Proc. Natl. Acad. Sci. USA
Vol. 73, No. 5, pp. 1689–1692, May 1976
Genetics

Evolution of continuous variation: Direct approach through joint distribution of genotypes and phenotypes*

(mutation/selection/heritability/environment/variance)

L. L. CAVALLI-SFORZA† AND M. W. FELDMAN‡

† Department of Genetics, Stanford University, School of Medicine, Stanford, California 94305; and ‡ Department of Biological Sciences, Stanford University, Stanford, California 94305

Communicated by Joshua Lederberg, February 11, 1976

ABSTRACT **The evolutionary dynamics of the joint distribution of genotypes and phenotypes is studied. The model, originally devised to study the joint effects of Mendelian and other types of transmissions, provides results of interest also to the theory of direct Mendelian transmission with natural selection. Assuming bivariate normal distributions, it is shown that in the latter case genotypic and phenotypic means and variances, and genotype–phenotype correlation can be expressed recursively as functions of the parameters for the selection, environmental, and mutation variance. Equilibria and rates of approach for these moments are calculated. It is also proved that in the presence of selection the heritability, defined as the ratio of expected genotypic to expected phenotypic variance after selection, is greater than that before selection by a predictable amount and that it can be greater than unity.**

In earlier papers (1, 2) we have considered the "phenotypic" transmission and evolution of continuous traits. We believe the models thus generated can be useful for the understanding of some aspects of cultural evolution. Genotypic transmission of the mode of reaction to environmental stimuli, accompanied by phenotypic transmission, has been the object of other studies, for a continuous trait (3) and for a discontinuous trait (4). In the latter case we also studied the equilibrium behavior of the phenotypes and genotypes.

This work has led us to introduce an approach in which the distributions of genotypes and phenotypes are considered *jointly* in problems of transmission and evolution. Such an approach should prove particularly fruitful in the study of transmission and evolution of traits observed phenotypically but with some genotypic basis. The purpose of the present paper, however, is limited to showing how this approach can be developed and applied to some classical problems of Mendelian transmission and Darwinian evolution.

The literature on the dynamical theory of evolution of continuous variation is not extensive. Two papers are of particular relevance to the study presented here. Kimura (5) has used a diffusion approximation argument to predict equilibrium values of a continuous polygenic trait under mutation and selection, assuming a quadratic deviation selection function. Slatkin (6) has proposed a general model producing a recurrence for the distribution of continuous "phenotypes" subject to an unspecified transmission rule and to selection. In neither of these studies was the contribution of environmental variation to overall phenotypic variation pursued. The fact that selection acts on phenotypes and not on genotypes, whereas Mendelian transmission involves genotypes only, suggests that for a complete treatment it is necessary to consider the joint distribution of genotypes and phenotypes. This we now proceed to do.

General model

Our general recurrence relation for the joint probability density function ϕ^* of phenotypes f and genotypes g after selection on the phenotypes takes the form

$$\phi^*_{t+1}(f,g) = \overline{W}^{-1} W(f) \int\int\int\int \psi_t(h,k,h',k')\, T(f,g;h,k,h',k')\, dh\,dk\,dh'\,dk' \quad [1]$$

where the subscripts denote the time in generations. The function $\phi^*_{t+1}(\cdot,\cdot)$ is the joint probability density of phenotypic values and genotypic values for a random individual from the reproducing adults of generation $t+1$ (i.e., after selection). On the right side of [1] the function $\psi_t(\cdot,\cdot,\cdot,\cdot)$ denotes the joint probability density of phenotypes and genotypes among mating pairs, the first and second variables (h, k) referring to the phenotype and genotypes of one member, and the third and fourth variables (h', k') referring to the phenotype and genotype of the other member, respectively. ψ_t can be expressed as

$$\psi_t(h,k,h',k') = \phi^*_t(h,k)\phi^*_t(h',k') + \chi_t(h,k,h',k') \quad [2]$$

where $\chi_t(\cdot,\cdot,\cdot,\cdot)$ is identically zero under random mating and different from zero under assortative and/or selective mating or differential fertility of the mating pairs. T is the transmission law, namely the probability that a young individual will be of phenotype f and genotype g if it is produced by a mating between parents (h, k) and (h', k'). $W(f)$ is a fitness function (i.e., the probability that individuals of phenotype f survive to maturity). $\overline{W}$ is the normalizer obtained by integrating the rest of the expression [1] with respect to f and g.

In [1] we consider a generation cycle from adult to adult. An equivalent expression can be given for a cycle from new zygote to new zygote, but is less useful for those cases where the population is examined at phenotypic maturity. Equivalent transformations are given elsewhere (4) for the discontinuous case.

