

IX 55C

Changes in the distribution of the genetic variance of a quantitative trait in small populations of *Drosophila melanogaster*

C.López-Fanjul, J. Guerra and A. García

Universidad Complutense, Facultad de Ciencias Biológicas, Departamento de Genética, 28040 Madrid, Spain

(received 28 March 1988, accepted 16 August 1988)

Summary — The coefficient of variation of the genetic variance at generation t , $CV(V_{A_t})$, of an additive trait among replicated unselected small populations has been theoretically shown to be largely due to linkage disequilibrium from sampling. Starting from a population in linkage equilibrium, $CV(V_{A_t})$ should initially increase, rapidly approaching an asymptotic value. However, when effective population size expands, $CV(V_{A_t})$ is expected to decrease. Experiments with *Drosophila melanogaster* were carried out to check these predictions. Inbred lines were established by brother-sister single pair matings and continued for 3 generations. Each line was then maintained with as many parents as possible up to generation 10. The trait considered was the total number of abdominal bristles on the 5th and the 6th sternites, in generations 0, 4 and 10. A single generation of divergent selection was carried out in each of 75 lines in the same generations. Assuming no differences in environmental variance between lines, $CV(V_{A_t})$ can then be estimated from the within-line phenotypic variances or from the responses to selection. As predicted by theory, the value of $CV(V_{A_t})$ substantially increased from generation 0 to 4. No reduction was detected afterwards, possibly because the trait was affected by blocks of genes. Other predictions made concerning the redistribution of the genetic variance have been substantiated.

Inbreeding – genetic variance – *Drosophila melanogaster*

Résumé — Changements de la distribution de la variance génétique d'un caractère quantitatif dans des populations d'effectif limité de *Drosophila melanogaster*. On a montré théoriquement que le coefficient de variation de la variance génétique d'un caractère additif dans la génération t , $CV(V_{A_t})$, dans un ensemble de lignées non sélectionnées ayant toutes le même effectif génétique, est dû essentiellement au déséquilibre de "linkage" généré par échantillonnage. Si la population est initialement en équilibre, $CV(V_{A_t})$ s'accroît au fil des générations, tendant rapidement vers une valeur asymptotique. Si, ensuite, l'effectif de la population augmente, $CV(V_{A_t})$ diminuera. Afin de contrôler ces prédictions, on a réalisé des accouplements consanguins frère x sœur, pendant 3 générations dans un ensemble de lignées de *Drosophila melanogaster*. Postérieurement, chaque lignée a été maintenue jusqu'à la génération 10 avec le plus grand nombre possible de géiteurs. Le caractère considéré a été le nombre total de soies dans les 5^e et 6^e segments abdominaux, évalué dans les générations 0, 4 et 10. Parallèlement, et dans ces mêmes générations, on a effectué une sélection divergente pendant une génération en 75 lignées. En supposant que la variance environnementale du caractère soit la même dans toutes les lignées considérées, $CV(V_{A_t})$ peut être évalué à partir des variances phénotypiques intra-lignée ou à partir des réponses à la

sélection. La valeur de $CV(V_A)$ a augmenté considérablement en passant de la génération 0 à la 4, ainsi que le prévoyait la théorie. Toutefois, elle ne s'est pas réduite ensuite, ce qui est probablement dû à l'existence de blocs géniques. Les prédictions théoriques relatives à la redistribution de la variance génétique ont été également vérifiées.

consanguinité – variance génétique – *Drosophila melanogaster*

Introduction

Consider a set of replicated unselected lines of the same effective size, originally sampled from a population in linkage equilibrium, and kept under the same environmental conditions. Random mating and no migration is assumed for each line and generations are discrete. For a quantitative trait, determined by many neutral loci all with additive gene action, the following predictions can be made regarding the temporal behaviour of the first two moments of the distributions of the mean and the genetic variance. It is well known (Wright, 1951) that the overall performance of the replicates will remain constant, but the variance among replicates will increase with time as $2F_t V_A$, where V_A is the additive variance in the base population and F_t the inbreeding coefficient at generation t . In parallel, the distribution of the genetic variance within lines will also change. Its mean will decrease with time as $(1 - F_t) V_A$, while its variance will increase and rapidly approach an asymptotic value, largely because of linkage disequilibrium built up by sampling (Bulmer, 1976; Avery and Hill, 1977). Similarly, when 2 additive traits are considered, the expected value of the within-line genetic covariance will decrease as $(1 - F_t) cov_A$, cov_A being the genetic covariance in the base population, and its variance will increase also towards an asymptotic value, due to disequilibrium (Avery and Hill, 1977).

