Seed and pollen flow and cline discordance among genes with different modes of inheritance

X-S Hu and B Li

Department of Forestry, North Carolina State University, PO Box 8002, Raleigh, NC 27695, USA

The relationships between seed and pollen flow and cline discordance/concordance between cytoplasmic and nuclear genes, with the incorporation of the effects of natural selection, are formulated for one locus with two alleles, under assumptions of random mating, no drift and no mutation. Results show that under certain conditions, the relative roles of seed and pollen flow in shaping cline discordance/concordance are very similar to their roles in influencing population differentiation for selectively neutral markers with different modes of inheritance. Where the disequilibria between cytoplasmic and nuclear genes are of the order similar to selection coefficient, cline discordance/concordance can be predicted from the relative values of the ratio of pollen to seed flow and the ratio of selection coefficients. Where the disequilibria attained by seed and pollen flow are significant, the integrated cytonuclear data are recommended for cline analysis. In both cases, the relative rates of selection coefficients between cytoplasmic and nuclear genes can be roughly estimated according to their characteristic length.
Heredity (2002) **88,** 212–217. DOI: 10.1038/sj/hdy/6800030

Introduction

The effectiveness of seed and pollen flow as agents of the interpopulation gene flow depends upon the mode of gene inheritance (Petit *et al*, 1993; Ennos, 1994). For nuclear and paternally inherited organelle genes, gene flow occurs in both pollen and seed; while for maternally inherited organelle genes, gene flow is mediated by seed only. Consequently, different levels of population differentiation at drift-migration equilibrium are expected for genes differing in mode of inheritance under different models of population genetic structure (Petit *et al*, 1993; Ennos, 1994; Hu and Ennos, 1997, 1999; Hu, 2000a). The relative rates of pollen and seed flow can be estimated using Wright's F-statistics and other genetic statistics (Ennos, 1994; Hu and Ennos, 1997; Hu, 2000b). However, those theoretical results are only appropriate for selectively neutral genes.

One important type of genetic structure observed in natural populations is clinal variation, where gene frequency exhibits a gradient change (increase or decrease) with geographical distance. Natural selection is often involved in cline formation (see Endler, 1977). Cline discordance (or concordance) between cytoplasmic and nuclear genes, which can be generally defined as the frequency clines with different (or the same) slopes (the rate for the change of frequency with distance) or positions (the geographical location for the same allele frequency) in space, was observed in practice (eg, Young, 1996). In theory, the impacts of seed and pollen flow, with incor-

poration of the effect of natural selection, were considered recently in a clinal situation (Nagylaki, 1997). However, haploid genes with paternal and maternal inheritance are not included. The relative roles of seed and pollen flow in bringing about cline discordance/concordance between cytoplasmic and nuclear genes remain to be addressed. Furthermore, the disequilibria between cytoplasmic and nuclear genes, generated by seed and pollen flow, are likely maintained, as was shown in theory in a population structure model of continent-island (Asmussen and Schnabel, 1991; Schnabel and Asmussen, 1992). Li and Nei (1973) also showed that stable linkage disequilibrium between nuclear loci without epistasis can be maintained in subdivided populations under certain conditions. Slatkin (1975) further demonstrated, when the recombination fraction between nuclear loci is of the same order as selection coefficient or smaller, a substantial amount of linkage disequilibrium can be present in a cline.

Unlike the linkage disequilibria between nuclear loci, organelle genomes are physically unlinked to the nuclear genome and the recombination fraction is $\frac{1}{2}$, greater than the weak selection coefficient that is assumed to involve in maintaining cline formation. The pre-existing theoretical results are not appropriate for elucidating the impacts of seed and pollen flow on cline discordance/concordance between cytoplasmic and nuclear genes. The purpose of this article is to develop further theory required for understanding and interpreting clinal variation of nuclear, paternally and maternally inherited organelle genes, and to formulate the relationships between seed and pollen flow and cline discordance/concordance between cytoplasmic and nuclear genes, with incorporation of the disequilibria between cytoplasmic and nuclear genes attained by seed and pollen flow. For the sake of simple expression, the term

Correspondence: *Xin-Sheng Hu, Department of Forest Sciences, The University of British Columbia, 3rd Floor, 3041–2424 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada.*
E-mail: *xhu@interchange.ubc.ca*

'cytonuclear' is used when cytoplasm and nuclear genes are interacted and considered jointly. It is necessary to note in advance that the cline discordance/concordance analyzed in this article refers to the specific case, where the same spatial selection pattern takes place between cytoplasmic and nuclear genes and the selection boundary points are assumed at the same position. Three different combinations between cytoplasmic and nuclear genes will be considered separately: nuclear *vs* paternally inherited organelle genes; nuclear *vs* maternally inherited organelle genes; and nuclear *vs* paternally and maternally inherited organelle genes. By analyzing these combinations, we look at the implications for the behaviour of nuclear, paternally, and maternally inherited organelle genes, and explore the inferences that can be drawn about the roles of seed and pollen flow in cline discordance/concordance from the data on these differently inherited genes.

