A STOCHASTIC MODEL CONCERNING THE MAINTENANCE OF GENETIC VARIABILITY IN QUANTITATIVE CHARACTERS*

By Motoo Kimura

NATIONAL INSTITUTE OF GENETICS, MISHIMA-SHI, SHIZUOKA-KEN, JAPAN

Communicated by James F. Crow, July 15, 1965

The mechanism by which genetic variability is maintained in natural populations for quantitative characters is not well understood. Many observations show that there is a considerable amount of genetic variability in a large population, and unless the character is closely correlated with fitness the optimum is usually near the mean, with fitness decreasing as the distance from the mean increases. Probably, Fisher[1] was the first to investigate a model in which the fitness was assumed to decrease in proportion to the squared deviation from the optimum. This model was also used by Haldane[2] and Wright.[3] Robertson[4] has shown that if genes are maintained by overdominance, but act additively with respect to a quantitative character, then the optimum is at the mean and the decrease of fitness is proportional to the squared deviation from the optimum.

In all these treatments the relation between the mutation rate and the amount of genetic variability maintained is either ambiguous or left out of consideration. The purpose of this paper is to propose a new model which enables one to make predictions about the relations between mutation rate, genotypic variance, and genetic load or amount of selective elimination involved in the maintenance of genetic variability.

Assumptions and Mathematical Formulation.—The basic assumptions are:

(1) At every locus involved with the quantitative character under discussion, mutation can produce an infinite sequence of alleles. Every mutation may produce a new allele different from the pre-existing ones.

(2) The effect of a new allele on the quantitative character is only slightly different from the parent allele from which it was derived by a single mutational step.

(3) The genes are additive with respect to their effect on the quantitative character.

(4) The optimum phenotype is fixed, and fitness decreases in proportion to the squared deviation from the optimum.

Consider a particular locus. We denote by x the average effect of an allele on the quantitative character, taking the optimum as the origin. We assume that by mutation an allele having an average effect x changes to another allele with an average effect $x + \xi$ with probability density given by $f(\xi)$.

If we denote by $p(x,t)$ the relative frequency of the alleles having an average effect x in a large population at time t (measured in generations), then the rate of change in p per generation is given by the sum of the following two components:

(a) *Change due to mutation:* If μ is the mutation rate per gene per generation (assumed to be constant), the contribution to $\partial p/\partial t$ by mutation is

$$-\mu p(x,t) + \mu \int_{-\infty}^{\infty} p(x - \xi,t) f(\xi) d\xi. \qquad (1)$$

The first term gives the rate of loss of alleles having an average effect x by mutation

to other alleles having a different effect. The second term gives the rate of gain from the mutation to alleles with average effect x from other alleles.

(b) *Change due to selection:* We assume that the fitness of an individual with total genotypic value Y with respect to the quantitative character is less on the average by an amount KY^2 in comparison with those having a total genotypic value of 0 (the optimum), fitness being measured in Malthusian parameters.

Consider an allele having an average effect x and let Y' be the value of the genetic background. Since the relative fitness of a genotype with genotypic value $x + Y'$ is $-K(x + Y')^2$, if Y' is distributed with mean M' and variance V', then the relative fitness of an allele having average effect x is $-K[(x + M')^2 + V']$. Therefore, the rate of change in p by selection is

$$p(x,t)[-Kx'^2 + K\int_{-\infty}^{\infty} x'^2 p(x,t)dx], \tag{2}$$

where K is a positive constant and $x' = x + M'$.

Thus, $\partial p(x,t)/\partial t$ is expressed as the sum of (1) and (2). We are mainly interested in the frequency distribution at equilibrium which is denoted by $p(x)$ and for which $\partial p/\partial t = 0$.

Let

$$\mu\int_{-\infty}^{\infty} \xi f(\xi)d\xi = -m, \tag{3}$$

$$\mu\int_{-\infty}^{\infty} \xi^2 f(\xi)d\xi = v \qquad (>0), \tag{4}$$

and assume that the terms,

$$\mu\int_{-\infty}^{\infty} \xi^n f(\xi)d\xi \qquad (n \geqq 3), \tag{5}$$

are negligible. Then the contribution from mutation, i.e., (1), may be reduced as follows:

$$\begin{aligned} &-\mu p(x,t) + \mu\int_{-\infty}^{\infty} p(x - \xi,t)f(\xi)d\xi \\ &= -\mu p + \mu\int_{-\infty}^{\infty}\left\{p - \xi\frac{\partial p}{\partial x} + \frac{\xi^2}{2}\frac{\partial^2 p}{\partial x^2} - \cdots\right\}f(\xi)d\xi \\ &= m\frac{\partial p(x,t)}{\partial x} + \frac{v}{2}\frac{\partial^2 p(x,t)}{\partial x^2}. \end{aligned}$$

Thus, at equilibrium when $\partial p/\partial t = 0$, we obtain the following ordinary differential equation:

$$\frac{v}{2}\frac{d^2p(x)}{dx^2} + m\frac{dp(x)}{dx} + K(k^2 - x'^2)p(x) = 0, \tag{6}$$

where

$$k^2 = \int_{-\infty}^{\infty} x'^2 p(x)dx. \tag{7}$$

The above equation may also be written in the form:

$$\frac{v}{2}\frac{d^2p}{dx'^2} + m\frac{dp}{dx'} + K(k^2 - x'^2)p = 0, \tag{6'}$$

where $x' = x + M'$.