Experimental evaluations of this theory are scarce, partial and generally inconclusive. For the most part, they have been restricted to small sets of lines, inbred by regular brother-sister matings, and refer to the evolution of the between- and within-line variance of a quantitative trait. In such analyses, the between-line variance, in the absence of maternal effects, is essentially genetic, although its estimate has been generally associated with a large sampling error as the number of lines involved was small. Moreover, the observed change of the within-line phenotypic variance may have resulted from 2 antagonistic processes: the genetic component should decrease as inbreeding progresses, but the environmental component may increase, due to a greater susceptibility of inbreds to environmental heterogeneity (see Falconer, 1981 for a review; pp. 243-246). Both causes may have contributed to the inconsistency of the studies carried out so far.

Several traits have been considered: abdominal and sternopleural bristle number and body weight in *Drosophila melanogaster* (Rasmuson, 1952; Kidwell and Kidwell, 1966); egg-laying of virgin females in *Iribolium castaneum* (López-Fanjul and Jódar, 1977); and litter size in mice (Bowman and Falconer, 1960). In these experiments, the within-line phenotypic variance oscillated more or less widely without showing a definite trend and only appeared to decline for the 2 bristle systems analysed by Rasmuson (1952). A reduction of the within-line genetic variance has been reported by Tantawy (1957) for wing and thorax length in *D. melanogaster*, but it was only detected in later stages of inbreeding. On the other hand, the between-line variance only appeared to increase for both bristle systems (Rasmuson, 1952) and egg-laying (López-Fanjul and Jódar, 1977), fluctuating over generations or even diminishing in the remaining instances. For perform-

ance traits in Hereford cattle, it has also been reported that the theoretical expectations for the redistribution of the genetic variance were not generally fulfilled (Russell *et al.*, 1984).

Part of the theory mentioned above has not yet been tested, specifically that concerned with the prediction of the changes of the variation among replicates in the within-replicate genetic variance. The present work has been designed to provide an experimental check of these expectations using *D. melanogaster*. Resulting data will also permit a further test of the remaining theoretical predictions.

Materials and Methods

The Consejo population was captured in southeast Spain 3 years prior to the start of these experiments. The flies were reared on baker's yeast-agar-saccharose standard medium, and all cultures were incubated at 25°C. The trait considered was the number of abdominal bristles on the 5th and the 6th sternites of females.

Two non-contemporaneous experiments (1 and 2) were carried out following the same design. Samples consisting of 4 pairs of parents taken from the base population were cultured in bottles (experiment 1: 205 samples; experiment 2: 99 samples) and the trait was scored on 20 female offspring per sample (generation 0). Independently, inbred lines were established from the same base population in separate vials by brother-sister single pair matings, and continued for 3 generations (experiment 1: 200 lines; experiment 2: 100 lines). At generation 3, 4 males and 4 virgin females were randomly taken from each survivor line to be the parents of generation 4 ($F_4 = 0.5$). Each line was continued thereafter in a bottle with as many parents as possible up to generation 9 in which, again, 4 males and 4 virgin females per line were taken to be the parents of generation 10. The trait was scored for each of 20 virgin females per line at generations 4 and 10. This procedure of restricting the number of parents was adopted in order to standardize culture density at those generations where the trait was to be scored, even though it would both slightly reduce the genetic variance in the offspring generation and increase the variance among lines.

From the virgin females scored at generations 4 and 10 in both experiments and at generation 0 in experiment 1, one generation of divergent individual selection for the total number of bristles was carried out on females with proportion 4/20 in a randomly chosen set of lines (experiment 1: 50 lines; experiment 2: 25 lines). For each line and direction of selection, the 4 females selected were mated to 4 males taken at random from that same line and generation and 20 female offspring were scored.

