Effect of dominance on heterozygosity and the fixation probability in a subdivided population

Jo Nishino and Fumio Tajima *

Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan.

(Received 10 October 2003, accepted 7 January 2004)

Under the assumptions of a subdivided population and the presence of dominance for fitness, the expected sum of heterozygosity in the total population during the lifetime of a mutant was investigated. It was shown analytically and by computer simulations that in the island model the effect of dominance on the expected sum of heterozygosity decreases as the migration rate decreases and is lost almost completely when the migration rate is very low. In addition to the expected sum of heterozygosity, the fixation probability of mutant was also investigated. The effect of dominance on the fixation probability also decreases as the migration rate decreases but is not completely lost when the migration rate is very low.

Key words: effect of dominance, fixation probability, heterozygosity, natural selection, subdivided population,

INTRODUCTION

A large amount of genetic variability is maintained in a population. Heterozygosity is a measure of genetic variability in the population. When there are two alleles in a locus, heterozygosity is defined as $2x(1-x)$, where $x$ is the frequency of one allele. On the other hand, molecular evolution is directly related with the fixation probability of a mutant. Let $\mu$, $u$, and $N_T$ be the mutation rate per gene per generation, the fixation probability of a newly arisen mutant and the size of population, respectively. Then, the rate of gene substitution per gene per generation is expressed as $2N_T\mu$. In brief, heterozygosity is a within-population diversity and the fixation of a mutant causes the difference between populations, and the both are determined by the same behavior of mutant in a population.

Denote by $H(1/2N_T)$ the expected sum of heterozygosity introduced by one mutant from the occurrence to the final fixation or loss, where $1/2N_T$ is the initial frequency of mutant at the time of mutation. $H(1/2N_T)$ is the expectation of $\sum_{t=0}^{\infty} 2x_t(1-x_t)$, where $x_t$ is the mutant frequency in the $t$th generation since the occurrence of mutation, especially $x_0 = 1/2N_T$. It is important to note that $2N_T\mu H(1/2N_T)$ gives the expectation of the average numbers of (pairwise) nucleotide differences, or $2N_T\nu H(1/2N_T)$ gives the expectation of nucleotide diversity, where $\nu$ is the mutation rate per nucleotide site. Kimura (1969) investigated the expected number of heterozygous sites per individual under the assumption of infinite site model with free recombination among those sites. Then, under random mating Kimura’s study shows the following. If a mutant is selectively neutral, $H(1/2N_T)$ is equal to 2. If a mutant is semidominant (i.e., genic selection) and advantageous, $H(1/2N_T)$ is at most 4. Namely, even if a semidominant mutant is sufficiently advantageous, the mutant contributes to heterozygosity only twice as much as a neutral one. Furthermore, according to the numerical calculations of (11) in Kimura (1969), a completely recessive mutant does not contribute to heterozygosity more than a neutral one, even if the recessive mutant is advantageous (See “panmictic” case in Table 1). On the other hand, a completely dominant mutant contributes to heterozygosity more than twice as much as a neutral one if the dominant mutant is advantageous. This means that the dominance of mutant affects the expected sum of heterozygosity substantially in a random mating population. A deleterious mutant does not contribute to heterozygosity more than a neutral one, regardless of the dominance. This is because a deleterious mutant tends to be removed from a population quickly and only few deleterious mutants can fix in a population. Of course, the dominance of deleterious mutant also affects the expected sum of heterozygosity greatly in a random mating population.

In nature, however, a population is often subdivided into many subpopulations. In this paper we will investigate the effect of dominance on the expected sum of heterozygosity in the subdivided population, using an island...
model. In addition to heterozygosity, we will also investigate the fixation probability of mutant although some of the results are essentially the same as those of Slatkin (1981) and Cherry (2003). We will use two methods. One is the birth-and-death approximation that can be used when the migration rate is very low (Slatkin 1981, Takahata 1991). Another is the diffusion approximation that can be used when the migration rate is not very low (Kimura 1962, 1969).

**MODEL**

The finite island model of population structure (for example, see Maruyama 1970 and Takahata 1991) is used in this study. The population consists of \( L \) demes (subpopulations), each of which has \( N \) monoeocious diploid individuals. There are \( NL \) or \( N_T \) individuals in total. The migration rate from one deme to the other demes per generation is denoted by \( m \). The fraction of immigrants in a recipient deme from a donor is \( m/(L–1) \) every generation.

