DIRECT EFFECTS OF SELECTION ON PHENOTYPIC VARIABILITY OF QUANTITATIVE TRAITS

W.G. Hill

Institute of Cell, Animal and Population Biology, University of Edinburgh,
West Mains Road, Edinburgh, EH9 3JT, UK

INTRODUCTION

Phenotypic and genetic variability in quantitative traits are ubiquitous in natural populations and in populations of livestock under artificial selection. Models to explain the level of genetic variation usually focus on the balance between its gain by mutation and loss by stabilising or directional selection and drift (see e.g. Falconer and Mackay, 1996; Bürger, 2000). The models do not explain, however, why the heritability of growth rate is typically c.25% in many species, even in highly selected modern broiler populations (Koerhuis and Thompson, 1997), or why the coefficient of variation (CV) of body size is typically c.10%, even in long-term selected mouse lines (Bünger *et al.*, 2001). As natural selection on fitness-associated traits such as litter size is presumably directional, it can account for their low heritability but not for their high variability and genetic CVs ('evolvability', Houle, 1992) typically as high as for other traits such as body size less clearly associated with fitness.

These problems indicate that there is a need to consider models of the genetic basis of variability *per se* and its relation to the population mean. It is customary in quantitative genetic models to assume that genotypes at loci differ in mean but not in variability of performance. Directional or stabilising selection acts through these differences, resulting in changes in gene frequency and consequential, rather than direct, effects on the genetic variance. The Bulmer effect generates negative correlations of gene frequency among loci, but the loss in variance is transient and is recovered if selection is relaxed (Bulmer, 1971).

Here the implications for directional selection are considered of a more general model in which both mean and variability differ between genotypes. SanCristobal-Gaudy *et al.* (1998) discuss such a model, but predict response to selection for canalisation, i.e. reduced variability of a trait, and review experimental evidence showing successful attempts to do so. A related problem in directional selection is the effect of selection among groups which differ in variance (Hill, 1984), such that individuals in the more heterogeneous groups (e.g. dairy herds) are more likely to be chosen if selection is intense.

ANALYSIS

Let phenotypic values of individuals of genotype A_iA_j be normally distributed with means m_{ij} and variances σ_{Pij}^2. For simplicity, assume a large population is fixed for A_2A_2, and a mutant A_1 appears. The previous mean and phenotypic variance are $\mu = m_{22}$ and $\sigma_P^2 = \sigma_{P22}^2$, and the increase for A_1A_2 individuals is $a = m_{12} - \mu$ and $b = \sigma_{P12} - \sigma_P$. Assume that truncation selection is practised on individual performance, with a proportion p of the population selected, all of which were initially A_2A_2 before A_1 appeared. Standard methods are used to derive the

selective value (Haldane, 1931; Latter, 1966; Falconer and Mackay, 1996). Assuming normality, the truncation point T is $1 - \Phi[(T - \mu)/\sigma_P] = p$ where $\Phi(y)$ is the standardised cumulative normal distribution function. The probability that an A_1A_2 individual is selected is $p(a, b) = 1 - \Phi[(T-\mu - a)/(\sigma_P + b)]$. Let $x = (T-\mu)/\sigma_P$ denote the truncation point and z the corresponding ordinate on the standardised normal, i.e. $d\Phi(y)/dy|_x = z$. Using Taylor series,

$p(a,b) = p + a(\partial p/\partial a) + b(\partial p/\partial b) + \ldots$ (1)

with the derivatives evaluated at $a = b = 0$. Hence

$p(a,b) = p + az/\sigma_P + bxz/\sigma_P + \ldots$

Scaling by the proportion selected to obtain the fitness, the selective value of the mutant is $s = p(a,b)/p - 1$. For normally distributed traits $z/p = i$ (selection intensity), so s is approximately

$s = ia/\sigma_P + ixb/\sigma_P$, (2)

or, if the effect is expressed as an increase in the variance to $\sigma_P^2 + c$, since $d\sigma_P/d\sigma_P^2 = \frac{1}{2}/\sigma_P$,

$s = ia/\sigma_P + \frac{1}{2}ixc/\sigma_P^2$. (2a)

The first term in eq. (2, 2a) is standard (e.g. Falconer and Mackay, 1996), and the second is derivable from Haldane (1931) who considers the effect of both different mean and variance. A term in ix also appears if an additional term in a^2 is included in (1) to account for genes of large effect (Latter, 1966). Note that if $p < 0.5$ then $x > 0$, i.e. with intense selection variable individuals are more likely to be selected; whereas if $p > 0.5$, $x < 0$, i.e. variable individuals are more likely to be rejected. The 'strength' of selection on variation is, however, very weak if the proportion selected is low, since i is small, but rises rapidly if selection becomes intense:

p = 0.80, 0.60, 0.50, 0.40, 0.20, 0.10, 0.05, 0.02, 0.01

i = -0.35, 0.64, 0.80, 0.97, 1.40, 1.75, 2.06, 2.42, 2.66

ix = -0.29, -0.16, 0.00, 0.24, 1.18, 2.25, 3.39, 4.97, 6.20

This implies, for example, that with intense selection on males and weak or no selection on females, there will still be selection for genes that increase the variance.