Realized heritabilities in one generation of divergent selection of the total number of bristles on both sternites were calculated in all cases, except in the base population in experiment 2, where daughter-dam regression was used.

Results

At generations 0, 4 and 10, the average within-line phenotypic variance of the total number of bristles was partitioned into 3 components: additive genetic, general environment and special environment. The first was calculated from the product of the phenotypic

variance and heritability, and the last estimated as twice the variance component within individuals derived from the corresponding hierarchical analysis of variance. Computation of environmental variance assumes a genetic correlation of one between the number of bristles on both sternites, as has often been reported (Caballero and López-Fanjul). A similar pattern appeared in the results of both experiments, as shown in Table I. As expected from theory, the additive variance at generation 4 was reduced to about one-half of its initial value. Although it experienced a further decrease at generation 10, the difference between the heritability estimates at generations 4 and 10 was non-significant in both experiments. Also, the component due to special environment was identical at generations 4 and 10, and somewhat larger than that of the base population (12-20%). This indicates a greater susceptibility of inbreds to localized circumstances operating during development. Finally, the general environmental variance was estimated by difference and, consequently, would have a larger error than the other components. Although its value fluctuated, it never exceeded 15% of the phenotypic variance.

The overall mean and the coefficient of variation, asymmetry and kurtosis of the distribution of the mean of the total number of bristles at generations 0, 4 and 10 are shown, for both experiments, in Table II. No appreciable change of the mean was detected when the 2 experiments were considered together. Nevertheless, the mean decreased in experiment 1 and increased in experiment 2 from generations 0 to 4, which may be attributed to unidentified environmental effects. No further changes of the mean were observed. The coefficient of variation of the mean at generation 0 was about double the value expected from sampling alone (1.7%). This discrepancy may be attributed to unforeseen environmental differences between cultures V_{Ec} . Estimates of V_{Ec} (experiment 1: 1.7; experiment 2: 1.6) were obtained by subtracting the expected value of the variance of the mean ($V_{Po}/20$) from that observed. Assuming V_{Ec} remains constant and that the increase in inbreeding coefficient from generation 4 to 10 is negligible, the expected value of the variance of the means at these generations will be given by $2F_t V_A + V_{Pt}/20 + V_{Ec}$, where V_{Pt} is the phenotypic variance at generation t . For the values of V_{Pt} and V_A in Table I, $F_t=0.5$ and $V_{Ec}=1.6$, the expected values of the coefficient of variation of the mean at generations 4 and 10 were, respectively, 6.8 and 6.2 percent in experiments 1 and 2. These figures agree with the observed results. The coefficients of asymmetry and kurtosis were not significantly different from zero in all cases, as predicted from the Central Limit theorem.

Table I. Phenotypic variance (V_P) and its components: additive genetic (V_A), special environment (V_{Es}) and general environment (V_E), and heritability (h^2) of the total number of bristles on both sternites, in experiments 1 and 2.

Generation	No. lines		V_P		V_A		V_{Es}		V_{Eg}		h^2	
	1	2	1	2	1	2	1	2	1	2	1	2
0	205	99	10.6	9.9	5.4	4.9	4.2	4.8	1.0	0.2	0.51 ± 0.06	0.50 ± 0.05
4	194	76	7.9	9.6	2.3	2.7	5.2	5.4	0.4	1.5	0.29 ± 0.05	0.28 ± 0.09
10	187	67	7.8	8.7	1.6	2.0	5.2	5.4	1.0	1.3	0.21 ± 0.04	0.23 ± 0.06

Table II. Overall mean and coefficient of variation (*CV*), asymmetry (g_1) and kurtosis (g_2) of the distribution of: 1) the mean; 2) the phenotypic variance of the total number of bristles on both sternites and; 3) the phenotypic covariance between the number of bristles on the 5th and the 6th sternites, among the lines in experiments 1 and 2.

