A Neutral Model With Fluctuating Population Size and Its Effective Size

Masaru Iizuka,*1 Hidenori Tachida† and Hirotsugu Matsuda†

*Division of Mathematics, Kyushu Dental College, Kitakyushu 803-8580, Japan, †Department of Biology, Graduate School of Sciences, Kyushu University, Fukuoka 810-8560, Japan and ‡Kyushu University, Fukuoka 812-8581, Japan

Manuscript received August 13, 2001
Accepted for publication February 11, 2002

ABSTRACT

We consider a diffusion model with neutral alleles whose population size is fluctuating randomly. For this model, the effects of fluctuation of population size on the effective size are investigated. The effective size defined by the equilibrium average heterozygosity is larger than the harmonic mean of population size but smaller than the arithmetic mean of population size. To see explicitly the effects of fluctuation of population size on the effective size, we investigate a special case where population size fluctuates between two distinct states. In some cases, the effective size is very different from the harmonic mean. For this concrete model, we also obtain the stationary distribution of the average heterozygosity. Asymptotic behavior of the effective size is obtained when the population size is large and/or autocorrelation of the fluctuation is weak or strong.

The average heterozygosity has been one of the most frequently used measures of genetic diversity. A large amount of data have been accumulated to estimate the average heterozygosity of various species using protein electrophoresis (see, for example, Nevo 1978; Nei and Graur 1984; Hamrick and Godt 1990). Recent developments in molecular techniques such as randomly amplified polymorphic DNA (Williams et al. 1990) and amplified fragment length polymorphism (Vos et al. 1995) enable us to carry out estimation of the average heterozygosity at the DNA level by randomly sampling many short sequences in the genome and examining their variation. Using the methods such as those developed by Clark and Lanigan (1993), Lynch and Milligan (1994), and Innan et al. (1999), it is now possible to estimate genome-wide heterozygosity fairly easily using these techniques (e.g., Miyashita et al. 1999).

One of the reasons that heterozygosity has been used for measuring genetic diversity is a simple relationship between its expectation and population genetic parameters under the neutrality assumption (Kimura 1968). For example, if we assume the infinite allele model in a haploid population of constant size N, the expected heterozygosity is expressed as $H = 2N\mu/(1 + 2N\mu)$, where $\mu$ is the mutation rate (Kimura and Crow 1964). Since $H$ is a monotone increasing function of $N\mu$, one can obtain information on $N\mu$ from estimates of $H$.

The size of the population is, however, hardly constant and it may fluctuate from generation to generation. In such cases, it is necessary to understand how fluctuation of population size would affect genetic diversity and summary statistics such as $H$. For this end, the effects of fluctuation of population size on the effective size of population $N^*$ must be clarified and a representation for $N^*$ must be obtained. By this representation, the expected heterozygosity may be expressed as $H = 2N^*\mu/(1 + 2N^*\mu)$ and it shows how the fluctuation of population size affects $H$.

Fluctuation of population size is not independent from generation to generation in general as in the case of stochastic selection (Gillespie 1972; Takahata et al. 1975; Gillespie and Guess 1978; Iizuka and Matsuda 1982; Seno and Shiga 1984; Iizuka 1987). In other words, the fluctuation of population size is most likely autocorrelated. In the literature, the effective size is said to be equal to the harmonic mean of the population size when population sizes are not constant (Wright 1938; Crow 1954; Nei et al. 1975; Gillespie 1998). Iizuka (2001) showed, however, that the effective size is not equal to the harmonic mean unless the fluctuation of population size is uncorrelated for the Wright-Fisher model with fluctuating population size. To obtain this result however, no mutation is assumed and the population size is fluctuating between two distinct states. These assumptions are restrictive. It is important to include the effect of mutation to investigate the influence of fluctuation of population size on genetic diversity of population. Further, two-state models may be very special and some of their conclusions do not hold for the general case. It is necessary to see whether or not the effective size is different from the harmonic mean for the general pattern of the fluctuation of population size.

In this article, we consider the diffusion model with neutral $K (2 \leq K \leq \infty)$ alleles whose population size is fluctuating randomly, incorporating the effect of mutation. First, we consider a general case with respect to the fluctuation of population size and we show that the effective size defined by the equilibrium average
heterozygosity is larger than the harmonic mean of population size but smaller than the arithmetic mean of population size. Then we consider a special case of a two-valued Markov chain as a model of the fluctuation of population size. This simplification enables us to obtain explicit formulas for the stationary distribution of the average heterozygosity and the effective size. We can see quantitatively how the effective size is different from the harmonic mean using the latter formula.