Equations (1) or (2) specify the fixation probability of a favourable mutant gene for artificial selection (Hill, 1982) (assuming $s > 1/2N_e$), taking into account any effect it has on the variance. Thus a mutant which increases the variance is more likely to be fixed if selection is intense. More generally, following from standard theory, the change in gene frequency (q) assuming additive gene action, i.e. $m_{12} - m_{22} = m_{11} - m_{12} = a$ and $\sigma_{12} - \sigma_{22} = \sigma_{11} - \sigma_{12} = b$, is given by $\Delta q = [\, i(a + xb)/\sigma_P]q(1 - q)$. One generation of selection will produce from this locus, to the same level of approximation, changes in mean and variance of

$\Delta\mu = (d\mu/dq)\Delta q = 2a\Delta q = 2[(ia^2 + ixab)/\sigma_P]q(1 - q)$ (3)

$\Delta\sigma_P^2 = 2\sigma_P(d\sigma_P/dq)\Delta q = 4b\sigma_P\Delta q = 4[iab + ixb^2]q(1 - q)$ (4)

or, equivalently,

$\Delta\sigma_P^2 = (d\sigma_P^2/dq)\Delta q = 2c\Delta q = [2iac/\sigma_P + ixc^2/\sigma_P^2]q(1 - q)$ (4a)

The formulae can readily be extended to take account of dominance. Inclusion of epistasis is more complicated, as a matrix of variance for each genotype is needed, and will be ignored. Let $\sigma_{A\mu}^2 = \Sigma_i 2a_i^2 q_i(1 - q_i)$ be the usual definition of additive variance in terms of effects of genes on the mean of the trait, with summation over loci. Similarly, let $\sigma_{A\sigma}^2 = \Sigma_i 2b_i^2 q_i(1 - q_i)$

and $cov_{A\mu\sigma} = \Sigma_i 2a_i b_i q_i(1 - q_i)$ be the equivalent additive genetic variance in the effects of genes on the phenotypic standard deviation and the covariance. Then

$$\Delta\mu = (i\sigma_{A\mu}^2 + ixcov_{A\mu\sigma})/\sigma_P \quad (5)$$

$$\Delta\sigma_P^2 = 2(icov_{A\mu\sigma} + ix\sigma_{A\sigma}^2) \quad (6)$$

If effects of genes on the mean and the variance of the trait are uncorrelated, i.e. $cov_{A\mu\sigma} = 0$, eq. (5) is the standard formula for the response (writing σ_A^2 rather than $\sigma_{A\mu}^2$). In a simple scale model, where $b_i = ka_i$ for all i, where k is the CV, then $cov_{A\mu\sigma} = k\sigma_{A\mu}^2$ and $\sigma_{A\sigma}^2 = k^2\sigma_{A\mu}^2$, and changes on the log scale follow standard predictions, $\Delta\sigma_P = k\Delta\mu$ or $\Delta\sigma_P^2 = 2k\sigma_P\Delta\mu$.

Assuming an infinitesimal model, the equations also apply in succeeding generations but the Bulmer effect has to be included and the change in variance it produces may swamp that from direct selection on the variance unless $\sigma_{A\sigma}^2$ is large. Many assumptions are, of course, needed for making long term predictions. In particular, issues of scale complicate the changes, which in the terms of the model imply that σ_{P22}^2, for example, depends on the mean performance. If there is already a satisfactory scale transformation, e.g. a log transformation to remove proportionality of phenotypic standard deviation to the mean (multiplicative effects), the analysis can be in those terms. The model then implies a difference between, say log σ_{P22} and log σ_{P12}, or equivalently, CV_{P22} and CV_{P12}. SanCristobal-Gaudy *et al*. (1998) undertook an analysis of canalising (stabilising) selection at the genotypic rather than gene locus level, and considered variability in log (environmental variance) among genotypic values. The present analysis can also be extended to stabilising selection, either by natural selection (fitness with e.g. nor-optimal form), or by artificial selection (intermediates selected), when genes which increase variance are at a selective disadvantage. With disruptive selection, when only high and low scoring extreme individuals are selected, such genes are at a selective advantage. With directional selection on fitness, e.g. regarding litter size of pigs as 'fitness', there is no selection on variation because there is a linear relation between mean performance and fitness.

