given in Asmussen and Schnabel (1991) and Schnabel and Asmussen (1992). Similar methods yield the explicit time-dependent solutions for the disequilibria needed to fully parameterize the three-locus dicytonuclear system. The expected form of these dynamical solutions with their multiple geometric terms shows that the resident population always converges to the unique three-locus equilibrium specified by the equilibrium equations given below.

We focus first on the outcome in mixed-mating populations receiving both pollen and seed migration ($0 < s, m, M < 1$); the distinctive features of populations that are purely selfing or random mating or experience only one form of gene flow are treated in the subsequent section as special cases. Two practical points should be kept in mind when interpreting the equilibrium values under these different biological conditions. First, although this is a continent-island model, the resident population does not simply become an exact replica of the source population because of the two distinct forms of gene flow and the various levels of nonrandom associations these can generate among the three markers. Second, we are particularly interested in when permanent disequilibria are produced. Although not strictly necessary for gene flow estimation, permanent disequilibria should increase the conditions under which the equilibrium for a cytonuclear or dictyonuclear system reflects, and can be used to estimate, the rate of pollen (M) or seed (m) migration.

Final cytonuclear variables and marginal frequencies: The equilibria for the two-locus subsystems are given in Asmussen and Schnabel (1991) and Schnabel and Asmussen (1992) and are provided in the current notation in appendix b, Equations B1–B4. Because of their key roles in the new three-locus formulas, we reiterate the equilibrium formulas for the marginal frequencies and the nuclear genotype frequencies here. The final nuclear and chloroplast allele frequencies,

$$\hat{P} = \frac{2m\bar{p} + (1 - m)M(1 - s)\bar{P}}{2m + (1 - m)M(1 - s)} \tag{27}$$

and

$$\hat{X}_C = \frac{m\bar{x}_C + (1 - m)M(1 - s)\bar{X}_C}{m + (1 - m)M(1 - s)}, \tag{28}$$

are weighted averages of the frequencies in migrant seeds and pollen, while the equilibrium mitochondrial frequency,

$$\hat{X}_M = \bar{x}_M, \tag{29}$$

is simply the frequency in migrant seeds, since only migrant seeds, and not migrant pollen, carry the mitochondrial locus. Note that in each case a polymorphic equilibrium is required to generate permanent disequilibria involving that marker. The frequency of nuclear heterozygotes converges to a complicated convolution of the frequency of heterozygotes in migrant seeds and the nuclear allele frequencies in the two migrant pools,

$$\hat{V} = \frac{2m\bar{v} + 2(1 - m)(1 - s)[M(\hat{P}\bar{Q} + \hat{Q}\bar{P}) + 2(1 - M)\hat{P}\hat{Q}]}{2 - (1 - m)s}, \tag{30}$$

with the final nuclear homozygote frequencies given by $\hat{U} = \hat{P} - \frac{1}{2}\hat{V}$ and $\hat{W} = \hat{Q} - \frac{1}{2}\hat{V}$.

The final marginal frequencies are those of the joint cytotypes, which are new to the three-locus system. From (20) and (23), we find that the equilibrium frequency of each joint cytotype

$$\hat{X}_{ij} = \frac{m\bar{x}_{ij} + (1 - m)(1 - s)\hat{X}_{iM}[M\bar{X}_{jC} + (1 - M)\hat{X}_{jC}]}{1 - (1 - m)s} \tag{31}$$

(where $X_{1M} = X_M$, $X_{2M} = Y_M$, $X_{1C} = X_C$, and $X_{2C} = Y_C$) is the weighted average of the corresponding frequency in migrant seeds ($m\bar{x}_{ij}$) and the frequency in resident outcrossed seeds formed using either migrant pollen $[(1 - m)M(1 - s)\hat{X}_{iM}\bar{X}_{jC}]$ or resident pollen $[(1 - m)(1 - M)(1 - s)\hat{X}_{iM}\hat{X}_{jC}]$.

Final two-way disequilibria: Turning to disequilibrium measures new to the three-locus system, we find that the equilibrium value for the cytoplasmic disequilibrium is

$$\hat{D}_{M/C} = \frac{m}{1 - (1 - m)s}\bar{d}_{M/C}. \tag{32}$$

From (32), we see that pollen migration has no effect on the final association between alleles at the two cytoplasmic loci since pollen carry only the chloroplast and not the mitochondrial marker. In fact, there will be permanent disequilibrium between the two cytoplasmic loci if and only if the mitochondrial and chloroplast cytotypes are nonrandomly associated in migrant seeds. Such nonrandom associations among these or other markers could reflect disequilibria generated by selection or other nonrandomizing forces in the source population; migrant seed (or pollen) disequilibria would also be expected if seeds (or pollen) are derived from multiple, genetically distinct sources, as would occur in hybrid zones or other areas of admixture (see Equations 15–18).

We now consider the representative three-locus disequilibria. At equilibrium, the joint allelic disequilibrium is a linear combination of six factors,

$$\begin{aligned}\hat{D}_{A/MC} = {} & c_1\bar{d}_{A/MC} + c_2\bar{d}_{A/M} + c_3\bar{d}_{A/C} + c_4\bar{D}_{A/C} \\ & + c_5(\bar{P} - \bar{p})\bar{d}_{M/C} + c_6(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C),\end{aligned} \tag{33}$$

where the c_i are constants that depend on the migration and selfing rates and the cytotype frequencies in the migrant seeds and pollen; these are given in (C1) in appendix c. We see immediately that joint allelic disequilibrium can be produced by the corresponding disequilibrium in migrant seeds ($\bar{d}_{A/MC}$) or by allelic cytonuclear disequilibria in the migrant pools within either the nuclear-mitochondrial or nuclear-chloroplast subsystems ($\bar{d}_{A/M}$, $\bar{d}_{A/C}$, $\bar{D}_{A/C}$). Joint allelic associations can also be generated by intermigrant interactions involving unequal nuclear allele frequencies in migrant pollen and seeds coupled either with cytoplasmic disequilibrium in migrant seeds ($\bar{P} \neq \bar{p}$ and $\bar{d}_{M/C} \neq 0$) or with unequal chloroplast frequencies in the two migrant pools ($\bar{P} \neq \bar{p}$ and $\bar{X}_C \neq \bar{x}_C$). Different nuclear or chloroplast frequencies in the migrant pollen and seeds could be a result of differences in average distance for these two types of gene flow, resulting in different sources for pollen and seeds or differences in the relative contribution of multiple source populations to the two migrant pools (*e.g.*, source 1 might contribute more pollen than seeds). Selection or other evolutionary forces within the life cycle of the source population(s) could also result in different nuclear or chloroplast frequencies in the two migrant pools. While joint nuclear-chloroplast allele frequency differences also generate nuclear-chloroplast allelic disequilibria [$\hat{D}_{A/C}$, which is $\hat{D}$ in Schnabel and Asmussen (1992)], the intermigrant interaction of nuclear allele frequency differences with cytoplasmic disequilibria in migrant seeds is a new factor unique to the three-locus system.

The final three-locus joint genotypic disequilibria can be written as

$$\begin{aligned}\hat{D}_{AA/MC} = {} & c_1\bar{d}_{AA/MC} + c_2\hat{D}_{Aa/MC} + c_3\bar{d}_{A/M} + c_4\bar{d}_{A/C} + c_5\bar{D}_{A/C} \\ & + c_6\bar{d}_{A/M}\bar{d}_{A/C} + c_7\bar{d}_{A/M}\bar{D}_{A/C} + c_8\bar{d}_{M/C} + c_9(\bar{X}_C - \bar{x}_C) \\ & + c_{10}(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C) + c_{11}(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M}\end{aligned} \tag{34a}$$

$$\begin{aligned}\hat{D}_{Aa/MC} = {} & c_1\bar{d}_{Aa/MC} + c_2\bar{d}_{A/M} + c_3\bar{d}_{A/C} + c_4\bar{D}_{A/C} + c_5\bar{d}_{A/M}\bar{d}_{A/C} \\ & + c_6\bar{d}_{A/M}\bar{D}_{A/C} + c_7\bar{d}_{M/C} + c_8(\bar{X}_C - \bar{x}_C) \\ & + c_9(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C) + c_{10}(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M},\end{aligned} \tag{34b}$$

where the c_i coefficients for each disequilibrium are given in (C2) and (C3), respectively. Similarly to the joint allelic disequilibrium, joint genotypic disequilibria can be produced by joint disequilibria in migrant seeds ($\bar{d}_{AA/MC}$ and $\bar{d}_{Aa/MC}$ for the homozygote disequilibrium, $\bar{d}_{Aa/MC}$ alone for the heterozygote disequilibrium) or by allelic cytonuclear disequilibria in the migrant pools ($\bar{d}_{A/M}$, $\bar{d}_{A/C}$, $\bar{D}_{A/C}$ and crossproducts of these). However, joint genotypic disequilibria can also be generated directly by cytoplasmic disequilibria in migrant seeds ($\bar{d}_{M/C}$) without the need for unequal nuclear allele frequencies in the two migrant pools. Finally, joint genotypic disequilibria can be generated via intermigrant interactions involving unequal chloroplast frequencies in migrant pollen and seeds ($\bar{X}_C \neq \bar{x}_C$), either alone or in conjunction with other intermigrant factors ($\bar{P} \neq \bar{p}$ with or without $\bar{d}_{A/M} \neq 0$).