**NEUTRAL MODEL WITH FLUCTUATING POPULATION SIZE**

Before we introduce the fluctuation of population size, we summarize some of the known results on the constant population model (Crow and Kimura 1970; Ewens 1979). We consider a neutral locus with K alleles A1, A2, . . . , Ak in a randomly mating haploid population of constant size N. The mutation rate from Ai to Aj with the rate u/(K − 1) per generation (i, j = 1, 2, . . . , K, i ̸= j). Under the diffusion approximation (diffusion model), let x(t) be the gene frequency of Ai at time t. We denote by E[·] the operation taking the expectation with respect to the random sampling drift. Then the average heterozygosity

\[ H(t) = 1 - \sum_{i=1}^{K} E[x_i(t)^2] \]  

at time t satisfies the differential equation

\[ \frac{dH(t)}{dt} = -\frac{1}{N}H(t) + 2u[1 - cKH(t)] \]  

(see appendix a), which has the solution

\[ H(t) = \frac{2Nu}{1 + 2NuS_k} + \left( H(0) - \frac{2Nu}{1 + 2NuS_k} \right) e^{-\frac{2uN(t+1)}{N(1)}} \]  

and

\[ H(\infty) = \lim_{t \to \infty} H(t) = \frac{2Nu}{1 + 2NuS_k} \]  

where we put

\[ c_k = \frac{K}{K - 1} \]  

Note that the case of K = \infty is the infinite allele model (Kimura and Crow 1964). Note also that c2 = 2 and c\infty = 1. By this formula, the population size N can be expressed as

\[ N = \frac{H(\infty)}{2u[1 - cKH(\infty)]} \]  

A large amount of ecological data suggest that numbers of individuals in natural populations fluctuate considerably in each epoch and from generation to generation (Elton and Nicholson 1942; Andrewartha and Birch 1954; Odum 1959). The variations in population size are influenced by such factors as climate, the abundance of available resources, fluctuation in prey-predator balance, and competition with other species using the same habitat (Nicholson 1957). In addition to those short-term changes demonstrated by ecological data, long-term changes of population size have been inferred from past climate and fossil data. It is well known that there were at least seven glacial and interglacial cycles with a period of \sim 100,000 years in the last 700,000 years. Organismal populations were thought to have responded to such climate shifts by changing their habitats (Webb and Bartlein 1992). For example, many plant and animal species retreated to a few refugia in the southern parts of Europe during the last glacial period (Bennett 1997; Hewitt 2000). Although these climate changes have strong cyclic components, biotic responses have many stochastic elements due to existence of physical barriers and species interaction. Thus, many species are thought to have experienced long-term stochastic changes of population size. Causes for long-term changes are not restricted to glacial cycles. Longer-term climate changes such as those in the last 3 million years (Webb and Bartlein 1992) and mountain building are among those. Thus, it is important to investigate fluctuation of population size in a general setting.

Now we consider the cases when population size fluctuates and let N(t) be the size of a haploid population at time t. In this article, we assume that \{N(t)\}_{t \in \mathbb{R}} is a stationary stochastic process that does not depend on gene frequencies \{(x1(t), x2(t), . . . , x_{K-1}(t))\}_{t \in \mathbb{R}}. In other words, the stochastic process that governs the change in population size is independent of the genetic structure of the population. We consider a diffusion model whose population size at time t is N(t) (for the precise meaning of this model, see appendix a). This model is referred to as the neutral diffusion model with fluctuating population size and the case of K = \infty is referred to as the infinite allele model with fluctuating population size. For this model, the average heterozygosity H(t) satisfies

\[ \frac{dH(t)}{dt} = 2u - \left( 2uS_k + \frac{1}{N(t)} \right) H(t) \]  

(see appendix a). The solution of this differential equation is

\[ H(t) = 2ue^{\int_0^t \frac{1}{N(s)} ds} + H(0)e^{\int_0^t \frac{1}{N(s)} ds} \]  

(8)

Note that \{H(t)\}_{t \geq 0} is a stochastic process induced by \{N(t)\}_{t \geq 0}

Let

\[ N_k = E[N(t)] \]  

and

\[ N_k = E[\frac{1}{N(t)}]^{-1} \]  

(9)