EXPERIMENTAL TESTS

A major question is whether the genetic variance in variability, $\sigma_{A\sigma}^2$, or covariance, $cov_{A\mu\sigma}$, is sufficient to matter in prediction equations. Whilst it seems unlikely that direct estimates could be made in unselected populations, residual variances of each genotype can be estimated for identified major genes (SanCristobal-Gaudy *et al*. (1998) give examples) or QTL. There would be some non-linearity of regression of breeding value on phenotype, possibly detectable by practising selection of different intensity. There is relatively lower response to weak selection in some experiments (Clayton *et al.*, 1957), but not in others (Hanrahan *et al.*, 1973).

In selection experiments, the change in variance is confounded by the Bulmer effect, by scale, and by changes in $\sigma_{A\mu}^2$ due to changes in gene frequency. Disruptive and stabilising selection should, however, lead to changes in variance but not in mean. The confounding Bulmer effect can be distinguished by relaxing selection and checking whether the variance returns to its original value, allowing time for linked loci to release variance. Sorensen and Hill (1983) undertook such an experiment using disruptive selection, but the results are not conclusive; and SanCristobal-Gaudy *et al*. (1998) cite results for canalising selection showing real change.

For directional selection, Falconer & Mackay (1996, p.221) comment: 'The loss of genetic variance expected by the theory [of quantitative genetics] should lead to a reduced phenotypic variance. The phenotypic variance, however, is seldom found to decline as expected; often it increases.' If variances in high and low selected lines are compared with the base or an unselected control population, then the base/control will have intermediate or higher variance (since not inbred) than the (log) mean of the selected lines if changes are due to scale. Consider two examples. Clayton and Robertson (1957) found the variance in both high and low long-term lines of *Drosophila* selected for abdominal bristle number greatly exceeded that in the base population. In a review of selection experiments for growth in the mouse, with responses of several SD, Bünger *et al.* (2001) reported almost constant values of the CV over 50 or more generations in high or low selection lines and control populations. Although not proof of selection response for the variance, the loss in genetic variability expected from selection and drift is not shown as reduced phenotypic variability on the log scale. The effects of selection from isogenic populations utilising mutations should be less confounded by Bulmer and scale effects. Mackay *et al.* (1994) found greater increases in variance in lines of *Drosophila* which did respond to selection (high abdominal and low sternopleural bristle number). Some of the increase may have been due to segregating semi-lethals; indeed conflicting natural and artificial selection may contribute to variance in other experiments.

Neither the theoretical analysis nor experimental results are conclusive, but a further development of the models and more thorough review of the literature seem needed to help understanding of the processes controlling variation in selected populations.

REFERENCES

Bulmer, M.G. (1971) *Amer. Nat.* **105** : 201-211.

Bünger, L., Renne, U. and Buis, R.C. (2001) In *Encyclopedia of Genetics*, p. 337-360, Editor E. C. Reeve, Fitzroy Dearborn, London.

Bürger, R. (2000) *The Mathematical Theory of Selection, Recombination and Mutation.* Wiley, Chichester, England.

Clayton, G.A. and Robertson, A. (1957) *J. Genet.* **55** :152-170.

Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics. 4th ed*. Longman, Harlow, England.

Haldane, J.B.S. (1930) *Proc. Cambridge Philos. Soc.* **27** : 131-136.

Hanrahan, J.P., Eisen, E.J. and Legates, J.E. (1973) *Genetics* **73** : 513-530.

Hill, W.G. (l982) *Proc. Natl. Acad. Sci. USA* **79** : l42-l45.

Hill, W.G. (l984) *Anim. Prod.* **39** : 73-477.

Houle, D. (1992) *Genetics* **130** : 195-204.

Koerhuis, A.N.M. and Thompson, R. (1997) *Genet., Sel., Evol.* **29** : 225-249.

Latter, B.D.H. (1966) *Aust. J. Biol. Sci.* **18** : 585-598.

Mackay, T.F.C., Fry, R.F., Lyman, R.F. and Nuzhdin, S.V. (1994) *Genetics* **136** : 937-951.

SanCristobal–Gaudy, M., Elsen, J.-M., Bodin, L. and Chevalet, C. (1998) *Genet., Sel., Evol.* **30** : 423-451.

Sørensen, D.A. and Hill, W.G. (l982) *Heredity* **48** : 27-33.