

## 集団遺伝学の拡散過程とブラウン運動

誌名	The Japanese journal of genetics
ISSN	0021504X
著者	丸山, 毅夫
巻/号	48巻3号
掲載ページ	p. 231-234
発行年月	1973年6月

農林水産省 農林水産技術会議事務局筑波産学連携支援センター  
Tsukuba Business-Academia Cooperation Support Center, Agriculture, Forestry and Fisheries Research Council  
Secretariat



## SHORT COMMUNICATION

### DIFFUSION MODELS AND BROWNIAN MOTION IN POPULATION GENETICS

TAKEO MARUYAMA

National Institute of Genetics, Misima 411

*Received May 28, 1973*

#### 1. INTRODUCTION

The stochastic models of gene frequency change put forward by Fisher (1930) and Wright (1931 and later) have become increasingly important in population genetics and evolutionary theory. For reviews, see Crow and Kimura (1970); Kimura (1964); Kimura and Ohta (1971); Wright (1969). The details of the models are given in these books. The model investigated in this paper, in brief, is as follows. The population consists of  $N$  breeding individuals. We consider one locus with two alleles ( $A$  and  $a$ ). Let  $s$  be the selective advantage of  $A$  over  $a$ , so that the relative fitnesses of genotypes  $AA$ ,  $Aa$  and  $aa$  are  $2s$ ,  $s$  and  $0$  respectively. It is also assumed that neither mutation nor migration from outside occurs.

The probabilistic change of the gene frequencies is of our interest and it is conveniently treated as a diffusion process. Under the assumption of random mating, the model has been investigated mathematically, (cf. Crow and Kimura 1970, chapt. 8 and 9; Kimura 1955a, b, 1964; Wright 1931, 1969, chapt. 13 and 14). However when we take into account the effect of geographical structure of the population which prevents random mating as a whole, the model becomes formidably difficult to handle mathematically. On the other hand, a Brownian motion is a simple process and therefore its properties are fully known. Thus it is of interest if some of the stochastic processes of population genetics can be reduced to Brownian motions, and this is the main topic of this paper.

#### 2. RANDOM MATING MODEL AND BROWNIAN MOTION

We denote by  $t$  the time measured in generations. Let  $P(t, X, Y)$  be the probability density that the frequency of  $A$  is  $Y$  at time  $t$ , given that it is  $X$  at time  $0$ . Then it is accepted that the transition probability  $P(t, X, Y)$  satisfies the differential equation

$$\frac{\partial P}{\partial t} = \frac{X(1-X)}{4N} \frac{\partial^2 P}{\partial X^2} + sX(1-X) \frac{\partial P}{\partial X}, \quad (1)$$

where  $P \equiv P(t, X, Y)$ , (cf. Crow and Kimura 1970, p. 396). Although the fundamental

solutions of this equation have been obtained by Kimura (1955a, b), they are quite complicated.

Throughout this paper, we consider a stochastic process as a collection of sample paths, i.e., a collection of populations evolving through time under the same conditions. For the sample path method of stochastic processes, the reader is referred to Ito and McKean (1965). We denote by  $\omega$  a particular path (the evolutionary course of a particular population) and by  $X_{t,\omega}$  the frequency of  $A$  in path  $\omega$  at time  $t$ . Let us define the following quantity for each path:

$$\tau_{\omega} \equiv \int_0^t 2X_{\xi,\omega}(1-X_{\xi,\omega}) d\xi. \quad (2)$$

Biologically, the  $\tau_{\omega}$  is the sum of the heterozygote frequencies that appeared in path  $\omega$ , from time 0 until  $t$ . Mathematically,  $\tau_{\omega}$  is a non-negative additive functional and therefore this can be used as a time parameter of the stochastic process. Note that this parameter depends on individual path, and therefore the speed of the clock depends on the location of the space. It is often called a *stochastic clock*.

Using this new clock, let  $Q(\tau, X, Y)$  be the probability density that the frequency of  $A$  is  $Y$  at time  $\tau$  given that it is  $X$  at  $\tau=0$ . It follows from Ito and McKean (1965, pp. 164-169) that the  $Q(\tau, X, Y)$  satisfies the differential equation;

$$\frac{\partial Q}{\partial \tau} = \frac{1}{8N} \frac{\partial^2 Q}{\partial X^2} + \frac{s}{2} \frac{\partial Q}{\partial X}. \quad (3)$$

Note that the coefficients in this equation are constant, and therefore the process governed by  $Q(\tau, X, Y)$  is a Brownian motion. (The procedure used for the substitution of (3) for (1) is a local time substitution which depends on individual path, but it is not a transformation of the time variable which is applicable to all paths simultaneously.)

The appropriate solution of equation (3) is

$$Q(\tau, X, Y) = 2e^{-2S(X-Y)} \sum_{n=1}^{\infty} e^{-(n^2\pi^2+4S^2)\tau/8N} \sin n\pi X \sin n\pi Y, \quad (4)$$

where  $S=Ns$ .