Final three-way disequilibria: The three-way associations are generated in fewer ways than the three-locus joint disequilibria. Allelic cytonuclear disequilibria in the migrant pools ($\bar{d}_{A/M}$, $\bar{d}_{A/C}$, or $\bar{D}_{A/C} \neq 0$) or "simple" intermigrant admixture effects due to allele frequency differences in migrant pollen and seeds ($\bar{X}_C \neq \bar{x}_C$, or $\bar{P} \neq \bar{p}$ and $\bar{X}_C \neq \bar{x}_C$) are insufficient to generate true three-way associations. For example, the three-way allelic disequilibrium is found at equilibrium to be

$$\hat{D}_{A/M/C} = c_1\bar{d}_{A/M/C} + c_2(\bar{P} - \bar{p})\bar{d}_{M/C} + c_3(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M}, \tag{35}$$

where the c_i are given in (C4). This can be generated in only three ways: by the corresponding disequilibria in migrant seeds ($\bar{d}_{A/M/C}$) or by either of two types of intermigrant interactions unique to the dicytonuclear system. The first couples unequal nuclear allele frequencies with cytoplasmic disequilibria in migrant seeds ($\bar{P} \neq \bar{p}$ and $\bar{d}_{M/C} \neq 0$), which also enters into the joint allelic disequilibrium (33). The second is a new, three-locus interaction involving unequal chloroplast frequencies in migrant pollen and seeds together with nuclear-mitochondrial allelic disequilibria in migrant seeds ($\bar{X}_C \neq \bar{x}_C$ and $\bar{d}_{A/M} \neq 0$).

Finally, at equilibrium, the three-way genotypic disequilibria are

$$\begin{aligned}\hat{D}_{AA/M/C} = {} & c_1\bar{d}_{AA/M/C} + c_2\hat{D}_{Aa/M/C} + c_3\bar{d}_{M/C} + c_4\bar{d}_{A/M}\,\bar{d}_{A/C} + c_5\bar{d}_{A/M}\bar{D}_{A/C} \\ & + c_6(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M} + c_7(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M} \\ & + c_8(\bar{X}_C - \bar{x}_C)\bar{d}_{AA/M}\end{aligned} \tag{36a}$$

$$\begin{aligned}\hat{D}_{Aa/M/C} = {} & c_1\bar{d}_{Aa/M/C} + c_2\bar{d}_{M/C} + c_3\bar{d}_{A/M}\bar{d}_{A/C} + c_4\bar{d}_{A/M}\bar{D}_{A/C} \\ & + c_5(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M} + c_6(\bar{P} - \bar{p})(\bar{X}_C - \bar{x}_C)\bar{d}_{A/M} \\ & + c_7(\bar{X}_C - \bar{x}_C)\bar{d}_{Aa/M},\end{aligned} \tag{36b}$$

where the c_i are given in (C5) and (C6), respectively. Paralleling the joint genotypic disequilibria in (34), these can be generated by three-way genotypic associations in migrant seeds ($\bar{d}_{AA/M/C}$ or $\bar{d}_{Aa/M/C}$ for the homozygote disequilibria, $\bar{d}_{Aa/M/C}$ alone for the heterozygote disequilibria) and by cytoplasmic disequilibria in migrant seeds ($\bar{d}_{M/C}$). However, as for the three-way allelic association, the allelic cytonuclear disequilibria in the migrant pools contribute only in conjunction with other terms. For example, nuclear-chloroplast allelic disequilibrium in either migrant pool can generate three-way genotypic associations only in conjunction with nuclear-mitochondrial allelic disequilibria in migrant seeds ($\bar{d}_{A/M} \neq 0$ plus $\bar{d}_{A/C} \neq 0$ or $\bar{D}_{A/C} \neq 0$). Additionally, nuclear-mitochondrial allelic disequilibria in migrant seeds can contribute when combined with unequal chloroplast frequencies in the migrant pools, with or without unequal nuclear allele frequencies ($\bar{X}_C \neq \bar{x}_C$ and $\bar{d}_{A/M} \neq 0$; or $\bar{P} \neq \bar{p}$, $\bar{X}_C \neq \bar{x}_C$, and $\bar{d}_{A/M} \neq 0$). Finally, another form of intermigrant interaction, new to the three-locus system, can also produce three-way genotypic disequilibria: unequal chloroplast frequencies in the migrant pools coupled with the corresponding nuclear-mitochondrial genotypic disequilibrium in migrant seeds (*e.g.*, $\bar{X}_C \neq \bar{x}_C$ and $\bar{d}_{AA/M} \neq 0$ for $\hat{D}_{AA/M/C}$). The dependence of the final three-way homozygote disequilibria on the heterozygote value ($\hat{D}_{Aa/M/C}$) means that the former can also be generated by unequal chloroplast frequencies in conjunction with nonrandom associations between heterozygotes and the mitochondrial marker in seeds ($\bar{X}_C \neq \bar{x}_C$ and $\bar{d}_{Aa/M} \neq 0$).

Equilibrium three-locus genotype frequencies: To

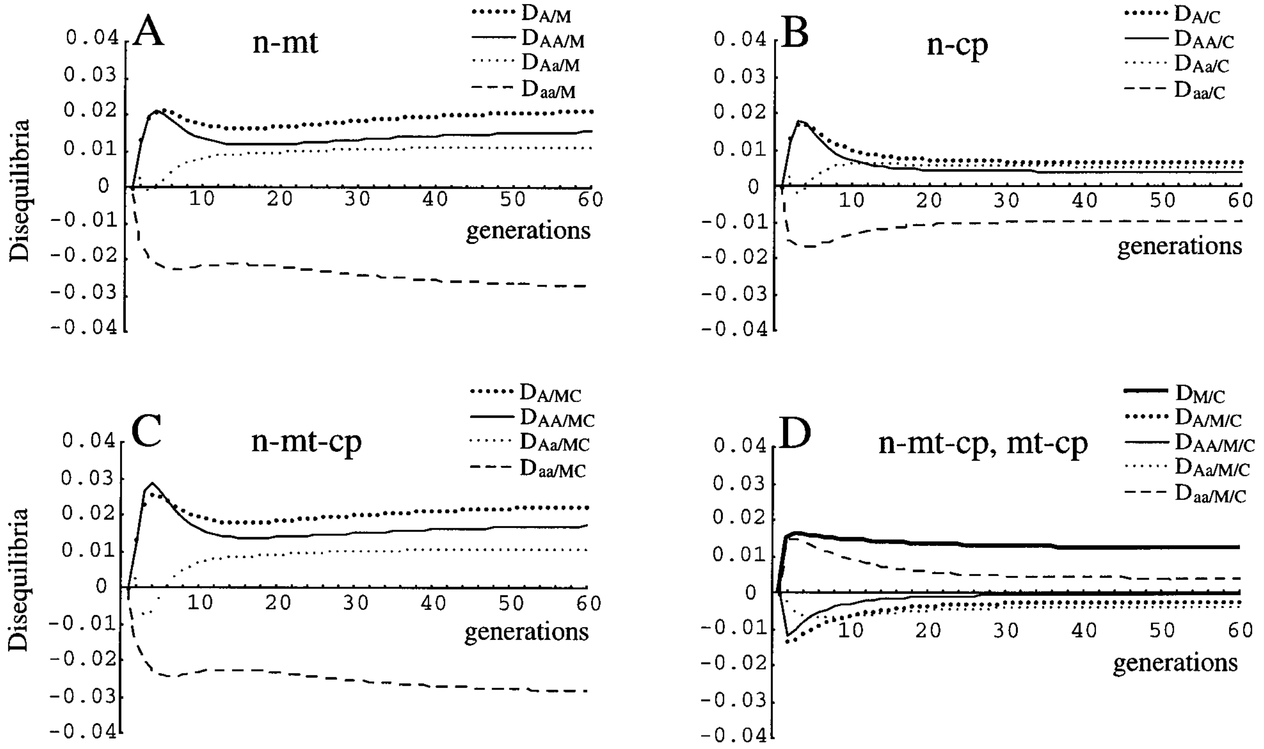

Figure 3.—Trajectories of permanent nonzero disequilibria. Initial conditions were $U_{11} = 1.0$; parameter values were $\overline{s} = 0.1$, $\overline{m} = 0.05$, $\overline{M} = 0.2$, $\overline{u}_{11} = 0.7$, $\overline{w}_{22} = 0.3$, and $\overline{Q}_{1C} = 1.0$. Disequilibria in the migrant pools were $\overline{d}_{M/C} = \overline{d}_{A/*} = \overline{d}_{AA/*} = -\overline{d}_{aa/*} = 0.21$, $\overline{d}_{Aa/*} = \overline{d}_{Aa/M/C} = 0$ (where * indicates M, C, or MC), $\overline{d}_{A/M/C} = \overline{d}_{AA/M/C} = -\overline{d}_{aa/M/C} = -0.084$, and $\overline{D}_{A/C} = 0$.

complete the specification of the equilibrium state for the dicytonuclear system, we must calculate the final three-locus genotype frequencies ($\hat{U}_{ij}$, $\hat{V}_{ij}$, $\hat{W}_{ij}$). These can be obtained by substituting the relevant equilibrium formulas [(27–30), (32–34), (B1–B4), (C1), and (C2)] into the decompositions given in Tables 6 and 7.

