Theor Appl Genet (1989) 77:142–148

Effect of mass selection on the within-family genetic variance in finite populations

E. Verrier*, J. J. Colleau and J. L. Foulley

INRA, Station de Génétique Quantitative et Appliquée, Centre National de Recherches Zootechniques, F-78350 Jouy-en-Josas, France

Received November 15, 1988; Accepted June 10, 1988
Communicated by L. D. Van Vleck

Summary. The adequacy of an expression for the within-family genetic variance under pure random drift in an additive infinitesimal model was tested via simulation in populations undergoing mass selection. Two hundred or one thousand unlinked loci with two alleles at initial frequencies of 1/2 were considered. The size of the population was 100 (50 males and 50 females). Full-sib matings were carried out for 15 generations with only one male and one female chosen as parents each generation, either randomly or on an individual phenotypic value. In the unselected population, results obtained from 200 replicates were in agreement with predictions. With mass selection, within-family genetic variance was overpredicted by theory from the 12th and 4th generations for the 1,000 and 200 loci cases, respectively. Taking into account the observed change in gene frequencies in the algorithm led to a much better agreement with observed values. Results for the distribution of gene frequencies and the within-locus genetic covariance are presented. It is concluded that the expression for the within-family genetic variance derived for pure random drift holds well for mass selection within the limits of an additive infinitesimal model.

Key words: Genetic variance – Selection – Inbreeding – Additive infinitesimal model

Introduction

Inbreeding and selection are two major factors affecting genetic variance in a population. Inbreeding leads to gene identity within a population, and so reduces genetic variance. Selection induces a change in gene frequencies, which is closely related to the magnitude of gene effects on the selected character; if the number of loci involved is large, the change in gene frequencies and its consequences on genetic variance can be neglected (Crow and Kimura 1970, pp 236–239). However, Bulmer (1971) showed that selection can induce an important and immediate change in genetic variance due to linkage disequilibrium. These two different possibilities of the impact of selection on genetic variance were first described but not quantified by Lush (1945, pp 141–143).

For predicting the joint effects of inbreeding and selection on additive genetic variance, it is useful to notice that the variance at a given generation depends on the variance of the selected parents and on the variance created by gene segregation during meiosis (Dempfle 1975; Chevalet 1987). The first term, the between family variance (VA_b), is related to the effect pointed out by Bulmer (1971); it is a function of the total variance of the previous generation, the selection intensity in each parental sex, and the accuracy of the selection criterion. Under the assumption of multivariate normality of breeding values and phenotypic values, VA_b is given by an equation of Pearson (1903). The second term is the within-family variance (VA_w). For a selected and randomly mated population of infinite size (and, therefore, non-inbred), the VA_w of a character determined by an infinite number of unlinked loci is constant over generations. In this case, this variance is equal to half the initial total genetic variance (V_0) in the population taken at equilibrium (Bulmer 1971). On the other hand, in an inbred population, the frequency of homozygotes increases at some loci, and for these loci segregation cannot regenerate any variance. Foulley and Chevalet (1981) proved that under an additive infinitesimal model, the expectation of VA_w in an inbred and unselected population is the following function of V_0 and of the sire (F_s) and dam (F_d) inbreeding coefficients:

$$E(VA_w) = (V_0/2)[1-(F_s+F_d)/2] \quad (1)$$

* Permanent address: INA-PG, Département des Sciences Animales, 16 rue Claude Bernard, F-75231 Paris Cedex 05, France

This formula was also provided by Quaas and Pollak (1980) in deriving the "reduced animal" model.

The purposes of this paper are: (i) to investigate by simulation whether equation (1) holds for selected populations of finite size, as conjectured by Sorensen and Kennedy (1984), and (ii) to study the genetic structure of such populations.

