Jour. Ind. Soc. Ag. Statistics
44 (2), 1992 : 171–192

Genetic Variability under Step-wise Discrete Mutation and Stabilizing Selection

P. Narain
Indian Agricultural Research Institute
Pusa, New Delhi - 110 012
(Received : May, 1992)

Summary

To explain the nature of genetic variability for quantitative traits in infinitely large natural populations, a model involving step-wise mutation with discrete allelic effects and stabilizing selection of optimal type is considered. When the number of alleles at a locus is taken as finite instead of an infinitely large number, the properties of the equilibrium seem to change. In particular, cases of up to fifteen alleles at the locus are discussed in detail. The results obtained are more general and encompass on the one hand Turelli's [22] findings based on the 'house of cards' approximation for strong selection and on the other the results of the normal approximation for weak selection. The results of Slatkin [21] based on a five alleles approximation for intermediate selection are also made more exact by solving the set of recurrence equations without assuming that the outermost alleles are negligible in frequency. The results obtained bring out clearly the behaviour of the genetic variability and heterozygosity at equilibrium as the ratio of mutation and selection parameters change from very low values to very high values. It seems the number of alleles considered at each locus could be a crucial factor in mutation-selection balance equilibria in large natural populations unless selection forces are sufficiently large that no more than two alleles can segregate at the locus.

Key words : Discrete mutation, stabilizing selection, genetic variability.

Introduction

Experimental investigations on natural populations of several organisms have indicated abundant genetic variation for most of quantitative traits. To explain the nature of this genetic variability, mathematical analyses based on different models have been attempted by different workers since 1920s. One of the mechanisms proposed is the balancing between forces of stabilizing selection and mutation. The former acts against deviants from an optimal value,

and so eliminates genetic variability. The latter provides new deviants restoring this variability and leading to an equilibrium. For infinitely large populations (i.e. with no random drift), the models differ in the number of alleles at a locus, the mutation scheme and the nature of time parameter. Di-allelic multi-locus models were investigated by Latter [11] [12], Bulmer [3] [4] and Barton [1]. This led to the conclusion that the equilibrium genetic variance is independent of the allelic effects but depends on the mutation rate, intensity of selection and the number of loci. The infinitely many alleles model introduced by Kimura [7], on the other hand, considered a continuous time parameter where alleles are distinguished according to the distribution of their additive effects and their frequencies do not enter into the analysis. At equilibrium the distribution of effects is Gaussian with genetic variance depending upon the mutation rate, the variance of the mutational change, the intensity of selection and the number of loci. On the basis of Kimura's [7] results, Lande [9] [10] assumed that the distribution of allelic effects at all loci is multivariate normal and analysed mutation-selection balance of a single character as well as multiple characters in terms of mean vector and covariance matrix of the distribution. Turelli [22] introduced an alternative 'house of cards' approximation for such a problem. Based on the premise that the variance of the mutational effects at a locus is much larger than the genetic variance at that locus, such an approximation led to the prediction of equilibrium genetic variance which is identical to that of the diallelic model, thus indicating that the equilibrium genetic variance is independent of the number of alleles considered at a locus. The mutation-selection model of Lande's kind was also considered by Fleming [5] who found an approximation to the equilibrium density of gametic types on the assumption that the forces of mutation and selection are weak relative to recombination. Nagylaki [13] calculated several functionals of this equilibrium solution and discussed the range of validity of the approximations.

A discretized version of the continuum-of-alleles model of Kimura [7] was given by Narain and Chakraborty [16] [17], Slatkin [21] and Narain [14] [15]. The model of Narain and Chakraborty [16] [17] was essentially a step-wise mutation model with discrete allelic effects and stabilizing selection. While the number of alleles considered was taken as infinitely large, it was indicated therein, in discussion, that when we consider a finite number of alleles, it is possible to produce Turelli's [22] results based on the 'house of cards' approximation. Slatkin [21] also used the step-wise mutation model but introduced useful approximations for most parameter

values of interest encompassing at one extreme, the 'house of cards' approximation for strong selection, and at the other, the normal approximation for weak selection relative to mutation. He also discussed a five alleles approximation for intermediate selection and explored the effects of varying degrees of dominance on a quantitative character as well as of directional selection imposed on a population already at equilibrium under stabilizing selection.

In this paper, we consider in general the case of a finite number of alleles at the locus, including cases up to fifteen alleles, and show how the published results fit into this general framework. We assume throughout that linkage disequilibria are negligible and mean of the character coincides with the optimum. The problem of mutation-selection balance for quantitative characters can then be reduced to that of a single locus, as adopted in this paper. A more recent analysis of mutation-selection balance for quantitative characters by Keightley and Hill [6] emphasises consideration of several characters simultaneously to include pleiotropy. However, we restrict ourselves in this paper to a single character and propose to take the multivariate problem later.

2. *On Methodology*

In order to study the statistical properties of equilibrium distribution under mutation-selection balance for quantitative traits, one could adopt several alternative approaches. One is to model the recursion equations for gene frequencies themselves and obtain the allele frequency profile. This was done in Narain and Chakraborty [16] [17] for the step-wise mutation model with one possible mutational step i.e. m = 1. The second approach to model the recursion equations for moments was also used in this work when m is greater than one. The moments of the allelic effects as well as genotypic effects were obtained, in particular, for the even order moments. It was noted then that the recursion equations for moments at a particular level depend on the higher order moments. For instance, the change in the second moment depends on the second as well as fourth moment. The second alternative approach was given a general treatment in Barton & Turelli [2] and came to be known as adaptive landscape approach (Turelli & Barton, [23]). A third and simpler alternative approach was provided by Price [20] using Price [18] [19] equation. However, it is now recognized that the second and third approaches, involving recursion equations for the moments of allelic effects, would be useful only when the higher order moments are expressible as simple functions of lower order

moments so that the recursion equations become a closed system. In Narain and Chakraborty [16] [17], Barton and Turelli [2] and Turelli and Barton [23], normal and house-of-cards approximations provided a way to close the system but then this approach obscures the dynamic nature of the process. Ultimately, therefore, one has to adopt the first approach and model the changes in allele frequencies directly. This is the method we have mainly used in this paper also.