Numerical examples: Numerical examples allow us to compare the cytoplasmic and three-locus disequilibria generated by pollen and seed migration in the full nuclear-mtDNA-cpDNA system with the disequilibria in each of the two cytonuclear subsystems. Previous work has found that 0.1 is roughly the minimal detectable level of two-way disequilibria, given reasonable sample sizes and marginal frequencies (Basten and Asmussen 1997). Generally, both permanent and transient disequilibria generated via pollen migration alone are smaller in magnitude than disequilibria generated via a comparable amount of seed migration alone, as was found previously for the cytonuclear subsystems by Schnabel and Asmussen (1992). Higher selfing rates increase the magnitude of disequilibria and slow the decay of transient disequilibria. Both permanent and transient three-way disequilibria tend to be smaller than the corresponding two-locus or joint disequilibria, as would be expected of measures for higher-order effects.

Figure 3 shows an example where the resident population is initially monomorphic at each marker ($U_{11} = 1.0$) and receives genetically distinct migrant pollen and seeds ($\overline{P} = 0, \overline{p} = 0.7, \overline{X}_C = 1.0, \overline{x}_C = 0.7$) with nonrandom cytonuclear associations in the seeds. Such a situation might arise, for example, if a genetically homogeneous population received migrant pollen from a source fixed for a different nuclear marker but received seeds from both this source and another source, differing in both nuclear and chloroplast frequencies. These differences in nuclear and chloroplast frequencies in the migrant pools and cytonuclear associations in migrant seeds generate permanent nonzero values for all the disequilibria. In the absence of actual values of seed and pollen migration rates and selfing rates, this example is used to illustrate cases with a larger pollen migration rate ($M = 0.2$) than seed migration rate ($m = 0.05$) as would be expected for most plant species with wind-dispersed pollen. The joint allelic and genotypic disequilibria for the three-genome system (Figure 3C) are similar in the shape of their trajectories to the disequilibria from the nuclear-mitochondrial subsystem (Figure 3A), although somewhat greater in magnitude. The nuclear-chloroplast and cytoplasmic disequilibria also have similar shapes but are smaller in magnitude (Figure 3, B and D). The three-way allelic and genotypic disequilibria for this example (Figure 3D) both have slightly smaller magnitudes and behave very differently than the two-way disequilibria. For instance, the three-way allelic and homozygote disequilibria have signs opposite those of the corresponding joint associations, while the three-way heterozygote disequilibrium differs from all other heterozygote disequilibria in this example by not changing sign along its trajectory. This example corresponds to a mainly outcrossing species ($s = 0.1$); increasing the selfing rate to $s = 0.9$ increases the magnitude of all the disequilibria except the heterozygote nuclear-mitochondrial ($D_{Aa/M}$) and joint ($D_{Aa/MC}$) disequilibria (all parameters other than s as in Figure 3, data not shown). Most of the other disequilibrium measures are increased by a factor of ~4 to 8, but some increase by a great deal more. For example, the homozygote nuclear-chloroplast disequilibrium ($D_{AA/C}$) increases by a factor of 20, from ~0.004 ($s = 0.1$) to 0.08 ($s = 0.9$).

Figure 4 gives a case where the resident population receives only migrant seeds and no migrant pollen ($M = 0$) and is initially fixed at the nuclear and mitochondrial loci ($P = X_M = 1.0$), but polymorphic at the chloroplast locus ($X_C = 0.8$). In contrast to the population shown in Figure 3, here the migrant seeds have no cytonuclear disequilibria, and there can be no intermigrant admixture effects with only one type of migration; therefore, no permanent nonzero disequilibria are generated. However, transient disequilibria are generated by differences in the nuclear, mitochondrial, and chloroplast frequencies between the original population and the migrant seeds ($\bar{p} = \bar{x}_M = 0.0$, $\bar{x}_C = 0.4$). In this example with high selfing ($s = 0.9$), the transient nuclear-mitochondrial disequilibria (Figure 4A) and joint dicytonuclear disequilibria (Figure 4C) that are generated can reach quite high values (>0.15 in magnitude) before dissipating; additionally, these can persist for relatively long periods of time (>0.05 in magnitude for almost 40 generations), increasing the probability of detection, particularly for long-lived organisms where generation times are long. However, if we consider lower selfing rates corresponding to predominantly outcrossing species, both the maximum level and the duration of the disequilibria are much reduced. For example, with a selfing rate of $s = 0.1$ (all other parameters as in Figure 4, data not shown), the transient nuclear-mitochondrial and joint dicytonuclear disequilibria do not reach 0.1 in magnitude before dissipating and remain above 0.05 in magnitude for <10 generations. Even with a high selfing rate ($s = 0.9$), neither the nuclear-chloroplast nor the cytoplasmic disequilibria (Figure 4, B and D) exceed 0.08 before dissipating. The three-way measures for this example are much smaller in magnitude than the other associations (Figure 4D) and are interesting in that all but the heterozygote association show a sign change in their trajectories.

DICYTONUCLEAR EQUILIBRIUM STRUCTURE FOR SPECIAL CASES

To further analyze the effects of the two different forms of gene flow and the mating system on the generation of permanent disequilibria, we consider the equilibrium structure under a series of important special cases subsumed within this general framework. These include (1) seed migration alone ($0 < m < 1$, $M = 0$), (2) pollen migration alone ($0 < M < 1$, $m = 0$), (3) complete random mating ($s = 0$), (4) complete self-fertilization with seed migration ($s = 1$, $0 < m < 1$), (5) equal nuclear allele frequencies in the two migrant pools ($\bar{P} = \bar{p}$), (6) equal chloroplast frequencies in the two migrant pools ($\bar{X}_C = \bar{x}_C$), (7) equivalent migrant pools ($\bar{P} = \bar{p}$, $\bar{X}_C = \bar{x}_C$, $\bar{D}_{A/C} = \bar{d}_{A/C}$), and (8) no migrant disequilibria ($\bar{D} = \bar{d} = 0$ for all migrant pollen and seed disequilibria). Since in each case the behavior of the disequilibria for both of the cytonuclear subsystems has been considered earlier (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992), details for these are not given unless necessary for clarification. In the following discussion, we focus on the distinctive features of each special case that differ from the general case presented above, with explicit equilibrium formulas given only when they represent significant simplifications.

Seed migration alone ($0 < m < 1$, $M = 0$): Populations that receive gene flow solely through seeds have a very different equilibrium structure from those receiving both forms of gene flow. With no pollen migration, all of the allele frequencies approach those in migrant seeds ($\hat{P} = \bar{p}$, $\hat{X}_M = \bar{x}_M$, and $\hat{X}_C = \bar{x}_C$) and, since there can be no intermigrant effects, permanent disequilibria are generated only by the presence of nonzero disequilibria in the migrant seed pool. These distinctions result in a much simpler form for the final joint allelic disequilibrium,

$$[1 - (1 - m)s]\hat{D}_{A/MC} = m\bar{d}_{A/MC} + \frac{m(1 - m)(1 - s)(\bar{x}_C\bar{d}_{A/M} + \bar{x}_M\bar{d}_{A/C})}{2 - (1 - m)(1 + s)}, \tag{37}$$

which now will be nonzero if and only if the nuclear alleles are nonrandomly associated with a joint cytotype or with one of the cytoplasmic markers in migrant seeds ($\bar{d}_{A/MC}$, $\bar{d}_{A/M}$, or $\bar{d}_{A/C} \neq 0$). The joint genotypic disequilibria now reach the values

$$\begin{aligned}[1 - (1 - m)s]\hat{D}_{AA/MC} = {} & m\bar{d}_{AA/MC} + \tfrac{1}{4}(1 - m)s\hat{D}_{Aa/MC} \\ & + \frac{2m(1 - m)(1 - s)\bar{p}(\bar{x}_C\bar{d}_{A/M} + \bar{x}_M\bar{d}_{A/C})}{2 - (1 - m)(1 + s)} \\ & + \frac{4m^2(1 - m)(1 - s)\bar{d}_{A/M}\bar{d}_{A/C}}{[2 - (1 - m)(1 + s)]^2} \\ & + \frac{m(1 - m)(1 - s)\bar{d}_A\bar{d}_{M/C}}{1 - (1 - m)s}\end{aligned} \tag{38a}$$