Methods

The simulated character considered is strictly additive; it is determined by a large number of loci with recombination rates of 1/2. At each locus there are 2 alleles having the same frequency (1/2) in the base population and with effects 0 and α, respectively: α is positive and such that V_0 is equal to 1. The individual genetic value is the product of the total number of "favorable" alleles times α. The phenotypic value is the sum of the genetic value and an independent environmental value which is normally distributed with mean 0 and variance 1. Thus, initial heritability is 0.5.

The initial generation included 100 individuals (50 males and 50 females) whose genes were chosen at random and independently. The population was maintained with full-sib matings during 15 generations; at each generation, one male and one female were chosen to procreate the next generation of 100 individuals. Under this special mating system, the parents' inbreeding coefficients are known exactly, and the total genetic variance amounts to the within-family variance. On the other hand, selection intensity is maximum, which is a stringent situation to test the hypothesis.

The number of loci must be large enough to assess the properties of the infinitesimal model, especially the stability of gene frequencies under selection. The expected change in frequency of a particular allele at one locus after one generation of selection is proportional to his average effect on the selected trait (Crow and Kimura 1970, formula 5.8.13, p 229) and consequently for a given genic variance, inversely proportional to the square root of the number of loci. A first model with 200 loci was simulated and used to compare two settings: choice of parents at random (population C) and choice of the best male and the best female on the basis of individual phenotypic values (population S). Using the formula previously quoted, adapted to this situation, it can be shown that in population S, selection leads to an expected initial increase in frequency of the favorable allele of 0.040 (+8%). To reduce this increase to 1%, about 12,600 loci would have been required. Because of computing constraints, only 1,000 loci could be simulated, leading to an expected initial increase of 3.5%; this setting was used also to obtain more detailed results from a mass selected population (population T). Each simulation set was run with 200 replicates.

At each generation and for each replicate, the genetic variance (VA) was computed from the individual genetic values, and the variance of gene effects (Va) was calculated using its definition based on gene effects and frequencies. The relationship between genetic and genic variances is given by:

$$VA = \mathrm{Var}\left[\sum_i (a_i + a_i^*)\right]$$

$$= \sum_i (\mathrm{Var}\, a_i + \mathrm{Var}\, a_i^*) \quad (2a)$$

$$+ 2\sum_i \mathrm{Cov}(a_i, a_i^*) \quad (2b)$$

$$+ 2\sum_{i\neq j} [\mathrm{Cov}(a_i, a_j^*)] + \sum_{i\neq j} [\mathrm{Cov}(a_i, a_j) + \mathrm{Cov}(a_i^*, a_j^*)] \quad (2c)$$

$$= Va + Chw + Cl \quad (2)$$

where a_i, a_i^* represent the effects of paternal (a_i) and maternal (a_i^*) genes at locus i. According to the terminology of Bulmer (1976), Chw (2b) is the covariance between allelic effects within-locus, and Cl (2c), the covariance between allelic effects in different loci. Chw and Cl were computed only for the 1,000 loci model. Under the assumptions of the simulations, the terms of (2) were directly computed as functions of α and of gene or genotypic frequencies:

$$Va = 2\alpha^2 \sum_i p_i(1-p_i) \quad (3a)$$

$$Chw = 2\alpha^2 \sum_i (f_i - p_i^2) \quad (3b)$$

$$Cl = VA - Va - Chw \quad (3c)$$

where p_i is the frequency of the favorable allele at the ith locus and f_i is the frequency of homozygotes for this allele.

The average frequency of the favorable allele within a replicate was computed as the mean of the p_i's. For the 1,000 loci model, the distribution of gene frequencies was derived by sorting the p_i's into 0.01 range classes and calculating the corresponding number of loci with that class.

