SHORT PAPERS

The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population*

By MOTOO KIMURA

National Institute of Genetics, Mishima, Japan

(*Received* 25 *August* 1969)

SUMMARY

Frequency distribution of the length of time until fixation (excluding the cases of eventual loss) of a selectively neutral mutant in a finite population was obtained. With $4N_e$ generations (N_e, effective population size) as the unit length of time, the distribution has mean $\mu_1 = 1$, standard deviation $\sigma = 0\cdot538$, skewness $\gamma_1 = 1\cdot67$ and kurtosis $\gamma_2 = 4\cdot51$ (see Fig. 1).

A general theory concerning the average number of generations until a mutant gene becomes fixed (established) in a finite population was developed by Kimura & Ohta (1969*a*) based on the diffusion models. For the special case of selectively neutral mutants, the problem was simplified and it was shown by them that the average length of time until fixation (excluding the cases of eventual loss) is approximately $4N_e$ generations, where N_e is the 'variance' effective number of the population (cf. Kimura & Crow, 1963). In addition it was shown by Narain (1969) and also by Kimura & Ohta (1969*b*) that for this case the standard deviation of the length of time until fixation is about $(2\cdot15)N_e$ generations.

The purpose of this note is to show that, for a neutral mutant, the entire probability distribution of the length of time until fixation can readily be derived from my previous results on the process of random genetic drift. Actually, I have shown (Kimura 1955) that if $f(p,1;t)$ is the probability of a selectively neutral allele reaching fixation by the tth generation, then

$$f(p,1;t) = p + \sum_{i=1}^{\infty} (2i+1)\, p(1-p)\, (-1)^i\, \mathrm{F}(1-i, i+2, 2, p)\, e^{-\lambda_i} \quad (t > 0), \tag{1}$$

where p is the initial frequency of the mutant allele and F denotes the hypergeometric function, and $\lambda_i = i(i+1)/4N_e$.

If we restrict our consideration to the cases in which the mutant allele is eventually fixed but disregard those in which it is eventually lost from the population, then the cumulative probability distribution of the length of time until fixation is given by $f(p,1;t)/p$, since the probability of eventual fixation is $f(p,\,1;\infty) = p$. The corresponding density function denoted by $y(p, t)$ may then be obtained by differentiating this with respect to t, namely

$$y(p, t) = (1-p) \sum_{i=1}^{\infty} (2i+1)\, (-1)^{i+1}\, \mathrm{F}(1-i, i+2, 2, p)\, \lambda_i\, e^{-\lambda_i t} \quad (t > 0). \tag{2}$$

9-2

We are particularly interested in the case in which the initial frequency of the mutant gene is very low. This corresponds to the situation in which the population is large while the mutant allele is represented only by one or two individuals at the moment of its appearance. The distribution for this case may be approximated by $\lim_{p \to 0} y(p, t)$, which we will denote by $y(t)$. Thus we have

$$y(t) = \sum_{i=1}^{\infty} (2i+1)(-1)^{i+1} \lambda_i e^{-\lambda_i t} \quad (t > 0). \tag{3}$$

Note that F is reduced to unity at $p = 0$. The moments of the length of time until fixation can then be evaluated by using this distribution. Namely, let μ'_n be the nth moment ($n \geqslant 1$) around zero, then

$$\mu'_n = \int_0^{\infty} t^n y(t)\, dt = (4N_e)^n n! \sum_{i=1}^{\infty} \frac{2i+1}{[i(i+1)]^n}(-1)^{i+1}. \tag{4}$$

In particular, the first four moments turn out to be as follows:

$$\mu'_1 = 4N_e$$
$$\mu'_2 = (4N_e)^2[(\tfrac{1}{3}\pi^2)-2] \approx 1\cdot29(4N_e)^2$$
$$\mu'_3 = (4N_e)^3(12-\pi^2) \approx 2\cdot13(4N_e)^3$$
$$\mu'_4 = (4N_e)^4[(\tfrac{7}{15}\pi^4)+8\pi^2-120] \approx 4\cdot41(4N_e)^4.$$

For n larger than 4, we have, approximately,

$$\mu'_n \approx (4N_e)^n n!\, \frac{3}{2^n},$$

the first term of the series.

Figure 1 illustrates this probability distribution in terms of $T = t/(4N_e)$, namely, taking $4N_e$ generations as the unit length of time. This has mean $\mu_1 = 1$, standard deviation $\sigma = 0\cdot538$, skewness $\gamma_1 = \mu_3/\sigma^3 = 1\cdot67$ and kurtosis $\gamma_2 = \mu_4/\sigma^4 - 3 = 4\cdot51$. It somewhat resembles the gamma distribution.
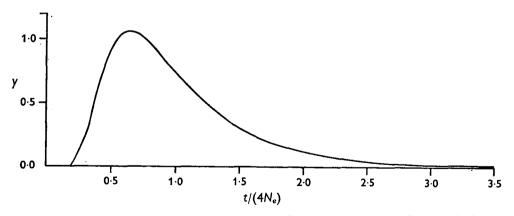

Fig. 1. Probability distribution of the length of time until fixation of a selectively neutral mutant. Abscissa: time measured with $4N_e$ generations as the unit length of time. N_e stands for the effective size of the population. Ordinate: probability density (y).

REFERENCES

KIMURA, M. (1955). Solution of a process of random genetic drift with a continuous model. *Proc. Natn. Acad. Sci., U.S.A.* **41**, 144–150.

KIMURA, M. & CROW, J. F. (1963). The measurement of effective population number. *Evolution* **17**, 279–288.

KIMURA, M. & OHTA, T. (1969*a*). The average number of generations until fixation of a mutant gene in a finite population. *Genetics, Princeton* **61**, 763–771.

KIMURA, M. & OHTA, T. (1969*b*). The average number of generations until extinction of an individual mutant gene in a finite population. *Genetics, Princeton* (in the Press).

NARAIN, P. (1969). A note on the diffusion approximation for the variance of the number of generations until fixation of a neutral mutant gene. *Genet. Res., Camb.* (in the Press).