Special assumptions

We proceed to incorporate into [1] the assumptions customary in the study of continuous variation. It is assumed that there is random mating with no fertility difference, as well as an additive polygenic system (that is, additivity over loci, no dominance, and no genotype–environment interaction). Selection is assumed to be "optimizing" (5–9) and expressed as

$$W(f) = \exp[-(f-\mu)^2/2S] \quad [3]$$

where μ is the optimum phenotype and S is an inverse measure of the strength of selection (i.e., S small implies very strong selection, $S \to \infty$ for a selectively neutral phenotype).

At time t the probability distribution of (f_t, g_t) is ϕ^*_t, assumed to be bivariate normal. $\mu^*_{g,t}$, $\mu^*_{f,t}$, G^*_t, F^*_t, and ρ^*_t, respectively, denote the genotypic and phenotypic means and

* This is paper no. 1 in a series.

variances and the correlation between phenotype and genotype in reproducing adults at time t; the asterisk indicates that they are taken after selection. These parameters completely define the distribution at time t, after selection.

The sequence of steps and symbols used in the model is outlined in Table 1. At a single locus with n alleles $A_1, A_2, \cdots, A_n$ of effects $a_1, a_2, \cdots, a_n$ and with frequencies $p_1, p_2, \cdots, p_n$ the total genotypic variance in the population is $2\sum_{i\neq j}\sum p_i p_j (a_i - a_j)^2$, whereas the average genotypic variance of the progeny of a random mating is exactly half as much. This motivates the assumption that for a sufficiently large number of loci with additive effects, if the genotypic variance in the parental population is G^*_t, then the variance of the progeny from a mating between parents with genotypic values k and k' is $G^*_t/2$ independent of k and k'.

It is clear that for a character controlled by a small number of genes, within family variances may differ between families. Also, in reality selection causes the progeny variance to vary from $G^*_t/2$. Our model assumes that the genotypic value of offspring from a mating of parents with genotypic values k and k' is a gaussian random variable with mean $[(k + k')/2] + m_g$ and variance $G^*_t/2 + M$. Here m_g is the change in the genotypic mean due to mutation analogous to $2m$ in Kimura's paper (5), whereas M is the amount added, because of mutation, to the genotypic variance of the offspring and is analogous to $2v$ in Kimura's treatment. M would be proportional to the (constant) mutation rate per zygote (twice that per gamete) and to the variance of changes in allelic effects determined by the mutation. We then assume that if an offspring has genotypic value g, his phenotypic value f is normally distributed around g with "environmental" variance E. The transmission function T in [1] is then completed by the assumption of independence of the genotypic and phenotypic transmission rules. Thus

$$T(f,g;h,k,h',k') = \frac{\exp\{-[g-(k+k')/2-m_g]^2/[2(G^*_t/2+M)]\}}{\sqrt{2\pi(G^*_t/2+M)}} \cdot \frac{\exp\{-(f-g)^2/2E\}}{\sqrt{2\pi E}} \quad [4]$$

It should be noted that T as written in [4] is independent of h and h'. A treatment incorporating phenotypic transmission along the lines of our previous work (3) is in preparation.

With T as in [4], the joint distribution of genotypic and phenotypic values before selection is completed by integrating T using the mating density [2], with $\chi_t = 0$, as in [1]. Our assumption on T entails that h and h' are essentially irrelevant in [1], which then reduces to a double integral over k and k'. When this integration is carried out, and before the insertion of $W(f)$, we have the joint phenotypic–genotypic density as

$$\phi_{t+1}(f,g) = N_f(g,E)\, N_g(\mu^*_{g,t} + m_g, G^*_t + M) \quad [5]$$

where in [5] the notation $N_f(g, E)$ denotes the normal phenotypic density with mean g and variance E, while $N_g(\mu^*_{g,t} + m_g, G^*_t + M)$ denotes the normal genotypic density with mean $\mu_{g,t+1} = \mu^*_{g,t} + m_g$ and genotypic variance $G_{t+1} = G^*_t + M$. The second factor on the right side of [5] arises from the fact that the offspring genotypic value, conditional on the parental genotypes, has variance $G^*_t/2 + M$, whereas the mid-parental genotype has variance $G^*_t/2$, their sum being $G^*_t + M$. At this stage the phenotypic variance is $F_{t+1} = G_{t+1} + E$, the phenotypic mean is $\mu_{f,t+1} = \mu_{g,t+1} = \mu^*_{g,t} + m_g$, whereas the correlation between phenotypic and genotypic values is $\rho_{t+1} = \sqrt{G_{t+1}/F_{t+1}}$.