(10)
be the arithmetic mean and the harmonic mean of \( N(t) \), respectively. Here \( E[\cdot] \) is the operation taking the expectation with respect to the probability law of \( N(t) \). On the other hand, the effective size of population \( N_e \) can be defined as

\[
N_e = \frac{E[H(\infty)]}{2q[1 - cE[H(\infty)]]} \tag{11}
\]

by (6). Then we have

\[
N_e < N_c < N, \tag{12}
\]

(see appendix b). The effective size of the population is larger than the harmonic mean but smaller than the arithmetic mean by (12). Note that this result holds for any fluctuation of \( N(t) \) as far as \( 1/N(t) \) is integrable and \( H(i) \) is described by (7). This result is general in this sense but we cannot see how much \( N_e \) is larger than \( N_c \) or how much it is smaller than \( N \). For this end, we must consider a concrete example of the fluctuation of \( N(t) \), which is discussed in the next section. As we noted before, the effective size is said to be equal to the harmonic mean of the population size when population sizes are not constant in the literature (Wright 1938; Crow 1954; Nei et al. 1975; Gillespie 1998). Our result shows, however, that this is not true for the neutral model with fluctuating population size. The fact that the effective size is not always the same as the harmonic mean was pointed out by Iizuka (2001) for the Wright-Fisher model with fluctuating population size.

**TWO-VALUED MARKOV CHAIN MODEL**

To see how the effective size of population \( N(t) \) depends on the probability law of \( \{N(t)\}_{t<\infty} \) and to what extent \( N_c \) is different from \( N \), and \( N_e \), we consider a special case of a continuous time two-valued Markov chain for \( \{N(t)\}_{t<\infty} \). Let \( \{N(t)\}_{t<\infty} \) be a Markov chain on \( \{N_1, N_2\} \) such that

\[
P(N(t + \Delta t) \neq N(t) \mid N(t) = N_i) = \gamma_i \Delta t + o(\Delta t), \tag{13}
\]

\[
P(N(t + \Delta t) = N(t) \mid N(t) = N_i) = 1 - \gamma_i \Delta t + o(\Delta t), \tag{14}
\]

\((\Delta t \downarrow 0)\), where \( N_1 < N_2 \). Here, \( \gamma_i \) is the jump rate from the state \( N_i \) to \( N_j \) \((j \neq i)\) and \( \lim_{\Delta t \downarrow 0} (\gamma_i \Delta t/\Delta t) = 0 \). Note that the stationary probabilities for \( N(t) = N_1 \) and \( N(t) = N_2 \) are

\[
p_1^{(i)} = 1 - r, \tag{15}
\]

\[
p_2^{(i)} = r, \tag{16}
\]

respectively, where we put

\[
r = \frac{\gamma_1}{\gamma_1 + \gamma_2}. \tag{17}
\]

Note that \( r = \frac{1}{2} \) for the symmetric case of \( \gamma_1 = \gamma_2 \). The autocorrelation of \( \{N(t)\}_{t<\infty} \) can be defined by

\[
\tau = \frac{1}{\text{Var}[N(0)]:0} \text{Cov}[N(0), N(\theta)]d\theta, \tag{18}
\]

where \( \text{Var}[N(0)] \) and \( \text{Cov}[N(0), N(\theta)] \) are the variance of \( N(0) \) and the covariance of \( N(0) \) and \( N(\theta) \), respectively. Since

\[
\text{Cov}[N(0), N(\theta)] = (N_2 - N_1)^2(r(1 - r)e^{-\gamma\theta}), \tag{19}
\]

where

\[
\gamma = \frac{\gamma_1 + \gamma_2}{2}, \tag{20}
\]

we have

\[
\tau = \frac{1}{2\gamma}. \tag{21}
\]

and \( N(t_1) \) and \( N(t_2) \) are positively autocorrelated \((t_1 \neq t_2)\).