3. *Step-Wise Discrete Mutation Model*

For a given locus, let A_i represent an allele occupying state i (any integer number from $-\infty$ to ∞) and having an allelic effect of ia, as shown in Fig. 1.

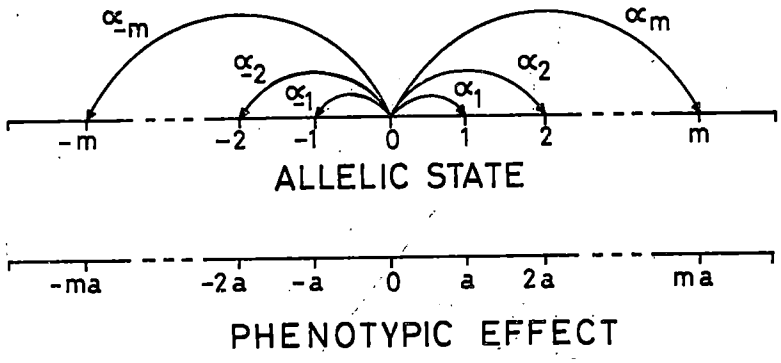

Fig. 1. Discrete alleiic-state model used in this paper. In this model allele A_i mutates to A_{i+r} with probability α_r (= α_{-r}). Allele A_i has a phenotypic effect of ai.

We assume that all allelic effects are additive with no dominance and no epistasis and that once A_i mutates, it changes to allelic state A_{i+r} with probability :

$$\alpha_r = \alpha_{-r} = \binom{2m}{m-r}\left(\frac{1}{2}\right)^{2m} \quad \text{for} \quad 0 \le r \le m$$

$$\alpha_r = 0, \text{ otherwise} \qquad (1)$$

where m is the possible number of mutational steps. The distribution is a shifted binomial with mean zero and variance m/2.

If v denotes the mutation rate, the absolute probability of such a mutation would be $v\alpha_r$. Thus, an allele that mutates has the same allelic effect as that of the original allele with probability $v\alpha_0$, so that in the conventional definition, the real mutation rate v' would be given by $v' = (1-\alpha_0)v$. The per generation increment of the variance of allelic effect by mutation is then $vma^2/2$. This is haploid variance, the corresponding value for diploids would be vma^2.

4. *Stabilizing Selection of Optimal Type*

Selection operates on the total phenotypic value x. In order to obtain the mean fitness of the individual with a given genotype, we assume that all genotypes experience the same environmental variance σ_e^2. The fitness function for the character value x is assumed to be Gaussian:

$$w(x) = x_{max} \exp\left[-\frac{(x-x_{opt})^2}{2\sigma_w^2}\right] \qquad (2)$$

where the character assumes the optimum fitness x_{max} at $x=x_{opt}$ and σ_w is the width of the function indicating the rate at which fitness declines with deviation of x from the optimum. Taking $w_{max}=1$, the mean fitness of the individuals with genotype A_iA_j having genotypic value a(i+j), would be

$$w_{ij} \alpha \exp\left[-s\left\{a(i+j)-x_{opt}\right\}^2\right] \qquad (3)$$

where $s=\frac{1}{2}(\sigma_w^2+\sigma_e^2)$ indicates the strength of the selection at the group level. A large σ_w means weak selection of the stabilizing type.

5. *Recurrence Relations and Equilibrium*

If $x_i(t)$ denotes the frequency of allele A_i in generation t with allelic effect ai and if we take the optimum phenotype to be at the origin i.e. $x_{opt} = 0$, an individual of genotype A_iA_j will have a mean reproductive fitness $w_{ij} = \exp[-sa^2(i+j)^2]$ and thus, the change in gene frequency of A_i from generation t to (t+1) is given by

$$\bar{w}(t)x_i(t+1) = (1-v+v\alpha_0)\sum_j x_i(t)\,x_j(t)\exp[-sa^2(i+j)^2]$$

$$+ v\sum_{r=1}^{m}\alpha_r\left[\sum_j x_j(t)\,[x_{i+r}(t)\exp\{-sa^2(i+j+r)^2\}\right.$$

$$\left. + x_{i-r}(t)\exp\{-sa^2(i+j+r)^2\}\,]\right] \qquad (4)$$

where $\bar{w}(t)$ is the mean fitness of individuals at the locus in the t-th generation so adjusted as to make $\sum_j x_j(t) = 1$.

In general, this recurrence relationship does not yield any explicit solution. However, for m=1 it is possible to derive the equilibrium allele frequency profile by neglecting powers of v and s. Then, the mean fitness $\bar{w}(t)$ is approximated as

$$\bar{w}(t) \simeq 1 - s\,\sigma_g^2(t) \qquad (5)$$

where $\sigma_g^2(t)$ is the total genotypic variance contributed by this locus at time t and is given by

$$\sigma_g^2(t) = a^2\sum_i\sum_j x_i(t)\,x_j(t)\,(i+j)^2 \qquad (6)$$

Since the optimum is at the origin, we have initially $x_i(0) = x_{-i}(0)$. Then for all i, $x_i(t) = x_{-i}(t)$ of each generation, this property being invariant to the transformation (4). From the symmetry of the model, we expect the equilibrium to be globally stable for m=1. In this case (4) reduces to

$$x_i(t+1) = \frac{v}{2}[x_{i+1}(t)+x_{i-1}(t)] + x_i(t)\left[1-v-s\left\{a^2i^2-\frac{\sigma_g^2(t)}{2}\right\}\right] \qquad (7)$$