Effect of selection

To complete the formulation [1] it remains to multiply [5] by $W(f)/\bar{W}$. Of course this does not alter the bivariate normal character of the distribution, but it does change all of the parameters. In fact we have

$$\phi^*_{t+1}(f,g) \propto \exp\left[\frac{-1}{2}\left\{\frac{(g-\mu_{g,t+1})^2}{G_{t+1}(1-\rho^2_{t+1})} + \frac{(f-\mu_{g,t+1})^2}{F_{t+1}(1-\rho^2_{t+1})} - \frac{2\rho_{t+1}(f-\mu_{g,t+1})(g-\mu_{g,t+1})}{(1-\rho^2_{t+1})\sqrt{F_{t+1}G_{t+1}}} - \frac{(f-\mu)^2}{S}\right\}\right]$$

$$\propto \exp\left[-\frac{1}{2[1-(\rho^*_{t+1})^2]}\left\{\frac{(g-\mu^*_{g,t+1})^2}{G^*_{t+1}} + \frac{(f-\mu^*_{g,t+1})^2}{F^*_{t+1}} - \frac{2\rho^*_{t+1}(f-\mu^*_{g,t+1})(g-\mu^*_{g,t+1})}{\sqrt{F^*_{t+1}G^*_{t+1}}}\right\}\right] \quad [6]$$

where the parameters $\mu^*_{g,t+1}$, $\mu^*_{f,t+1}$, G^*_{t+1}, F^*_{t+1}, and ρ^*_{t+1} are as in [7]

$$G^*_{t+1} = (G^*_t + M)(E + S)/(E + G^*_t + M + S) \quad [7a]$$

$$F^*_{t+1} = S(E + G^*_t + M)/(E + G^*_t + M + S) \quad [7b]$$

$$(\rho^*_{t+1})^2 = S(G^*_t + M)/(E + S)(E + G^*_t + M) \quad [7c]$$

$$\mu^*_{f,t+1} = \{S(\mu^*_{g,t} + m_g) + (E + G^*_t + M)\mu\}/(E + G^*_t + M + S) \quad [7d]$$

$$\mu^*_{g,t+1} = \{(E + S)(\mu^*_{g,t} + m_g) + (G^*_t + M)\mu\}/(E + G^*_t + M + S) \quad [7e]$$

In the above recursion system the parameters E, M, and S are written as constants. If they depend on time the analysis is substantially more complicated. From [7a], which is a linear fractional transformation, it follows that G^*_t converges geometrically fast to

$$\hat{G}^* = M\{\sqrt{1 + 4(E+S)/M} - 1\}/2 \quad [8a]$$

From [8a] in [7b] the phenotypic variance after selection F^*_t converges as $G^*_t \to \hat{G}^*$ to

$$\hat{F}^* = S(E + \hat{G}^* + M)/(E + \hat{G}^* + M + S) \quad [8b]$$

The convergence of G^*_t and F^*_t to $\hat{G}^*$ and $\hat{F}^*$ is at the geometric rate λ_1 with

$$\lambda_1 = (E + S - \hat{G}^*)/(E + \hat{G}^* + M + S) = [\hat{G}^*/(\hat{G}^* + M)]^2 \quad [9]$$

The correlation between genotype and phenotype, after selection, ρ^*_t, converges to the limit

$$(\hat{\rho}^*)^2 = S(\hat{G}^* + M)/(E + S)(E + \hat{G}^* + M) \quad [10]$$

along with G^*_t at the geometric rate λ_1. The phenotypic and genotypic means converge to

$$\hat{\mu}^*_f = \mu + \frac{Sm_g}{\hat{G}^* + M} \text{ and } \hat{\mu}^*_g = \mu + \frac{E+S}{\hat{G}^* + M} m_g \quad [11]$$

respectively, at the asymptotic rate.

$$\lambda_2 = (E + S)/(E + \hat{G}^* + M + S) = \hat{G}^*/(\hat{G}^* + M) \quad [12]$$

per generation.

