

Genetic divergence between transposable elements

BRIAN CHARLESWORTH

Department of Biology, The University of Chicago, 1103 East 57th Street, Chicago, IL 60637, USA

(Received 6 December 1985 and in revised form 20 May 1986)

Summary

The probabilities of genetic identity between different members of the same family of transposable elements in a randomly mating host population are determined, under the assumption of statistical equilibrium between neutral mutations, random genetic drift, transposition and unbiased gene conversion. The method allows for variation in numbers of copies of the element between individuals within the host population, and for dependence of the probability of identity between a pair of elements on the frequencies of elements at the sites from which they were drawn. It is shown that, for the range of parameters under which the approximations used are valid, the effects of gene conversion on identity probabilities are relatively small, as are the effects of copy number variation and variation between sites in element frequencies.

1. Introduction

Several recent papers have been concerned with modelling the extent of genetic divergence between different members of the same family of transposable elements, due to the accumulation of selectively neutral mutations (Ohta, 1984, 1985; Slatkin, 1985; Brookfield, 1986; Hudson & Kaplan, 1986). Such models are of considerable importance in relation to the interpretation of data on DNA sequence divergence between related transposable elements (Rubin, 1983), more of which will undoubtedly become available in the future. Each of the papers cited above has used a different approach to the problem, although they all share the idea that the degree of genetic divergence between elements belonging to the same family is the product of a statistical balance between factors such as mutation (promoting divergence), and random genetic drift, transposition and gene conversion (promoting similarity). A common assumption of a fixed number of copies of members of a given family per host individual has also been made, except for Hudson & Kaplan (1986).

The purpose of this paper is to re-examine the problem, using an approach that does not require the assumption of a fixed number of elements per individual, or the implicit assumption of a lack of dependence of probability of divergence on element frequencies at the sites in question which was made by Hudson & Kaplan (1986). The procedure is to use a version of Ohta's (1984, 1985) identity coefficient methods, but taking into account the fact that the frequencies of elements at individual chromosomal

sites follow the stationary probability distribution of Langley, Brookfield & Kaplan (1983) and Charlesworth & Charlesworth (1983).

2. The Model and its Analysis

(i) Parameters and variables

A diploid, randomly mating host population will be assumed, and the model and notation of Charlesworth & Charlesworth (1983), as modified by Charlesworth (1985), will be employed. Members of the family of elements are assumed to be able to insert at random into unfilled but potentially occupable chromosomal sites by means of replicative transposition. Independent element frequencies at different sites are assumed, so that the state of a population at a given time is characterized by the vector of element frequencies at each site, $\{x_i\}$ ($i = 1, 2, \dots, m$). A site is thus equivalent to a locus in conventional genetic models. It is also assumed that the number of occupable sites in a haploid genome (m) is very large, so that insertions mostly take place into sites where elements were previously absent. Justifications for these assumptions are provided by Langley *et al.* (1983) and Charlesworth & Charlesworth (1983).

Element frequencies at each site are assumed to follow a stationary probability distribution, generated by the interaction of the forces of selection and transposition with the effects of drift in a monoecious population of size N , with $N_e = N$. The expected number of elements per individual is \hat{n} . The probability of transposition per generation of an element may

depend on the number of elements in the same individual (Langley *et al.* 1983; Charlesworth & Charlesworth, 1983). By taking the expectation of the rate of transposition over the stationary probability distribution of numbers of copies of the element per host individual, we obtain a probability \hat{u} of transposition per element per generation. To a good approximation, \hat{n} and \hat{u} can be equated to their values at equilibrium in an infinite population (Charlesworth & Charlesworth, 1983). Let the stationary probability of element frequency x be $P(x)$ ($x = 0, (2N)^{-1}, \dots, 1$). When all evolutionary forces are weak, a continuous approximation can be employed, such that the probability of element frequency over a given interval is approximated by the integral over this interval of a probability density $\phi(x)$. Writing $\Theta = 4N\hat{u}$, for large m the probability density for the case when copy numbers are controlled by weak selection, or self-regulated transposition, is

$$\phi(x) = \alpha x^{\alpha-1} (1-x)^{\Theta-1}, \tag{1}$$

where $\alpha = \Theta \hat{n} (2m)^{-1} \ll 1$ (Charlesworth & Charlesworth, 1983). The mean frequency is $\hat{x} = \hat{n} (2m)^{-1}$, and the variance is $\sigma_x^2 \approx \hat{x} (1 + \Theta)^{-1}$. This continuous approximation is used extensively in the Appendix.

Each element is assumed to have a probability ν per generation of mutating to a previously unobserved allelic form that is selectively equivalent to existing alleles (the infinite-allele model of Kimura & Crow, 1964). The description of gene conversion and its effects will be given later. N^{-1} , \hat{u} , ν and the probability of gene conversion are each assumed to be sufficiently small that their second-order terms can be neglected. It will also be assumed that Θ and $\hat{n} \gg 1$. The data of Montgomery & Langley (1983) on the distribution of 3 families of *copia*-like elements in a *Drosophila* population suggest that these conditions are frequently met.