On re-arrangement, we get

$$\Delta x_i(t) = x_i(t+1) - x_i(t)$$

$$= -v\left[x_i(t) - \frac{x_{i+1}(t)+x_{i-1}(t)}{2}\right] + sx_i(t)\left[\frac{\sigma_g^2(t)}{2} - a^2i^2\right] \qquad (8)$$

When the population reaches equilibrium under the opposing pressures of mutation and selection, $\Delta x_i = 0$. In this case, a complete solution for the allele frequency profile is given in the Appendix A of Narain and Chakraborty [17]. The heterozygosity (H_e) is then given by

$$H_e = 1 - x_0^2 - 2\sum x_i^2 \tag{9}$$

m-step mutational changes

In the general case of m-step mutational changes, the moments of the allelic effects as well as those of genotypic effects can be obtained analytically under optimum selection and the same assumptions for v and s as those used for m=1. Denoting the k-th moment of the distribution of allelic effects at a locus in the t-th generation by

$$M_k(t) = \sum_{i=-\infty}^{\infty} a^k i^k x_i(t), \tag{10}$$

the recurrence relationship for the even order moments is approximately given by

$$\begin{aligned}\Delta M_{2k}(t) \simeq\ & -v\left[1 - \binom{2m}{m}\left(\frac{1}{2}\right)^{2m}\right] M_{2k}(t-1) \\ & + s\left[M_2(t-1)M_{2k}(t-1) - M_{2k+2}(t-1)\right] \\ & + 2v\sum_{l=0}^{k}\binom{2k}{2l} M_{2l}(t-1)\sum_{r=1}^{m}\binom{2m}{m-r}\left(\frac{1}{2}\right)^{2m}(ar)^{2k-2l}\end{aligned} \tag{11}$$

The above result on recursion equations for moments of allelic effects can be directly obtained with the help of Price[18][19] equation. If $\bar{z}$ denotes the population average of a character or for that matter z can be any other function (square, cubic, quadratic etc.) of a character value, the Price equation is given by

$$\bar{w}\,\Delta\,\bar{z} = \mathrm{Cov}(w, z) + E(w\Delta z) \tag{12}$$

where the first term on the right hand side indicates the effect of selection (the relationship between fitness and character value) whereas the second term gives the effect of mutation (due to transmission of characteristic from parents to offspring). For the dynamics of the 2k-th non-central moment of the character z, this becomes

$$\overline{w}\,\Delta\,\overline{z^{2k}} = \mathrm{Cov}(w, z^{2k}) + E(w\Delta\, z^{2k}) \tag{13}$$

In the context of the problem under consideration, z is ai for effect of allele A_i and, assuming additivity of effects, is a (i+j) for the genotype A_iA_j, whereas w is approximated by $1 - sa^2 (i+j)^2$. It can be shown that

$$\mathrm{Cov}(w, z^{2k}) = s\left[\sigma_g^2\,\overline{z^{2k}} - \overline{z^{2k+2}}\right] \tag{14}$$

$$E(w\Delta\, z^{2k}) = -v(1-ao)\overline{z^{2k}} + 2v\sum_{r=1}^{m} a_r \sum_{l=0}^{k} \binom{2k}{2l}\overline{z^{2l}}(ar)^{2k-2l} \tag{15}$$

Combining the two, using (13), approximating, replacing z^{2k} by M_{2k} for t-th and (t–1)-th generations, and noting the value of α_r given by (1), we get (11).

The change of variance of allelic effects at generation t, $\Delta M_2(t)$ is then given by

$$\Delta M_2(t) = \frac{mva^2}{2} + s[M_2^2(t) - M_4(t)], \tag{16}$$

indicating that it depends on the fourth moment and therefore cannot lead to any solution. The system could, however, be closed by normal approximation where we have

$M_4 = 3M_2^2$ leading to

$$\Delta M_2(t) \simeq \frac{mva^2}{2} - 2sM_2^2(t)$$

$$= 2s\left[\frac{mva^2}{4s} - M_2(t)\right] \tag{17}$$

At equilibrium in such a case, therefore,

$$\hat{M}_2 = \left[\frac{mva^2}{4s}\right]^{1/2} \tag{18}$$

The genotypic variance at a locus in the equilibrium population, being $2\hat{M}_2$, is then

$$\hat{\sigma}_g^2 = \left[\frac{mva^2}{s}\right]^{1/2} = \left[\frac{\sigma_m^2}{s}\right]^{1/2} \tag{19}$$

σ_m^2 being the per generation incremental variance due to mutation.

This result is identical to that of Kimura [7] even though his model assumes continuous time parameter and a continuous distribution of allelic effects. Such an equivalence lends faith in the step-wise mutation model as a more realistic model even though it may not have any intrinsic biological interest.