Let \( \rho(h, N_1) \) and \( \rho(h) \) be the stationary probability density functions of \( (H(t), N(t)) \) and \( H(t) \), respectively \((i = 1, 2)\). Applying the results of Matsuda and Ishii (1981), we have

\[
\rho(h, N_1) = \frac{(\alpha_1 + \alpha_2)H(h - H_1)^{\alpha_1 - 1}(H_2 - h)^{\alpha_2}}{B(\alpha_1, \alpha_2)(\alpha_1H_1 + \alpha_2H_2)(H_2 - H_1)^{\alpha_1 + \alpha_2}} \tag{22}
\]

\[
\rho(h, N_2) = \frac{(\alpha_1 + \alpha_2)H(h - H_2)^{\alpha_1}(H_1 - h)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)(\alpha_1H_1 + \alpha_2H_2)(H_2 - H_1)^{\alpha_1 + \alpha_2}} \tag{23}
\]

and

\[
\rho(h) = \rho(h, N_1) + \rho(h, N_2) = \frac{(\alpha_1 + \alpha_2)H(h - H_1)^{\alpha_1 - 1}(H_2 - h)^{\alpha_2} - (\alpha_1 + \alpha_2)H(h - H_2)^{\alpha_1}(H_1 - h)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)(\alpha_1H_1 + \alpha_2H_2)(H_2 - H_1)^{\alpha_1 + \alpha_2}} \tag{24}
\]

\((H_1 < h < H_2)\), where

\[
H_i = \frac{2N_i u}{1 + 2N_i u}, \tag{25}
\]

\[
\alpha_i = \frac{N_i}{1 + 2N_i u}, = \frac{\gamma_i H_i}{2a} \tag{26}
\]

and

\[
B(\beta, \theta) = \int_0^\beta \theta^{\beta-1}(1 - \theta)^{\gamma-1}d\theta \tag{27}
\]

is the beta function \((\beta, \gamma > 0); \text{ see appendix c})\). The stationary probability density function \( \rho(h) \) of the average heterozygosity is presented in Figure 1 for the case of \( K = 10 \), \( 2N_1 u = 0.1 \) (\( H_1 = 0.09 \)), and \( 2N_1 u = 1 \) (\( H_2 = 0.5 \)). The solid line, the dotted line, and the dashed line represent the cases of \( \alpha_1 = \alpha_2 = 0.5, \alpha_1 = \alpha_2 = 1, \) and \( \alpha_1 = \alpha_2 = 2 \), respectively. The stationary probability density function \( \rho(h) \) is bell shaped if \( \alpha_1, \alpha_2 > 1 \) (the jump rates \( \gamma_1 \) and \( \gamma_2 \) are large) and it is U shaped if \( 0 < \alpha_1, \alpha_2 < 1 \) (the jump rates \( \gamma_1 \) and \( \gamma_2 \) are small). Note that the equilibrium heterozygosity \( H(\infty) \) is a random variable due to the stochastic fluctuation of population size and the distribution function of \( H(\infty) \) is
As noted above, the variance is due to the stochastic fluctuation of population size. Thus, the variance quantifies the extent of variation expected among average heterozygosities of independent populations whose sizes obey the same probability law. The variance should not be mixed up with that among the average heterozygosities of respective loci in one species, because those loci are considered to have experienced the same population history. The conditional stationary probability density function of $H(t)$ given $N(t) = N_i$ is

$$p(h|N_i) = \frac{p(h, N_i)}{p^{(o)}(N_i)}$$

$$= \frac{(\alpha_1 + \alpha_2)H_1(h - H_1)^{\alpha_1 - 1}(H_2 - h)^{\alpha_2}}{(1 - r)B(\alpha_1, \alpha_2)(\alpha_1 H_1 + \alpha_2 H_2)(H_2 - H_1)^{\alpha_1 + \alpha_2}}$$

(32)

and that given $N(t) = N_2$ is

$$p(h|N_2) = \frac{p(h, N_2)}{p^{(o)}(N_2)}$$

$$= \frac{(\alpha_1 + \alpha_2)H_2(h - H_2)^{\alpha_1}(H_3 - h)^{\alpha_2 - 1}}{rB(\alpha_1, \alpha_2)(\alpha_1 H_1 + \alpha_2 H_2)(H_3 - H_2)^{\alpha_1 + \alpha_2}}$$

(33)

Then the conditional expectation of $H(\infty)$ given the population size being $N_i$ is

$$E[H(\infty)|N_i] = \int_{R} h p(h|N_i) dh$$

$$= \frac{\alpha_2}{(1 - r)(\alpha_1 + \alpha_2 + 1)} \left[ 1 + \frac{H_1}{\alpha_1 H_1 + \alpha_2 H_2} \right]$$

(34)

and that given the population size being $N_2$ is

$$E[H(\infty)|N_2] = \int_{R} h p(h|N_2) dh$$

$$= \frac{\alpha_1}{r(\alpha_1 + \alpha_2 + 1)} \left[ 1 + \frac{H_2}{\alpha_1 H_1 + \alpha_2 H_2} \right]$$