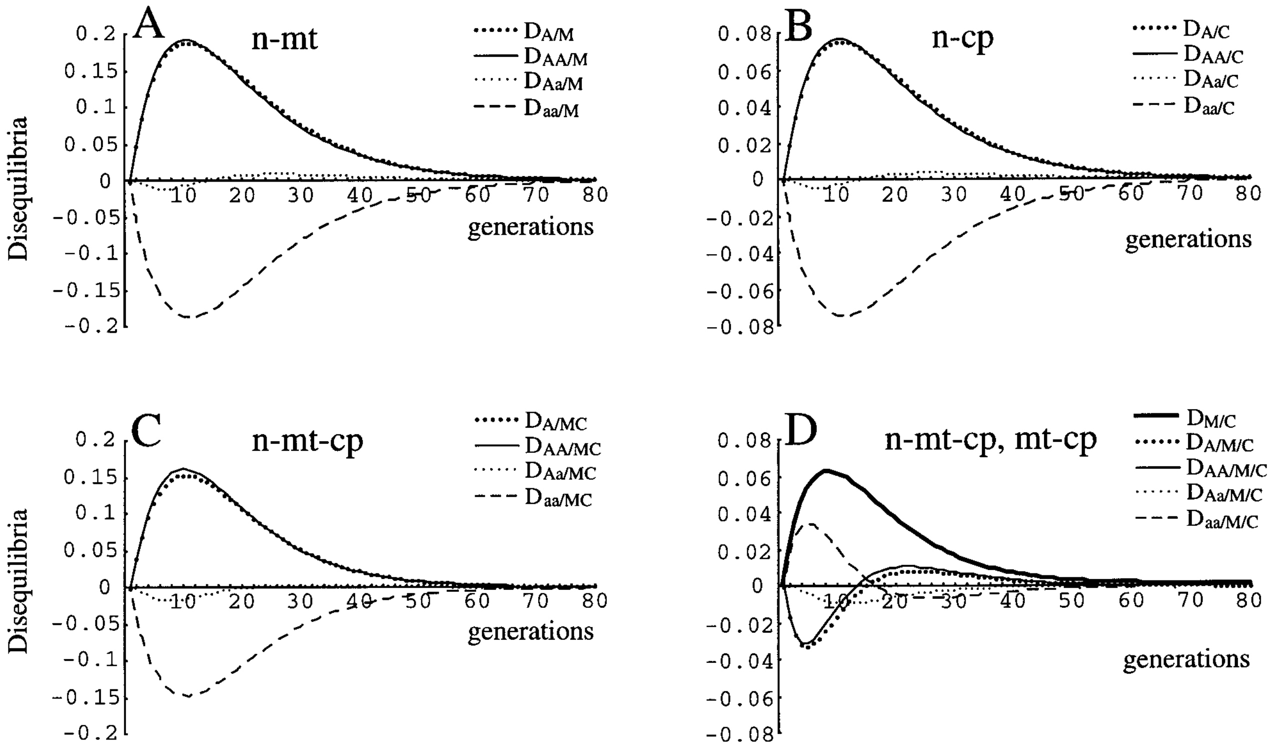

Figure 4.—Trajectories of transient nonzero disequilibria. Initial conditions were $U_{11} = 0.8$, $U_{12} = 0.2$; parameter values were $s = 0.9$, $m = 0.05$, $M = 0$, $\overline{w}_{21} = 0.4$, $\overline{w}_{22} = 0.6$, $\overline{Q}_{2C} = 0.6$, and $\overline{Q}_{1C} = 0.4$. All disequilibria in the migrant pools are zero. Note differences in scale between left and right.

$$[2 - (1 - m)s]\hat{D}_{Aa/MC} = 2m\overline{d}_{Aa/MC} + \frac{4m(1 - m)(1 - s)(\overline{q} - \overline{p})(\overline{x}_C\overline{d}_{A/M} + \overline{x}_M\overline{d}_{A/C})}{2 - (1 - m)(1 + s)} - \frac{16m^2(1 - m)(1 - s)\overline{d}_{A/M}\overline{d}_{A/C}}{[2 - (1 - m)(1 + s)]^2} - \frac{4m(1 - m)(1 - s)\overline{d}_A\overline{d}_{M/C}}{1 - (1 - m)s}, \tag{38b}$$

where $\overline{d}_A = \overline{u} - \overline{p}^2$ is the Hardy-Weinberg disequilibrium (Weir 1996) in migrant seeds. Note that, in contrast to the general case, the cytoplasmic disequilibrium only contributes to joint genotypic associations if the migrant seeds are not in Hardy-Weinberg equilibrium ($\overline{d}_A$ and $\overline{d}_{M/C} \neq 0$), implying that deviations from random mating or other causes of Hardy-Weinberg disequilibrium in the source population may be necessary to generate these types of associations when there is only seed migration.

The final three-way allelic disequilibrium,

$$\hat{D}_{A/M/C} = \frac{m\overline{d}_{A/M/C}}{1 - (1 - m)s}, \tag{39}$$

is now nonzero if and only if there is corresponding disequilibrium in migrant seeds ($\overline{d}_{A/M/C} \neq 0$). In contrast, seed migration alone can generate permanent three-way genotypic disequilibria

$$[1 - (1 - m)s]\hat{D}_{AA/M/C} = m\overline{d}_{AA/M/C} + \tfrac{1}{4}(1 - m)s\hat{D}_{Aa/M/C} + \frac{4m^2(1 - m)(1 - s)\overline{d}_{A/M}\overline{d}_{A/C}}{[2 - (1 - m)(1 + s)]^2} + \frac{m(1 - m)(1 - s)\overline{d}_A\overline{d}_{M/C}}{1 - (1 - m)s} \tag{40a}$$

$$[2 - (1 - m)s]\hat{D}_{Aa/M/C} = 2m\overline{d}_{Aa/M/C} - \frac{16m^2(1 - m)(1 - s)\overline{d}_{A/M}\overline{d}_{A/C}}{[2 - (1 - m)(1 + s)]^2} - \frac{4m(1 - m)(1 - s)\overline{d}_A\overline{d}_{M/C}}{1 - (1 - m)s} \tag{40b}$$

by three factors: three-way genotypic disequilibria in migrant seeds, allelic cytonuclear disequilibria in migrant seeds for both subsystems ($\overline{d}_{A/M}$ and $\overline{d}_{A/C} \neq 0$), or cytoplasmic associations coupled with Hardy-Weinberg disequilibria in migrant seeds ($\overline{d}_{M/C}$ and $\overline{d}_A \neq 0$).

Pollen migration alone ($0 < M < 1$, $m = 0$): An even more distinctive equilibrium structure results when pollen is the only form of gene flow, as might arise in wind-pollinated and other species where pollen travels much farther than seeds. When there is no seed migration, both the nuclear and chloroplast allele frequencies approach those in migrant pollen ($\hat{P} = \overline{P}$, $\hat{X}_C = \overline{X}_C$), while the mitochondrial frequency remains at the initial value of the resident population ($\hat{X}_M = X_M^{(0)}$). The cytoplasmic disequilibrium and the three-way disequilibria all become zero ($\hat{D}_{M/C} = \hat{D}_{N/M/C} = 0$), while the final joint allelic and joint genotypic disequilibria are constant multiples of (and thus solely generated by) the nuclear-chloroplast allelic cytonuclear disequilibrium in migrant pollen ($\bar{D}_{A/C}$),

$$\hat{D}_{A/MC} = \frac{MX_M^{(0)}\bar{D}_{A/C}}{1 + M} \tag{41}$$

$$\hat{D}_{AA/MC} = \frac{M[4(1 - s)\bar{P} + s]X_M^{(0)}\bar{D}_{A/C}}{(1 + M)(2 - s)} \tag{42a}$$

$$\hat{D}_{Aa/MC} = \frac{4M(1 - s)(\bar{Q} - \bar{P})X_M^{(0)}\bar{D}_{A/C}}{(1 + M)(2 - s)}. \tag{42b}$$

Note that these are the equilibrium values for the corresponding two-locus nuclear-chloroplast disequilibria [(B2) and (B4)] multiplied by the initial mitochondrial frequency ($\hat{D}_{N/MC} = X_M^{(0)}\hat{D}_{N/C}$). Another notable feature of pollen migration alone is that the final joint allelic disequilibrium is independent of the selfing rate.