(ii) *Changes in identity probabilities*

Two measures of genetic similarity will be used here. The first is the probability that two elements from the same site are identical (Ohta, 1984, 1985). Conditioned on the element frequency at the site in question, this probability is written as $f(x)$, where x is the element frequency. (Since $f(x)$ is undefined for sites where less than two elements are present in the population, we have $N^{-1} \leq x \leq 1$. The unconditional probability of identity (f_0) for two elements drawn from a site where elements are segregating can be obtained by noting that the chance of drawing an element from a site with element frequency x is $Q(x)$, where

$$Q(x) = x P(x) / \sum_{x=N^{-1}}^1 x P(x) \tag{2}$$

so that

$$f_0 = \sum_{x=N^{-1}}^1 f(x) Q(x). \tag{3}$$

The second identity probability is that for a pair of elements, each drawn from a different site. This will be written as $C(x, y)$ for sites with frequencies x and y ($(2N)^{-1} \leq x, y \leq 1$). The unconditional probability of identity for a pair of elements, drawn from a pair of segregating sites, is thus

$$C_0 = \sum_x \sum_y C(x, y) Q'(x) Q'(y), \tag{4}$$

where $Q'(x)$ is defined analogously to $Q(x)$, except that $(2N)^{-1}$ is used as the lower limit of summation. (For large N , as here, this difference is unimportant, and $Q'(x)$ can be equated to $Q(x)$ for all practical purposes.)

The mean conditional frequency of elements at segregating sites is

$$F = \sum x Q(x) \approx (1 + \Theta)^{-1} \tag{5}$$

using the continuous approximation of equation (1).

(a) *Changes in $f(x)$ under mutation, drift and gene conversion.* The effect of drift can be modelled by noting that the number of elements at a site with current frequency x is $2Nx$. Taking mutation into account, the probability of identity for pairs of elements from sites with current element frequency x is thus changed to

$$(1 - \nu)^2 \left\{ \left(1 - \frac{1}{2Nx} \right) f(x) + \frac{1}{2Nx} \right\} \left(x \geq \frac{1}{N} \right). \tag{6}$$

The effect of unbiased gene conversion can be obtained as follows. No distinction is made between conversion events involving different sites on the same chromosome, and sites on different chromosomes. Furthermore, conversions involving a pair of randomly chosen elements at the same site can only occur if the pair comes from the same individual, which has a probability N^{-1} . If the probability of gene conversion is small, as is assumed here, the net chance that such a pair of random elements have converted each other can be neglected.

If a pair of elements is chosen at random from the same site i , their identity probability is thus altered only if one of them has been converted by an element located at another site in the individual from which it originated (multiple events will be ignored here). Following Ohta (1984, 1985) and Slatkin (1985), conversion is assumed to take place by the asymmetric heteroduplex mode, in which one of the two elements is changed to the allelic state of the other, which remains unchanged. The chance of such an event will in general depend on the expected number of elements at other sites in the individual in question, which can be obtained as follows. Consider first the haploid genome containing the sampled element. The expected number of elements at other sites in this genome is the expected number per haploid genome, conditioned on the presence of at least one element, minus one. For a population with mean number of elements per indi-

vidual \bar{n} and a Poisson distribution of copy number, as is the case with $\theta \gg 1$ (Charlesworth & Charlesworth, 1983), the mean haploid number conditioned on the presence of at least one element is close to $\frac{1}{2}\bar{n}(1 - \exp - \frac{1}{2}\bar{n})^{-1}$. To a good approximation, the expected number of elements in the same haploid genome as the sampled element is then

$$\frac{1}{2}\bar{n} = \frac{1}{2}\hat{n}(1 - \exp - \frac{1}{2}\hat{n})^{-1} - 1. \tag{7}$$

For large θ and \hat{n} , $\bar{n} \approx \hat{n}$.

Consider next the other haploid genome. There is a chance x that site i is also occupied, given a current element frequency of x for i , in which case the expected number of elements at other sites is again $\frac{1}{2}\hat{n}$. If this site is unoccupied, the number is approximately $\frac{1}{2}\hat{n}$. Let πn be the probability that the sampled element is converted in a genome with n elements at different sites from the sampled element. If $\hat{n}\pi$ is $\ll 1$, and $\hat{n} \gg 1$ the net probability of conversion of the sampled element is thus approximated by

$$\lambda = \pi\hat{n}. \tag{8}$$

There is a probability 2λ that one of the two elements drawn from a site with element frequency x experiences conversion by an element at a different site, in which case their identity probability changes from $f(x)$ to $C_0(x)$ where

$$C_0(x) = \sum C(x, y) Q'(y). \tag{9}$$

As a result of these forces, $f(x)$ is changed by the quantity

$$\Delta_1 f(x) = -f(x) \left\{ 2v + 2\lambda + \frac{1}{2Nx} \right\} + \frac{1}{2Nx} + 2\lambda C_0(x). \tag{10}$$

We also have to take into account the fact that element frequencies themselves change, so that sites with current element frequencies x differ from those with frequency x in the new generation. Each element at a given site in the new generation is sampled independently from the $2Nx$ present at that site in the current generation; it follows that the probability of identity of a pair drawn from this site in the new generation is independent of the new number of elements at the site. The process of change in element frequency may thus be treated separately from that of change in identity probability. Writing $g(x - \delta x, \delta x)$ for the probability of a transition from $x - \delta x$ to x , and $\tilde{f}(x)$ for $f(x) + \Delta_1 f(x)$, we obtain the new value of $f(x)$ resulting from mutation, drift and gene conversion as

$$f^*(x) = P(x)^{-1} \sum_{\delta x} \tilde{f}(x - \delta x) P(x - \delta x) g(x - \delta x, \delta x). \tag{11}$$

A diffusion approximation to this is given in Appendix

1; further analysis will be deferred until the consideration of equilibrium in section 2(iii).