(35)

Note that $E[H(\infty)|N_i]$ is the expectation of the equilibrium average heterozygosity knowing that the population size at the observation time is $N_i$. It is easy to see

$$H_1 < E[H(\infty)|N_i] < E[H(\infty)|N_2] < H_2$$

(36)

as we expect. On the other hand, the conditional stationary probabilities of $N(t) = N_i$ and $N(t) = N_2$ given $H(t) = h$ are

$$p(N_i|h) = \frac{b(h, N_i)}{p(h)} = \frac{H_i(H_2 - h)}{(H_2 - H_i) h}$$

(37)

and

$$p(N_2|h) = \frac{b(h, N_2)}{p(h)} = \frac{H_2(h - H_1)}{(H_2 - H_1) h}$$

(38)

respectively. Note that $p(N|h)$ is the stationary probability that the population size is $N$ knowing that the average heterozygosity is $h$.

Now we have an explicit expression for $N_c$. By (11) and (29), we have

$$N_c = N_b + \frac{(N_b - N_b)(N_b - N_b)}{2(\gamma + u_0)N_b N_b + N_b}$$

(39)
The dependence of \( \log_{10}(\frac{N}{N_h}/\frac{N_1}{N_0}) \) in the two-valued Markov chain model for the case of \( \gamma_1 = \gamma_2, K = \infty \) and \( \theta_2 = 1 \). The solid line, the dotted line, and the dashed line represent the cases of \( \Gamma_2 = 0.1, \Gamma_2 = 1, \) and \( \Gamma_2 = 10 \), respectively.

Note that \( N_e \) depends on the mutation rate \( u \). By (39), \( N_e \) is an increasing function of the measure of autocorrelation \( \tau \) (a decreasing function of \( \gamma \) if we fix the value of \( r \) (note that \( N_e \) does not depend on \( \gamma \) if we fix the value of \( r \)). To see how much \( N_e > N_h \), we consider the ratio of the effective size to the harmonic mean

\[
\frac{N_e}{N_h} = 1 + \frac{r(1 - r)(N_h/N_0) - 1)^2}{(N_h/N_1) + [r + (1 - r)(N_h/N_0)](\Gamma_2 + \theta_2 e^\theta)}.
\]

(41)

where

\[
\Gamma_2 = 2N_e \gamma,
\]

and

\[
\theta_2 = 2N_e u.
\]

The dependence of \( \log_{10}(N_e/N_h) \) on \( \log_{10}(N_e/N_0) \) is presented in Figure 2 for the case of \( \gamma_1 = \gamma_2, K = \infty \), and \( \theta_2 = 1 \). The solid line, the dotted line, and the dashed line represent the cases of \( \Gamma_2 = 0.1, \Gamma_2 = 1, \) and \( \Gamma_2 = 10 \), respectively. The ratio \( N_e/N_h \) can be very large if \( N_1 \ll N_0 \) and the jump rate \( \gamma \) is not large. For example, \( N_e/N_h = 4.73 \) if \( N_2/N_1 = 100 \) and \( \Gamma_2 = 10 \). On the other hand, \( N_e/N_h = 161.91 \) if \( N_2/N_1 = 1000 \) and \( \Gamma_2 = 0.1 \). Note that \( N_e/N_h \) and \( \log_{10}(N_e/N_0) \) are decreasing functions of the scaled mutation rate \( \theta_2 \) and the values of \( \log_{10}(N_e/N_i) \) for \( \theta_2 < 1 \) are larger than those in Figure 2 if we fix \( \log_{10}(N_0/N_1) \) and \( \Gamma_2 \).

Note that

\[
N_e = (1 - r)N_1 + rN_2,
\]

(44)

we can express \( N_e \) in terms of \( N_1 \) and \( N_h \),

\[
N_e = R_k N_h + (1 - R_k)N_1,
\]

(45)

where

\[
R_k = \frac{N_h V}{N_h V + (1 + (u/\gamma) \theta)(N_e/N_0) - 1},
\]

(46)

and

\[
V = \int_0^\infty \text{Cov}\left[ \frac{1}{N(0)}, \frac{1}{N(t)} \right] dt = \tau r(1 - r)(\frac{1}{N_1} - \frac{1}{N_e})^2
\]

(47)

(0 < R_k < 1). Note that we have obtained explicit expressions for \( N_e \) [i.e., (39) and (45)] by introducing the two-valued Markov chain model.