Complete random mating ($s = 0$): The overall effects of the mating system are illustrated by the extremes with either complete random mating or complete selfing. In a completely random mating population, all equilibrium values simplify somewhat but retain the same overall form as in the general case (*e.g.*, $\hat{D}_{M/C} = m\bar{d}_{M/C}$), with the exception that the joint and three-way homozygote disequilibria do not depend on the corresponding heterozygote associations in migrant seeds ($\bar{d}_{Aa/MC}$ or $\bar{d}_{Aa/M/C}$). The overall equilibrium structure in mixed-mating populations is thus, qualitatively, that due to its random outcrossing component.

Complete self-fertilization with seed migration ($s = 1$, $0 < m < 1$): Completely selfing populations are distinct, as a result of having both a unique mating system and being closed to pollen flow. Those receiving seed flow are like mixed-mating populations with seed migration alone in that allele frequencies converge to those in migrant seeds ($\hat{P} = \bar{p}$, $\hat{X}_M = \bar{x}_M$, $\hat{X}_C = \bar{x}_C$), and since there can be no intermigrant effects, permanent disequilibria require disequilibria in the migrant seed pool. Purely selfing populations, however, have further distinctive features. The final cytoplasmic, joint allelic, and three-way allelic disequilibria are now simply the corresponding values in migrant seeds ($\hat{D}_{M/MC} = \bar{d}_{M/C}$, $\hat{D}_{A/MC} = \bar{d}_{A/MC}$, $\hat{D}_{A/M/C} = \bar{d}_{A/M/C}$). The joint genotypic disequilibria become

$$\hat{D}_{AA/MC} = \bar{d}_{AA/MC} + \frac{(1 - m)\bar{d}_{Aa/MC}}{2(1 + m)} \tag{43a}$$

and

$$\hat{D}_{Aa/MC} = \frac{2m\bar{d}_{Aa/MC}}{1 + m} \tag{43b}$$

and the formulas for the three-way genotypic disequilibria are equivalent to these with MC replaced by M/C throughout. In each case, the homozygote disequilibrium is a linear combination of the corresponding homozygote and heterozygote disequilibria in migrant seeds, while the heterozygote disequilibrium is a simple fraction of the migrant seed value.

Equal nuclear allele frequencies in migrant pollen and seeds ($\bar{P} = \bar{p}$): The opportunity for permanent disequilibria is reduced if there are no intermigrant allele frequency differences, as when both migrant pollen and migrant seeds come from the same source population, and no selection or other forces change allele frequencies between the adult and seed stages in the source. Moreover, with different source populations for migrant pollen and seeds, we might expect more differentiation across populations in cytoplasmic markers (since they have smaller effective population sizes) than in nuclear markers, such that only the chloroplast marker differs between migrant pollen and seeds. In these and other cases where the two migrant pools have equal nuclear allele frequencies, the nuclear allele frequency reaches the common migrant value ($\hat{P} = \bar{P} = \bar{p}$), and it is more difficult to generate three-locus disequilibria. We find, for example, that permanent joint and three-way allelic disequilibria are no longer generated by cytoplasmic disequilibrium in migrant seeds ($\bar{d}_{M/C}$). In addition, the final joint allelic disequilibrium ($\hat{D}_{A/MC}$) loses all intermigrant effects and is generated if and only if nuclear alleles are nonrandomly associated with a joint cytotype ($\bar{d}_{A/MC} \neq 0$) or with at least one of the cytoplasmic markers ($\bar{d}_{A/M}$, $\bar{d}_{A/C}$, or $\bar{D}_{A/C} \neq 0$) in the migrant pools. As with only seed migration, Hardy-Weinberg disequilibrium in migrant seeds plays a critical role, being needed when $\bar{P} = \bar{p}$ for unequal chloroplast frequencies in the migrant pools to produce joint genotypic disequilibria ($\bar{d}_A \neq 0$ and $\bar{X}_C \neq \bar{x}_C$) and for cytoplasmic disequilibrium to produce joint or three-way genotypic disequilibria ($\bar{d}_{M/C}$ and $\bar{d}_A \neq 0$).

Equal chloroplast frequencies in migrant pollen and seeds $\bar{X}_C = \bar{x}_C$): With equal chloroplast frequencies in the two migrant pools, the chloroplast frequency approaches the common migrant value ($\hat{X}_C = \bar{X}_C = \bar{x}_C$). Permanent joint allelic disequilibrium ($\hat{D}_{A/MC}$) can be generated by all of the factors in the general case except for unequal nuclear and chloroplast frequencies in the migrant pools. Note that, unlike the case of equal nuclear allele frequencies, the joint allelic disequilibrium does not lose all dependence on intermigrant admixture effects, as it can still be generated by intermigrant differences in nuclear allele frequencies coupled with cytoplasmic disequilibrium in migrant seeds ($\bar{P} \neq \bar{p}$ and $\bar{d}_{M/C} \neq 0$). The joint genotypic disequilibria lose all intermigrant admixture effects involving unequal allele frequencies in the migrant pools, although they retain the interaction between nuclear-mitochondrial allelic disequilibrium in the migrant seed pool and nuclear-chloroplast allelic disequilibrium in the migrant pollen pool ($\bar{d}_{A/M} \neq 0$ and $\bar{D}_{A/C} \neq 0$). The final distinction is that three-way allelic disequilibria can no longer be generated by nuclear-mitochondrial associations in migrant seeds ($\bar{d}_{A/M}$).

Equivalent migrant pools ($\bar{P} = \bar{p}$, $\bar{X}_C = \bar{x}_C$, $\bar{D}_{A/C} = \bar{d}_{A/C}$): The equilibrium for completely equivalent migrant pools is similar to that for seed migration alone. The allele frequencies reach the same values, which are now the common frequencies in the two migrant pools, and permanent disequilibria are generated by the same migrant disequilibrium factors found in (37–40); the final three-way allelic disequilibrium is, in fact, given by (39) in both cases. However, unlike the case of seed migration alone, the three-way allelic disequilibrium is the only three-locus association that loses its dependence on the pollen migration rate (M). For the other three-locus disequilibria, the pollen migration rate enters into the terms involving the nuclear-chloroplast allelic disequilibrium ($\bar{d}_{A/C}$).

No migrant disequilibria ($\bar{D} = \bar{d} = 0$, for all migrant pollen and seed disequilibria): With no nonrandom associations in the migrant pools, the cytoplasmic and all of the three-way disequilibria become zero ($\hat{D}_{M/C} = \hat{D}_{N/M/C} = 0$, where $N = A, AA, Aa, aa$). Permanent joint allelic disequilibrium will be generated if and only if nuclear and chloroplast frequencies both differ in the migrant pools ($\bar{P} \neq \bar{p}$ and $\bar{X}_C \neq \bar{x}_C$), while joint genotypic disequilibria can also be generated directly by unequal chloroplast frequencies in the migrant pools ($\bar{X}_C \neq \bar{x}_C$).

DISCUSSION

Unidirectional pollen migration and seed migration have distinctive effects on the dicytonuclear structure of plant populations in which mitochondria are inherited maternally and chloroplasts paternally, as appears true in many conifer species. We have provided an extensive analysis of such systems by a direct extension of previous pollen and seed dispersal models for two-locus cytonuclear systems with either maternal or paternal cytoplasmic inheritance (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992). Explicit analytic formulas are given for the effects of population admixture (via seeds or individuals) upon cytoplasmic and cytonuclear disequilibria, as well as for the full dicytonuclear equilibrium structure for selectively neutral markers in random mating, mixed mating, and purely selfing populations receiving gene flow through pollen and seeds. We are particularly concerned with the extent to which nonrandom associations among the three loci can be used as indicators of pollen and seed flow.

Permanent two- and three-locus disequilibria are generated in three main ways in this system. First, nonrandom associations are generated in the resident population if there are nonzero disequilibria in the migrant pools, especially in migrant seeds. Because of this, seed migration is of special importance for the generation of permanent disequilibria. Pollen migration alone can only generate permanent nuclear-chloroplast and joint three-locus disequilibria, and only if there are nonrandom associations between nuclear and chloroplast alleles in migrant pollen ($\bar{D}_{A/C} \neq 0$). The critical disequilibria can arise in several ways in either migrant pool; perhaps the most common way (and also the one of greatest practical importance) is where migrants are derived from multiple source populations with different genetic makeups, such as often happens in hybrid zones and other zones of admixture.