In the population shown above, we consider the process that one mutant occurs at generation 0 (\( t = 0 \)), and selection, migration and random mating change the frequency of mutant in this order every generation. Denote the mutant and wild type by \( a \) and \( A \), respectively. It is assumed that the fitnesses of genotypes \( AA \), \( Aa \) and \( aa \) are 1, \( 1 + hs \) and \( 1 + s \), respectively, where \( s \) is the selec-

---

### Table 1  The expected sum of heterozygosity in the total population, \( H(1/2N_T) \).

<table>
<thead>
<tr>
<th>( N_T )</th>
<th>( L )</th>
<th>( m )</th>
<th>( s )</th>
<th>Sim</th>
<th>Diff</th>
<th>Sim</th>
<th>Diff</th>
<th>Sim</th>
<th>Diff</th>
<th>Low migr</th>
<th>hline</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>1</td>
<td>Panmictic</td>
<td>-0.002</td>
<td>1.92</td>
<td>1.91</td>
<td>1.38</td>
<td>1.37</td>
<td>1.02</td>
<td>1.02</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2.04</td>
<td>2.00</td>
<td>1.97</td>
<td>2.00</td>
<td>1.97</td>
<td>2.00</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>1.96</td>
<td>1.98</td>
<td>2.26</td>
<td>2.32</td>
<td>2.66</td>
<td>2.75</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>1.92</td>
<td>1.93</td>
<td>2.59</td>
<td>2.62</td>
<td>3.62</td>
<td>3.61</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>1.70</td>
<td>1.74</td>
<td>3.20</td>
<td>3.22</td>
<td>6.06</td>
<td>5.93</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>1.51</td>
<td>1.50</td>
<td>3.62</td>
<td>3.58</td>
<td>7.82</td>
<td>7.86</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>1.25</td>
<td>1.25</td>
<td>3.79</td>
<td>3.76</td>
<td>9.41</td>
<td>9.42</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>-0.002</td>
<td>2.24</td>
<td>2.35</td>
<td>1.88</td>
<td>1.94</td>
<td>1.58</td>
<td>1.61</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2.83</td>
<td>2.81</td>
<td>2.71</td>
<td>2.81</td>
<td>2.83</td>
<td>2.81</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>2.94</td>
<td>2.97</td>
<td>3.33</td>
<td>3.27</td>
<td>3.60</td>
<td>3.61</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>3.08</td>
<td>3.06</td>
<td>3.73</td>
<td>3.69</td>
<td>4.52</td>
<td>4.46</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>3.19</td>
<td>3.13</td>
<td>4.56</td>
<td>4.53</td>
<td>6.51</td>
<td>6.54</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>3.10</td>
<td>3.04</td>
<td>5.00</td>
<td>5.04</td>
<td>7.98</td>
<td>8.05</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>2.94</td>
<td>2.90</td>
<td>5.26</td>
<td>5.29</td>
<td>9.16</td>
<td>8.95</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>-0.002</td>
<td>56.98</td>
<td>-</td>
<td>53.84</td>
<td>-</td>
<td>54.35</td>
<td>-</td>
<td>53.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>82.32</td>
<td>-</td>
<td>82.28</td>
<td>-</td>
<td>82.46</td>
<td>-</td>
<td>81.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>98.71</td>
<td>-</td>
<td>97.65</td>
<td>-</td>
<td>93.55</td>
<td>-</td>
<td>95.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>112.81</td>
<td>-</td>
<td>112.23</td>
<td>-</td>
<td>109.46</td>
<td>-</td>
<td>108.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>138.38</td>
<td>-</td>
<td>136.98</td>
<td>-</td>
<td>136.07</td>
<td>-</td>
<td>135.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>154.27</td>
<td>-</td>
<td>155.75</td>
<td>-</td>
<td>155.50</td>
<td>-</td>
<td>151.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>158.66</td>
<td>-</td>
<td>163.28</td>
<td>-</td>
<td>162.27</td>
<td>-</td>
<td>159.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>30</td>
<td>0.01</td>
<td>-0.002</td>
<td>1.37</td>
<td>1.43</td>
<td>0.99</td>
<td>0.97</td>
<td>0.67</td>
<td>0.69</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2.82</td>
<td>2.94</td>
<td>2.70</td>
<td>2.94</td>
<td>2.93</td>
<td>2.94</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>3.54</td>
<td>3.41</td>
<td>4.97</td>
<td>4.90</td>
<td>6.81</td>
<td>7.00</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>3.29</td>
<td>3.32</td>
<td>5.54</td>
<td>5.46</td>
<td>8.69</td>
<td>8.59</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>-0.002</td>
<td>28.86</td>
<td>-</td>
<td>28.95</td>
<td>-</td>
<td>28.98</td>
<td>-</td>
<td>28.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>93.12</td>
<td>-</td>
<td>93.27</td>
<td>-</td>
<td>96.28</td>
<td>-</td>
<td>93.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>157.62</td>
<td>-</td>
<td>157.33</td>
<td>-</td>
<td>161.90</td>
<td>-</td>
<td>158.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>178.87</td>
<td>-</td>
<td>181.01</td>
<td>-</td>
<td>181.12</td>
<td>-</td>
<td>176.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panmictic means the random mating population. The values of Sim were obtained by the simulations. The values of Diff and Low migr were obtained by the diffusion approximation and the birth–and-death approximation respectively. The number of replications in each simulation is 1,000,000.
tion coefficient and $h$ is the degree of dominance of mutant. It is also assumed that selection and random mating are practiced independently in each deme, and that selection and migration change the allele frequency deterministically.