Generation	Mean		CV (%)		g_1		g_2	
	1	2	1	2	1	2	1	2
1) Distribution of the mean								
0	41.4	41.3	3.6	3.5	0.18	0.27	-0.02	-0.25
4	40.5	42.4	6.4	6.7	0.09	-0.07	0.06	0.41
10	40.3	42.6	6.9	6.7	0.21	0.14	0.45	0.17
2) Distribution of the phenotypic variance								
0	10.6	9.9	41.4	37.5	1.26	1.36	2.46	4.27
4	7.9	7.8	49.7	51.5	1.84	3.06	4.76	15.19
10	7.8	8.7	43.2	46.7	1.05	1.59	1.18	3.20
3) Distribution of the phenotypic covariance								
0	1.7	1.4	67.4	71.8	1.13	0.97	2.86	3.59
4	0.8	1.2	120.1	100.3	1.58	2.82	4.26	14.10
10	0.8	1.0	106.9	100.0	0.78	1.32	0.54	3.34

The mean and the coefficient of variation, asymmetry and kurtosis of the distribution of the phenotypic variance of the total number of bristles on both sternites of the lines at generations 0, 4 and 10 are shown in Table II for both experiments. Identical parameters for the distribution of the phenotypic covariance between the number of bristles on the 5th and the 6th sternites are also shown in Table II. The overall variance and covariance decreased from generation 0 to 4, without any significant changes thereafter. The coefficient of variation of the variance or the covariance at generation 4 were larger than at generations 0 and 10. The coefficients of asymmetry and kurtosis were largest at generation 4 in both cases.

The coefficient of variation of the additive variance of the total number of bristles on both sternites at generation t , $CV(V_{A_t})$, can be estimated from our data in 3 different ways:

1) Assuming no differences in environmental variance between the lines, the variances of the phenotypic and additive variances of the lines will be the same and, therefore:

$$CV(V_{A_t}) = CV(V_{P_t}) / h^2_t,$$

where $CV(V_{P_t})$ and h^2_t are the coefficient of variation of the phenotypic variance and the heritability of the total number of bristles at generation t , respectively;

2) In the same situation as before, Avery and Hill (1977) have shown that approximately :

$$CV(V_{At}) = CV(R_t) / (1 - h_t^2 / 2),$$

where $CV(R_t)$ is the coefficient of variation of the response to one generation of divergent selection for the total number of bristles, starting at generation t ;

3) If the environmental covariance between the number of bristles on the 5th and the 6th sternites remains constant over lines, the variance of the phenotypic and the genetic covariances between these 2 traits will be the same and, therefore, their corresponding coefficients of variation ($CV(cov_{Pt})$, $CV(cov_{At})$) relate as follows:

$$CV(cov_{At}) = CV(cov_{Pt}) r_{Pt} / (r_{At} h_{Ft} h_{St}),$$

where h_{Ft}^2 and h_{St}^2 are the heritabilities of the number of bristles on the 5th and the 6th sternites, and r_{At} and r_{Pt} the genetic and phenotypic correlations between these 2 traits, all at generation t . Assuming that the genetic correlation r_{At} is one, it follows that $cov_{At} = V_{At}$, and if the phenotypic variances of both traits are equal:

$$h_{Ft}^2 = h_{St}^2 = h_t^2 (1 + r_{Pt}) / 2,$$

and then

$$CV(V_{At}) = 2CV(cov_{Pt}) r_{Pt} / (h_t^2 (1 + r_{Pt})).$$

The different estimates of $CV(V_{At})$ at generations 0, 4 and 10, all of them extremely large, are shown in Table III. Close agreement was found among these estimates in both experiments. The value of $CV(V_{At})$ substantially increased from generation 0 to 4, as predicted by theory. A further small increment was then observed in experiment 1, and the opposite in experiment 2.

Estimates of $CV(V_{At})$ are associated with large errors, even in the simplest case (1). Assuming normality of the distribution of the trait, the relationship between the variance

Table III. The coefficient of variation of the additive variance of the total number of bristles on both sternites among the lines at generation t , $CV(V_{At})$, estimated from the coefficients of variation of: 1) the phenotypic variances $CV(V_{Pt})$ of total bristle number; 2) the response to selection $CV(R_t)$ for total bristle number; 3) the phenotypic covariances $CV(cov_{Pt})$ between the number of bristles on the 5th and the 6th sternites (h_t^2 = heritability of the total number of bristles; r_{Pt} = phenotypic correlation between the number of bristles on each sternite) (see text for further explanation).