6. *Equilibrium with a Finite Number of Alleles*

Let us now consider, for m=1, a finite number (2k+1) alleles A_{-k}, $A_{-(k-1)}$, ..., A_{-1}, A_0, A_1, ..., $A_{(k-1)}$, A_k at each locus. The change in the frequency of i-th allele then becomes

$$\Delta x_i(t) = -v\left[x_i(t)\left\{1-\frac{x_k(t)+x_{-k}(t)}{2}\right\}-\frac{x_{i+1}(t)+x_{i-1}(t)}{2}\right]$$

$$+ sx_i(t)\left[\sum_{j=-k}^{k} a^2j^2x_j(t)+2\left(\sum_{j=-k}^{k} ajx_j(t)\right)^2 - a^2i^2-2ai\sum_{j=-k}^{k} ajx_i(t)\right] \tag{20}$$

When the population reaches equilibrium, $\Delta x_i(t) = 0$ giving $\hat{x}_i = \hat{x}_{-i}$ for each i, the equilibrium mean is zero and we have with $\hat{x}_0 + 2\sum_{i=1}^{k} \hat{x}_i = 1$,

$$v\left[\hat{x}_i(1-\hat{x}_k)-\frac{\hat{x}_{i+1}+\hat{x}_{i-1}}{2}\right] = s\hat{x}_i\left[\sum_{j=-k}^{k} a^2j^2\hat{x}_j - a^2i^2\right],\ i=0, 1, 2, \ldots,k \tag{21}$$

This general result can be studied either for a general k-value or for each value of k separately.

In the former case, equation (21) can be re-arranged to give

$$(\hat{x}_{i+1} - 2\hat{x}_i + \hat{x}_{i-1}) + 2\hat{x}_i\hat{x}_k = \frac{2s}{v}[a^2i^2 - \sigma_h^2]\hat{x}_i \tag{22}$$

where σ_h^2 is the haploid equilibrium variance at the locus, being $\frac{\sigma_g^2}{2}$ and is therefore given by

$$\sigma_h^2 = a^2 \sum_j i^2 x_i = a^2 \sum_j j^2 x_j \tag{23}$$

The first term on the left of (22) can be approximated as $\frac{d^2\hat{x}_i}{di^2}$ so as to give the following differential equation

$$\frac{d^2\hat{x}_i}{di^2} + 2\left[\hat{x}_k + \frac{s}{v}(\sigma_h^2 - a^2 i^2)\right]\hat{x}_i = 0 \tag{24}$$

For the infinitely many alleles case, there is no term involving $\hat{x}_k$ in this equation which then reduces to Weber equation given by Kimura (1965) for the continuum-of-alleles model. The solution of the differential equation, in this case, gives rise to a normal distribution for equilibrium allele frequency given by

$$\hat{x}_i = \frac{1}{(2\pi\sigma_h^2)^{1/2}} \exp\left(-\frac{a^2i^2}{2\sigma_h^2}\right) \tag{25}$$

This result of Kimura [7] is, however, true for $a^2 << (v/2s)$ when we have weak selection and mutation is a much stronger force. If we put

$$\beta = \left(\frac{v}{2sa^2}\right), \tag{26}$$

this condition becomes $\beta >> 1$. The equilibrium value of σ_g^2 is then approximately

$$\hat{\sigma}_g^2(GA) = 2\left(a^2\frac{v}{2s}\right)^{1/2}$$

$$= \left(a^2\frac{2v}{s}\right)^{1/2} = 2a^2\sqrt{\beta}$$

where GA in parenthesis stands for 'Gaussian Approximation'.

Expressing in units of $4a^2$,

$$\frac{\hat{\sigma}_g^2(GA)}{4a^2} = \frac{\sqrt{\beta}}{2} \qquad (27)$$

The house-of-cards approximation due to Turelli [22] in which selection is a much stronger force than mutation, is based on the reverse condition of $a^2 >> (v/2s)$ i.e. $\beta << 1$ and leads to the equilibrium genetic variance

$$\hat{\sigma}_g^2(HCA) = \frac{2v}{s}$$

where HCA in parenthesis stands for 'House-of-Cards Approximation'.

Expressing in units of $4a^2$

$$\frac{\hat{\sigma}_g^2(HCA)}{4a^2} = \beta \qquad (28)$$

which agrees with the result of diallelic case (Bulmer, [3]) as well as triallelic case (Turelli [22]; Narain & Chakraborty [17]; Slatkin[21]).

(a) Three alleles

For k=1, we have three alleles A_{-1}, A_0, A_1 with respective equilibrium frequencies $\hat{x}_1 = \hat{x}_{-1}$ and $\hat{x}_0$ so that in view of $\hat{x}_0 + 2\hat{x}_1 = 1$ we have only one unknown say $\hat{x}_1$ to be determined from equations (21) with k=1 i.e.

$$v[\hat{x}_0(1-\hat{x}_1) - \hat{x}_1] = s\hat{x}_0[2a^2\hat{x}_1]$$

$$v\left[\hat{x}_1(1-\hat{x}_1) - \frac{\hat{x}_0}{2}\right] = s\hat{x}_1[2a^2\hat{x}_1 - a^2] \qquad (29)$$

Eliminating $\hat{x}_0$ leads to the quadratic equation in $\hat{x}_1$ given by

$$2(1+\beta)\hat{x}_1^2 - (1+4\beta)\hat{x}_1 + \beta = 0 \tag{30}$$

This gives two roots

$$\hat{x}_1 = \frac{(1+4\beta) \pm (1+8\beta^2)^{1/2}}{4(1+\beta)} \tag{31}$$

the permissible root being one with $\hat{x}_1 \leq \frac{1}{2}$ otherwise $\hat{x}_0$ would be negative.