**BIRTH-AND-DEATH APPROXIMATION**

A simple way to examine the change of gene frequency in the island model is to assume that the migration rate is very low ($Nm \ll 1$). When the migration rate is very low, there are at most two states in any deme, either “mutant” or “non-mutant”. In such a situation, birth-and-death process may describe the change of gene frequency. Slatkin (1981) and Takahata (1991) used this process for solving the fixation probability. Here, we will show the approximation for the expected sum of heterozygosity, assuming very low migration rate ($Nm \ll 1$).

First, the derivation of fixation probability will be briefly described in the following, although the result is included in Slatkin (1981). The reason is that the derivation of fixation probability is important for obtaining the expected sum of heterozygosity. The important simplification in the following derivation is the independence between the effects of migration in one generation and those in the other generation on the fixation or extinction of mutant in each deme. This simplification is considered valid from the following two facts. First, under the assumption of random mating the fixation probability of mutant is linear with the initial frequency of mutant when the frequency is nearly 0. Similarly, the extinction probability of mutant is linear with the initial frequency of mutant when the frequency is nearly 1. Second, under the finite island model, the mutant frequency in each deme is nearly to 0 or 1 in almost all generations when the migration rate is very low ($Nm \ll 1$). Thus, the effects of migration in some generations on the fixation or extinction of mutant in each deme are almost additive and independent of each other when the migration rate is very low.

**Fixation probability** Now suppose that at a given generation, there are $i$ mutant and $L - i$ non-mutant demes. Denote this state by $S_i$ ($i = 0, 1, 2, 3, ..., L$). We denote the probabilities that $S_i$ moves to $S_{i+1}$ and $S_{i-1}$ in one generation by $f_i$ and $d_i$, respectively ($i = 1, 2, ..., L - 1$). Especially, denote the probability that a newly arisen mutant in one deme fixes in this deme by $f_0$. If the newly arisen mutant fixes in the first deme with the probability $f_0$, then the state moves $S_1$ and the process continues until the state becomes $S_0$ or $S_L$.

In the very low migration rate ($Nm \ll 1$), the fixation rate per generation per non-mutant deme is considered to be proportional to the number of mutant demes, $i$, approximately. This is because the amount of mutant migration per generation in a non-mutant deme is proportional to the number of mutant demes, and the fixation rate per generation per non-mutant deme is also proportional to the amount of mutant migration per generation in the non-mutant deme. Similarly, the extinction probability of mutant per generation per mutant deme is considered to be proportional to the number of non-mutant demes, $L - i$. Let $f$ be the probability that one mutant deme causes to change one non-mutant deme into a mutant deme by migration per generation, and $d$ be the probability that one non-mutant deme causes to change one mutant deme into a non-mutant deme by migration per generation. The mathematical definitions of $f$ and $d$ are shown later. Note that $i$ mutant demes affect $L - i$ non-mutant demes and that $L - i$ non-mutant demes affect $i$ mutant demes. Then, we have

$$f_i = i(L - i) f, \quad d_i = (L - i) i d.$$  \hspace{1cm} (1)

The state $S_i$ changes to either $S_{i+1}$ or $S_{i-1}$. The former probability is $f_i / (f_i + d_i) = f / (f + d)$, and the latter probability is $d / (f_i + d_i) = d / (f + d)$. Thus, the fixation probability for the entire population in the very low migration ($Nm \ll 1$) is the same as that in the random walk, where the boundaries are at states $S_0$ and $S_L$, and the transition probabilities from $S_i$ to $S_{i+1}$ and $S_{i-1}$ are $f / (f + d)$ and $d / (f + d)$, respectively (Feller 1957).