The size of population \( N(t) \) may be very large in natural populations. Further, the autocorrelation of \( \{N(t)\}_{t=0} \) may be very weak or strong. In such cases, we can consider the asymptotic behavior of \( N_e \). For this end, we parameterize \( N_1, \gamma_2, N_0, N_w, N_r, \gamma_1, \gamma_2, \gamma, V, u, \) and \( R_k \) by \( \epsilon \) (\( \epsilon \to 0 \)) such as \( N_1, N_0, N_2, N_w, N_r, \gamma_1, \gamma_2, \gamma, \epsilon, V, u, \) \( \gamma \), \( \epsilon \), etc., \( R_k \). Here, we assume that \( r = \gamma_1/(\gamma_1 + \gamma_2) \) and \( b = N_w/N_1 \) do not depend on \( \epsilon \) for simplicity. Note that \( N_e/N_h \) does not depend on \( \epsilon \) in this case. For simplicity, we assume that the limits

\[
\Gamma = \lim_{\epsilon \to 0} 2N_e \gamma
\]

(48)

and

\[
\theta = \lim_{\epsilon \to 0} 2N_e u
\]

(49)

exist (0 ≤ \( \Gamma \), \( \theta \) ≤ \( \infty \)). By

\[
R_k = \frac{1}{1 + 2(N_e \gamma + N_w u + \theta)([r + b(1 - r)]/[1 - r + b\theta])}
\]

(50)

and (45), it is easy to see that

\[
\lim_{\epsilon \to 0} \frac{N_e}{N_h} = 1
\]

(51)

if \( \Gamma = \infty \) or \( \theta = \infty \),

\[
\lim_{\epsilon \to 0} \frac{N_e}{N_h} = 1
\]

(52)

if \( \Gamma = 0 \) and \( \theta = 0 \). The cases that \( \Gamma = \infty, 0 < \Gamma < \infty, \) and \( \Gamma = 0 \) are referred to as weak autocorrelation, moderate autocorrelation, and strong autocorrelation limits, respectively. In the same way, the cases that \( \theta = 0, 0 < \theta < \infty, \) and \( \theta = \infty \) are referred to as weak mutation, moderate mutation, and strong mutation limits, respectively. By (51) and (52), the effective size is asymptoti-
cally equal to the harmonic mean in the case of weak autocorrelation or strong mutation and the effective size is asymptotically equal to the arithmetic mean in the case of strong autocorrelation and weak mutation. Note that, in the case that \(r\) and \(b\) do not depend on \(\varepsilon\), weak, moderate, and strong autocorrelation limits correspond to \(\lim_{\varepsilon \to 0} N^k V^\varepsilon = 0\), \(0 < \lim_{\varepsilon \to 0} N^k V^\varepsilon < \infty\), and \(\lim_{\varepsilon \to 0} N^k V^\varepsilon = \infty\), respectively.

The Wright-Fisher model with fluctuating population size is investigated by Iizuka (2001). Let \(N^{(k)}(0)\) be the size of haploid population in generation \(k\), where \(\{N^{(k)}(t)\} = 0, \pm 1, \pm 2\) is a two-valued Markov chain. This model is defined as the Wright-Fisher model with no mutation and no selection whose population size in generation \(k\) is \(N^{(k)}(0)\). For this model, the effective size \(N^e\) can be defined by

\[
1 - \frac{1}{N^e} = \lim_{\varepsilon \to 0} \left[ E\left[ \prod_{j=0}^{k-1} \left( 1 - \frac{1}{N^{(j)}(t)} \right) \right] \right]^{1/2}. \tag{53}
\]

The effective size \(N^e\) is smaller than the harmonic mean if the change in population size is negatively autocorrelated, it is larger than the harmonic mean if the change is positively autocorrelated, and it is equal to the harmonic mean if the change is uncorrelated. These results are consistent with those for the diffusion model in this section. Note that the change in population size is positively autocorrelated in our formulation through the diffusion model in this section. In other words, the introduction of the fluctuation of population size after the diffusion approximation implies that this fluctuation cannot be uncorrelated or negatively autocorrelated. There are some properties on asymptotic behavior of the Wright-Fisher model with fluctuating population size that are consistent with those of the diffusion model in this section. For example, the effective size is asymptotically equal to the harmonic mean in the case of weak autocorrelation for both models (see IZUKA 2001).