The second general cause of permanent nonrandom associations among the three loci is through intermigrant admixture effects due to differing cytonuclear compositions in the two migrant pools (Schnabel and Asmussen 1992). For example, joint differences in nuclear and chloroplast allele frequencies between migrant pollen and seeds ($\bar{P} \neq \bar{p}$ and $\bar{X}_C \neq \bar{x}_C$) generate allelic and genotypic disequilibria in the nuclear-chloroplast subsystem as well as joint allelic and joint genotypic disequilibria in the three-locus system. Such allele frequency differences are expected when migrant pollen and seeds are derived from different, genetically distinct source populations, when migrant pollen and seeds are derived from the same (multiple) sources whose relative contributions vary with the type of gene flow, or when selection or other evolutionary forces act in the life cycle of the source population(s), leading to genetic differentiation of the adult and seed stages. The generation of disequilibria in these ways obviously requires the presence of both pollen and seed migration and cannot be generated in the absence of either. Moreover, such simple intermigrant frequency differences cannot produce permanent nuclear-mitochondrial, cytoplasmic, or three-way disequilibria. These nonrandom associations involve the mitochondrial locus separately from the chloroplast locus and not as a joint cytotype; no analogous intermigrant interactions exist that could generate such associations since the mitochondrial marker is found in only one of the migrant pools.

The third and final general source of permanent disequilibria comes from the complex intermigrant effects that generate joint and three-way three-locus disequilibria. These new interactions involve all three genomes and require either allelic disequilibria in both migrant pools or differences in allele frequencies in the two migrant pools plus nonrandom associations in migrant seeds. For example, three-way genotypic associations can be generated by nuclear-mitochondrial allelic disequilibrium in migrant seeds in conjunction with nuclear-chloroplast allelic disequilibrium in migrant pollen ($\bar{d}_{A/M} \neq 0$ and $\bar{D}_{A/C} \neq 0$), but not by either alone, while both joint and three-way allelic disequilibria can be generated via intermigrant nuclear allele frequency differences coupled with cytoplasmic disequilibria in migrant seeds ($\bar{P} \neq \bar{p}$ and $\bar{d}_{M/C} \neq 0$). Yet another new intermigrant factor enters into the formation of three-way disequilibria; this involves an interaction between intermigrant allele frequency differences for the chloroplast marker and two-locus nuclear-mitochondrial disequilibria ($\bar{X}_C \neq \bar{x}_C$ and $\bar{d}_{N/M} \neq 0$, where $N = A, AA, Aa,$ or aa).

Even when no permanent disequilibria are generated, transient disequilibria can be produced by admixture between the migrants and the resident population. Such transient associations can reach quite high values and persist for relatively large numbers of generations. In long-lived organisms such as trees, generation times can be quite long, indicating that measurable nonrandom associations may be found even in situations where permanent disequilibria are not possible.

The magnitude of disequilibria formed by gene flow (and thus the ability to detect such nonrandom associations in natural populations) depends on whether populations are censused prior to or after mating and segregation can break up associations formed by admixture. We have focused here on censusing adults; in general, assaying three markers from adult tissues will be easier than doing so from seeds, especially for species whose seeds are small. However, the model presented here can be readily extended to populations of species such as conifers, which are easily censused at the seed stage (Conkle 1971), following the same approach used in the prior two-locus models (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992). The equilibrium frequencies and disequilibria in seeds are readily obtained by substituting the adult equilibria into the right-hand sides of the interim seed values given in appendix a. For either adult or seed census data, the statistical significance of the joint disequilibria can be determined using the procedures for multiallelic markers developed by Basten and Asmussen (1997), treating the four joint cytotypes as four alleles at a single cytoplasmic marker. The three-way disequilibria require the development of new statistical theory, which will be presented elsewhere.

Three-locus nuclear-dicytoplasmic data from plants with both paternally and maternally inherited organelles represent a unique opportunity to more accurately estimate both the absolute and relative rates of pollen- and seed-mediated gene flow. Analysis of joint genotype frequencies and nonrandom associations among all three loci can allow a better understanding of the relative roles that the mating system and gene flow play in shaping the genetic structure of such populations than has previously been possible. The dicytonuclear migration model presented here, along with previous work on plant cytonuclear systems (Schnabel and Asmussen 1989, 1992; Asmussen and Schnabel 1991), completes the necessary background for understanding the effects of asymmetrical gene flow upon the cytonuclear and dicytonuclear structure of plant populations. Together, these results provide the theoretical foundation for a new method for estimating pollen and seed gene flow from such joint cytonuclear and dicytonuclear data, which is presented in a companion article (Orive and Asmussen 2000).

We thank Chris Babcock, Mike Goodisman, Deb Overath, María Sánchez, Janet Hatt, and two anonymous reviewers for their helpful comments and suggestions on this work, and Sally Otto for her notational advice. Part of this work was completed while M.E.O. was a fellow at the Institute of Cell, Animal and Population Biology, University of Edinburgh, and we thank Nick Barton and his lab for their assistance. This research was supported in part by National Institutes of Health grant GM-48528 and National Science Foundation grant DEB 92-10895 to M.A.A. and by a National Science Foundation NATO postdoctoral fellowship and a University of Kansas New Faculty General Research Fund grant to M.E.O.

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., and C. J. Basten, 1996 Constraints and normalized measures for cytonuclear disequilibria. Heredity **76:** 207–214.

Asmussen, M. A., and A. Schnabel, 1991 Comparative effects of pollen and seed migration on the cytonuclear structure of plant populations. I. Maternal cytoplasmic inheritance. Genetics **128:** 639–654.

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.

Babcock, C. S., and M. A. Asmussen, 1996 Effects of differential selection in the sexes on cytonuclear polymorphism and disequilibria. Genetics **146:** 839–853.

Babcock, C. S., and M. A. Asmussen, 1998 Effects of differential selection in the sexes on cytonuclear dynamics: life stages with sex differences. Genetics **149:** 2063–2077.

Basten, C. J., and M. A. Asmussen, 1997 The exact test for cytonuclear disequilibria. Genetics **146:** 1165–1171.

Bennett, J. H., 1954 On the theory of random mating. Ann. Eugen. **18:** 311–317.

Birky, C. W., 1988 Evolution and variation in plant chloroplast and mitochondrial genomes, pp. 23–53 in *Plant Evolutionary Biology*, edited by L. D. Gottlieb and S. K. Jain. Chapman & Hall, London.

Clegg, M. T., 1980 Measuring plant mating systems. Bioscience **30:** 814–818.

Conkle, M. T., 1971 Isozyme specificity during germination and early growth of knobcone pine. For. Sci. **17:** 494–498.

Goodisman, M. A. D., and M. A. Asmussen, 1997 Cytonuclear theory for haplodiploid species and X-linked genes. I. Hardy-Weinberg dynamics and continent-island, hybrid zone models. Genetics **147:** 321–338.

Harrison, R. G., and J. J. Doyle, 1990 Redwoods break the rules. Nature **344:** 295–296.

Kirk, J. T. O., and R. A. E. Tilney-Bassett, 1978 *The Plastids: Their Chemistry, Structure, Growth and Inheritance*. Elsevier/North-Holland, Amsterdam.

Medgyesy, P., A. Pay and L. Marton, 1986 Transmission of paternal chloroplasts in *Nicotiana*. Mol. Gen. Genet. **204:** 195–198.

Mogensen, H. L., 1996 The hows and whys of cytoplasmic inheritance in seed plants. Am. J. Bot. **83:** 383–404.

Neale, D. B., and R. R. Sederoff, 1989 Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. Theor. Appl. Genet. **77:** 212–216.

Neale, D. B., N. C. Wheeler and R. W. Allard, 1986 Paternal inheritance of chloroplast DNA in Douglas-fir. Can. J. For. Res. **16:** 1152–1154.

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 using joint nuclear-cytoplasmic data. Genetics **155:** 833–854.

Schnabel, A., and M. A. Asmussen, 1989 Definition and properties of disequilibria within nuclear-mitochondrial-chloroplast and other nuclear-dicytoplasmic systems. Genetics **123:** 199–215.

Schnabel, A., and M. A. Asmussen, 1992 Comparative effects of pollen and seed migration on the cytonuclear structure of plant populations. II. Paternal cytoplasmic inheritance. Genetics **132:** 253–267.

Wagner, D. B., G. R. Furnier, M. A. Saghai-Maroof, S. M. Williams, B. P. Dancik *et al.*, 1987 Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc. Natl. Acad. Sci. USA **84:** 2097–2100.

Wagner, D. B., J. Dong, M. R. Carlson and A. D. Yanchuk, 1991 Paternal leakage of mitochondrial DNA in *Pinus*. Theor. Appl. Genet. **82:** 510–514.

Weir, B. S., 1996 *Genetic Data Analysis II*. Sinauer, Sunderland, MA.