t	$CV(V_{Pt})/h_t^2$ (%)		$CV(R_t)/(1 - h_t^2/2)$ (%)		$2CV(cov_{Pt}) r_{Pt} / (h_t^2 (1 + r_{Pt}))$ (%)	
	1	2	1	2	1	2
0	80	75	101	—	84	82
4	166	191	125	171	175	178
10	196	173	161	153	214	200

of the phenotypic variance of independent lines $\widehat{V}(V_P)$, each based on n data, and its true value $V(V_P)$ is, as shown in the Appendix:

$$\widehat{V}(V_P) = V(V_P) + 2(\widehat{V}(V_P) + \overline{V_P})/(n-1),$$

where $\overline{V_P}$ is the average within-line phenotypic variance. When estimating $CV(V_{At})$ it follows that:

$$\widehat{CV}^2(V_{At}) = \widehat{CV}^2(V_{Pt})/h_t^4 = CV^2(V_{At}) + 2(1 + \widehat{CV}^2(V_{Pt}))/((n-1)h_t^4) \quad (1)$$

assuming h_t^2 is estimated without error.

From the values of h_t^2 and $CV(V_{Pt})$ in Tables I and II, the bias involved in the estimation of $CV_2(V_{At})$ can be calculated from expression (1) and used to correct the corresponding values of $CV(V_{At})$ in Table III (first two columns). Corrected estimates of $CV^2(V_{At})$ pooled over experiments were 0.12, 1.11 and 1.33 at generations 0, 4 and 10, respectively. The first coincides with its predicted value of $1/N$ (0.12 for $N=8$), as obtained from normal sampling theory (Avery and Hill, 1977). On the contrary, the last two were much larger than the expected asymptotic value of $2/3N$ (0.27 for $N=2.5$), as obtained for unlinked loci (Avery and Hill, 1979).

Discussion

The theoretical predictions that we intended to contrast in this work have been developed under certain assumptions that conditioned the choice of experimental design, trait and organism. In the first place, the effective population size of the lines was set up at its minimum value, as changes in $CV(V_{At})$ would be easier to detect in a small number of generations. Second, the genetic variance of abdominal bristle number is known to be essentially due to the segregation of additive genes, and most of the environmental variation of this trait can be ascribed to developmental noise (Robertson, 1955). The trait is also peripheral with respect to fitness (Robertson, 1955), and therefore segregating loci affecting bristle number can be assumed to be initially in linkage equilibrium. Furthermore, losses of lines during the period of inbreeding can be considered to occur at random with respect to the mean of the trait. Nevertheless, *Drosophila* has obvious disadvantages, as its small number of chromosomes renders likely the existence of blocks of genes affecting bristle number. As a result, the disequilibrium generated during the period of inbreeding will decrease very slowly after the size of the lines is expanded, since the recombination fraction between adjacent loci will be small and there is no crossing-over in males. Information on the number of loci affecting the trait is virtually non-existent (Hammond and James, 1972), although all 4 chromosomes have been found to carry them (Mather and Harrison, 1949). Nevertheless, it has been shown by Avery and Hill (1977) that unless the number of loci is very small, the fluctuation in additive variance among inbred replicates is mainly due to random departures from linkage equilibrium.

Previous experiments have been faced with 2 limitations. No predictions could be made on the evolution of the within-line phenotypic variance, since the magnitude of the special environment component is generally dependent on the degree of inbreeding. Furthermore, large sampling errors occurred. However, our data clearly show that when the

within-line component of the genetic variance is directly estimated, its value diminishes with increasing inbreeding. In parallel, the between-line component increases. A good fit between observed and expected values was obtained in both cases.

The variability of the genetic variance within lines in different generations has been represented by the corresponding coefficient of variation, as the magnitude of the genetic variance will change as inbreeding progresses. The value of $CV(V_{At})$ has been calculated indirectly from phenotypic data, as it would have been prohibitive to obtain accurate estimates of the additive variance in a sufficiently large number of lines. As indicated by theory, $CV(V_{At})$, estimated from within-line phenotypic variances or covariances, or from response to selection, has clearly increased with the degree of inbreeding. In the first 2 instances, its value at generation 4 was double that of the base population. In general, the values of $CV(V_{At})$ at generation 10 were similar to those obtained at generation 4. Although some reduction of $CV(V_{At})$ might be expected after a period of expanded population size, the decline of the previously generated linkage disequilibrium will be slow and most likely not observed in the short-term, as blocks of tightly linked loci have been identified affecting bristle number.