(b) *Changes in $f(x)$ due to transposition.* Transposition produces an effect on element frequency at a given site of order $\hat{u}m^{-1}$ (Charlesworth & Charlesworth, 1983), which is ignored in the diffusion approximation to equation (11) (equation [A 3]) and so must be considered separately. Given large m and small \hat{u} , at most one new insertion will occur at a given site in the population each generation, producing an increase of $(2N)^{-1}$ in element frequency. For all sites with element frequency $x - 1/2N$, the expected number of such events is $N\hat{u} P(x - 1/2N)$ in a given generation. The number of sites with element frequency x is $mP(x)$; hence, the probability that a site in this interval has experienced an insertion is $N\hat{u} P(x - 1/2N)m^{-1} P(x)^{-1}$. The probability that a random element, drawn from a site with element frequency x , is derived by transposition is thus

$$N\hat{u} P(x - 1/2N) m^{-1} P(x)^{-1} x^{-1} (2N)^{-1} = \hat{u} \hat{x} x^{-1} P(x - 1/2N) P(x)^{-1} \quad (x \geq N^{-1}). \tag{12}$$

The probability that the transposed element originated from the same site as its present location, conditioned on the prevailing set of element frequencies, is $(x - 1/2N) / \sum_i x_i$. The net probability of this event is thus

$$\psi(x - 1/2N) = (x - 1/2N) E \left\{ \frac{1}{\sum x_i} \right\} \approx 2(x - 1/2N) \hat{n}^{-1}. \tag{13}$$

There is thus a probability $1 - \psi$ that it came from another site, chosen at random from the sites segregating for elements in the previous generation. Noting that the assumption of weak evolutionary forces implies additivity of changes in element frequencies, we obtain the following expression for the contribution of transposition to the change in $f(x)$:

$$\Delta_2 f(x) = 2\hat{u} \hat{x} x^{-1} \left\{ 1 - \psi(x - 1/2N) \right\} \left\{ C_0(x - 1/2N) - f(x - 1/2N) \right\} P(x - 1/2N) P(x)^{-1}. \tag{14}$$

The net change in $f(x)$ is given by

$$\Delta f(x) = f^*(x) + \Delta_2 f(x) - f(x). \quad \left(x \geq \frac{1}{N} \right) \tag{15}$$

(c) *Changes in $C(x, y)$ due to mutation and gene conversion.* Mutation changes $C(x, y)$ by $-2vC(x, y)$. The effect of gene conversion can be found as follows. Consider a pair of elements drawn from two distinct sites i and j , with element frequencies x and y respectively. $C(x, y)$ will be affected by conversion of the element at site j if it is converted by an element at site i in the same individual, or by an element at any other site. In the first case, the expected number of elements at site i is $2x$, so that the probability of the event is $2\pi x$, with an associated change in identity

probability from $C(x, y)$ to $f(x)$. Using the notation of section 2(i).a, the probability of the second kind of event is $\lambda - 2\pi x$ and the identity probability becomes $C_0(x)$. Symmetrical relations hold for conversion of the element at site i . The change in $C(x, y)$ due to mutation and gene conversion is thus

$$\Delta_1 C(x, y) = -2(v + \lambda) C(x, y) + 2\pi \{x[f(x) - C_0(x)] + y[f(y) - C_0(y)]\} + \lambda \{C_0(y) + C_0(x)\}. \quad (16)$$

The new value of $C(x, y)$ resulting from drift, mutation and gene conversion, writing $\tilde{C}(x, y) = C(x, y) + \Delta_1 C(x, y)$, is equal to $C^*(x, y)$, where

$$C^*(x, y) = P(x)^{-1} P(y)^{-1} \sum_{\delta x} \sum_{\delta y} \tilde{C}(x - \delta x, y - \delta y) P(x - \delta y) g(x - \delta x, \delta x) g(y - \delta y, \delta y). \quad (17)$$