Communicating editor: A. G. Clark

APPENDIX A: INTERIM SEED VALUES

Marginal frequencies and disequilibria in cytonuclear subsystems: The interim nuclear and cytotype frequencies in seeds are

$$v = s(\tfrac{1}{2}V) + (1 - s)[M(P\overline{Q} + Q\overline{P}) + 2(1 - M)PQ] \tag{A1}$$

$$p = \tfrac{1}{2}M(1 - s)\overline{P} + [1 - \tfrac{1}{2}M(1 - s)P] \tag{A2}$$

$$x_M = X_M \tag{A3}$$

$$x_C = M(1 - s)\overline{X}_C + [1 - M(1 - s)]X_C \tag{A4}$$

with $u = p - \tfrac{1}{2}v$ and $w = 1 - p - \tfrac{1}{2}v$. The interim two-locus disequilibria in the nuclear-mitochondrial subsystem are

$$d_{A/M} = \tfrac{1}{2}(1 + s)D_{A/M} \tag{A5}$$

$$\begin{aligned} d_{AA/M} &= \tfrac{1}{2}sD_{AA/M} + \{\tfrac{1}{2}s + (1 - s)[M\overline{P} + (1 - M)P]\}D_{A/M} \\ d_{Aa/M} &= \tfrac{1}{2}sD_{Aa/M} + (1 - s)\{M(\overline{Q} - \overline{P}) + (1 - M)(Q - P)\}D_{A/M}, \end{aligned} \tag{A6}$$

and those in the nuclear-chloroplast subsystem are

$$\begin{aligned} d_{A/C} = {} & \tfrac{1}{2}[1 + s - M(1 - s)]D_{A/C} + \tfrac{1}{2}M(1 - s)\overline{D}_{A/C} \\ & + \tfrac{1}{2}M(1 - s)[1 - M(1 - s)](P - \overline{P})(X_C - \overline{X}_C) \end{aligned} \tag{A7}$$

$$\begin{aligned} d_{AA/C} = {} & \tfrac{1}{2}sD_{AA/C} + [\tfrac{1}{2}s + (1 - M)(1 - s)P]D_{A/C} + M(1 - s)P\overline{D}_{A/C} \\ & + M(1 - s)(u - P\overline{P})(X_C - \overline{X}_C) \\ d_{Aa/C} = {} & \tfrac{1}{2}sD_{Aa/C} + (1 - s)(Q - P)[M\overline{D}_{A/C} + (1 - M)D_{A/C}] \\ & + M(1 - s)(v - P\overline{Q} - Q\overline{P})(X_C - \overline{X}_C). \end{aligned} \tag{A8}$$

Three-locus disequilibria: The interim three-locus joint allelic disequilibrium in seeds is

$$\begin{aligned} d_{A/MC} = {} & sD_{A/MC} + \tfrac{1}{2}Ms(1 - s)(P - \overline{P})D_{M/C} \\ & + \tfrac{1}{2}(1 - s)[M(X_M\overline{D}_{A/C} + \overline{X}_C D_{A/M}) \\ & \qquad + (1 - M)(X_M D_{A/C} + X_C D_{A/M})] \\ & + \tfrac{1}{2}M(1 - s)[1 - M(1 - s)](P - \overline{P})X_M(X_C - \overline{X}_C), \end{aligned} \tag{A9}$$

and the corresponding joint genotypic associations are

$$\begin{aligned} d_{AA/MC} = {} & s(D_{AA/MC} + \tfrac{1}{4}D_{Aa/MC}) \\ & + s(1 - s)[P - \tfrac{1}{4}V - M\overline{P}P - (1 - M)P^2]D_{M/C} \\ & + M(1 - s)(\overline{PX}_C D_{A/M} + PX_M\overline{D}_{A/C} + D_{A/M}\overline{D}_{A/C}) \\ & + (1 - M)(1 - s)[P(X_C D_{A/M} + X_M D_{A/C}) + D_{A/M}D_{A/C}] \\ & + M(1 - s)[s(P - \tfrac{1}{4}V - P\overline{P}) + (1 - s)(1 - M)P(P - \overline{P})] \\ & \times X_M(X_C - \overline{X}_C) \\ d_{Aa/MC} = {} & \tfrac{1}{2}sD_{Aa/MC} + s(1 - s)[\tfrac{1}{2}V - M(P\overline{Q} + Q\overline{P}) \\ & \qquad - 2(1 - M)PQ]D_{M/C} \\ & + M(1 - s)[(Q - P)X_M\overline{D}_{A/C} + (\overline{Q} - \overline{P})\overline{X}_C D_{A/M} - 2D_{A/M}\overline{D}_{A/C}] \\ & + (1 - M)(1 - s)[(Q - P)(X_C D_{A/M} + X_M D_{A/C}) \\ & \qquad - 2D_{A/M}D_{A/C}] \\ & + M(1 - s)[s(\tfrac{1}{2}V - P\overline{Q} - Q\overline{P}) + (1 - s)(1 - M) \\ & \qquad \times (2PQ - P\overline{Q} - Q\overline{P})]X_M(X_C - \overline{X}_C). \end{aligned} \tag{A10}$$

Finally, we can write the interim three-way, three-locus allelic disequilibrium in seeds as

$$\begin{aligned} d_{A/M/C} = {} & sD_{A/M/C} + \tfrac{1}{2}Ms(1 - s)(P - \overline{P})D_{M/C} \\ & + \tfrac{1}{2}Ms(1 - s)(X_C - \overline{X}_C)D_{A/M} \end{aligned} \tag{A11}$$

and the three-way genotypic disequilibria as

$$\begin{aligned} d_{AA/M/C} = {} & s(D_{AA/M/C} + \tfrac{1}{4}D_{Aa/M/C}) + (1 - s)D_{A/M}[M\overline{D}_{A/C} + (1 - M)D_{A/C}] \\ & + s(1 - s)[P - \tfrac{1}{4}V - MP\overline{P} - (1 - M)P^2]D_{M/C} \\ & + M(1 - s)\{\tfrac{1}{2}s - [1 - M(1 - s)]\overline{P} + (1 - M)(1 - s)P\} \\ & \times (X_C - \overline{X}_C)D_{A/M} + \tfrac{1}{2}Ms(1 - s)(X_C - \overline{X}_C)D_{AA/M} \\ d_{Aa/M/C} = {} & \tfrac{1}{2}sD_{Aa/M/C} - 2(1 - s)D_{A/M}[M\overline{D}_{A/C} + (1 - M)D_{A/C}] \\ & + s(1 - s)[\tfrac{1}{2}V - M(P\overline{Q} + Q\overline{P}) - 2(1 - M)PQ]D_{M/C} \\ & + M(1 - s)\{(1 - M)(1 - s)(Q - P) \\ & \qquad - [1 - M(1 - s)](\overline{Q} - \overline{P})\}(X_C - \overline{X}_C)D_{A/M} \\ & + \tfrac{1}{2}Ms(1 - s)(X_C - \overline{X}_C)D_{Aa/M}. \end{aligned} \tag{A12}$$

If there is no pollen migration ($M = 0$), equations (A9–A12) simplify to those in Schnabel and Asmussen (1989) for $D'_{A/MC}$, $D'_{AA/MC}$, $D'_{Aa/MC}$, $D'_{A/M/C}$, $D'_{AA/M/C}$, and

$D'_{Aa/M/C}$ (where $'$ indicates values after mating) for an isolated mixed-mating population with no gene flow, as would be expected.

APPENDIX B: EQUILIBRIA FOR THE TWO-LOCUS CYTONUCLEAR DISEQUILIBRIA

The final nuclear-mitochondrial disequilibria are

$$\hat{D}_{A/M} = \frac{2m\overline{d}_{A/M}}{2 - (1 - m)(1 + s)} \tag{B1}$$

$$\hat{D}_{AA/M} = \frac{2m\overline{d}_{AA/M} + (1 - m)\{s + 2(1 - s)[M\overline{P} + (1 - M)\hat{P}]\}\hat{D}_{A/M}}{2 - (1 - m)s}$$

$$\hat{D}_{Aa/M} = \frac{2m\overline{d}_{Aa/M} + 2(1 - m)(1 - s)[M(\overline{Q} - \overline{P}) + (1 - M)(\hat{Q} - \hat{P})]\hat{D}_{A/M}}{2 - (1 - m)s} \tag{B2}$$

and correspond to $\hat{D}$, $\hat{D}_1$, and $\hat{D}_2$ for the case of strictly maternal cytoplasmic inheritance (Asmussen and Schnabel 1991; Schnabel and Asmussen 1992). The final nuclear-chloroplast disequilibria

$$\hat{D}_{A/C} = \frac{2m\overline{d}_{A/C} + (1 - m)M(1 - s)\overline{D}_{A/C} + K_1(\overline{P} - \overline{p})(\overline{X}_C - \overline{x}_C)}{2 - (1 - m)[1 + s - M(1 - s)]} \tag{B3}$$