After selection the parametric "heritability" is

$$(h^*)^2 = \hat{G}^*/\hat{F}^* = \frac{M(E+S)[1 + \sqrt{1 + 4(E+S)/M}]}{S[2E + M + M\sqrt{1 + 4(E+S)/M}]} \quad [13]$$

Table 1. Joint distributions of genotype (g) and phenotype (f): A generation cycle

Variables used in text	Density function	Symbols of bivariate normal parameters: Genotype Mean	Genotype Variance	Phenotype Mean	Phenotype Variance	Phenotype–genotype correlation
Parent 1 phenotype h genotype k; Parent 2 phenotype h' genotype k'	$\phi^*_t(f,g)$	$\mu^*_{g,t}$	G^*_t	$\mu^*_{f,t}$	F^*_t	ρ^*_t
mutation, mating, segregation						
zygote → phenotype (phenotype expression)	$\phi_{t+1}(f,g)$	$\mu_{g,t+1}$	G_{t+1}	$\mu_{f,t+1}$	F_{t+1}	ρ_{t+1}
selection						
phenotype f genotype g	$\phi^*_{t+1}(f,g)$	$\mu^*_{g,t+1}$	G^*_{t+1}	$\mu^*_{f,t+1}$	F^*_{t+1}	ρ^*_{t+1}

* Entries refer to values after selection.

Thus, if $M[1 + \sqrt{1 + 4(E + S)/M}] > 2S$, this heritability is greater than unity. This condition reduces to $S^2 < M(E + 2S)$, or, if $E \gg M$, $S < \sqrt{ME}$. Note that the equilibrium values before selection are

$$\hat{G} = \hat{G}^* + M, \hat{F} = E + G^* + M, \hat{\rho}^2 = \hat{G}/\hat{F} \quad [14]$$

and the ratio of the heritability after selection [13] to that before selection ($\hat{\rho}^2$ in [14]) is

$$(h^*)^2/\hat{\rho}^2 = 1 + E/S \quad [14a]$$

Also $(\hat{\rho}^*)^2$ from [10] is equal to $S\hat{\rho}^2/(E + S)$ and therefore the genotype–phenotype correlation is reduced after selection by the factor $S/(E + S)$. The latter and [14a] are valid for all t.

Remarks. If $M \ll S$ and $E = 0$, the formula for the genotypic variance reduces to $\sqrt{MS}$, which is identical to the result obtained (5) by the diffusion approximation method. Our $S = 1/(2K)$ in Kimura's notation (5) for the haploid case. Equilibrium values of the genotypic variance similar to [8a] have been obtained by Latter (7) and Lande (8) using approaches quite different from the above.

Applications

Two examples of optimizing selection analyzed in ref. 9, stature and birth weight in man, may be further explored using the above model. It was found that the variance of male stature observed before marriage is $\hat{F} = 42.3$ cm^2, and that of married males is $\hat{F}^* = 35.8$ cm^2, with negligible difference between the means. If we define R to be the ratio of the variance $\hat{F}^*$ after selection (in this case, due to mating) to that before selection $\hat{F}$ ($R = \hat{F}^*/\hat{F}$), R at any generation t is, from [7], equal to $S/(E + \hat{G}^* + M + S) = (\hat{F} + S)$. Here $R = 0.846$, so that $S = R\hat{F}/(1 - R) = 5.5\hat{F}$. At equilibrium, with H the heritability before selection, we derive from [8a] the mutation variance M relative to the genotypic variance before selection ($\hat{G}^* + M$), or to the phenotypic one, as

$$\frac{M}{\hat{G}^* + M} = H(1 - R) \quad [15a]$$

$$\frac{M}{\hat{F}} = \frac{M}{E + \hat{G}^* + M} = H^2(1 - R) \quad [15b]$$

If we assume equilibrium, additivity, and a heritability before selection of $H = 0.5$, the mutation variance is found from [15b] to be about $1/26$ of the total phenotypic variance before selection. The only source of selection estimated here is that due to mating; other sources, probably acting earlier, would decrease the total value of S and therefore the value of R; the relative estimate of the mutation variance would therefore increase. Assuming other sources of selection to be negligible, the estimates of E, M, S relative to $\hat{F}$ are $E = 0.5\hat{F}$, $M = 0.04\hat{F}$, $S = 5.5\hat{F}$.

For human birth weights, the variance before selection (due to neonatal mortality) is about 1.7 pounds^2 (0.35 kg^2), and the variance S from neonatal mortality is about 16 pounds^2 (3.3 kg^2), or 9.4 times the phenotypic variance. Again, assuming this as the sole source of selection, etc., $R = 16/(1.7 + 16) = 0.904$. The heritability here is probably extremely low, say 0.05, giving a mutation variance 4000 times smaller than the total phenotypic variance.