General assumptions

A single locus with two alleles is considered for each of the three modes of gene inheritance. Two alleles are denoted by A_1 and A_2 for biparentally inherited diploid nuclear genes, B_1 and B_2 for paternally inherited haploid organelle genes, and M_1 and M_2 for maternally inherited haploid organelle genes. A hermaphrodite plant species is distributed in an infinite chain of equally spaced populations. Population size is assumed so large that the influence of genetic drift can be ignored. Symmetrical migration rate is assumed between any two populations for either seed or pollen flow. The life cycle within a short time interval, Δt, follows a sequence of events; pollen flow, random mating, seed flow, and selection. The population distribution is assumed uniform after selection, and hence migration rate between any two populations represents the probability of migration from one population to another within the time, Δt. Density-independent selection in the offspring takes place independently in each population after seed flow.

Nuclear *vs* paternally inherited organelle genes

General case

Let $p(i,t)$ be the frequency of allele A_1 in population i at time t, and $q(i,t)$ for allele A_2 ($p(i,t) + q(i,t) = 1$), $u(i,t)$ and $v(i,t)$ ($u(i,t) + v(i,t) = 1$) be the frequencies of alleles B_1 and B_2, respectively. Define $D_1(i,t)$ as freq.$(A_1B_1) - p(i,t) q(i,t)$, the gametic (allelic) disequilibrium between A_1 and B_1 in population i. According to Asmussen *et al* (1987), the gametic disequilibria between any other paternally inherited organelle and nuclear alleles are summarized in Table 1. Use Wrightian fitnesses, $1 + s_1\Delta t g_1(i)$, $1 +$

Table 1 Gametic disequilibria between nuclear (A_1, A_2) and paternally inherited organelle (B_1, B_2) alleles in population i

	A_1	A_2	Total
B_1	$p(i,t)u(i,t) + D_1(i,t)$	$q(i,t)u(i,t) - D_1(i,t)$	$u(i,t)$
B_2	$p(i,t)v(i,t) - D_1(i,t)$	$q(i,t)v(i,t) + D_1(i,t)$	$v(i,t)$
Total	$p(i,t)$	$q(i,t)$	

$hs_1\Delta t g_1(i)$, and $1 - s_1\Delta t g_1(i)$ for genotypes, A_1A_1, A_1A_2, and A_2A_2, respectively, where h is the degree of dominance and $g_1(i)$ indicates the pattern for the change of selection coefficient with geographical distance (Nagylaki, 1975). Similarly, define $1 + s_2\Delta t g_2(i)/2$ and $1 - s_2\Delta t g_2(i)/2$ as the fitnesses for alleles B_1 and B_2, respectively. The multiplication viability model that omits the product form for selection coefficient (selection epistasis) is used so that the fitness for any combination of paternal and nuclear genotype can be simply calculated (Clark, 1984).

Scale the population location by $x = i\epsilon$, where ϵ is the spacing between colonies. Use the notations of $\dot{f} = \partial f/\partial t$, $f' = \partial f/\partial x$, and $f'' = \partial^2 f/\partial x^2$ for function f. Following the sequence of events mentioned in the general assumptions, we can derive by diffusion approximation:

$$\dot{p} = \frac{\sigma_1^2}{2} p'' + pq(1 + h(q - p))g_1(x)s_1 + D_1g_2(x)s_2/2 , \quad (1)$$

$$\dot{u} = \frac{\sigma_2^2}{2} u'' + (1 + h(q - p))g_1(x)D_1s_1 + uvg_2(x)s_2 , \quad (2)$$

where p, q, u, v and D_1 refer to $p(x,t)$, $q(x,t)$, $u(x,t)$, $v(x,t)$ and $D_1(x,t)$, respectively, and $\sigma_1^2 = \sigma_S^2 + \sigma_P^2/2$, $\sigma_2^2 = \sigma_S^2 + \sigma_P^2$. σ_S^2 and σ_P^2 are defined respectively as the dispersal variances of seed and pollen.