Let $U_i$ be the probability that the mutant spreads to the entire population, given that there are $i$ fixed demes. Then, we have

$$U_i = \frac{f}{f + d} U_{i+1} + \frac{d}{f + d} U_{i-1} (1 \leq i \leq L - 1)$$  \hspace{1cm} (2)

with the boundary conditions, $U_0 = 0$ and $U_L = 1$. Solving the above equation, we have

$$U_i = \frac{1 - \left(\frac{d}{f}\right)^i}{1 - \left(\frac{d}{f}\right)^L},$$  \hspace{1cm} (3)

where $f \neq d$.

Now we consider the value of $d/f$. Denote by $u(p)$ the fixation probability with the initial frequency of $p$ under the random mating population with $N$ individuals, $u(p)$ is given by (13) in Kimura (1962). One non-mutant deme receives mutants from one mutant deme with the rate of $m/L - 1$ per generation. Thus $f$ is given by $u(m/L - 1)$. One mutant deme receives wild types from one non-mutant deme with the rate of $m/L - 1$ per generation. Note that the fixation of wild type is the extinction of mutant. Thus $d$ is given by $1 - u(1 - m/L - 1)$ Therefore $d/f$ can be given by

$$d/f = \frac{d}{f} = \frac{1 - u(1 - m/L - 1)}{u(m/L - 1)}.$$
of heterozygosity, \( H_{\text{H}} \) (Slatkin 1981).

When the migration rate is very low, the migration rate does not affect the fixation probability in the entire population. In addition, when the migration rate is very low (\( Nm \ll 1 \)), the relationship of \( \tilde{f}_0 : f = \frac{1}{2N} : \frac{m}{L-1} \) holds, we have \( \tilde{f}_0 = \frac{L-1}{2Nm} f \) approximately. Thus, \( H(1/2N_T) \) becomes

\[
H(1/2N_T) = \frac{L-1}{2Nm} f \times \frac{2(-1+(\tilde{f}^2) + L-(\tilde{f}^2)L)}{L^2(f-d)(1-(\tilde{f}^2))} = \frac{L-1}{L^2Nm} \frac{1-(\tilde{f}^2) + L(1-(\tilde{f}^2))}{(1-(\tilde{f}^2))(1-(\tilde{f}^2))}.
\]

Next, \( df/dt \) is given by (4), and we have

\[
H(1/2N_T) = \frac{L-1}{LNm(1-e^{-2NS})} \left( U_1 - \frac{1}{L} \right)
\]

where \( U_1 \) is given by (5). Therefore, \( H(1/2N_T) \) is independent of the degree of dominance, \( h \). The expression of \( H(1/2N_T) \) is similar to (15) of Kimura (1969), and a similar discussion is possible as shown below. At the limit of \( s \to 0 \), we have

\[
H(1/2N_T) = \frac{(L-1)^2}{2L^3Nm}.
\]

If the mutant is advantageous such that \( 2Ns \gg 1 \), we have

\[
H(1/2N_T) = \frac{(L-1)^2}{L^3Nm},
\]

since \( 1 - e^{-2NS} \approx 1 \) and \( U_1 \approx 1 \). Therefore, if the migration rate is very low (\( Nm \ll 1 \)) and mutation is advantageous, \( H(1/2N_T) \) is at most twice as high as that of neutral mutant.

**DIFFUSION APPROXIMATION**

In the previous section the very low migration rate (\( Nm \ll 1 \)) was assumed. The approximation described in this section is appropriate to the moderate to high migration (\( Nm \ll 1 \)). The expected sum of heterozygosity, \( H(1/2N_T) \), as well as the fixation probability, is expressed in terms of the mean and the variance of the
effect of dominance on heterozygosity

change of mutant frequency per generation, using the diffusion method (Kimura 1962, 1969). Note that $H(1/2N_T)$ is given by (11)-(14) of Kimura (1969), substituting 1 into $v_m$, where $v_m$ is the average number of mutation per gamete per generation.