Putting

$$\left.\begin{aligned} p &= e^{-\alpha x'}U(x') \\ x' &= \beta y \end{aligned}\right\} \tag{8}$$

in which $\alpha = m/v$ and $\beta^4 = v/8K$, equation (6′) reduces to

$$\frac{d^2U}{dy^2} + \left(A - \frac{y^2}{4}\right)U = 0, \tag{9}$$

where

$$A = \frac{k^2}{4\beta^2} - \alpha^2\beta^2. \tag{10}$$

Equation (9) is a Weber equation. The solution of this which is nonnegative and satisfies the conditions $U(-\infty) = U(\infty) = 0$ is uniquely determined, except for the scaling factor C (>0), as

$$U(y) = Ce^{-y^2/4}. \tag{11}$$

Therefore,

$$p(x) = Ce^{-\alpha x' - x'^2/4\beta^2}. \tag{12}$$

Furthermore, it is known in the theory of the Weber equation that for the solution given by (11), the constant A in (9) is equal to $^1/_2$. This leads to

$$k^2 = \beta^2(2 + 4\alpha^2\beta^2). \tag{13}$$

On the other hand, substituting (12) in the right side of (7), we obtain

$$k^2 = C\sqrt{4\pi}\beta^3(2 + 4\alpha^2\beta^2)e^{\alpha^2\beta^2}. \tag{14}$$

Comparing (13) and (14), we get

$$C = e^{-\alpha^2\beta^2}/\sqrt{4\pi\beta^2},$$

and (12) becomes

$$p(x) = \frac{1}{\sqrt{4\pi\beta^2}}\, e^{-(x' + 2\alpha\beta^2)^2/4\beta^2}. \tag{15}$$

Thus, x is normally distributed with mean

$$-2\alpha\beta^2 - M' = -\frac{m}{\sqrt{2vK}} - M', \tag{16}$$

and variance

$$2\beta^2 = \sqrt{\frac{v}{2K}}. \tag{17}$$

Since a diploid organism has two genes at each locus, the distribution of the genotypic value among individuals is given by the normal distribution with mean

$$-\frac{2m}{\sqrt{2vK}} - M'' \tag{18}$$

and variance

$$2\sqrt{\frac{v}{2K}}, \tag{19}$$

where $-m$ and v are the amount of increase in mean and variance of the genotypic value per gene per generation. M'' is the mean of the genotypic value produced by the remaining loci.

Considering all relevant loci, the genotypic value is distributed normally with mean

$$M = -\Sigma_i \frac{2m_i}{\sqrt{2v_iK}} \tag{20}$$

and variance

$$V_G = \Sigma_i \sqrt{\frac{2v_i}{K}}, \tag{21}$$

where the subscript i refers to the ith locus.

The Genetic Load Associated with This Mechanism.—Since the selective disadvantage of a genotype with genotypic value Y is KY^2, the total genetic load for a diploid organism with respect to this quantitative character is

$$\begin{aligned} L &= KY^2 = K(V_G + M^2) \\ &= \Sigma_i\sqrt{2Kv_i} + 2\left(\Sigma_i \frac{m_i}{\sqrt{v_i}}\right)^2, \end{aligned} \tag{22}$$

where $m_i = -\mu\bar{\xi}_i$ and $v_i = \mu\overline{\xi_i^2}$.

In expression (22), the first term is the component caused by the increase in variance, while the second is caused by the mean being shifted away from the optimum. As might be expected, this is proportional to the mutation rate.

Relations among Observable Quantities.—If we denote by r the ratio of the new genotypic variance produced by mutation in one generation to the total genotypic variance, then

$$r = \frac{\Sigma 2v_i}{\Sigma\sqrt{\frac{2v_i}{K}}}. \tag{23}$$

From (22) we let $L_v = \Sigma\sqrt{2Kv_i}$, this being the genetic load associated with increased variance. Substituting this into (23) leads to

$$r = \frac{L_v\Sigma v_i}{(\Sigma\sqrt{v_i})^2} \geqq \frac{L_v}{n},$$

where n is the number of loci involved, so that

$$n \geqq L_v/r, \tag{24}$$

where the inequality changes to an equality when the v_i's are all equal.

The following relation also follows from this model. From (21) and (22) we obtain

$$V_G = L_v/K.$$

Therefore,

$$\frac{V_G}{V_T} = \frac{L_v}{KV_T} = \frac{L_v}{-\log_e w(\sqrt{V_T})} \tag{25}$$

where $w(\sqrt{V_T})$ is the relative selective value of an individual whose phenotype deviates from the optimum by a standard deviation of total phenotypic value.