The change for the gametic disequilibrium is given by:

$$\dot{D}_1 = -\frac{1}{2} D_1 + \frac{1}{2}\frac{\sigma_2^2}{2} D_1'' + \sigma_1^2 p'u'$$
$$+ s_1g_1(x)((q - p) (1 + (q - p)h) - 2pqh)D_1/2$$
$$+ s_2g_2(x)(v - u)D_1/2 . \quad (3)$$

where D_1 refers to $D_1(x,t)$. Equation (3) is comparable with (7) of Kruuk *et al* (1999). It can be viewed from (3) that there are three components underlying the change of D_1. The first component is induced by the system of random mating that reduces D_1 in half per time interval, Δt. The second component is the change due to migration, and the third is associated with selection.

Letting $h = 0$ and considering that the second term on right side of (3) is negligible (see also Kruuk *et al*, 1999), we can simply obtain

$$D_1 \approx 2\sigma_1^2 (1 + s_1g_1(x)(q - p) + s_2g_2(x)(v - u)) p'u' \quad (4)$$

If $2\sigma_1^2p'u' \leq s_1$ or $2\sigma_1^2p'u' \leq s_2$, the case may likely occur when clines are very smooth. D_1 is then of the order similar to selection coefficient and D_1 terms in (1) and (2) can be ignored. If $2\sigma_1^2p'u' > s_1$ and $2\sigma_1^2p'u' > s_2$, which may take place when clines are very steep, then D_1 is of the order $2\sigma_1^2p'u'$ and its effect is likely significant. In this case, it is not reasonable to remove D_1 terms from (1) and (2).

D_1 with the order similar to selection coefficient

In this case, the clines for nuclear and paternally inherited organelle genes can be approached independently, and analytic solutions for p and u at steady state can be solved separately using the method introduced by Haldane (1948). According to Slatkin's (1973) result, the characteristic length is defined as $w_1 = \sqrt{\sigma_1^2/s_1}$ for nuclear genes and $w_2 = \sqrt{\sigma_2^2/s_2}$ for paternally inherited organelle genes. Thus, the relationship between their characteristic lengths is obtained by:

$$w_1 = \sqrt{\frac{1}{2}\left(1 + \frac{1}{1 + r}\right) \cdot \frac{s_2}{s_1}} \, w_2 \quad (5)$$

where $r = \sigma_P^2/\sigma_S^2$, the ratio of pollen to seed flow. When, $s_1 = s_2$, a large value of the ratio r can enlarge cline discordance.

D_1 with significant effect

In this case, the combination between nuclear and paternally inherited organelle genes can be considered jointly as a whole system. Assume that the same spatial selection pattern is possessed between nuclear and paternally inherited organelle genes, ie $g_1(x) = g_2(x) = g(x)$, where $g(x) = 1$ when $x < 0$, and $g(x) = -\beta^2$ when $x > 0$ (Haldane, 1948). Combining equations (1) and (2) at steady state and using (4), we can obtain a two dimensional nonlinear differential equation,

$$\frac{\sigma_1^2}{2} p'' + \frac{\sigma_2^2}{2} u'' + 2\sigma_1^2 g(x)(s_2/2 + s_1)$$

$$p'u' + pqg(x)s_1 + uvg(x)s_2 = 0 \qquad (6)$$

When a complete concordance takes place, let $p' = u'$, $p = u$ and $\chi = (p')^2$, and then $p'' = \frac{1}{2} \partial \chi/\partial p$. We can derive:

$$(p')^2 = -\frac{c_2}{c_1} (p(1-p) - (q-p)/c_1 - 2/c_1^2$$
$$+ (1 + 2/c_1) \exp(-c_1 p)/c_1), \qquad (7a)$$

when $x > 0$, where $c_1 = -4\beta^2(2s_1 + s_2)\sigma_1^2/(\sigma_1^2 + \sigma_2^2)$ and $c_2 = -4\beta^2(s_1 + s_2)/(\sigma_1^2 + \sigma_2^2)$;

$$(p')^2 = -\frac{c_2}{c_1} (p(1-p) - (q-p)/c_1 - 2/c_1^2$$
$$+ (1 + 2/c_1) \exp(c_1 q)/c_1), \qquad (7b)$$