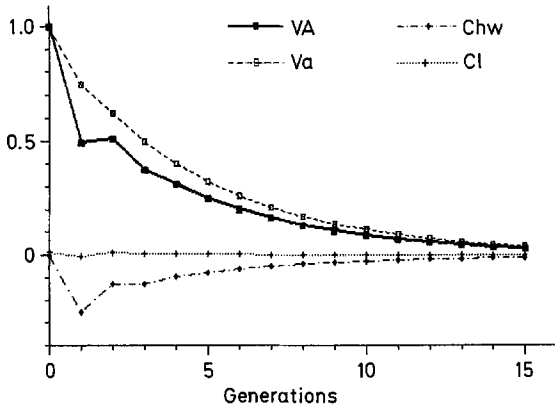

Fig. 1. Evolution of the genetic (*VA*) and genic (*Va*) variances, and covariance components (*Chw*, *Cl*), in a selected population under full-sib mating. 1,000 loci model, mean of 200 replicates

Results

Figure 1 shows the evolution of the genetic variance and of its components, for the 1,000 loci model. The genetic variance (VA) decreased drastically in the first generation; in fact, the between family variance (half of the initial variance) is immediately lost under this selection scheme. After a stabilization in generations 1 and 2, VA decreased regularly down to nearly zero. The genic variance (Va) was always larger than the genetic variance. The difference between VA and Va results from a negative within-locus covariance (Chw), whereas the covariance between loci (Cl) was significantly different from zero only at generation 2, 6, and 7.

The evolution of genetic variance (VA) in the three simulated populations and a comparison between ob-

served and predicted values from (1) are shown in Table 1. The decrease in the genetic variance over time was larger in the selected populations than in the unselected one, as expected. For 200 loci with no selection, it is an exception when the observed variance is different from the expected value (generation 3). For 200 loci under selection, the observed and predicted values differed beyond the 4th generation. In the 1,000 loci case, significant differences were observed only from generation 12 onwards. These differences, expressed as a percentage of the expected value of VA, increased regularly (in absolute value) to

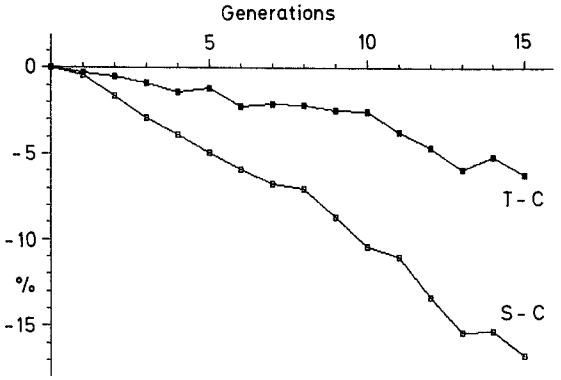

Fig. 2. Differences between values of the genic variance (*Va*) observed in the selected (*S, T*) and unselected (*C*) populations under full-sib mating, as a percentage of observed values in *C*. Mean of 200 replicates

−14.5% and −4.5% in populations S and T, respectively, in generation 15.

The genic variance (Va) decreased more quickly under selection than under pure drift (Fig. 2); this difference in variance was larger for the S than for the T population, i.e., when the number of loci involved was smaller.

The average frequency of the favorable allele did not deviate from its initial value in the unselected population (Fig. 3). On the other hand, significant increases were observed in the selected populations for this frequency and for the genetic mean of the trait (of about 7 times the initial genetic standard deviation in the two populations). As expected, the smaller the number of loci, the larger was this increase in gene frequencies. The fast change in the early generations in the distribution of gene frequency among the 1,000 loci in the T population is shown in Fig. 4. In generation 0, the distribution of the frequency for the favorable allele was unimodal, with mean 1/2. Later on, five characteristic peaks appeared: three centered at frequencies of 1/4, 1/2, and 3/4 and two at extreme frequencies of 0 and 1. The location of these peaks is due to the sampling of 4 genes at each locus among 2 parents in each line, which imposes discrete variation from 0 to 1 with steps of 1/4. On the other hand, the variability around the intermediate peaks, at 1/4 and 3/4, results from the sampling of genes among a limited number of candidates for selection; this number is large enough to make the original binomial distribution