The Wright-Fisher model with fluctuating population size is more fundamental than the diffusion model with fluctuating population size since the stochastic effect of fluctuation of population size is introduced after the diffusion approximation for the latter model. It seems not to be easy, however, to incorporate the effect of mutation into the Wright-Fisher model with fluctuating population size. It is easy to incorporate the effect of mutation into the diffusion model with fluctuating population size since the differential equation for \(H(t)\) is linear. Furthermore, we can consider a very general pattern of fluctuation of population size for the diffusion model with fluctuating population size as we have shown in this article.

**DISCUSSION**

Here, we discuss some biological relevance of our results. Suppose that we are interested in effects of selection on genetic variation in a species. Effects of weak selection depend on population size (Ohta 1973, 1992). Thus, it is important to obtain information on population size although this is usually very difficult by nongenetic means (see BASSER et al. 2001 for the problems that have to be dealt with by researchers when trying to estimate effective size using demographic parameters). One of the easiest things we can do is to estimate the effective size defined here by measuring the average heterozygosity at neutral loci such as those of pseudogenes. Then, we can guess what variation pattern would be expected for alleles with a selection coefficient \(s\). In fact, effects of selection depend not only on the effective size but also on the details of how population size changes (see OHTA 1997a, 1998). For example, the behavior of TAJIMA’s (1989) \(D\) as a function of the intensity of selection is very different if the change rate of population size differs with the effective size being kept constant (TACHIDA 2000). Nevertheless, we can use (39) to know what parameter combinations lead to the effective population size under the assumption of the two-state model and then examine effects of selection on the basis of this information. Although the two-state model is unrealistic and we need to extend theoretical studies for more general cases, at least we can obtain a rough idea as to how selection affects genetic variation in the species by measuring the effective population size. For the inference of the mechanism of molecular evolution under fluctuating population size, see also ARAKI and TACHIDA (1997) and OHTA (1997b).

TACHIDA (1985) developed a method to calculate the probabilities that two neutral genes taken at random from a population have certain allelic states, which is called the joint frequencies of alleles (see also GRIFFITHS 1981). Using (29), we can extend this method to the case where population size is fluctuating. Let \(q(k)\) be the probability that two neutral genes taken at random from a population have \(k\) mutations since they diverged from their most recent common ancestor. Then the probability generating function of \(q(k)\)

\[
\Theta(z) = \sum_{k=0}^{\infty} q(k) z^k \tag{54}
\]

is expressed as

\[
\Theta(z) = 1 - E[H(\infty)] e^{-u(1-z)}, \tag{55}
\]

where \(E[H(\infty)] e^{-u(1-z)}\) is the value of \(E[H(\infty)]\) with the total mutation rate \(u\) being replaced by \(u(1-z)\). Then we can obtain the joint frequencies of alleles by (6) of TACHIDA (1985). We denote the effective size by \(N_e(u)\) as a function of mutation rate \(u\). It is easy to see that the expectation of \(k\) is

\[
\langle k \rangle = 2N_e(0) u, \tag{56}
\]

since
\[ \langle k \rangle = \frac{d\Theta(z)}{dz}\bigg|_{z=1} = \frac{dE[H(\infty)]}{du}\bigg|_{u=0} u \quad (57) \]

and

\[ dE[H(\infty)] \bigg|_{u=0} = 2N(z,0) \quad (58) \]

where \( (d\Theta(z)/dz)\bigg|_{z=1} \) is the value of \( d\Theta(z)/dz \) when \( z = 1 \) and \( (dE[H(\infty)]/du)\bigg|_{u=0} \) is the value of \( dE[H(\infty)]/du \) when \( u = 0 \). The same method may be applied to microsatellite allele frequencies. For example, (11a) of Slatkin (1995) may be extended for the case where population size is fluctuating as follows. By (36) and (7) of Slatkin (1995), the right-hand side of (11a) of Slatkin (1995) can be expressed as \( 2N(z,0)\mu\sigma^2_u \), where \( \mu \) and \( \sigma^2_u \) are the parameters describing the mutation scheme for microsatellite alleles.

We thank A. Shimizu for noting the result of Dudley (1989) on a modification of Jensen’s inequality and two anonymous reviewers for valuable comments. M.I. was partially supported by a grant-in-aid (no. 12680391) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. H.T. was supported in part by a grant from Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRID) and a grant from Uehara Memorial Foundation.