when $x < 0$, where $c_1 = 4(2s_1 + s_2)\sigma_1^2/(\sigma_1^2 + \sigma_2^2)$ and $c_2 = 4(s_1 + s_2)/(\sigma_1^2 + \sigma_2^2)$. The characteristic length for the cytonuclear genes as a whole system under concordance, denoted by w_{12D}, is thus equal to $\sqrt{c_1/c_2} = \sigma_1/\sqrt{1 - s_1/(2s_1 + s_2)}$. It can be viewed that $w_{12D} < w_1$ and $w_{12D} < w_2$. With the same order of selection coefficient between cytoplasmic and nuclear genes, the cline for the whole system under concordance is distributed within the range of the clines for its two components that evolve independently (Figure 1).
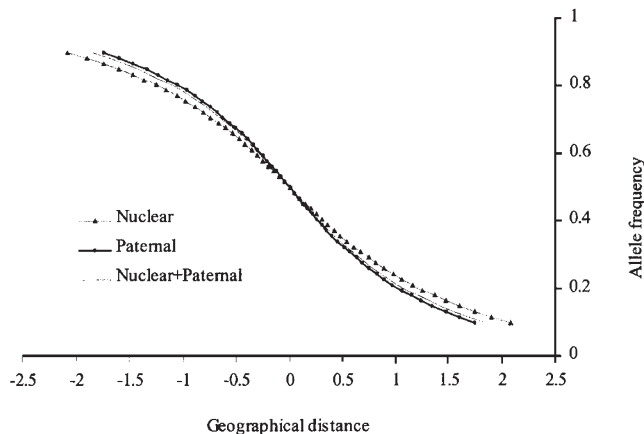

Figure 1 The cline for the combination between nuclear and paternal organelle genes under concordance is compared with those of its two components that evolve independently in a given frequency interval [0.1, 0.9]. Parameter settings are $h = 0.0$, $\beta^2 = 1.0$, $\sigma_P^2 = 0.07$, $\sigma_S^2 = 0.05$, $s_1 = 0.03$, $s_2 = 0.06$.

Assume that $g(0)$ exists. According to Slatkin (1973), the characteristic length under the impacts of cytonuclear disequilibria can also be calculated by $w_{1D} = (p'|_{p=0.5})^{-1}$ and w_{2D} $(u'|_{u=0.5})^{-1}$. Thus, the value of D_1 at cline centre (maximum point) is:

$$D_1(x = 0) = \frac{2\sigma_1^2}{1 + hs_1g_1(0)/2} \cdot \frac{1}{w_{1D}w_{2D}}$$

$$= \frac{2\sigma_1^2}{1 + hs_1g(0)/2} \cdot \frac{1}{w_{12D}^2}, \text{ when } w_{1D} = w_{2D} = w_{12D}. \qquad (8)$$

If $h = 0$, (8) reduces to the result obtained by Barton and Gale (1993).

Barton (personal communication) indicated that a small linear perturbation analysis can be used as an approach to cytonuclear combination when D_1 is significant. Let $\lambda = (p + u)/2$, $s = (s_1 + s_2)/2$, and $\sigma^2 = (\sigma_1^2 + \sigma_2^2)/2$. We can approximate an average equation even if there are some differences between cytonuclear genes, $\dot{\lambda} = \frac{\sigma^2}{2} \lambda'' + \lambda(1 - \lambda)s + 3sD/2$. Firstly, consider the case without the D effect, ie $\dot{\lambda} = \frac{\sigma^2}{2} \lambda'' + \lambda(1 - \lambda)s$. At steady state, it becomes

$$\dot{\lambda}_\epsilon = \frac{\sigma^2}{2} \lambda_\epsilon'' + \lambda_\epsilon (1 - \lambda_\epsilon)s = 0 \qquad (9)$$

λ_ϵ' can be solved analytically. Next consider the impacts of small linear perturbations, γ, and let $\lambda = \lambda_\epsilon + \gamma\lambda_\epsilon'$. Following the same method used by Barton (1979), we can obtain

$$\dot{\gamma} = \frac{\sigma^2}{2} \gamma'' + \sigma^2 \frac{\lambda_\epsilon''}{\lambda_\epsilon'} \gamma' - \frac{\lambda_\epsilon (1 - \lambda_\epsilon)}{\lambda_\epsilon'} s'\gamma \qquad (10)$$