Kimura (1962, 1969) focused on the case of a random mating population. In the following, we approximate Kimura’s expression includes only the mutant frequency in the total population. In the following, we approximate $vm$ is given by (11)-(14) of Kimura (1969), substituting 1 into fusion method (Kimura 1962, 1969). Note that change of mutant frequency per generation, using the difference in the total population, is $\mu = \frac{1}{2} \frac{\delta}{x} x \{ 1 - 3h + 3x(2h - 1) \} + \mu_s (2h - 1)$. The third moment of $x$, about $x$, $\mu_3$, when a mutant is neutral, becomes

$$\mu_3 = \frac{x(x - 1)(2x - 1)}{(1 + 4Nm)(1 + 4Nm)}$$

(19)

Note that the assumption of the infinite number of demes means that $x$ is fixed, whereas in the model of finite number of demes $x$ fluctuates with time. In addition, in the model of infinite number of demes, one deme receives the mutants with the rate of $mx$ from $x$ by migration, while in the model of finite number of demes the $i$th deme receives mutants with the rate of $\frac{1}{2} \frac{\delta}{x} x \{ 1 - 3h + 3x(2h - 1) \} + \mu_s (2h - 1)$. The variance in the total population, $V_n = \sum_{s=1}^{L} \frac{1}{L} V_{n, s}$, is

$$V_n = \frac{(1 - F_{ST}) \frac{x(1-x)}{2N_T}}{1 - \frac{x(x - 1)(2x - 1)}{(1 + 4Nm)(1 + 4Nm)}}$$

(20)

Substituting $\mu_3$ of (20) and $\frac{\delta}{x} x \{ 1 - 3h + 3x(2h - 1) \} + \mu_s (2h - 1)$ into $\mu_3$ of (15) and $F_{ST}$ of (16) respectively, we obtain $M_n$ and $V_n$ in terms of only $x$ and parameters $(L, m, N)$. Thus, $H(1/2N_T)$ can be calculated using (11)-(14) of Kimura (1969). Similarly, the fixation probability is calculated using (3) and (4) of Kimura (1962). Note that the approximation by using (20) and (17) is considered to be valid for moderate to high migration ($Nm \ll 1$), weak selection ($Ns \ll 1$) and a large number of demes. The reason is as follows. First, (20) and (17) are based on the assumption of neutral mutation. Second, $\mu_3$ and $F_{ST}$ fluctuate with time. If the number of demes is large, the fluctuation is relatively small. Third, when $\mu_3$ and $F_{ST}$ deviate from the values of (20) and (17), the deviation can disappear quickly if the migration rate is moderate to high

$F_{ST}$ and $\mu_3$ in terms of some parameters $(L, m, N)$ and/or the mutant frequency in the total population, $x$. Let $\tilde{F}_{ST}$ be the value of $F_{ST}$ when a mutant is neutral. Then, the expectation of $\tilde{F}_{ST}$ is given by

$$E(\tilde{F}_{ST}) = \frac{1}{1 + 4Nm(\frac{L}{N_T})^2}$$

(17)

(for example, see Takahata 1983). If the selection is weak ($Ns \ll 1$), the value of $F_{ST}$ is approximately given by (17). When the selection is strong, however, $F_{ST}$ tends to deviate from the value of (17).

Next, we assume that the number of demes, $L$, is infinite and a mutant is neutral. Then, the probability density of $x$, is given by

$$\frac{\Gamma(Nm)}{\Gamma(Nm x) \Gamma(Nm(1-x))} x^{4Nm-1}(1-x)^{4Nm(1-x)-1}$$

(18)

where $\Gamma()$ represents gamma function (Wright 1931). The third moment of $x$, about $x$, $\mu_3$, when a mutant is neutral, becomes

$$\mu_3 = \frac{x(x - 1)(2x - 1)}{(1 + 4Nm)}$$

(19)

Note that the assumption of the infinite number of demes means that $x$ is fixed, whereas in the model of finite number of demes $x$ fluctuates with time. In addition, in the model of infinite number of demes, one deme receives the mutants with the rate of $mx$ from $x$ by migration, while in the model of finite number of demes the $i$th deme receives mutants with the rate of $\frac{1}{2} \frac{\delta}{x} x \{ 1 - 3h + 3x(2h - 1) \} + \mu_s (2h - 1)$. The variance in the total population, $V_n = \sum_{s=1}^{L} \frac{1}{L} V_{n, s}$, is

$$V_n = \frac{(1 - F_{ST}) \frac{x(1-x)}{2N_T}}{1 - \frac{x(x - 1)(2x - 1)}{(1 + 4Nm)}}$$

(20)