(d) *Changes in $C(x, y)$ due to transposition.* Using the same argument as for $f(x)$ in section 2(ii)b, the probability that a transposed element at the site i of section 2(ii)c comes from site i and is identical with the element at site i is (approximately) $\psi(x)f(x)$, and so we obtain a contribution of $\hat{u}\hat{x}y^{-1}P(y - 1/2N)P(y)^{-1}\psi(x)f(x)$ to the new value of $C(x, y)$. There is a probability of approximately $1 - \psi(x)$ that a transposed element at site i comes from a site different from i , in which case there is a probability $C_0(x)$ of identity between the elements at sites i and j (where $C_0(x)$ is defined by equation (9)). Symmetric results hold in the case when the transposed element is at site i . Collecting terms and using the argument that led to equation (14), the net effect of transposition is given by

$$\begin{aligned} \Delta_2 C(x, y) = & \hat{u}\hat{x} \left\{ x^{-1} P\left(x - \frac{1}{2N}\right) P(x)^{-1} [\psi(y)f(y) \right. \\ & \left. + (1 - \psi(y)) C_0(y) - C(x, y)] + y^{-1} P\left(y - \frac{1}{2N}\right) \right. \\ & \left. P(y)^{-1} [\psi(x)f(x) + (1 - \psi(x)) C_0(x) - C(x, y)] \right\} \\ & (x, y > (2N)^{-1}). \quad (18) \end{aligned}$$

For consistency, $f[(2N)^{-1}]$ is defined as unity in this equation

The net change in $C(x, y)$ is given by

$$\Delta C(x, y) = C^*(x, y) + \Delta_2 C(x, y) - C(x, y) \quad (19)$$

(iii) *Equilibrium results*

(a) *Approximate analyses.* The approximate results derived below are valid when θ and $\hat{n} \gg 1$ as has been assumed above. The details of the derivations are given in the Appendices. Equations (A 7), (A 9), and (A 11) yield the following approximate formula relating the equilibrium values of f_0 and C_0 , \hat{f} and \hat{C} :

$$\hat{f} \{ \hat{u}^{-1}(v + \lambda) [1 + \frac{1}{2} \hat{u}^{-1}(v + \lambda)] + 1 \} \approx 1 + \frac{1}{2} \hat{u}^{-1}(v + \lambda) + \frac{\lambda}{2} \hat{u}^{-1} [1 + \hat{u}^{-1}(v + \lambda)] \hat{C}. \quad (20)$$

The analysis following equation (A 16) yields the formula

$$\hat{C} = (1 + \gamma)^{-1} \hat{f} \quad (21)$$

where $\gamma = 2N\hat{n}v\hat{u}(\hat{u} + \lambda)^{-1}$.

Hence,

$$\hat{f} = \{1 + \frac{1}{2} \hat{u}^{-1}(v + \lambda)\} \{1 + \hat{u}^{-1}(v + \lambda) [1 + \frac{1}{2} \hat{u}^{-1}(v + \lambda)] - \frac{\lambda \hat{u}^{-1}}{2(1 + \gamma)} [1 + \hat{u}^{-1}(v + \lambda)]\}^{-1}. \quad (22)$$

If $\hat{u}^{-1}(v + \lambda)$ and γ are $\ll 1$, these results can be approximated further. Writing $\hat{H}_1 = 1 - \hat{f}$ and $\hat{H}_2 = 1 - \hat{C}$ for the probabilities of genetic dissimilarity between elements sampled from the same and different sites, respectively, we obtain

$$\hat{H}_1 \approx N\hat{n}v\lambda(\hat{u} + \lambda)^{-1}, \quad (23)$$

$$\hat{H}_2 \approx 2N\hat{n}v\{1 - \frac{1}{2}\lambda(\hat{u} + \lambda)^{-1}\}. \quad (24)$$

When conversion rates are small relative to transposition rates, this formula implies that divergence is decreased by an increased rate of conversion, as might be expected intuitively. Equation (24) displays an effect of the gene conversion rate on \hat{H}_2 , in contrast to the conclusion of Slatkin (1985). This effect arises from the inclusion of the term involving the integral of $f(x)(1 - \theta x)\xi(x) \approx -\theta(df/dx)_F$ in equation (A 7); if this term is neglected, as well as the terms which were shown to be negligible, we find

$$\hat{H}_1 \approx 2N\hat{n}v\lambda(\hat{u} + \lambda)^{-1}, \quad (25)$$

$$\hat{H}_2 \approx 2N\hat{n}v. \quad (26)$$

Equation (26) is similar to equation (9) of Slatkin (1985), and shows a complete independence of the frequency of gene conversion. Variation in element frequencies between sites thus has a qualitative effect on the probabilities of identity.

3. Discussion

The methods described above allow the calculation of the identity probability for a pair of homologous transposable elements sampled from the same site (\hat{f}), and for a pair sampled from two different sites (\hat{C}), under the assumption of statistical equilibrium between neutral mutation, random genetic drift, transposition, and unbiased gene conversion. Loose linkage between elements at different sites, and large values of the mean number of elements per host individual and the parameter θ are also assumed. These assumptions seem biologically realistic, as discussed earlier. Although the results are approximate, variation in copy number between individuals, and dependence of identity probabilities on element frequencies at the sampled sites, are allowed for. The present results are thus more general in this respect than those derived previously (Ohta, 1984, 1985; Slatkin, 1985; Brookfield, 1986; Hudson & Kaplan, 1986).