where $s' = 0$ for either $x > 0$ or $x < 0$. Assume that D effect is equivalent to the influence by linear perturbation that eventually induces a non-zero steady flux into the system described by (9). Let $3sD/2 = \dot{\lambda} - \dot{\lambda}_\epsilon = \dot{\gamma}\lambda_\epsilon'$. Since $|D| \leq 1/4$ according to Asmussen and Basten (1996), then $|3sD/2| = |\dot{\gamma}\lambda_\epsilon'| \leq 3s/8$, which limits arbitrary choice of the value γ. An integral that is invariant during cline shift due to perturbation can be configured, $I = \int_{-\infty}^{+\infty} g\lambda_\epsilon'\gamma dx$. We obtain, $\dot{I} = \int_{-\infty}^{+\infty} (g\lambda_\epsilon')\dot{\gamma}dx = \int_{-\infty}^{+\infty} \gamma \left(\frac{\sigma^2}{2} (g\lambda_\epsilon')'' - \sigma^2(g\lambda_\epsilon'')' \right) dx$. I is invariant for all possible γ when $g = \lambda_\epsilon'$. Thus, the eventual movement of cline due to linkage disequilibria can be approached by $-(\int_{-\infty}^{+\infty} g\lambda_\epsilon'\gamma dx)/(\int_{-\infty}^{+\infty} g\lambda_\epsilon' dx)$ (Barton, 1979).

Nuclear *vs* maternally inherited organelle genes

Let $y(i,t)$ be the frequency of allele M_1, in population i at time t, and $z(i,t)$ $(y(i,t) + z(i,t) = 1)$ of allele M_2. Define $D_2(i,t)$ as freq.$(A_1M_1) - p(i,t) y(i,t)$, the gametic disequilibrium between A_1 and M_1 in population i. Similarly, denote the fitness by $1 + s_3\Delta tg_3(i)/2$ for M_1 and $1 - s_3\Delta tg_3(i)/2$ for M_2. The partial differential equations can be readily gained by replacing D_1, s_2, $g_2(x)$, σ_2^2, u and v in (1) and (2) with D_2, s_3, $g_3(x)$, σ_3^2, y and z, respectively. The relationship between their relative characteristic lengths is given by

$$w_1 = \sqrt{(1 + r/2)s_3/s_1}\, w_3. \qquad (11)$$

The change for D_2 is given by

$$\dot{D}_2 = -\frac{1}{2} D_2 + \frac{1}{2} \frac{\sigma_s^2}{2} \tag{12a}$$

$$D_2'' + \sigma_S^2 p'y' + s_1 g_1(x)((q-p)(1+(q-p)h) - 2pqh)D_2/2 + s_3 g_3(x)(z-y)D_2/2 .$$

The magnitude of D_2 at centre is

$$D_2(x=0) = \frac{2\sigma_S^2}{1 + h s_1 g_1(0)/2} \frac{1}{w_{1D} w_{3D}} \tag{12b}$$

where w_{3D} is the characteristic length for maternally inherited organelle genes. Under concordance, the solution for the first order p' can be obtained immediately by substituting $c_1 = 4g(x)(2s_1 + s_3)\sigma_s^2/(\sigma_1^2 + \sigma_s^2)$ and $c_2 = 4g(x)(s_1 + s_3)/(\sigma_1^2 + \sigma_s^2)$ into (7a, b). It can be viewed that $w_{13D} < w_1$ and $w_{13D} < w_3$, where $w_{13D} = \sqrt{c_1/c_2}$. With the same order of selection coefficient, the cline is wider for the whole cytonuclear system under concordance, than for maternally inherited organelle genes but narrower than for nuclear genes (Figure 2).

Paternally *vs* maternally inherited organelle genes

According to the general assumptions the two haploid organelle alleles B and M are shown to be independent. The relationship between their characteristic lengths is given by

$$w_2 = \sqrt{(1+r)s_3/s_2}\, w_3 . \tag{13}$$

Nuclear *vs* paternally and maternally inherited organelle genes

Similarly, define $D(i,t)$ as freq.$(A_1 B_1 M_1) - p(i,t) u(i,t) y(i,t)$, the gametic disequilibrium among alleles A_1, B_1, and M_1 in population i. Let $D(i,t) = D_1(i,t)y(i,t) + D_2(i,t)u(i,t)$. The gametic disequilibria for any other combinations among

Figure 2 The cline for the combination between nuclear and maternal organelle genes under concordance is compared with those of its two components that evolve independently in a given frequency interval [0.1, 0.9]. Parameter settings are $h = 0.0$, $\beta^2 = 1.0$, $\sigma_P^2 = 0.07$, $\sigma_S^2 = 0.05$, $s_1 = 0.03$, $s_3 = 0.06$.

three genomes are summarized in Table 2. The previous definitions for D_1 and D_2 still hold, and so do the partial differential equations for either paternally or maternally inherited organelle genes.