Substituting $\mu_3$ of (20) and $\frac{\delta}{x} x \{ 1 - 3h + 3x(2h - 1) \} + \mu_s (2h - 1)$ into $\mu_3$ of (15) and $F_{ST}$ of (16) respectively, we obtain $M_n$ and $V_n$ in terms of only $x$ and parameters $(L, m, N)$. Thus, $H(1/2N_T)$ can be calculated using (11)-(14) of Kimura (1969). Similarly, the fixation probability is calculated using (3) and (4) of Kimura (1962). Note that the approximation by using (20) and (17) is considered to be valid for moderate to high migration ($Nm \ll 1$), weak selection ($Ns \ll 1$) and a large number of demes. The reason is as follows. First, (20) and (17) are based on the assumption of neutral mutation. Second, $\mu_3$ and $F_{ST}$ fluctuate with time. If the number of demes is large, the fluctuation is relatively small. Third, when $\mu_3$ and $F_{ST}$ deviate from the values of (20) and (17), the deviation can disappear quickly if the migration rate is moderate to high

$F_{ST}$ and $\mu_3$ in terms of some parameters $(L, m, N)$ and/or the mutant frequency in the total population, $x$. Let $\tilde{F}_{ST}$ be the value of $F_{ST}$ when a mutant is neutral. Then, the expectation of $\tilde{F}_{ST}$ is given by

$$E(\tilde{F}_{ST}) = \frac{1}{1 + 4Nm(\frac{L}{N_T})^2}$$

(17)

(for example, see Takahata 1983). If the selection is weak ($Ns \ll 1$), the value of $F_{ST}$ is approximately given by (17). When the selection is strong, however, $F_{ST}$ tends to deviate from the value of (17).

Next, we assume that the number of demes, $L$, is infinite and a mutant is neutral. Then, the probability density of $x$, is given by

$$\frac{\Gamma(Nm)}{\Gamma(Nm x) \Gamma(Nm(1-x))} x^{4Nm-1}(1-x)^{4Nm(1-x)-1}$$

(18)

where $\Gamma()$ represents gamma function (Wright 1931). The third moment of $x$, about $x$, $\mu_3$, when a mutant is neutral, becomes

$$\mu_3 = \frac{x(x - 1)(2x - 1)}{(1 + 4Nm)}$$

(19)

Note that the assumption of the infinite number of demes means that $x$ is fixed, whereas in the model of finite number of demes $x$ fluctuates with time. In addition, in the model of infinite number of demes, one deme receives the mutants with the rate of $mx$ from $x$ by migration, while in the model of finite number of demes $x$ because $x$ includes the fraction $x$ itself. Thus, more appropriate approximation in the model of finite number of demes may be given by

$$\mu_3 = \frac{x(x - 1)(2x - 1)}{(1 + 4Nm)}$$

(20)

Substituting $\mu_3$ of (20) and $\frac{\delta}{x} x \{ 1 - 3h + 3x(2h - 1) \} + \mu_s (2h - 1)$ into $\mu_3$ of (15) and $F_{ST}$ of (16) respectively, we obtain $M_n$ and $V_n$ in terms of only $x$ and parameters $(L, m, N)$. Thus, $H(1/2N_T)$ can be calculated using (11)-(14) of Kimura (1969). Similarly, the fixation probability is calculated using (3) and (4) of Kimura (1962). Note that the approximation by using (20) and (17) is considered to be valid for moderate to high migration ($Nm \ll 1$), weak selection ($Ns \ll 1$) and a large number of demes. The reason is as follows. First, (20) and (17) are based on the assumption of neutral mutation. Second, $\mu_3$ and $F_{ST}$ fluctuate with time. If the number of demes is large, the fluctuation is relatively small. Third, when $\mu_3$ and $F_{ST}$ deviate from the values of (20) and (17), the deviation can disappear quickly if the migration rate is moderate to high
NUMERICAL CALCULATION & DISCUSSION

Table 1 shows the expected sums of heterozygosity in random mating populations and under the island model, which were obtained by numerical calculations and computer simulations. In the computer simulations, the pseudosampling method (Kimura and Takahata 1983) was used to perform the random mating in each deme. The values in abbreviation Low migr were obtained by the birth-and-death approximation described in this paper. The values in abbreviation Diff were obtained by the diffusion approximation described in this paper, except that the value under the assumption of random mating was obtained by Kimura (1969). The values obtained by approximations (the diffusion approximation and the birth-and-death approximation) are consistent with those obtained by simulations. Under the assumption of random mating, we see distinct differences in the values of expected sums of heterozygosity among the recessive ($h = 0$), semidominant ($h = 1/2$), and dominant ($h = 1$) mutants. In the case of semidominant mutant, $H(1/2N_T)$ is at most twice as large as that of neutral mutation. In the case of dominant mutant, $H(1/2N_T)$ can become much larger than twice of that of neutral mutation. In the case of recessive mutant, $H(1/2N_T)$ does not become larger than that of neutral mutation. In the case of moderate migration ($m = 0.01$ or $Nm = 0.5$), the difference becomes a little smaller than that of random mating population. In very weak migration ($m = 0.0001$, $Nm = 0.005$), there seems to be no difference among the recessive, semidominant, and dominant mutants. Thus, we can conclude that the effect of dominance on the ability to produce heterozygosity becomes small as the migration rate becomes small, and is almost completely lost when the migration rate is very low ($Nm \ll 1$).