Discussion

The approach we have taken allows the complete exposition of the time-dependent behavior of the bivariate normal phenotype–genotype distribution. The approach of Prout (10) in an investigation of problems in progeny testing is similar in some respects to ours, although our primary interest is the evolutionary process rather than the short term variance analysis. The central quantity for these dynamics is the genotypic variance. In Kimura's (5) treatment the distribution of the genotypic values approached normality. When this is true and the distribution of environmental effects on phenotypes within genotypes as well as the selection function are normal, the bivariate normal distribution must emerge. Our treatment takes as its point of departure that the phenotypes and genotypes have a

bivariate-normal joint distribution. Clearly, the conclusions we have reached may be modified if the selection function departs from that postulated in Eq. [3] or if the phenotypes are not normally distributed within genotypes. One may conjecture, for example, that if there were sufficient fluctuation in the environmental or selection variance over time, an equilibrium distribution might not be attained. In spite of these caveats optimal selection of the type postulated by [3] has been documented in several cases (9). A classical example is that of human birth weight (9, 11).

The recursion system [7] and its equilibrium [8], [10], and [11] demonstrate certain unexpected features. One striking conclusion is that heritability, defined as the ratio of the genotypic to phenotypic variance, when it is calculated after selection, $\hat{G}^*/\hat{F}^*$, can be greater than unity. In fact this heritability after selection is greater than that before selection by a factor $1 + E/S$, where E and S are the environmental and selection variances, respectively. (When selection is strong, S is small.) The role of environmental variation in amplifying this selective adjustment to the heritability is greater the more intense the selection. Common estimates of heritability (based on correlations between relatives) that ignore selection will have a bias that depends on E and S. This appears difficult to estimate without an evaluation of the selective regimes involved and requires further theoretical studies.

The apparent paradox of the heritability greater than one can be resolved intuitively by noting that if selection is so intense (i.e., S so small) that only one phenotype can survive, the phenotypic variance F^*_t can be smaller than the genotypic G^*_t (i.e., many genotypes can produce the same phenotype). From [13], the heritability after selection becomes unbounded as $S \to 0$.

It is clear from [7] that the fate of the genotypic variance depends on environmental parameters. Thus, if the mutation variance M is small, the genotypic variance at equilibrium is approximately $\sqrt{M(E + S)}$. [See also Kimura (5) where $E = 0$.]

It is of interest to develop a partition, analogous to the classical one, of the total phenotypic variance (at any generation) before selection. Suppressing the time subscript, write $F = F^* + F'$ with F^* the phenotypic variance remaining after selection and F' that lost because of selection. Then from [7] we have

$$F' = F(E + G)/(E + G + S)$$

where G is the genotypic variance before selection. The phenotypic variance remaining after selection, $F^* = FS/(E + G + S)$, can be partitioned into the genotypic G^* and the environmental E^* variance after selection, plus a negative genotype–environment interaction term, i.e.,

$$F^* = G^* + E^* + 2W^*_{EG}$$

where

$$E^* = \frac{E(G + S)}{E + G + S} \qquad [15]$$

and

$$W_{GE} = \frac{-EG}{E + G + S}$$

Clearly this is one of many possible partitions.

It is of considerable interest to augment the basic scheme presented above to include dominance, interaction between genotype and phenotype, and other types of transmission and selection. The effects of these on common estimates of heritability are important.

The authors are grateful to Drs. F. B. Christiansen, E. Dempster, M. Kimura, D. K. Wagener, and G. A. Watterson for their valuable criticism of an early draft. This research was supported in part by National Institutes of Health Grant USPHS 10452-10-11, U.S. Atomic Energy Commission Grant AT(04-3)-326, and National Science Foundation Grant GB 37835.

1. Cavalli-Sforza, L. L. & Feldman, M. W. (1973) *Theor. Pop. Biol.* **4**, 42–55.
2. Feldman, M. W. & Cavalli-Sforza, L. L. (1975) *Ann. Hum. Biol.* **2**, 215–226.
3. Cavalli-Sforza, L. L. & Feldman, M. W. (1974) *Am. J. Hum. Genet.* **25**, 618–637.
4. Feldman, M. W. & Cavalli-Sforza, L. L. (1976) *Theor. Pop. Biol.*, in press.
5. Kimura, M. (1965) *Proc. Natl. Acad. Sci. U.S.A.* **54**, 731–736.
6. Slatkin, M. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 87–93.
7. Latter, B. D. H. (1970) *Genetics* **66**, 165–186.
8. Lande, R. (1976) *Genet. Res.*, in press.
9. Cavalli-Sforza, L. L. & Bodmer, W. F. (1971) in *The Genetics of Human Populations* (Freeman, San Francisco), pp. 612–615.
10. Prout, T. (1958) *J. Genet.* **56**, 1–7.
11. Karn, M. N. & Penrose, L. S. (1951) *Ann. Eugen.* **15**, 206–233.