Fo $\beta << 1$, we have approximately one permissible root as

$$\hat{x}_1 = \beta = \frac{v}{2sa^2} \tag{32}$$

The equilibrium genetic variance at the tri-allelic locus can then be shown to be equal to

$$\hat{\sigma}_g^2 = 2a^2\left[\hat{x}_0(1-\hat{x}_0) + 4\hat{x}_1\hat{x}_{-1}\right]$$

$$= 4a^2\hat{x}_1 = \frac{2v}{s} \tag{33}$$

Expressed in units of $4a^2$,

$$\frac{\hat{\sigma}_g^2}{4a^2} = \hat{x}_1$$

$$= \frac{v}{2sa^2}$$

$$= \beta \tag{34}$$

The heterozygosity in this case is given by

$$H_e = 2\beta(2-3\beta)$$

$$\simeq 4\beta \tag{35}$$

For n loci, under the assumption of approximate global linkage equilibria, we then have

$$V_g \simeq 4 \sum_{i=1}^{n} v_i (\sigma_w^2 + \sigma_e^2) = \sum_{i=1}^{n} 2(v_i/s) \tag{36}$$

In the notation of Turelli [22], $V_s = \sigma_w^2 + \sigma_e^2 = 1/2s$ so that (36) corresponds exactly to the relation (3.29) of his paper. Further, the condition $\beta << 1$ amounts to $v << 2sa^2 = a^2/V_s$ which corresponds exactly to the relation (3.27) in Turelli [22]. This is the same condition which leads to the house-of-cards approximation already stated in (28). As such, the analysis of a model with tri-allelic loci reported in Turelli [22], giving the house-of-cards prediction for the equilibrium genetic variance, happens to be a particular case of the step-wise discrete mutation model with m=1 and k=1.

(b) *Five alleles*

For k = 2, we have now five alleles A_{-2}, A_{-1}, A_0, A_1 and A_2 with respective equilibrium frequencies $\hat{x}_2 = \hat{x}_{-2}$, $\hat{x}_1 = \hat{x}_{-1}$ and $\hat{x}_0$ so that in view of

$$\hat{x}_0 + 2(\hat{x}_1 + \hat{x}_2) = 1, \tag{37}$$

we have two unknowns say x_1 and x_2 to be determined from equations (21) with k = 2 i.e.

$$v[\hat{x}_0 (1 - \hat{x}_2) - \hat{x}_1] = s\hat{x}_0 [2a^2 (\hat{x}_1 + 4\hat{x}_2)] \tag{38}$$

$$v\left[\hat{x}_1 (1 - \hat{x}_2) - \frac{\hat{x}_2 + \hat{x}_0}{2}\right] = s\hat{x}_1 [2a^2 (\hat{x}_1 + 4\hat{x}_2) - a^2] \tag{39}$$

$$v\left[\hat{x}_2 (1 - \hat{x}_2) - \frac{\hat{x}_1}{2}\right] = s\hat{x}_2 [2a^2 (\hat{x}_1 + 4\hat{x}_2) - 4a^2] \tag{40}$$

Expressing $\hat{x}_0$ in terms of $\hat{x}_1$ and $\hat{x}_2$ in (38), gives

$$2\hat{x}_1 + 2(\beta + 4)\hat{x}_2^2 + 2(\beta + 5)\hat{x}_1\hat{x}_2 - (1 + 3\beta)\hat{x}_1 - (4 + 3\beta)\hat{x}_2 + \beta = 0 \tag{41}$$

whereas from (40), $\hat{x}_1$ can be expressed in terms of $\hat{x}_2$ as

$$\hat{x}_1 = \frac{(\beta+2)\hat{x}_2 - (\beta+4)\hat{x}_2^2}{\frac{\beta}{2} + \hat{x}_2} \tag{42}$$

Eliminating $\hat{x}_1$ between (41) and (42) leads to a cubic equation in $\hat{x}_2$ given by

$$4(\beta^3+6\beta^2+12\beta+12)\hat{x}_2^3 - 2(6\beta^3+23\beta^2+20\beta+12)\hat{x}_2^2 + \beta(9\beta^2+14\beta+4)\hat{x}_2 - \beta^3 = 0 \tag{43}$$

We examine the positive and permissible roots of this equation which are consistent with (37). It is obvious from (37) that, in order that $\hat{x}_0$ should be positive, the sum of the roots, $\hat{x}_1$ and $\hat{x}_2$ should be less than 0.5. This is satisfied for that pair of roots, out of the two possible pairs, for which $\hat{x}_1$ has a value less than 0.5. This therefore, gives rise to only one unique root for $\hat{x}_1$ and $\hat{x}_2$. For different values of β say from β=0.02 to β=100, such values of the roots as well as the corresponding genetic variance in equilibrium, expressed in units of $4a^2$ and given by

$$\gamma = \frac{\hat{\sigma}_g^2}{4a^2} = \hat{x}_1 + 4\hat{x}_2 \tag{44}$$

are worked out numerically. The heterozygosity given by

$$H_e = 1 - \hat{x}_0^2 - 2(\hat{x}_1^2 + \hat{x}_2^2) \tag{45}$$

is also determined numerically. The results are given in Table 1.

It is apparent from this table that the equilibrium genetic variance is close to β only when β is very small. The value of β as we have seen is the genetic variance with three alleles as well as under house-of-cards approximation given by Turelli [22]. In fact for β close to zero, this result holds almost exactly for five alleles also. But when β takes moderate values greater than 0.1 but still less than 1.0, the equilibrium genetic variance under five alleles case is much smaller than that predicted under three allele case. For β=0.9, this reduction is as high as 55 per cent. The results in this table also indicate the

Table 1. Gene frequencies, Genetic variance [γ] and Heterozygosity [H_e] at equilibrium with five alleles