The quantitative effects of including these factors

Table 1. Equilibrium identity probabilities

$\theta = 5$ $\hat{u} = 10^{-4}, \hat{C}_0 = 0.988$						$\hat{u} = 10^{-5}, \hat{C}_0 = 0.889$					
λ	\hat{C}_1	\hat{f}_2	\hat{C}_2	\hat{f}_3	\hat{C}_3	λ	\hat{C}_1	\hat{f}_2	\hat{C}_2	\hat{f}_3	\hat{C}_3
0	0.985	1.000	0.988	1.000	0.988	0	0.868	0.999	0.888	0.995	0.888
5×10^{-5}	0.985	0.996	0.988	0.997	0.989	5×10^{-6}	0.869	0.962	0.888	0.977	0.902
1×10^{-4}	0.985	0.994	0.988	0.996	0.990	1×10^{-5}	0.870	0.944	0.888	0.962	0.905
2×10^{-4}	0.985	0.992	0.988	0.994	0.990	2×10^{-5}	0.871	0.925	0.888	0.943	0.905
5×10^{-4}	0.985	0.990	0.988	0.991	0.989	5×10^{-5}	0.873	0.907	0.888	0.919	0.901
$\theta = 20$ $\hat{u} = 10^{-4}, \hat{C}_0 = 0.952$						$\hat{u} = 10^{-5}, \hat{C}_0 = 0.667$					
λ	\hat{C}_1	\hat{f}_2	\hat{C}_2	\hat{f}_3	\hat{C}_3	λ	\hat{C}_1	\hat{f}_2	\hat{C}_2	\hat{f}_3	\hat{C}_3
0	0.950	1.000	0.952	1.000	0.952	0	0.655	0.999	0.666	1.000	0.666
5×10^{-5}	0.950	0.984	0.952	0.990	0.958	5×10^{-6}	0.655	0.969	0.666	0.983	0.676
1×10^{-4}	0.950	0.976	0.952	0.984	0.960	1×10^{-5}	0.656	0.833	0.666	0.882	0.705
2×10^{-4}	0.950	0.968	0.952	0.976	0.960	2×10^{-5}	0.657	0.777	0.666	0.823	0.705
5×10^{-4}	0.951	0.956	0.952	0.966	0.958	5×10^{-5}	0.658	0.722	0.666	0.752	0.694
$\theta = 100$ $\hat{u} = 10^{-4}, \hat{C}_0 = 0.800$						$\hat{u} = 10^{-5}, \hat{C}_0 = 0.286$					
λ	\hat{C}_1	\hat{f}_2	\hat{C}_2	\hat{f}_3	\hat{C}_3	λ	\hat{C}_1	\hat{f}_2	\hat{C}_2	\hat{f}_3	\hat{C}_3
0	0.798	1.000	0.800	1.000	0.800	0	0.283	0.999	0.285	1.000	0.286
5×10^{-5}	0.798	0.959	0.800	0.959	0.822	5×10^{-6}	0.284	0.761	0.286	0.842	0.316
1×10^{-4}	0.799	0.931	0.800	0.931	0.828	1×10^{-5}	0.284	0.642	0.286	0.729	0.324
2×10^{-4}	0.799	0.896	0.800	0.896	0.828	2×10^{-5}	0.284	0.524	0.286	0.594	0.324
5×10^{-4}	0.800	0.854	0.800	0.854	0.820	5×10^{-5}	0.285	0.405	0.286	0.442	0.312

$\hat{n} = 50$ and $\nu = 10^{-8}$ for all entries. \hat{C}_0 is the value of \hat{C} from Slatkin (1985); \hat{C}_1 is the value from Hudson & Kaplan (1986); \hat{f}_2 and \hat{C}_2 are from equations (27) and (21), and \hat{f}_3 and \hat{C}_3 are from equations (22) and (21).

are displayed in Table 1, which gives examples of equilibrium identity probabilities computed by several different methods. The first column of each section shows the \hat{C} values obtained using Hudson & Kaplan's (1986) equations (14) and (15). (These assume fixed copy number per individual when there is gene conversion.) The next two columns show the \hat{f} and \hat{C} values given by the present approach, but neglecting the effects of variation in element frequencies between sites. Equation (A 9) implies that, in this case, equation (22) is replaced by

$$\hat{f} = \{1 + u^{-1}(\nu + \lambda)[1 + (1 + \gamma)^{-1}]\}^{-1} \quad (27)$$

while equation (22) is unchanged. The last two columns in each section display the results from equations (21) and (22), which allow for dependence of identity probabilities on element frequencies. The value of \hat{C} from Slatkin's (1985) equation (9) is given at the head of each section; his analysis (which assumed a fixed copy number per individual) suggested that identity probabilities were nearly independent of the rate of gene conversion, for the range of parameter values considered here.

It will be seen that the different formulae for \hat{C} produce rather similar numerical results. The probability of identity between two elements from the same

site is always larger than that for elements from different sites, and is considerably more sensitive to the effect of gene conversion, as would be expected intuitively. Given the fact that the mean number of elements per individual is typically fairly large, at least for *Drosophila* (Rubin, 1983), so that it is unlikely that two randomly sampled elements come from the same site, the identity measure \hat{C} is of much greater interest biologically than \hat{f} . Table 1 shows that \hat{C}_3 , the value which takes variation in element frequencies between sites into account, varies with the rate of gene conversion more than the others, but the effect is small, and of little significance when the problems involved in comparisons of theoretical values with experimental estimates are borne in mind. For most practical purposes, Slatkin's equation (9) can probably be safely used, unless the rate of gene conversion is high compared with the rate of transposition.