The validity of our estimates of $CV(V_{At})$ rests on the assumption of both general and special environmental variances being the same in all lines at a given generation, although they may change from one generation to another. Given that it is practically impossible to obtain identical experimental conditions, the estimates of $CV(V_{At})$ will be upwardly biased in any given generation. Notwithstanding, a positive bias can also be expected from sampling alone, as shown in the Appendix. However, if the biases have similar magnitude in different generations, the basic qualitative conclusion of $CV(V_{At})$ increasing with inbreeding from generation 0 to 4 and stabilizing when population size was expanded from generation 4 to 10, is not affected.

In the present experiment, the value of $CV(V_{At})$ after a severe bottleneck was found to be large. This result agrees with the theoretical prediction of Avery and Hill (1977), insofar as the expected value of the within-line additive variance of populations, previously subjected to a bottleneck, cannot be inferred from the values of the additive variance estimated in too small a number of replicates, if an acceptable degree of precision is sought.

Acknowledgments

We are grateful to Drs. A. Cuevas, A. Gallego, A. García-Dorado, W.G. Hill and M.A. Toro for their helpful comments and advice. This work was supported by a grant from Comisión Asesora de Investigación Científica y Técnica.

References

- Avery P.J. & Hill W.G. (1977) Variability in genetic parameters among small populations. *Genet. Res.* 29, 193-213
- Avery P.J. & Hill W.G. (1979) Variance in quantitative traits due to linked dominant genes and variance in heterozygosity in small populations. *Genetics* 91, 817-844
- Bowman J.C. & Falconer D.S. (1960) Inbreeding depression and heterosis of litter size in mice. *Genet. Res.* 1, 262-274

- Bulmer M.G. (1976) The effect of selection on genetic variability: a simulation study. *Genet. Res.* 28, 101-117
- Caballero A. & López-Fanjul C. (1987) An experimental evaluation of the usefulness of secondary traits in index selection, using *Drosophila melanogaster*. *J. Anim. Breed. Genet.* 104, 175-179
- Falconer D.S. (1981) *Introduction to Quantitative Genetics*, Longman, London, 2nd edn.
- Hammond K. & James J.W., (1972) The use of higher degree statistics to estimate the number of loci which contribute to a quantitative character. *Heredity* 28, 146-147
- Hill W.G. (1974) Variability of response to selection in genetic experiments. *Biometrics* 30, 363-366
- Kidwell J.F. & Kidwell M.M. (1966) The effects of inbreeding on body weight and abdominal chaeta number in *Drosophila melanogaster*. *Can. J. Genet. Cytol.* 8, 207-215
- López-Fanjul C. & Jodar B. (1977) The genetic properties of egg laying of virgin females of *Tribolium castaneum*. *Heredity* 39, 251-258
- Mather K. & Harrison B.J. (1949) The manifold effects of selection. *Heredity* 3, 1-52, 131-162
- Rasmuson M. (1952) Variation in bristle number of *Drosophila melanogaster*. *Acta Zool.* 33, 277-307
- Robertson A. (1955) Selection in animals: synthesis. *Cold Spring Harbor Symp. Quant. Biol.* 20, 225-229
- Russell W.C., Brinks J.S. & Richardson G.V. (1984) Changes in the genetic variances with increased inbreeding of beef cattle. *J. Hered.* 75, 8-10
- Tantawy A.O. (1957) Genetic variance of random-inbred lines of *Drosophila melanogaster* in relation to coefficients of inbreeding. *Genetics* 42, 121-136
- Wright S. (1951) The genetical structure of populations. *Ann.Eugen.* 15, 323-354