β	$\hat{x}_1$	$\hat{x}_2$	γ	H_e
0.02	0.0192	0.0001	0.0196	0.0746
0.03	0.0284	0.0002	0.0293	0.1082
0.04	0.0367	0.0004	0.0382	0.1395
0.05	0.0453	0.0006	0.0475	0.1687
0.06	0.0531	0.0008	0.0563	0.1959
0.07	0.0608	0.0011	0.0651	0.2213
0.08	0.0681	0.0014	0.0735	0.2450
0.09	0.0751	0.0017	0.0818	0.2670
0.10	0.0818	0.0020	0.0899	0.2876
0.20	0.1347	0.0066	0.1611	0.4322
0.30	0.1685	0.0122	0.2172	0.5083
0.40	0.1906	0.0179	0.2621	0.5513
0.50	0.2053	0.0234	0.2991	0.5773
0.60	0.2154	0.0287	0.3302	0.5942
0.70	0.2226	0.0336	0.3569	0.6058
0.80	0.2278	0.0381	0.3803	0.6141
0.90	0.2316	0.0423	0.4009	0.6203
1.00	0.2344	0.0462	0.4194	0.6250
2.00	0.2427	0.0731	0.5363	0.6436
2.30	0.2429	0.0784	0.5565	0.6456
2.40	0.2429	0.0800	0.5627	0.6462
2.50	0.2429	0.0814	0.5686	0.6467
2.60	0.2428	0.0828	0.5742	0.6471
3.00	0.2425	0.0878	0.5936	0.6486
5.00	0.2402	0.1030	0.6522	0.6522
10.00	0.2370	0.1171	0.7053	0.6547
20.00	0.2348	0.1251	0.7353	0.6558
40.00	0.2335	0.1295	0.7513	0.6563
60.00	0.2330	0.1309	0.7568	0.6565
80.00	0.2328	0.1317	0.7595	0.6566
100.00	0.2326	0.1321	0.7612	0.6566

values of β for which the five alleles case provides values close to the house-of-cards approximation by using the criterion that $\hat{x}_2$ must be sufficiently small as adopted by Slatkin [21]. With $\beta \geq 1$ corresponding to the situation when mutation is a much stronger force than selection in maintaining the genetic variability, it is found that the equilibrium frequency of A_1 (i.e. $\hat{x}_1$), which usually increases with β, becomes approximately stationary around β close to 2.5 and starts decreasing thereafter. The values of $\hat{x}_2$ are now not negligible and the equilibrium genetic variance for five alleles case is distinctly different from the three alleles case. This is also not close to $\sqrt{\beta}/2$, the value predicted under the normal approximation due to Kimura [7].

(c) *More than five alleles*

For k>2, the algebraic manipulation of equation (21) becomes formidable. We therefore have to resort to numerical iterations of the equations. This was done for k = 3, 4, 5, 6 and 7. The numerical results for the equilibrium genetic variance (γ) and heterozygosity (H_e) are shown in Figures 2 and 3 using a log scale for β. The graphs for k = 1 and 2 corresponding to 3 and 5 alleles cases are also drawn for the sake of comparison.

It is apparent from Figure 2 that in all the cases of 3 to 15 alleles considered, the equilibrium genetic variance increases with increase in the value of β. Initially, for smaller values of β, such increases are linear with a slope equal to β itself irrespective of the number of alleles. This confirms the contention of Turelli [22]. But for higher values of β, the curves are characteristic in increases and depend on the number of alleles considered. The curve for 3 alleles case departs from the rest of the curves around $\beta = 0.1$. The curve for 5 alleles case departs from the rest of 5 curves for 7, 9, 11, 13 and 15 alleles at around $\beta = 0.5$. A similar pattern is observed as we move on to curves for higher number of alleles. After the bifurcation each curve increases in a characteristic manner. In the literature β=1.0 is taken as the threshold below which, two alleles approximation holds and above which normal approximation is assumed. Our results however show that close to β=1, even three alleles approximation differs from the two alleles one and there are significant variations in the behaviour above β=1. The highest values are found for the 15 alleles case and the lowest for the 3 alleles case. The values for k=7 corresponding to 15 alleles case indicate closeness to the limiting value of $\sqrt{\beta}/2$ given by Kimura [7].

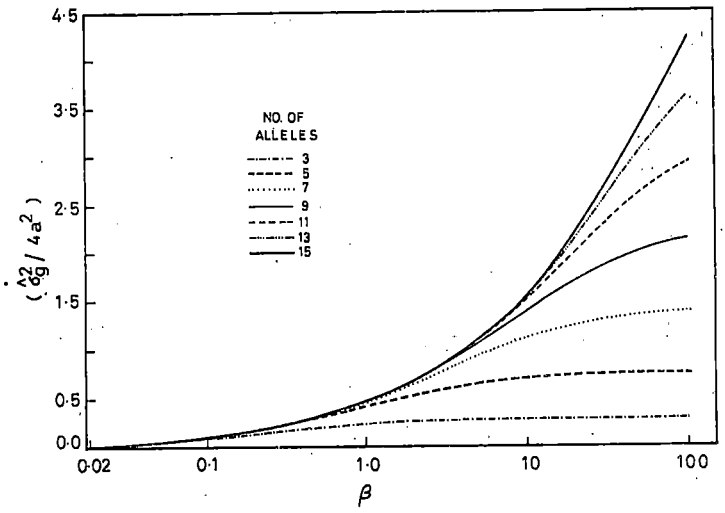

Fig. 2. Variation of equilibrium genetic variance in units of $4a^2$ with β (plotted on a log scale) for different number of alleles.