Table 2 shows the fixation probabilities. The values under a random mating population were obtained by Kimura (1962). For the other notations, refer to that of Table 1. The number of replications in each simulation is 1,000,000.

<table>
<thead>
<tr>
<th>$N_T$</th>
<th>$L$</th>
<th>$s$</th>
<th>$m$</th>
<th>Sim</th>
<th>Diff</th>
<th>Sim</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>10</td>
<td>0.005 (Low migr)</td>
<td>–</td>
<td>(4.63)</td>
<td>–</td>
<td>(5.44)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.001</td>
<td>4.48</td>
<td>4.05</td>
<td>5.64</td>
<td>6.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.002</td>
<td>4.15</td>
<td>3.95</td>
<td>5.94</td>
<td>6.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.005</td>
<td>3.68</td>
<td>3.67</td>
<td>6.61</td>
<td>6.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.01</td>
<td>3.41</td>
<td>3.38</td>
<td>7.06</td>
<td>7.07</td>
<td></td>
</tr>
<tr>
<td>Panmictic</td>
<td></td>
<td>2.57</td>
<td>2.53</td>
<td>8.57</td>
<td>8.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01  (Low migr)</td>
<td>–</td>
<td>(8.46)</td>
<td>–</td>
<td>(11.63)</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>7.82</td>
<td>7.43</td>
<td>12.24</td>
<td>12.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>7.43</td>
<td>7.15</td>
<td>12.93</td>
<td>13.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>6.68</td>
<td>6.45</td>
<td>14.05</td>
<td>13.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>5.84</td>
<td>5.69</td>
<td>15.07</td>
<td>15.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panmictic</td>
<td></td>
<td>3.63</td>
<td>3.57</td>
<td>18.66</td>
<td>18.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td>100</td>
<td>0.002 (Low migr)</td>
<td>–</td>
<td>(1.93)</td>
<td>–</td>
<td>(2.06)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.001</td>
<td>1.78</td>
<td>1.79</td>
<td>2.22</td>
<td>2.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.002</td>
<td>1.70</td>
<td>1.67</td>
<td>2.35</td>
<td>2.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.005</td>
<td>1.36</td>
<td>1.40</td>
<td>2.67</td>
<td>2.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.01</td>
<td>1.10</td>
<td>1.14</td>
<td>3.05</td>
<td>2.95</td>
<td></td>
</tr>
<tr>
<td>Panmictic</td>
<td></td>
<td>0.50</td>
<td>0.50</td>
<td>3.94</td>
<td>3.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005 (Low migr)</td>
<td>–</td>
<td>(4.60)</td>
<td>–</td>
<td>(5.40)</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>4.14</td>
<td>4.44</td>
<td>5.88</td>
<td>5.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>4.00</td>
<td>4.10</td>
<td>6.31</td>
<td>5.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>3.29</td>
<td>3.35</td>
<td>6.83</td>
<td>6.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>2.58</td>
<td>2.63</td>
<td>7.66</td>
<td>7.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panmictic</td>
<td></td>
<td>0.78</td>
<td>0.80</td>
<td>9.82</td>
<td>9.85</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The notations are the same as those in Table 1. The number of replications in each simulation is 1,000,000.
Effect of dominance on heterozygosity

Table 1. The most values obtained by approximations (the diffusion and the birth-and-death approximation) are consistent with those obtained by simulations. As Slatkin (1981) discussed by using the results of computer simulations and the analytical values for low migration ($Nm \ll 1$), Table 2 shows that in the case of dominant mutant the lower the migration rate is, the smaller the fixation probability is, whereas in the case of recessive mutant the lower the migration rate is, the larger the fixation probability is. The effect of dominance on the fixation probability decreases but is not completely lost, even if the migration rate is very low ($Nm \ll 1$). The effect of dominance in the very low migration rate reflects the term of $f_0$ in (6). Whereas, the fixation probability of semidominant mutant is independent of the migration rate (Maruyama 1970, Cherry 2003).