Appendix

Consider a set of l independent lines. The phenotypic variance of the trait in the i th line is V_i , and the mean and the variance of the distribution of the variances of the lines are V and $V(V_i)$, respectively. In each line, the phenotypic variance estimated from n data is \hat{V}_i . Therefore, V and $V(V_i)$ are estimated by:

$$\hat{V} = \sum_1^l \hat{V}_i / l, \quad \hat{V}(V_i) = \sum_1^l (\hat{V}_i - \hat{V})^2 / l.$$

$\hat{V}(V_i)$ is not an unbiased estimate of $V(V_i)$, and the bias involved can be calculated as follows. In each line:

$$E(\hat{V}_i - \hat{V})^2 = E(\hat{V}_i - V_i + V_i - \hat{V})^2 = E(\hat{V}_i - V_i)^2 + E(V_i - \hat{V})^2 + 2E((\hat{V}_i - V_i)(V_i - \hat{V})), \quad (1)$$

where,

$$E(\hat{V}_i - V_i)^2 = V(\hat{V}_i). \quad (2)$$

$$E(V_i - \hat{V})^2 = E(V_i - V + V - \hat{V})^2 = (V_i - V)^2 + E(V - \hat{V})^2 + 2(V_i - V)E(V - \hat{V}) = (V_i - V)^2 + V(\hat{V}), \quad (3)$$

since $E(V - \hat{V}) = 0$.

$$\begin{aligned} E((\hat{V}_i - V_i)(V_i - \hat{V})) &= V_i E(\hat{V}_i - V_i) - E(\hat{V}(\hat{V}_i - V_i)) = \\ &= -E(\hat{V}(\hat{V}_i - V_i)) = -E(\hat{V}_i(\hat{V}_i - V_i)) / l = V(\hat{V}_i) / l, \end{aligned} \quad (4)$$

since $E(\widehat{V}_i - V_i) = 0$.

Substituting (2), (3) and (4) in (1), we obtain:

$$E(\widehat{V}_i - \widehat{V})^2 = (V_i - V)^2 + V(\widehat{V}) + (l-2) V(\widehat{V}_i)/l.$$

Averaging over lines,

$$\widehat{V}(V_i) = V(V_i) + V(\widehat{V}) + (l-2) \sum_1^l V(\widehat{V}_i)/l^2,$$

$$\text{where } V(\widehat{V}) = \sum_1^l V(\widehat{V}_i)/l^2.$$

Therefore, the bias is always positive and is given by:

$$\widehat{V}(V_i) - V(V_i) = (l-1) \sum_1^l V(\widehat{V}_i)/l^2. \quad (5)$$

Assuming that phenotypic values are distributed normally, it is well known that:

$$\sum_1^l V(\widehat{V}_i) = 2 \sum_1^l \widehat{V}_i^2/(n-1) = 2l(\widehat{V}(V_i) + \widehat{V}^2)/(n-1),$$

and the bias is now given by:

$$\widehat{V}(V_i) - V(V_i) = 2(l-1)(\widehat{V}(V_i) + \widehat{V}^2)/(l(n-1)).$$

When the number of lines is large,

$$\widehat{V}(V_i) - V(V_i) \sim 2(\widehat{V}(V_i) + \widehat{V}^2)/(n-1).$$

The same procedure can be used to compute the bias involved in estimating $CV(V_{At})$ from the coefficient of variation of the responses to selection. The response to one generation of divergent selection in the i th line is R_i , and the mean and the variance of the distribution of the responses obtained in different independent lines are R and $V(R_i)$, respectively. In each line, N individuals are selected in each direction from M scored, and the estimated response is \widehat{R}_i . In a set of l lines, R and $V(R_i)$ are estimated by:

$$\widehat{R} = \sum_1^l \widehat{R}_i/l, \quad \widehat{V}(R_i) = \sum_1^l (\widehat{R}_i - \widehat{R})^2/l.$$

From expression (5), the bias is now:

$$\widehat{V}(R_i) - V(R_i) = (l-1) \sum_1^l V(\widehat{R}_i)/l^2 \sim \sum_1^l V(\widehat{R}_i)/l,$$

and from Hill (1974),

$$V(\widehat{R}_i) = 2\widehat{V}_i(h_i^2(1-h_i^2) + h_i^4(p+0.2) + p)/N,$$

where $p = NM$ and h_i^2 is the heritability of the selected trait in the i th line.