### 3. GEOGRAPHICALLY STRUCTURED POPULATION

We assume that the population consists of a finite number of colonies, and they are connected by migration. We denote by  $N_{t,i}$  and  $x_{t,i}$  respectively the colony size and the frequency of  $A$  in colony  $i$  at time  $t$ . Let

$$X_t = \frac{1}{N} \sum_i x_{t,i} N_{t,i},$$

where  $N = \sum_i N_{t,i} = \text{constant}$  by assumption. We first consider the change of the  $X_t$  as a stochastic process. Let  $P(t, X, Y)$  be the density of the transition probability, then it satisfies the diffusion equation;

$$\frac{\partial P}{\partial t} = \frac{1}{4N^2} \left\{ \sum_i x_{t,i} (1-x_{t,i}) N_{t,i} \right\} \frac{\partial^2 P}{\partial X^2} + \frac{s}{N} \left\{ \sum_i x_{t,i} (1-x_{t,i}) N_{t,i} \right\} \frac{\partial P}{\partial X}, \tag{5}$$

where  $P \equiv P(t, X, Y)$ , (see Maruyama 1972). Let

$$\tau_\omega \equiv \frac{1}{N} \int_0^t \sum_i 2x_{\xi,i,\omega} (1-x_{\xi,i,\omega}) N_{\xi,i} d\xi, \tag{6}$$

where  $\omega$  indicates a path. As before, the  $\tau_\omega$  is the sum of the heterozygosity that appeared in path  $\omega$ . Just as equation (3) is substituted for (1), (5) is replaced by

$$\frac{\partial Q}{\partial \tau} = \frac{1}{8N} \frac{\partial^2 Q}{\partial X^2} + \frac{s}{2} \frac{\partial Q}{\partial X}, \tag{7}$$

where  $Q \equiv Q(\tau, X, Y)$  is the probability density that the frequency of  $A$  is  $X$  at  $\tau=0$  and it is  $Y$  at time  $\tau$ . Note that equations (3) and (7) are identical. Thus the process treated by the new measure ( $\tau$ ) of time is independent of the population structure, and the solutions of (7) are given by (4). It is interesting that the quantities given in (2) and (6) are respectively proportional to the additive genetic variance of the whole population and the weighted mean of the variances in colonies, and therefore time  $\tau$  is proportional to the total gain of fitness made by selection in each population.

Some of the consequences following from the new process  $X_\tau$  being independent of the geographical structure are as follows. (i) The total number of heterozygotes formed while the population is polymorphic is independent of the population structure. That due to a mutant gene is  $2N$  if  $s=0$ , and  $4N$  if  $Ns \gg 1$  and  $0 < s \ll 1$ . The higher moments of these quantities are also independent of the population structure. (ii) The number of heterozygotes summed over those generations in which the gene frequency in the entire population is specified is independent of the population structure. If we denote by  $Y$  the specified frequency, it is  $(1 - \exp(-4Ns(1-Y)) (1 - \exp(-2s)) / s(1 - \exp(-4Ns))$  for  $Y=1/2N, 2/2N, 3/2N, \dots, (2N-1)/2N$ . In particular, if  $s=0$ , it is  $2(1-Y)$ , while if  $Ns \gg 1$  and  $0 < s \ll 1$ , it is approximately 2 for all  $Y$ . These results follows from the differential equation  $(1/8N)f_{XX} + (s/2)f_X + \delta(X-Y) = 0$ , where  $\delta(\cdot)$  is the Dirac delta function. (iii) The ultimate fixation probability of a mutant gene is independent of the population structure. The special case  $s=0$  of (i) was found by Maruyama (1971), (iii) was suggested by Maruyama (1970), and the other properties were given in Maruyama (1972) based on the method of Maruyama and Kimura (1971).

ACKNOWLEDGMENTS

I am greatly indebted to Dr. M. Kimura for encouragement and helps offered at all times.

LITERATURE CITED

Crow, J. F., and M. Kimura, 1970 An introduction to population genetics theory. Harper and Row, New York.

- Fisher, R. A., 1930 The genetical theory of Natural selection. Clarendon Press, Oxford.
- Ito, K., and H. P. McKean, Jr., 1965 Diffusion processes and their sample paths. Springer, Berlin.
- Kimura, M., 1955a Solution of a process of random genetic drift with a continuous model. Proc. Nat. Acad. Sci. USA **41**: 144-150.
- Kimura, M., 1955b Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harbor Symp. Quant. Biol. **20**: 33-53.
- Kimura, M., 1964 Diffusion models in population genetics. Jour. Appl. Probability **1**: 177-232.
- Kimura, M., and T. Ohta, 1971 Theoretical aspects of population genetics. Princeton University Press, Princeton.
- Maruyama, T., 1970 On the fixation probability of mutant genes in a subdivided population. Genetical Research (Camb.) **15**: 221-225.
- Maruyama, T., 1971 An invariant property of a geographically structured population. Genetical Research (Camb.) **18**: 81-84.
- Maruyama, T., 1972 Some invariant properties of a geographically structured finite population: distribution of heterozygotes under irreversible mutation. Genetical Research (Camb.) **20**: 141-149.
- Maruyama, T., and M. Kimura, 1971 Some methods for treating continuous stochastic processes in population genetics. Japan. J. Genetics **46**: 407-410.
- Wright, S., 1931 Evolution in Mendelian populations. Genetics **16**: 97-159.
- Wright, S., 1969 Evolution and the genetics of populations. Vol. 2. Theory of gene frequencies. University of Chicago Press, Chicago.