Table 1. Evolution of within-family genetic variance (VA_w) in the simulated populations. Comparison with predictions from Foulley-Chevalet's equation

Generation	Predicted		Observed within-family genetic variance								
			200 loci						1,000 loci and selection		
			No selection			selection					
	F[a]	VA_w	$\bar{X}$[b]	σ[c]		$\bar{X}$	σ		$\bar{X}$	σ	
1	0.0000	0.5000	0.5051	0.0753	NS	0.5031	0.0725	NS	0.4903	0.0739	NS
2	0.2500	0.5000	0.4978	0.0773	NS	0.4974	0.0711	NS	0.5067	0.0746	NS
3	0.3750	0.3750	0.3850	0.0558	*	0.3711	0.0592	NS	0.3744	0.0573	NS
4	0.5000	0.3125	0.3122	0.0532	NS	0.3092	0.0480	NS	0.3131	0.0518	NS
5	0.5938	0.2500	0.2493	0.0424	NS	0.2392	0.0405	***	0.2502	0.0363	NS
6	0.6719	0.2031	0.2030	0.0396	NS	0.1917	0.0368	***	0.2011	0.0306	NS
7	0.7344	0.1641	0.1634	0.0335	NS	0.1574	0.0297	**	0.1631	0.0260	NS
8	0.7852	0.1328	0.1334	0.0280	NS	0.1239	0.0265	***	0.1309	0.0218	NS
9	0.8262	0.1074	0.1087	0.0223	NS	0.0977	0.0215	***	0.1052	0.0158	NS
10	0.8594	0.0869	0.0868	0.0199	NS	0.0780	0.0192	***	0.0850	0.0150	NS
11	0.8862	0.0703	0.0701	0.0170	NS	0.0626	0.0172	***	0.0686	0.0127	NS
12	0.9080	0.0569	0.0571	0.0146	NS	0.0500	0.0154	***	0.0553	0.0099	*
13	0.9255	0.0460	0.0468	0.0142	NS	0.0400	0.0135	***	0.0441	0.0086	**
14	0.9398	0.0372	0.0374	0.0131	NS	0.0318	0.0117	***	0.0354	0.0071	***
15	0.9513	0.0301	0.0312	0.0120	NS	0.0262	0.0101	***	0.0287	0.0055	***

[a] Average inbreeding coefficient
[b] Mean of 200 replicates
[c] Standard deviation between replicates, significance levels: * $P<0.05$, ** $P<0.01$, *** $P<0.001$

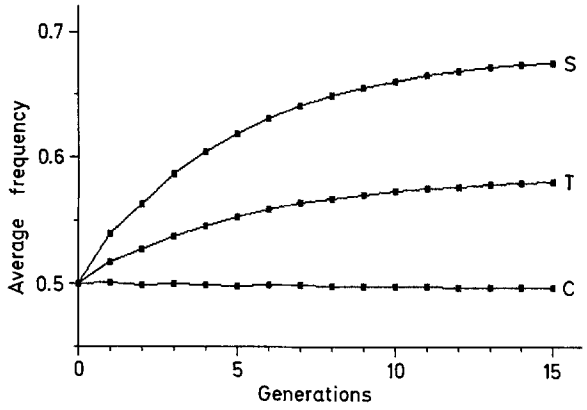

Fig. 3. Evolution of the average frequency of the favorable allele in the simulated populations. Mean of 200 replicates