The partial differential equation for the frequency p is given by

$$\dot{p} = \frac{\sigma_1^2}{2} p'' + pq(1 + h(q-p))g_1(x)s_1 + D_1 g_2(x)s_2/2 + D_2 g_3(x)s_3/2 . \tag{14}$$

It has been shown that the partial differential equations for D_1 and D_2 remain unaltered; while the result for the change of $D(= D_{A_1 B_1 M_1})$ becomes more complicated,

$$\dot{D} = -D/2 + \frac{\sigma_S^2}{2} D'' + \sigma_S^2(pu'y' + up'y' + yp'u')$$

$$+ \frac{\sigma_p^2}{4}(D_2 u'' + y D_1'' + 2yp'u')$$

$$+ (D((q-p)(1+(q-p)h) - 2pqh)/2 \tag{15}$$

$$+ (1-h)D_1 D_2 s_1 g_1(x)$$

$$+ (uvD_2 + y(v-u)D_1)s_2 g_2(x)/2$$

$$+ (yzD_1 + u(z-y)D_2)s_3 g_3(x)/2$$

where D represents $D(x,t)$. It is further derived that the following relationship holds,

$$\dot{D}_{A_i B_j} = \dot{D}_{A_i B_j M_1} + \dot{D}_{A_i B_j M_2} \quad (i,j = 1,2) \tag{16a}$$

$$\dot{D}_{A_i M_j} = \dot{D}_{A_i B_1 M_j} + \dot{D}_{A_i B_2 M_j} \quad (i,j = 1,2) \tag{16b}$$

Thus, according to equations (15) and (16a,b) a partial differential equation for any other linkage disequilibrium among the three genomes can be readily worked out.

Under cline concordance among the three genomes, the solution for the first order of cline gradient can be obtained by substituting:

$$c_1 = \frac{4g(x)((2s_1 + s_2)\sigma_1^2 + (2s_1 + s_3)\sigma_s^2)}{\sigma_1^2 + \sigma_2^2 + \sigma_s^2}, \tag{17a}$$

$$c_2 = \frac{4g(x)(s_1 + s_2 + s_3)}{\sigma_1^2 + \sigma_2^2 + \sigma_s^2}, \tag{17b}$$

in (7a) when $x > 0$ and (7b) when $x < 0$. Define w_{123D} as the characteristic length for the cline of three genomes under concordance. It can be viewed that the relationship of $w_{123D}^2 < w_{12D}^2 + w_{13D}^2$ holds. The cline for the three genomes under concordance is distributed within the range of the clines for the two genomes under concordance, if the same order of selection coefficient is possessed for each genome (Figure 3).

Discussion

This paper has explicitly formulated the relationships between the seed and pollen flow and cline discordance/ concordance between cytoplasmic and nuclear genes. It can be concluded that the integrated effects of gene flow (seed and pollen flow) with selection shape cline discordance/ concordance. The effect of seed flow is not the same as that of pollen flow because of the asymmetric migration rates that are utilized by the three plant genomes with contrasting modes of inheritance. Previous studies show that for the selectively neutral markers, a large migration rate of seeds can reduce the level of
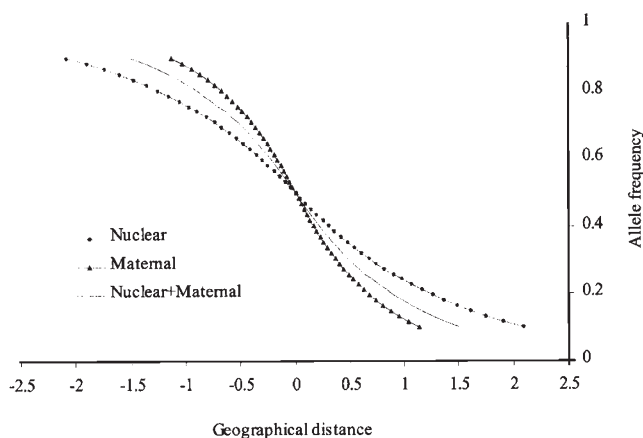
Table 2 Gametic disequilibria among nuclear (A_1, A_2) and paternally (B_1, B_2) and maternally inherited organelle (M_1, M_2) alleles in populataion i

| | A_1 | | A_2 | |
	B_1	B_2	B_1	B_2
M_1	$puy + D$	$pvy + D_2 - D$	$quy - D$	$qvy - D_2 + D$
M_2	$puz + D_1 - D$	$pvz - D_1 - D_2 - D$	$quz - D_1 + D$	$qvz + D_1 + D_2 - D$