LITERATURE CITED


Communicating editor: W. Stephan
APPENDIX A

Let \( f(x_1(t), x_2(t), \ldots, x_{K-1}(t)) \) be an arbitrary function of gene frequencies in the neutral diffusion model. We denote by \( \mathbb{E}[\cdot] \) the operation taking the expectation with respect to the random sampling drift. By the general theory of diffusion processes (see Ewens 1979, pp. 136–137, or Karlin and Taylor 1981, pp. 213–216), the expectation of \( f(x_1(t), x_2(t), \ldots, x_{K-1}(t)) \) satisfies

\[
\frac{d\mathbb{E}[f(x_1(t), x_2(t), \ldots, x_{K-1}(t))]}{dt} = \mathbb{E}[Lf(x_1(t), x_2(t), \ldots, x_{K-1}(t))],
\]

where

\[
Lf = \frac{1}{2} \sum_{i=1}^{K-1} x_i (\delta_{ij} - x_j) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{K-1} \frac{u}{K-1} (1 - Kx_i) \frac{\partial f}{\partial x_i},
\]

\( \delta_{ij} = 1 \) if \( i = j \) and \( \delta_{ij} = 0 \) if \( i \neq j \).

Substituting

\[
h(x_1, x_2, \ldots, x_{K-1}) = 1 - \sum_{i=1}^{K} x_i^2
\]

for \( f(x_1, x_2, \ldots, x_{K-1}) \) in (A1), we have

\[
Lh = -\frac{1}{N} h + 2u(1 - c_\theta h).
\]

Noting that

\[
H(t) = \mathbb{E}[h(x_1(t), x_2(t), \ldots, x_{K-1}(t))],
\]

we have (2) by (A1).

The neutral diffusion model with fluctuating population size is defined as a diffusion model with a random parameter by replacing \( N \) in (A2) with the stationary stochastic process \( N(t) \). In other words, this model is defined as a diffusion process in random environments. We have (7) in the same way as we have (2).

APPENDIX B

For a function \( f(x) \) with

\[
f(tx + (1 - t)y) < tf(x) + (1 - t)f(y)
\]

for any \( x, y \), and \( t \) \((x \neq y, 0 < t < 1)\), and for a random variable \( X \), we have

\[
f(\mathbb{E}[X]) < \mathbb{E}[f(X)]
\]

By applying (B2) for \( f(x) = e^{-x} \) \((x > 0)\), we have

\[
E[H(t)] > \frac{2N_0u}{1 + 2N_0uc_h} \left[ 1 - e^{-[(2u + 1)/N_0]t} \right] + H(0)e^{-[(2u + 1)/N_0]t}
\]

(B3)

and

\[
E[H(\infty)] > \frac{2N_0u}{1 + 2N_0uc_h}.
\]

(B4)

By (B4) and the definition of \( N_0 \) in (11), we have

\[
N_0 > N_c.
\]

(B5)

By using the inequality (B2) again for \( f(x) = -e^{-1/x} \) \((x > \frac{1}{2})\), we have

\[
E[H(t)] < \frac{2N_0u}{1 + 2N_0uc_h} \left[ 1 - e^{-[(2u + 1)/N_0]t} \right]
\]

+ \( H(0)e^{-[(2u + 1)/N_0]t} \)

(B6)

and

\[
E[H(\infty)] < \frac{2N_0u}{1 + 2N_0uc_h}.
\]

(B7)

Then we have

\[
N_0 < N_c.
\]

(B8)

by (B7) and (11).

APPENDIX C

For the two-valued Markov chain model, (7) can be expressed as

\[
\frac{dH(t)}{dt} = F_t(H(t))
\]

(C1)

if \( N(t) = N_0 \), where

\[
F_t(h) = -\frac{h}{N_0} + 2u(1 - c_\theta h) = -\frac{2u}{H_0}(h - H).
\]

(C2)

We can regard (C1) as an ordinary differential equation, one of whose parameters stochastically changes. Note that \( H_t \leq H(t) \leq H_z \) if \( H_t \leq H(0) \leq H_z \).

By (2.17) of Matsuda and Ishi (1981), the stationary probability density function \( p(h, N(t)) \) of \( (H(t), N(t)) \) is given by

\[
p(h, N) = \frac{C}{[F_t(h)]} \exp \left\{ -\int_{H_t}^h \left[ \frac{\gamma_1}{F_t(t)} + \frac{\gamma_2}{F_2(t)} \right] dt \right\},
\]

(C3)

\((H_t < h < H_z)\), where \( C \) is the normalization constant.

By this formula, we have (22) and (23).