This work was supported in part by a grant from the Louis Block Fund of the University of Chicago, and by grant BSR-8516629 from the National Science Foundation. I thank T. Nagylaki, W. Stephan, and an anonymous reviewer for their comments on the manuscript. I am grateful to T. Nagylaki for pointing out to me the need to consider the effects of variation in element frequencies between sites.

References

Brookfield, J. F. Y. (1986). A model for DNA sequence evolution within transposable element families. *Genetics* **112**, 393–407.

Charlesworth, B. (1985). The population genetics of transposable elements. In *Population Genetics and Molecular Evolution* (ed. T. Ohta and K. -I. Aoki), pp. 213–232. Berlin: Springer-Verlag.

Charlesworth, B. & Charlesworth, D. (1983). The population dynamics of transposable elements. *Genetical Research, Cambridge* **42**, 1–27.

Hudson, R. R. & Kaplan, N. L. (1986). On the divergence of members of a transposable element family. *Journal of Mathematical Biology* (In the Press.)

Kimura, M. & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.

Langley, C. H., Brookfield, J. F. Y. & Kaplan, N. L. (1983). Transposable elements in Mendelian populations. I. A theory. *Genetics* **104**, 457–471.

Montgomery, E. A. & Langley, C. H. (1983). Transposable elements in Mendelian populations. II. Distribution of three *copia* -like elements in a natural population of *Drosophila melanogaster*. *Genetics* **104**, 473–483.

Ohta, T. (1984). The population genetics of transposable elements. *Mathematics Applied in Medicine and Biology* **1**, 17–29.

Ohta, T. (1985). A model of duplicative transposition and gene conversion of repetitive DNA families. *Genetics* **110**, 513–524.

Rubin, G. M. (1983). Dispersed repetitive DNAs in *Drosophila*. In *Mobile Genetic Elements* (ed. J. A. Shapiro), pp. 299–328. New York: Academic Press.

Slatkin, M. (1985). Genetic differentiation of transposable elements under mutation and unbiased gene conversion. *Genetics* **110**, 145–158.

Appendix 1

Diffusion approximations

We start by replacing the discrete probability formulation of the text with the continuous approximation of equation (1). $Q(x)$ and $Q'(x)$ are replaced by $\xi(x) = x\phi(x)/\int_0^1 x\phi(x) dx$; $f(x)$ and $C(x, y)$ are now defined for x and y over the closed interval $[0, 1]$, and summations are replaced by integration over this interval. Applying the usual Taylor series expansion to the right-hand side of equation (11), we obtain the diffusion approximation

$$f^*(x)\phi(x) = \tilde{f}(x)\phi(x) - \frac{\partial f(x)\phi(x)M}{\partial x} \delta x + \frac{1}{2} \frac{\partial^2 f(x)\phi(x)V_{\delta x}}{\partial x^2}, \quad (A 1)$$

where $M_{\delta x} = (4N)^{-1}\alpha(1-x) - \hat{u}x$ and $V_{\delta x} = x(1-x)/2N$ are respectively the mean and variance of the change in element frequency per generation (Charlesworth & Charlesworth, 1983).

But $\phi(x)$ satisfies the stationarity condition

$$-\frac{\partial \phi(x)M_{\delta x}}{\partial x} + \frac{1}{2} \frac{\partial^2 \phi(x)V_{\delta x}}{\partial x^2} = 0$$

so that equation (A 1) yields

$$f^*(x)\phi(x) = \tilde{f}(x)\phi(x) + \frac{\partial f(x)}{\partial x} \left\{ \frac{\partial \phi V_{\delta x}}{\partial x} - \phi M_{\delta x} \right\} + \frac{1}{2} \frac{\partial^2 f(x)}{\partial x^2} \phi(x) V_{\delta x}. \quad (A 2)$$

Using equations (1) and (15), and neglecting the term in α (which is $0(\hat{u}m^{-1})$) in $M_{\delta x}$, we obtain the relation

$$\Delta f(x) = \Delta_1 f(x) + \Delta_2 f(x) - \hat{u}x \frac{\partial f(x)}{\partial x} + \frac{x(1-x)}{4N} \frac{\partial^2 f(x)}{\partial x^2} \quad (A 3)$$

A similar analysis can be carried out for $C(x, y)$, yielding the equation

$$\Delta C(x, y) = \Delta_1 C(x, y) + \Delta_2 C(x, y) - \hat{u} \left\{ x \frac{\partial C(x, y)}{\partial x} + y \frac{\partial C(x, y)}{\partial y} \right\} + \frac{1}{4N} \left\{ x(1-x) \frac{\partial^2 C(x, y)}{\partial x^2} + y(1-y) \frac{\partial^2 C(x, y)}{\partial y^2} \right\} \quad (A 4)$$