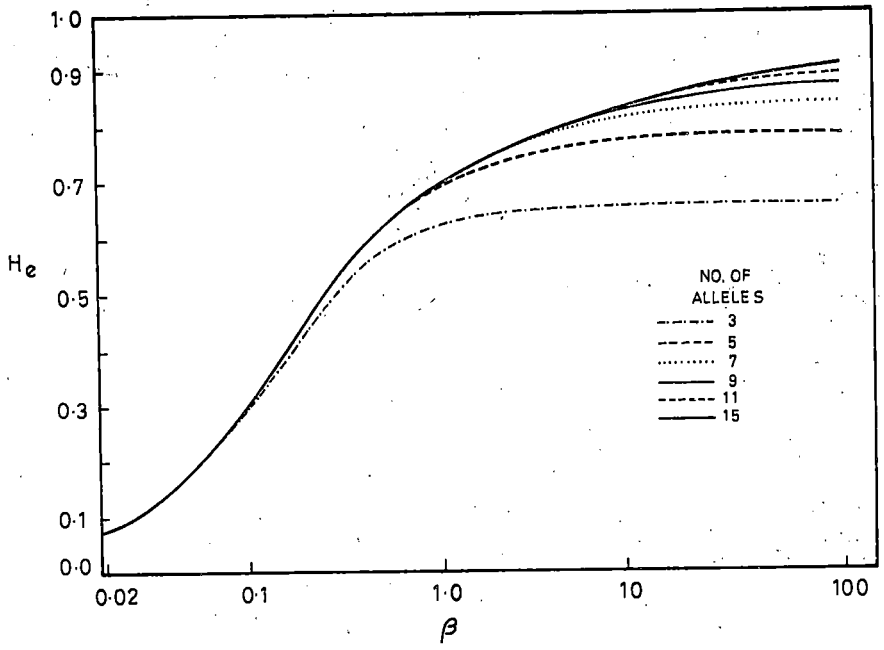

Fig. 3. Variation of heterozygosity at equilibrium with β (plotted on a log scale) for different number of alleles.

To see it more clearly we present the variation in equilibrium variance with increase in the number of alleles for a few selected values of β in Table 2.

Table 2. Equilibrium genetic variance [γ] and heterozygosity [H_e] for different number of alleles at few selected values of β

Number of alleles	β					
	0.1	1.0	5.0	20.0	50.0	100.0
3	0.0820 **0.2876**	0.2500 **0.6250**	0.2843 **0.6522**	0.2907 **0.6558**	0.2920 **0.6564**	0.2925 **0.6566**
5	0.0900 **0.2939**	0.4194 **0.6934**	0.6522 **0.7651**	0.7353 **0.7799**	0.7546 **0.7828**	0.7612 **0.7837**
7	0.0901 **0.2940**	0.4603 **0.7001**	0.9330 **0.7984**	1.2573 **0.8304**	1.3563 **0.8373**	1.3929 **0.8395**
9	0.0902 **0.2940**	0.4647 **0.7004**	1.0510 **0.8058**	1.7171 **0.8530**	2.0151 **0.8658**	2.1412 **0.8703**
11	0.0902 **0.2940**	0.4649 **0.7005**	1.0803 **0.8070**	2.0090 **0.8617**	2.6128 **0.8811**	2.9249 **0.8885**
13	0.0902 **0.2940**	0.4649 **0.7005**	1.0852 **0.8071**	2.1426 **0.8644**	3.0511 **0.8886**	3.6407 **0.8993**
15	0.0902 **0.2940**	0.4649 **0.7005**	1.0858 **0.8071**	2.1885 **0.8650**	3.3086 **0.8917**	4.2016 **0.9053**

The figures in bold face refer to heterozygosity [H_e]

The equilibrium variance of 0.46 for 15 alleles when β is 1.0 is close to $\sqrt{\beta}/2=0.50$ as predicted under the infinite alleles case by Kimura [7]. Similarly, for β as high as 100.0, the variance for 15 alleles is 4.20, smaller than $\sqrt{\beta}/2=5.0$. The difference, however, is not very great.

In so far as heterozygosity is concerned, Figure 3 indicates clearly that it also increases with the increase in the value of β but in a characteristic manner depending on the number of alleles. Initially, the increase is rapid and does not depend on the number of alleles considered. But for higher values of β, the curves behave in a characteristic manner. The pattern of one curve bifurcating from the rest of the remaining curves at a particular value of β noticed for the equilibrium genetic variance is also observed for the heterozygosity. The curves seem to stabilize at the values depending

upon the number of alleles considered. This can also be seen from Table 2. For instance, when we consider 7 alleles, the heterozygosity is close to 83-84 per cent for β lying between 20 and 100. For the 3 alleles case, however, it is around 66 per cent for the same range of values of β. When we consider smaller values of β say β=0.1, the heterozygosity is 29 per cent throughout, irrespective of the number of alleles considered.

7. *Discussion*

The dynamics of genetic variability for quantitative traits is an outstanding problem for which various mathematical models have been invoked and alternative approximations have been discussed. Mutation-selection equilibria in such a case determined so far assume either two alleles per locus or an infinitely large number of alleles per locus, with either continuous or discrete allelic effects. Different approximations lead to different results and their relative merits depend on empirical results pertaining to the magnitude of mutation rates and selection coefficients for quantitative characters as well as to the distribution of effects of new mutation. In this context recently interest has been generated on the step-wise discrete mutation model [Narain & Chakraborty [16] [17]]; Slatkin [21] in which at the underlying locus, an infinite or finite number of alleles A_i ($i=0, \pm 1, \pm 2, \ldots$) is possible with allele A_i having additive effect ai on a quantitative character and mutation occurs according to a shifted binomial distribution. While studying such a model with stabilizing selection, Narain & Chakraborty [17] could reproduce the results of Turelli [22] based on house-of-cards approximation by considering three alleles at each locus. In this paper, this approach has been extended up to fifteen alleles case.

The numerical results obtained here are more general in that these are not restricted to a given number of alleles and cover all the cases of selection, weak or strong. When the selection is stronger than mutation i.e. $\beta \ll 1$, and we consider five alleles case, we find from Table 1 that the equilibrium frequency of $\hat{x}_2$ is very close to zero and the five alleles case reduces to the three alleles case. The equilibrium genetic variance is found to be close to the value of β which it should take from (34). This is also found to be true for the cases of 7 to 15 alleles. In fact, when the force of selection is quite strong, no more than two alleles can segregate at the locus. These results, as already noted, are similar to those of Turelli [22] who used 'house-of-cards' approximation after Kingman's [8] house-of-cards model of mutation to produce results of Kimura [7]

and of three alleles approximation. Since the equilibrium genetic variance is found to be the same as with two alleles case, he came to the conclusion that the variance at equilibrium is independent of the number of alleles considered. Our results confirm this conclusion. But when β takes larger values, we find that we cannot afford to neglect frequency of alleles other than $\hat{x}_1$ at equilibrium. The five or more alleles case cannot then reduce to the three alleles case. The equilibrium genetic variance is now lower than β substantially.