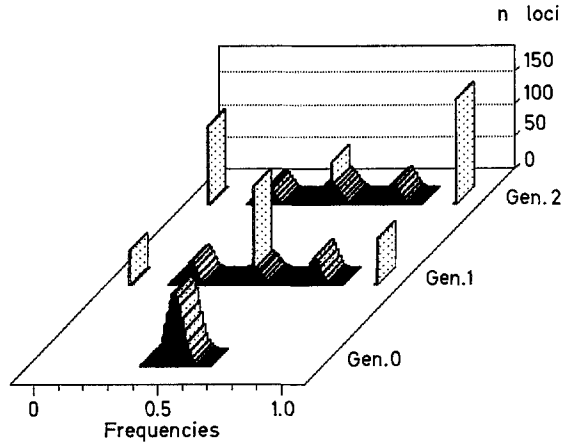

Fig. 4. Distribution of frequencies of the favorable allele in the genome during the first generations of selection and full-sib mating. 1,000 loci model, mean of 200 replicates

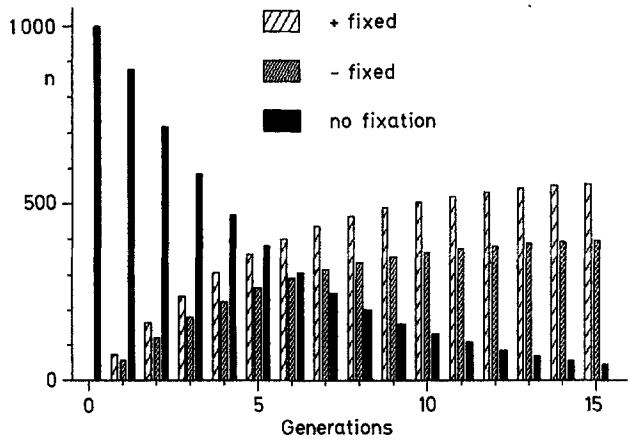

Fig. 5. Evolution of number (n) of loci where the favorable (+) or unfavorable (−) allele is fixed, under selection and full-sib mating. 1,000 loci model, mean of 200 replicates

around the peak very close to a normal one. The distribution around 1/2 is more complex: it involves a frequency of 1/2 without any variation due to progeny out of homozygous parents for either the favorable or the unfavorable allele at each locus; the other case corresponds to a gene frequency varying around 1/2 in progeny from matings of heterozygotes. As expected, the peak at 1 was larger than at 0, due to the effect of selection. The evolution of the number of loci undergoing fixation is reported in Fig. 5. The number of segregating loci decreased very quickly from 1,000 to 44 at generation 15; in this generation 398 unfavorable and 558 favorable alleles were fixed.

Discussion

The evolution of genetic variance and its prediction

The derivation of Eq. (1) by Foulley and Chevalet (1981) is based on two major assumptions: an infinite number of unlinked loci with strictly additive genes, and no selection. The simulation of an unselected population provided a numerical check of this equation. Results for selected populations show differences in the middle- and long-term, which may have several possible causes: inadequacy of the algorithm in the selection situation, or incorrect simulation of the infinitesimal model.

The number of loci plays a role as observed differences were larger with 200 loci than with 1,000. As a consequence of the mating system, the inbreeding coefficients at a given generation were the same in the three populations and for all individuals. Therefore, the difference between observed and predicted variances registered in selected populations can be ascribed to differences in genic variance due to significant changes in gene frequencies. Actually, the experimental design used in this study (initial gene frequencies at 0.5 and high selection intensity) provided conditions for maximum variation in genic variance (Va). A simple way to accomodate Eq. (1) for these changes in frequencies consists in replacing V_0 in (1) by the genetic variance in a non-inbred population at equilibrium, and where the frequency of the favorable allele at each locus is the observed average frequency in the parental generation. This can be written as follows:

$$E[VA_w^{[t+1]}] = \alpha^2 \, n \, p^{[t]}(1-p^{[t]})[1-(F_s^{[t]}+F_d^{[t]})/2] \qquad (4)$$

where n is the number of loci and $p^{[t]}$, the observed average frequency of the favorable allele at generation t in which the parents are born. Obviously, this modification would not be feasible in practice because of the difficulty in assessing gene frequencies. Anyhow, after this modification, the predictions agree much better with the observed genetic variance, as observed differences are almost always not significant (Fig. 6).