Appendix 2

The equilibrium solution to equation (A 3) and (A 4) can be approximated as follows, assuming $\Theta \gg 1$. Terms of order v, \hat{u}, π and N^{-1} will be denoted by $0(\epsilon)$, when convenient. Consider first equation (A 3). Setting $\Delta f(x) = 0$, multiplying by $x\xi(x)$ and integrating, we find (after neglecting $0(\epsilon^2)$ terms) that

$$\int_0^1 x \left\{ \Delta_1 f(x) + \Delta_2 f(x) \right\} \xi(x) dx - \hat{u} \int_0^1 x^2 \frac{df(x)}{dx} \xi(x) dx + (4N)^{-1} \int_0^1 x^2(1-x) \frac{d^2 f(x)}{dx^2} \xi(x) dx = 0. \quad (A 5)$$

For large Θ , integrating by parts and neglecting $0(\epsilon^2)$ terms we have

$$\int_0^1 x^2 \frac{d^2 f}{dx^2} \xi(x) dx \approx - \int_0^1 x(2-\Theta x) \frac{df}{dx} \xi(x) dx, \quad (A 6a)$$

$$\int_0^1 x \frac{df}{dx} \xi(x) dx \approx - \int_0^1 (1-\Theta x) f(x) \xi(x) dx. \quad (A 6b)$$

Substituting into equation (A 5), multiplying by $2N$, and noting that $\Theta(4N)^{-1} = \hat{u}$, we find that

$$2N \int_0^1 x \{ \Delta_1 f(x) + \Delta_2 f(x) \} \xi(x) dx + \int_0^1 (1-\Theta x) f(x) \xi(x) dx = 0. \quad (A 7)$$

For large, Θ , values of x for segregating sites with significant probability will be close to $x = F \approx \Theta^{-1}$. We can thus approximate $f(x)$ and $C_0(x)$ in equations (3) and (4) to high accuracy by the first two terms of their Taylor's expansion about F . From equation (1), it follows that

$$\int_0^1 x^n \xi(x) dx = n! / (1+\Theta)(2+\Theta) \dots (n+\Theta). \quad (A 8)$$

Using this in the Taylor approximation to the left-

hand side of equation (A 7) around $x = \Theta^{-1}$, and neglecting the contribution from $\Delta_2 f(x)$ (which is of order $\Theta \hat{x}$ and is thus $\ll 1$ for sufficiently large m), the first pair of terms in equation (A 7) can be written as

$$-\hat{f}\hat{u}^{-1}[v+\lambda][1+0(\Theta^{-1})]+1\}+1+\hat{u}^{-1}\lambda\hat{C}\{1+0(\Theta^{-1})\} \tag{A 9}$$

where \hat{f} and \hat{C} are the equilibrium values of the quantities defined by equations (3) and (4). The third term is approximately equal to $-\Theta^{-1}(df/dx)_F$ and will be considered below.

We now consider the question of the magnitude of the derivatives of $f(x)$ and $C(x, y)$ in relation to the approximations made above. An estimate of the values of the derivatives of $f(x)$ can be obtained as follows. Setting $\Delta f(x)$ in equation (A 3) to zero, multiplying by $2Nx$ and differentiating with respect to x gives the equation

$$-f(x)4N(v+\lambda)+4N\left\{\lambda C_0(x)+x\lambda\frac{dC_0(x)}{dx}\right\} \approx \left(\frac{df}{dx}\right)\left\{4Nx[\hat{u}+v+\lambda]+1\right\}-x\left\{\frac{d^2f}{dx^2}+\frac{x}{2}\frac{d^3f}{dx^3}\right\} \tag{A 10}$$

assuming $x \ll 1$.

Neglecting $dC_0(x)/dx$ and the higher order derivatives of $f(x)$, we obtain the following expression

$$\left(\frac{df}{dx}\right)_F \approx \frac{4N\{\hat{f}v+(\hat{f}-\hat{C})\lambda\}}{\{2+\hat{u}^{-1}(v+\lambda)\}} \tag{A 11}$$

Substituting this into the approximation to $-\Theta^{-1}(df/dx)_F$, equation (A 7), we obtain equation (20) of the text.

The validity of neglecting the term in $dC_0(x)/dx$ will be examined below; the inaccuracies introduced by neglecting the higher order derivatives of $f(x)$ can be determined as follows. Differentiating equation (A 10) again, we obtain

$$-\left(\frac{d^2f(x)}{dx^2}\right)_F \approx \left(\frac{df}{dx}\right)_F \frac{4N\{\hat{u}+2(v+\lambda)\}}{\{2+\hat{u}^{-1}(v+\lambda)\}}, \tag{A 12}$$

$$-\left(\frac{d^3f(x)}{dx^3}\right)_F \left(\frac{d^2f}{dx^2}\right)_F \frac{4N\{\hat{u}+3(v+\lambda)\}}{\{2+\hat{u}^{-1}(v+\lambda)\}}. \tag{A 13}$$

Combining these results, it is found that

$$-x\left\{\left(\frac{d^2f}{dx^2}\right)_F+\frac{x}{2}\left(\frac{d^3f}{dx^3}\right)_F\right\}\approx\left(\frac{df}{dx}\right)_F\frac{\{1+2\hat{u}^{-1}(v+\lambda)\}}{\{2+\hat{u}^{-1}(v+\lambda)\}}\left\{1-\frac{[2+3\hat{u}^{-1}(v+\lambda)]}{2[2+\hat{u}^{-1}(v+\lambda)]}\right\}. \tag{A 14}$$

It is apparent from this that equation (A 11) overestimates $-(df/dx)_F$ unless $\hat{u} \ll v+\lambda$. If $\hat{u} \gg v+\lambda$, the addition of the above term yields an estimate which is $\frac{3}{8}$ times the earlier one. The inaccuracy in the estimate of df/dx from the neglect of the higher order derivatives therefore does not seem serious.