Effect of Continued Truncation Selection.—If the quantitative character is subject to continued truncation selection, the relative fitness for this selection of alleles having an average effect x is $1 + sx$, where $s = z/b\sqrt{V_T}$. V_T is the total phenotypic variance, z is the ordinate at the truncation point, and b is the proportion saved, the latter two quantities being defined for the normal distribution with mean zero and unit variance. The contribution of this selection to $\partial p/\partial t$ is $s(x - \bar{x})$, where

$$\bar{x} = \int_{-\infty}^{\infty} xp(x,t)dx.$$

Thus, the equation corresponding to (6) is now

$$\frac{v}{2}\frac{d^2p(x)}{dx^2} + m\frac{dp(x)}{dx} + [K(k^2 - x'^2) + s(x - \bar{x})]p(x) = 0. \tag{26}$$

It is possible to show then that x is distributed normally with mean

$$\frac{s}{2K} - \frac{m}{\sqrt{2vK}}$$

and variance

$$\sqrt{\frac{v}{2K}}.$$

It follows that, under continued truncation selection, the mean of the genotypic value of an individual increases by an amount

$$\frac{zn}{Kb\sqrt{V_T}}$$

more than without the truncation selection, but the genotypic variance remains the same. Thus, the variance is more stable than the mean.

Summary.—A new model was proposed to explain the maintenance of genetic variability in quantitative characters. The model assumes that at every locus involved with the quantitative character, mutation can produce an infinite sequence of alleles and the effect of a new allele is only slightly different from the parental allele from which it was derived by a single mutational step. The new model is in sharp contrast with the conventional models in which mutation is assumed to occur only between a pair of alleles, say A or a. Together with the additional assumptions that the genes are additive with respect to the quantitative character, that the optimum phenotype is fixed, and that fitness decreases in proportion to the squared deviation from the optimum, the properties of the model were worked out, enabling one to make predictions about the relation between mutation rates, genotypic variance, and mutational load.

The author would like to thank Dr. J. F. Crow for his stimulating discussions, for help in the completion of the manuscript, and for suggesting the idea of assuming small mutational steps in quantitative characters. Thanks are also due to Dr. Felix M. Arscot for his help in the solution of a differential equation.

* Contribution no. 535 from the National Institute of Genetics.

[1] Fisher, R. A., *The Genetical Theory of Natural Selection* (Oxford: Clarendon Press, 1930).
[2] Haldane, J. B. S., *The Causes of Evolution* (London: Longmans, 1932).
[3] Wright, S., *J. Genet.*, **30**, 243–256 (1935).
[4] Robertson, A., *J. Genet.*, **54**, 236–248 (1956).

---

*FERTILIZATION IN DROSOPHILA, II. TIME OF INACTIVATION OF A GENE EFFECT**

BY PHILIP E. HILDRETH

LAWRENCE RADIATION LABORATORY AND DEPARTMENT OF ZOOLOGY, UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated by Curt Stern, July 2, 1965

The discovery that X irradiation of *Drosophilia melanogaster* embryos can inactivate the effect of a suppressor gene and thus influence developmental processes that are to unfold several days later was reported by Glass and Plaine in 1950.[1] In the same report it was also shown that irradiation of fertilized eggs, still in meiotic stages, produced similar results, the effect of the suppressor gene having been negated by X rays. Specifically, the experiments involved the gene *er* (erupt eyes) and its suppressor *Su-er*. Flies homozygous for *er* in most cases have large outgrowths of nonfaceted material near the center of one or both eyes, and in some cases several bristles may be present; in weaker manifestations of this character the facets may be merely disarranged, or there may be only an extra bristle at the anterior margin of the eye. When *er/er* individuals are homozygous for *Su-er*, the effect of the erupt gene is suppressed almost completely, but occasionally flies may show either the weak or extreme erupt phenotype. The above authors observed that when eggs or larvae from parents homozygous for both *er* and *Su-er* were irradiated with 1000 r of X rays, the effect of the suppressor gene was inactivated. As a result, about 90–100 per cent of the treated individuals had erupt eyes. Even when eggs only 8 ± 8 min old were irradiated, there was 100 per cent inactivation of the *Su-er* effect, but irradiation of eggs and sperm prior to fertilization failed to cause such inactivation. Assuming that these 8 ± 8-min-old eggs "could hardly have reached the first cleavage division, on the average," Glass and Plaine concluded that the entrance of the sperm into the egg immediately activates it to produce "at least one specific gene-initiated substance or morphogenetic system," and that this is inactivated by X rays. During the second and third larval instars, X rays had a progressively diminishing inhibitory effect on the action of *Su-er*, indicating "either that the primary gene product or substrate is gradually being used up, or that the morphogenetic system in which the product (or substrate) participates has advanced beyond the stage at which the product or substrate can modify it." Since this report, *Su-er* has been often considered the earliest-acting gene in