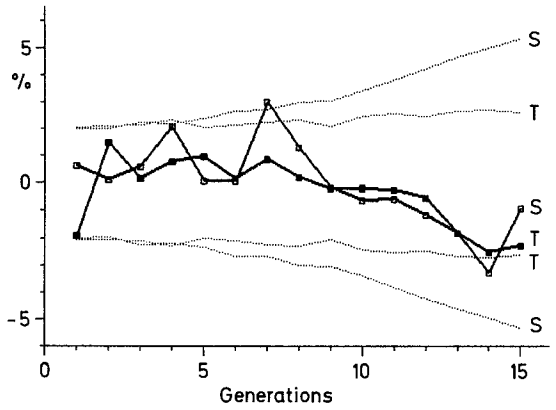

Fig. 6. Check of Foulley-Chevalet's equation after adjusting for the observed evolution of gene frequencies under selection. *Solid lines* represent differences between observed and expected values of VAw, as a percentage of (4), in the two selected populations (*S*, *T*). *Dotted lines* represent critical values for statistical significance at the 5% level, computed in the two populations using the observed standard deviation between replicates

The evolution of covariance components

Observed values of genetic variance (VA) in the three populations were always smaller than those of genic variance (Va), which implies negative covariances between gene effects. Results in the 1,000 loci model showed that only the covariance within-locus (Chw) was affected, since the covariance between loci (Cl) was null. In the situation of pure drift, a null expectation for Chw was implicitly assumed by Jacquard (1974, p 213), and also for the sum (Chw + Cl) by Bulmer (1980, pp 229–231). On the other hand, Bulmer (1976) postulated that Chw is negative, a priori, depending on Va and on the effective population size (Ne) as:

$$\mathrm{Chw} = -\mathrm{Va}/2\,\mathrm{Ne} \qquad (5)$$

Starting from expression (3b), it is possible to show (see Appendix) that under the assumed determinism of the trait, and with equal numbers of male and female parents, the observed value of Chw in the offspring born at generation t is the following function of differences in gene frequencies in the parental groups:

$$\mathrm{Chw}^{[t]} = -2\sum_i \alpha_i^2\,[p_{i,m}^{*[t-1]} - p_{i,f}^{*[t-1]}]^2 \qquad (6)$$

where $p_i^{*[t-1]}$ is the frequency of the favorable allele at locus i in the parents (males and females for indices m and f respectively) born and chosen at generation $t-1$, and α_i is the effect of this allele, assuming that the other one has a null effect (α_i was equal to a constant α in the simulations).

Assuming no selection has occurred, an expression for the expectation of Chw is (see Appendix):

$$E[\mathrm{Chw}^{[t]}] = -E[\mathrm{Va}^{[t-1]}]/2\,(N_m + N_f) \qquad (7)$$

where N_m, N_f are the numbers of male (N_m) and female (N_f) parents. Equation (7) differs from (5) because it does not use Ne, and it involves genic variance in the previous generation. Formula (7) shows that the within-locus covariance between gametes, due to a departure from Hardy-Weinberg proportions, decreases when the population size or the number of generations increases. When $N_m + N_f$ or t tend to infinity, E(Chw) is null.

In a selected line, the conditional expectations of $p_{i,m}^*$ and $p_{i,f}^*$ in (6), given the gene frequency in the pool from which genes are sampled, may differ due to different selection practices in the two sexes (Ziehe and Gregorius 1981) resulting in an increase in E(Chw) compared to a pure random drift situation. On the contrary, if the same selection rules are used in males and females, selection will make male and female gene frequencies more homogeneous and, consequently, will reduce E(Chw). Gene frequencies in the parental groups were not computed in the present study. However, in the T population, values of (Chw) averaged over replicates were about 22% smaller than those predicted with (7). It is difficult to put more explicitly the joint effects of selection and drift on the within-locus covariance, but in the middle-term, effects of drift are likely to prevail leading to a gradual exhaustion of genic variance within a population and, consequently, to an exhaustion of Chw.