Equation (A 4) can be analysed in a similar way. Multiplying both sides by $\xi(x)\xi(y)$, integrating and neglecting $0(\varepsilon^2)$ terms, we obtain the equilibrium expression

$$\int_0^1\{A_1C(x,y)+A_2C(x,y)\}\xi(x)\xi(y)dx dy-2\hat{u} \int_0^1x\frac{dC_0(x)}{dx}\xi(x)dx+ (2N)^{-1}\int_0^1x(1-x)\frac{d^2C_0(x)}{dx^2}\xi(x)dx=0. \tag{A 15}$$

This can be simplified by the method used for equation (A 3) to yield

$$\int_0^1\{A_1C(x,y)+A_2(x,y)\}\xi(x)\xi(y)dx dy -(2N)^{-1}\int_0^1\frac{dC(x)}{dx}\xi(x)dx. \tag{A 16}$$

Providing that $\partial^2C(x, y)/\partial x \partial y$ and $\partial^2C(x, y)/\partial x^2$ at $x, y = F$ are of order Θ or less, we can approximate $C_0(x)$, $C_0(y)$ and C_0 using the first terms in the Taylor expansion of $C(x, y)$ around $x, y = F$, to accuracy of order Θ^{-1} . $C_0(x)$ and C_0 can then be approximated by $C(x, F)$ and $C_0(F)$ respectively. On this basis, the term involving A_1 can be written as $-2\{v\hat{C}-2\pi F(\hat{f}-\hat{C})\}+0(\hat{u}\Theta^{-2})$. Neglecting terms $0(m^{-1})$, the integral of A_2 is equal to $4\hat{u}\hat{n}^{-1}F(\hat{f}-\hat{C})+0(\hat{u}\Theta^{-2})$. This yields equation (21) of the text if the last term in equation (A 16) is neglected.

A similar procedure can be applied in order to estimate the derivatives of $C(x, y)$, by neglecting them in equation (A 4) and setting $\Delta C(x, y) = 0$. Implicit differentiation of the resulting equation then yields approximate values of the derivatives.

The derivatives of $C(x, y)$ at $x, y = F$ are found to be

$$\left(\frac{\partial C(x, F)}{\partial x}\right)_F \approx \frac{2\pi(\hat{f}-\hat{C})+2\Theta^{-1}\{\pi+\hat{u}\hat{x}\hat{n}^{-1}\Theta^{-1}\}}{2\{(v+\lambda)+\hat{u}\hat{x}\Theta\}}\left(\frac{df}{dx}\right)_F, \tag{A 17}$$

$$\left(\frac{\partial^2 C(x, F)}{\partial x^2}\right)_F \approx \frac{\Theta}{2}\left(\frac{\partial C(x, F)}{\partial x}\right)_F + \frac{\Theta\hat{n}^{-1}}{2}\left(\frac{df}{dx}\right)_F\{2+\hat{u}^{-1}(v+\lambda)\}^{-1}, \tag{A 18}$$

$$\left(\frac{\partial^2 C(x, y)}{\partial x \partial y}\right)_F \approx -\frac{3\Theta}{2}\left(\frac{\partial C(x, F)}{\partial x}\right)_F. \tag{A 19}$$

Using equation (A 11) to eliminate df/dx from equation (A 17), we find that $(\partial C/\partial x)_F \approx -4Nv\hat{f}\Theta^{-3}\hat{n}^{-1}$ (if $\pi \approx 0$) or $\approx \hat{n}^{-1}(\hat{f}-\hat{C})$ (if $\pi > 0$). Reference back to equation (A 10) shows that the contribution of $(dC_0/dx)_F$ to $(df/dx)_F$ can thus be neglected if $\Theta^{-1}\hat{n}^{-1} \ll 1$, as is the case under the assumed conditions. Similarly, the contribution of the last term in equation (A 16) can be neglected.

The second derivatives of $C(x, y)$ are of order Θ or less at $x, y = F$ under these conditions. Equation (21)

is then valid to a relative accuracy of θ^{-1} , given the magnitudes of the multiplicands of \hat{f} and \hat{C} .

Finally, we must consider whether there is an inconsistency in neglecting the derivatives of $C(x, y)$ in

equation (A 4) in order to estimate their values. At $x, y = F$ we require that $\partial C(x, y)/\partial x$ be $\ll 1$, and that $\partial^2 C(x, y)/\partial x^2 \ll \frac{1}{2}\theta$. These conditions are satisfied in $\hat{n}^{-1}(\hat{f} - \hat{C})$ and $\hat{n}^{-1}\theta^{-1}(df/dx)_F$ are both $\ll 1$.