*D, D_1, D_2, p, q, u, v, y, and z represent $D(i,t)$, $D_1(i,t)$, $D_2(i,t)$, $p(i,t)$, $q(i,t)$, $u(i,t)$, $v(i,t)$, $y(i,t)$, and $z(i,t)$, respectively. $D(i,t) = D_1(i,t)y(i,t) + D_2(i,t)u(i,t)$.
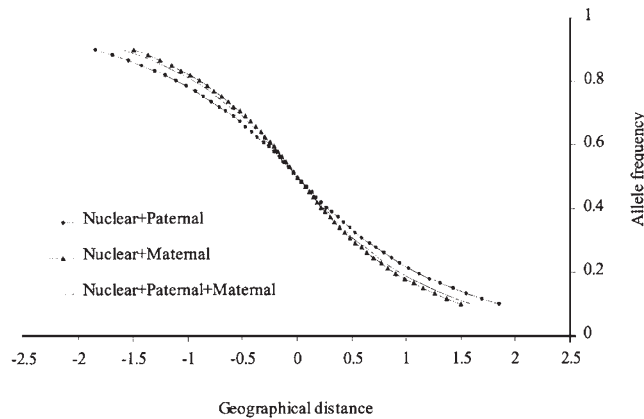

Figure 3 Under concordance the cline for the combination of three genomes is compared with those of two other combinations (nuclear vs paternal, nuclear vs maternal) in a given frequency interval [0.1, 0.9]. Parameter settings are $h = 0.0$, $\beta^2 = 1.0$, $\sigma_P^2 = 0.07$, $\sigma_S^2 = 0.05$, $s_1 = 0.03$, $s_2 = s_3 = 0.06$.

population differentiation among these three genomes, whereas a large migration of pollen grains can increase the divergence in population differentiation between maternally inherited organelle genes and the other two counterparts (Ennos, 1994; Hu and Ennos, 1997, 1999). The results similar to the neutral case are further proven in the clinal situation if weak selection coefficients are of the same order among the genomes. Large seed flow can reduce the clinal differential among these three genomes, and hence enhance the level of cline concordance. A large migration rate of pollen grains can increase cline concordance between nuclear and paternally inherited organelle genes, but reduce cline concordance between the maternally inherited organelle genes and the other two counterparts.

The cline concordance/discordance between cytoplasmic and nuclear genes should be different from that between linked nuclear genes. Since the cytoplasmic genomes are biologically not linked to the nuclear genome, the differential in gene introgression will be greater between cytoplasimic and nuclear genes than between many linked nuclear genes under comparable selection intensity. Moreover, compared with the selection acting on those linked nuclear genes, the selection acting on cytoplasmic and nuclear genes with disequilibria is usually weak (Barton and Hewitt, 1985, 1989). Thus, it is expected that cline discordance between cytoplasmic and nuclear genes occur more often than that between linked nuclear genes. Complete cline concordance between cytoplasimic and nuclear genes cannot occur frequently and

hence the concordance case generally reflects a limit situation.

The mechanism responsible for the discrepancy in the speed of gene introgression between cytoplasimic and nuclear genes is difficult to address inclusively. Of many factors, such as asymmetric mating and founder effects, the differentiation in dispersal of male and female parents is probably primary. Owing to the different vectors by which migration of cytoplasmic and nuclear genes is realized, the impact of the differentiation in gene dispersal should be substantial and genes with localized introgression will lag behind those with dispersed introgression. Several plant species display a large ratio of pollen to seed flow (Ennos, 1994; Ennos *et al*, 1999), implying that the speed of gene introgression is slower for maternally inherited organelle genes than for nuclear and paternally inherited organelle genes. Observations on the differential patterns in introgression between cytoplasmic and nuclear genes have been recorded in the literature (eg Harrison, 1989).

In addition to the asymmetric migration among nuclear and uniparentally inherited organelle genes, selection as another force to operate in cline formation should also be attentive. It can act against nuclear genes but not against cytoplasmic genes, or against cytoplasmic genes but not against nuclear genes. Several reports on roughly equivalent frequencies for nuclear and cytoplasmic gene flow indirectly indicate that diverse natural selection intensities can act on them (Rieseberg and Wendel, 1993). Those nuclear genes that affect reproductive isolation between hybridizing populations or taxa will much likely lag behind those organelle genes that do not seriously concern the survival of plants. Thus, the differential in gene introgression between cytoplasmic and nuclear genes is also attributable to diverse selection intensities.