$$\begin{aligned} K_2\hat{D}_{AA/C} = {} & 2m\overline{d}_{AA/C} + (1 - m)[s + 2(1 - M)(1 - s)\hat{P}]\hat{D}_{A/C} \\ & + 2(1 - m)M(1 - s)\hat{P}\overline{D}_{A/C} \\ & + \frac{2m(1 - m)M(1 - s)}{m + (1 - m)M(1 - s)}(\hat{P}\overline{P} - \overline{u})(\overline{X}_C - \overline{x}_C) \\ K_2\hat{D}_{Aa/C} = {} & 2m\overline{d}_{Aa/C} + 2(1 - m)(1 - s)(\hat{Q} - \hat{P})[M\overline{D}_{A/C} + (1 - M)\hat{D}_{A/C}] \\ & + \frac{2m(1 - m)M(1 - s)}{m + (1 - m)M(1 - s)}(\hat{P}\overline{Q} + \hat{Q}\overline{P} - \overline{v})(\overline{X}_C - \overline{x}_C) \end{aligned} \tag{B4}$$

similarly correspond to $\hat{D}$, $\hat{D}_1$, and $\hat{D}_2$ for the case of strictly paternal cytoplasmic inheritance (Schnabel and Asmussen 1992), where

$$K_1 = \frac{2m(1 - m)M(1 - s)}{2m + (1 - m)M(1 - s)} \quad \text{and} \quad K_2 = 2 - (1 - m)s.$$

APPENDIX C: EQUILIBRIA FOR THE THREE-LOCUS DISEQUILIBRIA

We give here the details of the full forms for the final three-locus disequilibria. These are given in terms of the six constant factors,

$$\begin{aligned} k_1 &= 1 - (1 - m)s \\ k_2 &= 2 - (1 - m)(1 + s) \\ k_3 &= m + (1 - m)M(1 - s) \\ k_4 &= 2 - (1 - m)[1 + s - M(1 - s)] \\ k_5 &= 2m + (1 - m)M(1 - s) \\ k_6 &= 2 - (1 - m)s, \end{aligned}$$

which are functions of the migration and mating system parameters (m, M, and s). Using these, the constant coefficients of the various terms for the three-locus joint allelic disequilibrium, $\hat{D}_{A/MC}$, in (33) are

$$\begin{aligned} c_1 &= \frac{m}{k_1} \\ c_2 &= \frac{m(1 - m)(1 - s)\{m(1 - M)\overline{x}_C + M[1 - (1 - m)s]\overline{X}_C\}}{k_1k_2k_3} \\ c_3 &= \frac{m(1 - m)(1 - M)(1 - s)\overline{x}_M}{k_1k_4} \\ c_4 &= \frac{(1 - m)M(1 - s)\overline{x}_M}{k_4} \\ c_5 &= -\frac{m(1 - m)M(1 - s)}{k_1k_5} \\ c_6 &= \frac{2m(1 - m)M(1 - s)\overline{x}_M}{k_4k_5}. \end{aligned} \tag{C1}$$

For the homozygote three-locus joint genotypic disequilibrium, $\hat{D}_{AA/MC}$, the constants in (34a) are

$$\begin{aligned} c_1 &= \frac{m}{k_1} \\ c_2 &= \frac{(1 - m)s}{4k_1} \\ c_3 &= \frac{2m(1 - m)(1 - s)[M\overline{P}\,\overline{X}_C + (1 - M)\hat{P}\hat{X}_C]}{k_1k_2} \\ c_4 &= \frac{2m(1 - m)(1 - M)(1 - s)\hat{P}\overline{x}_M}{k_1k_4} \\ c_5 &= \frac{2(1 - m)M(1 - s)\hat{P}\overline{x}_M}{k_4} \\ c_6 &= \frac{4m^2(1 - m)(1 - M)(1 - s)}{k_1k_2k_4} \\ c_7 &= \frac{4m(1 - m)M(1 - s)}{k_2k_4} \\ c_8 &= \frac{m(1 - m)(1 - s)[\overline{u} - M\hat{P}\overline{P} - (1 - M)\hat{P}^2]}{(k_1)^2} \\ c_9 &= \frac{m(1 - m)M(1 - s)(\hat{P}\overline{P} - \overline{u})\overline{x}_M}{k_1k_3} \\ c_{10} &= \frac{2m(1 - m)^2M(1 - M)(1 - s)^2\hat{P}\overline{x}_M}{k_1k_4k_5} \\ c_{11} &= \frac{4m^2(1 - m)^2M(1 - M)(1 - s)^2}{k_1k_2k_4k_5}. \end{aligned} \tag{C2}$$

For the heterozygote three-locus joint genotypic disequilibrium, $\hat{D}_{Aa/MC}$, the constants in (34b) are

$$
\begin{aligned}
c_1 &= \frac{2m}{k_6}\\
c_2 &= \frac{4m(1-m)(1-s)[M(\overline{Q}-\overline{P})\overline{\mathrm{X}}_C + (1-M)(\hat{Q}-\hat{P})\hat{X}_C]}{k_2k_6}\\
c_3 &= \frac{4m(1-m)(1-M)(1-s)(\hat{Q}-\hat{P})\overline{x}_M}{k_4k_6}\\
c_4 &= \frac{4(1-m)M(1-s)k_1(\hat{Q}-\hat{P})\overline{x}_M}{k_4k_6}\\
c_5 &= -\frac{16m^2(1-m)(1-M)(1-s)}{k_2k_4k_6}\\
c_6 &= -\frac{16m(1-m)M(1-s)k_1}{k_2k_4k_6}\\
c_7 &= \frac{2m(1-m)(1-s)[\overline{v} - MK - 2(1-M)\hat{P}\hat{Q}]}{k_1k_6}\\
c_8 &= \frac{2m(1-m)M(1-s)(K-\overline{v})\overline{x}_M}{k_3k_6}\\
c_9 &= \frac{4m(1-m)^2M(1-M)(1-s)^2(\hat{Q}-\hat{P})\overline{x}_M}{k_4k_5k_6}\\
c_{10} &= -\frac{16m^2(1-m)^2M(1-M)(1-s)^2}{k_2k_4k_5k_6},
\end{aligned}
\tag{C3}
$$

where $K = \hat{P}\overline{Q} + \hat{Q}\overline{P}$.

The constants multiplying the various terms for the three-locus, three-way allelic disequilibrium, $\hat{D}_{A/M/C}$, in (35) are

$$
\begin{aligned}
c_1 &= \frac{m}{k_1}\\
c_2 &= -\frac{m(1-m)M(1-s)}{k_1k_5}\\
c_3 &= -\frac{m(1-m)M(1-s)}{k_2k_3}.
\end{aligned}
\tag{C4}
$$

The constants for the three-locus, three-way homozygote disequilibrium, $\hat{D}_{AA/M/C}$, in (36a) are

$$
\begin{aligned}
c_1 &= \frac{m}{k_1}\\
c_2 &= \frac{(1-m)s}{4k_1}\\
c_3 &= \frac{m(1-m)(1-s)[\overline{u} - M\hat{P}\overline{P} - (1-M)\hat{P}^2]}{(k_1)^2}\\
c_4 &= \frac{4m^2(1-m)(1-M)(1-s)}{k_1k_2k_4}\\
c_5 &= \frac{4m(1-m)M(1-s)}{k_2k_4}\\
c_6 &= \frac{2m^2(1-m)M(1-s)\overline{P}}{k_1k_2k_3}\\
c_7 &= \frac{4m^2(1-m)^2M(1-M)(1-s)^2}{k_1k_2k_4k_5}\\
c_8 &= -\frac{m(1-m)M(1-s)}{k_1k_3}.
\end{aligned}
\tag{C5}
$$

Finally, the constants for the three-locus, three-way heterozygote disequilibrium, $\hat{D}_{Aa/M/C}$, in (36b) are

$$
\begin{aligned}
c_1 &= \frac{2m}{k_6}\\
c_2 &= \frac{2m(1-m)(1-s)[\overline{v} - MK - 2(1-M)\hat{P}\hat{Q}]}{k_1k_6}\\
c_3 &= -\frac{16m^2(1-m)(1-M)(1-s)}{k_2k_4k_6}\\
c_4 &= -\frac{16m(1-m)M(1-s)k_1}{k_2k_4k_6}\\
c_5 &= \frac{4m^2(1-m)M(1-s)(\overline{Q}-\overline{P})}{k_2k_3k_6}\\
c_6 &= -\frac{16m^2(1-m)^2M(1-M)(1-s)^2}{k_2k_4k_5k_6}\\
c_7 &= -\frac{2m(1-m)M(1-s)}{k_3k_6}.
\end{aligned}
\tag{C6}
$$