The five alleles approximation of Slatkin [21] is based on the situation when selection and mutation balance in such a manner that $\hat{x}_2 << 1$ at equilibrium. In the present paper, this restriction is not imposed to solve the set of recurrence relations. We are therefore able to extend Slatkin's [21] analysis and get results as presented in Table 1 for the intermediate situation of selection when both the forces are of comparable magnitude. For such a case, if we assume that $\hat{x}_2 << 1$ at equilibrium as in Slatkin [21] we reproduce, using (38)-(40), his results for $\hat{x}_1$ which clearly increases from a smaller value to about 0.23. But as against Slatkin's [21] analysis in terms of bounds of $\hat{x}_2$, $\hat{x}_1$ and (v/s), we have now numerical results indicating how the equilibrium genetic variance behaves as β changes. This seems to point out that the number of alleles considered at each locus could be a crucial factor in mutation-selection balance equilibria unless β is very small.

For higher values of β, the genetic variance increases with the number of alleles, and approaches asymptotically the value predicted under the infinite alleles case. The equilibrium heterozygosity in such a case also increases. For β=100.0 and 15 alleles, as seen from Table 2, it can be as high as 90 per cent.

The results obtained in this paper thus indicate how the approximate solutions of Kimura [7], Turelli [22], Narain & Chakraborty [17] and Slatkin [21] fit in the general case of any number of alleles and weak as well as strong selection.

ACKNOWLEDGEMENTS

I am grateful to Shri Mahesh Kumar, IASRI, N. Delhi for help in computations and to the referee for constructive suggestions.

REFERENCES

[1] Barton, N.H., 1986. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet. Res. Camb.*, **47**, 209-216.

[2] Barton, N.H. and Turelli, M., 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res. Camb.*, **49**, 157-173.

[3] Bulmer, M.G., 1972. The genetic variability of polygenic characters under optimizing selection of mutation and drift. *Genet. Res. Camb.*, **19**, 17-25.

[4] Bulmer, M.G., 1980. The Mathematical Theory of Quantitative Genetics, Clarendon Press, Oxford.

[5] Fleming, W.H., 1979. Equilibrium distributions of continuous polygenic traits. *SIAM J. Appl. Math.* **36**, 148-168.

[6] Keightley, P.D. and Hill, W.G., 1990. Variation maintained in quantitative characters with mutation-selection balance : pleiotropic side-effects on fitness traits. *Proceedings of the Royal Society of London, Series B*, **242**, 95-100.

[7] Kimura, M., 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci., U.S.* **54**, 731-736.

[8] Kingman, J.F.C., 1978. A simple model for the balance between selection and mutation., *J. Appl. Pro.* **15**, 1-12.

[9] Lande, R., 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci., *Genet. Res. Camb.*, **26**, 221-235.

[10] Lande, R., 1980. The genetic covariance between characters maintained by pleiotropic mutations., *Genetics*, **94**, 203-215.

[11] Latter, B.D.H., 1960. Natural selection for an intermediate optimum., *Aust. J. Biol. Sci.*, **13**, 30-35.

[12] Latter, B.D.H., 1970. Selection in finite populations with multiple alleles. II. Centripetal selection, mutation and isoallelic variation. *Genetics*, **66**, 165-186.

[13] Nagylaki, T., 1984. Selection on a quantitative character. In: Chakravarti, A., (ed.) Human Population Genetics. The Pittsburgh Symposium (pp 275-306) Van Nostrand Reinhold, New York.

[14] Narain, P., 1988. The evolutionary dynamics of quantitative characters. *Indian J. Pure Appl. Math.* **19(2)**, 125-144.

[15] Narain, P., 1990. Genetic models of evolution for polygenic traits. In : Bahardur, R.R.(ed.) Probability, statistics and design of experiments. Proceedings R.C. Bose Symposium (pp. 527-549), Wiley Eastern.

[16] Narain, P. and Chakraborty, R., 1983. Genetic differentiation of quantitative characters between populations or species. II. Optimal selection in infinite populations. Unpublished (cited in Turelli, M. (1984) TPB **25**, 138-193).

[17] Narain, P. and Chakraborty, R., 1987. Genetic differentiation of quantitative characters between populations or species. II. Optimal selection in infinite populations. *Heredity*, **59**, 199-212.

[18] Price, G.R., 1970. Selection and covariance. *Nature*, **227**, 520-521.

[19] Price, G.R., 1972. Extension of covariance selection mathematics. *Annals of Human Genetics*. **35**, 485-490.

[20] Price, G.R., 1990. The distribution of allelic effects under mutation and selection. *Genet. Res. Camb.* **55**, 111-117.

[21] Slatkin, M., 1987. Heritable variation and heterozygosity under a balance between mutation and stabilizing selection. *Genet. Res. Camb.* **50**, 53-62.

[22] Turelli, M., 1984. Heritable Genetic Variation via Mutation-Selection Balance: Lerch's Zeta meets the Abdominal Bristle. *Theo Pop. Biol.* **25**, 138-193.

[23] Turelli, M. and Barton, N. H., 1990. Dynamics of polygenic characters under selection. *Theo. Pop. Biol.* **38**, 1-57.