The diffusion method described in this paper is based on essentially the same idea as shown in Cherry (2003). The both methods simplify a multi-dimensional diffusion equation (equations for all demes) to a one-dimensional equation. There is a difference between the two methods. In this paper we take into account the number of demes in the expression of quantities which are related with the moments of the distribution of $x_n$ such as $\mu_n$ and $F_{ST}$, whereas the method of Cherry (2003) does not take the number of demes into consideration in the expression of the moments. Since the moments do not include the number of demes in the method of Cherry (2003), the expressions of $M_n$ and $V_n$ in Cherry can be simpler than those of the present method. On the other hand, in the cases where the number of demes is relatively small our method seems to give better values of the fixation probability and the expected sum of heterozygosity. For example, on the expected sum of heterozygosity, in the case where $L = 10$, $N = 50$, $m = 0.01$, $s = 0.001$, Cherry’s method gives 3.22 ($h = 0$) and 3.79 ($h = 1$), whereas our method gives 2.97 ($h = 0$) and 3.61 ($h = 1$) and the simulation gives 2.94 ($h = 0$) and 3.60 ($h = 1$). On the fixation probability, in the case where $L = 10$, $N = 50$, $m = 0.01$, $s = 0.01$, Cherry’s method gives 0.00633 ($h = 0$) and 0.01415 ($h = 1$), whereas our method gives 0.00569 ($h = 0$) and 0.01506 ($h = 1$) and the simulation gives 0.00584 ($h = 0$) and 0.01507 ($h = 1$). There may, however, be little difference in the two methods when the number of demes is large.

APPENDIX

As shown in the following, (7) is expressed as a matrix form.

$$
\begin{pmatrix}
1 - \frac{f}{\tau_{ED}} & 0 \\
-\frac{f}{\tau_{ED}} & 1 \\
\vdots & \vdots \\
1 - \frac{f}{\tau_{ED}} & 0 \\
0 & -\frac{d}{\tau_{ED}} & 1
\end{pmatrix}
= \begin{pmatrix}
H_1 \\
H_L-1
\end{pmatrix}
$$

(A1)

Using Cramer’s rule, $H_1$ is given by the two $(L-1)$ dimensional determinants,

$$
H_1 = \frac{\begin{vmatrix}
\frac{2}{\tau_{ED}} & -\frac{f}{\tau_{ED}} & 0 \\
\vdots & 1 & \vdots \\
\vdots & -\frac{d}{\tau_{ED}} & \vdots \\
\vdots & \vdots & \vdots \\
0 & -\frac{d}{\tau_{ED}} & 1
\end{vmatrix}}{\begin{vmatrix}
1 - \frac{f}{\tau_{ED}} & 0 \\
-\frac{d}{\tau_{ED}} & 1 \\
\vdots & \vdots \\
1 - \frac{f}{\tau_{ED}} & 0 \\
0 & -\frac{d}{\tau_{ED}} & 1
\end{vmatrix}}
$$

(A2)

Denote by $D_n$ the n-dimensional determinants which have the same form as the denominator in (A2). By cofactor expansion, we have

$$
D_n = D_{n-1} - \frac{df}{(f+d)^n} D_{n-2}.
$$

(A3)

By using

REFERENCES


Wright, S. (1931) Evolution in Mendelian populations. Genetics 16, 97–159
\[ D_1 = 1 \text{ and } D_2 = 1 - \frac{df}{(f + d)^2} \quad \text{(A4)} \]

as the initial conditions, we obtain

\[ D_n = \frac{f^{n+1} - d^{n+1}}{(f - d)(f + d)^n}. \quad \text{(A5)} \]

Similarly, denote by \( E_n \) the \( n \)-dimensional determinants which have the same form as the numerator in (A2). By cofactor expansion, we have

\[ E_n = \frac{2}{L^2(f + d)} D_{n-1} + \frac{f}{f + d} E_{n-1}. \quad \text{(A6)} \]

By using

\[ E_1 = \frac{2}{L^2(f + d)} \text{ and } E_2 = \frac{2(d + 2f)}{L^2(f + d)^2} \quad \text{(A7)} \]

as the initial conditions, we obtain

\[ E_n = \frac{2(d^{n+1} - f^n(d + dn - fn))}{L^2(f - d)^2(f + d)^n}. \quad \text{(A8)} \]

Calculating \( H_1 = E_{L-1}/D_{L-1} \), we have (8).