It can be predicted in theory that the validity of equations (5) and (11) indirectly indicates no significant cytonuclear disequilibria involving the formation of cline discordance/concordance. Once all genetic marker data in clinal situations are available, the disequilibria between cytoplasmic and nuclear genes, in fact, can be tested using the methods introduced by Basten and Asmussen (1997). Characteristic length can also be estimated, as demonstrated in one typical dispersal-selection cline recently reported in *Castanea sativa* Mill (Villani *et al*, 1999). If the ratio of pollen to seed flow can be approximately estimated by using neutral markers (Ennos, 1994), equations (5) or (11) provides a convenient way to estimate the relative selection coefficients between cytoplasmic and nuclear genes. If the disequilibria between

cytoplasmic and nuclear genes are significant, use of the integrated cytonuclear data is recommended. Under this case, the characteristic length for combined data can be used to roughly estimate the relative contributions to selection, ie $s_1/(s_1 + s_2) = w_{12D}^2/\sigma_1^2 - 1$, or $s_1/(s_1 + s_3) = w_{3D}^2/\sigma_s^2 - 1$. Application of these techniques to real populations awaits accumulation of appropriate data collections.

Acknowledgements

References

Asmussen MA, Arnold J, Avise JC (1987). Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic and genotypes. *Genetics* **115**: 755–768.

Asmussen MA, Basten CJ (1996). Constraints and normalized measures for cytonuclear disequilibria. *Heredity* **76**: 207–214.

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

Barton NH (1979). The dynamics of hybrid zones. *Heredity* **43**: 341–359.

Barton NH, Gale KS (1993). Genetic analysis of hybrid zones. In: Harrison RG (ed) *Hybrid Zones and the Evolutionary Process*, Oxford University Press: New York, pp 13–45.

Barton NH, Hewitt GM (1985). Analysis of hybrid zones. *Ann Rev Ecol Syst* **16**: 113–148.

Barton NH, Hewitt GM (1989). Adaptation, speciation and hybrid zones. *Nature* **341**: 497–503.

Basten C, Asmussen MA (1997). The exact test for cytonuclear disequilibria. *Genetics* **146**: 1165–1171.

Clark AG (1984). Natural selection with nuclear and cytoplasmic transmission. I. A deterministic model. *Genetics* **107**: 679–701.

Endler JA (1977). *Geographic Variation, Speciation, and Clines*. Princeton University Press: Princeton.

Ennos RA (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**: 250–259.

Ennos RA, Sinclair WT, Hu XS, Langdon A (1999). Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: Hollingsworth PM, Bateman RM, Gornall RJ (eds) *Molecular Systematics and Plant Evolution*, Taylor & Francis: London, pp 1–19.

Haldane JBS (1948). The theory of a cline. *J Genet* **28**: 277–284.

Harrison RG (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol Evol* **4**: 6–11.

Hu XS (2000a). Impacts of stochastical migration of seed and pollen flow on differentiation of plant population genetic structure. *Acta Genetica Sinica* **27**: 351–360.

Hu XS (2000b). A preliminary approach to the theory of geographical gene genealogy for plant genomes with three different modes of inheritance and its application. *Acta Genetica Sinica* **27**: 440–448.

Hu XS, Ennos RA (1997). On estimation of the ratio of pollen to seed flow among plant populations. *Heredity* **79**: 541–552.

Hu XS, Ennos RA (1999). Impacts of seed and pollen flow on population differentiation for plant genomes with three contrasting modes of inheritance. *Genetics* **152**: 441–450.

Kruuk LEB, Baird SJE, Gale KS, Barton NH (1999). A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics* **153**: 1959–1971.

Li WH, Nei M (1974). Stable linkage disequilibrium with out epistasis in subdivided populations. *Theor Pop Biol* **6**: 173–183.

Nagylaki T (1975). Conditions for the existence of clines. *Genetics* **80**: 595–615.

Nagylaki T (1997). The diffusion model for migration and selection in a plant population. *J Math Biol* **35**: 409–431.

Petit RJ, Kremer A, Wagner DB (1993). Finite island model for organelle and nuclear genes in plants. *Heredity* **71**: 630–641.

Rieseberg LH, Wendel JF (1993). Introgression and its consequences in plants. In: Harrison RG (ed) *Hybrid Zones and the Evolutionary Process*, Oxford University Press: New York, pp 70–109.

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

Slatkin M (1973). Gene flow and selection in a cline. *Genetics* **75**: 733–756.

Slatkin M (1975). Gene flow and selection in a two-locus system. *Genetics* **81**: 787–802.

Villani F, Sansotta A, Cherubini M, Cesaroni D, Sbordoni V (1999). Genetic structure of natural populations of *Castanea sativa* in Turkey: evidence of a hybrid zone. *J Evol Biol* **12**: 233–244.

Young ND (1996). Concordance and discordance: a tale of two hybrid zones in the Pacific Coast irises (Iridaceae). *Am J Bot* **83**: 1623–1629.