When the number of parents is not too small, Chw can be neglected, as shown by simulation results of Mueller and James (1983). In such a situation, the difference between genetic (VA) and genic (Va) variances is due mainly to a negative covariance between loci (Cl), arising from linkage disequilibrium. This effect of selection occurs only on the between full-sib families component of variance as proved by Bulmer (1971); under the present mating system, linkage disequilibrium cannot be generated by selection, but only by drift. Therefore, the values of Cl different from zero have not the same sign and result only from sampling.

The evolution of the distribution of gene frequencies

The fast evolution of the distribution of gene frequencies is a consequence of the very small effective size of the simulated populations, which provides extreme conditions for drift. An important fact is that fixation occurs as early as in the first generation, when inbreeding has not yet appeared. Also, the proportion of fixed loci was always larger ($P < 0.001$) than the coefficient of inbreeding. Because gene sampling occurs from the first generation and because of the absence of self-fertilization, there isn't a one-to-one parallelism between inbreeding and allelic fixation in a dioecius population of limited size.

Results from the 1,000 loci model also illustrated the theory of conflict between selection and drift introduced by Wright (1931), as indicated by the typical pattern of

curves in Fig. 4. Robertson (1960) developed a theory for a quantitative trait using Kimura's formula (1957) for the probability of fixation of a gene under selection. If this formula is applied to the situation of this study, it gives a probability of 0.571 that a favorable allele becomes ultimately fixed. This value is in agreement with the proportion of favorable alleles fixed among non-segregating loci, which was observed at generation 15 (i.e. 0.584). If it is possible to say that after 15 generations of selection 398 unfavorable alleles were fixed by chance, it can also be said that the same number of favorable alleles were fixed by chance, and only 160 out of 558 (29%) by selection. This extreme situation is due to the size of the parental population which, in this case, was the smallest one possible. Simulation results of Sirkkomaa and Lindstrom (1981) showed the effect of effective size on allelic fixation. For the same selection intensity and with a larger effective population size (10 males and 20 females selected out of 750 individuals in each sex) than in the present study, they found that the proportion of loci with the unfavorable allele fixed was only 1% (over 36 loci), and the proportion of loci segregating still 46% after 20 generations of mass selection.

Impact on BLUP techniques

The assumptions underlying formula (1) are implicitly made when breeding values are predicted with BLUP (Best Linear Unbiased Prediction) under an "animal" model. Using statistical arguments, it has been shown that BLUP is unaffected by selection decisions provided these are based on translation invariant functions of the data, all records are available, pedigree information is complete up to the base population, and the variance covariance parameters are known at least to proportionality in the base population (Henderson 1975; Goffinet 1983). The condition of translation invariance has been recently relaxed by Fernando et al. (1985).

Foulley and Chevalet's algorithm is based on the assumption that the conditional distribution of the additive genetic value of an offspring (A_o) given that of his sire (A_s) and dam (A_d) has mean $(A_s+A_d)/2$ and variance $V_0(1-\bar{F})/2$, where $\bar{F}$ is the average of parental inbreeding coefficients. This property was already implied by Bulmer (1971), Dempfle (1975), and Thompson (1977), and was used implicitly in the direct method of computing the inverse of a relationship matrix proposed by Henderson (1976) and Quaas (1976) as shown by Sorensen and Kennedy (1984). However, the moments of the distribution of $A_o|A_s, A_d$ were explicitly formulated in the previous form only when the "reduced animal" model was proposed by Quaas and Pollak (1980). This highlights the fact that the appealing properties of predicting individual genetic merits with BLUP techniques under the animal model rely on the genetic infinitesimal model, with strict additivity of gene effects. This basic assumption should be kept in mind even if simulation results (Maki-Tanila and Kennedy 1986) suggest that mixed model methodology may be robust to a small number of loci and dominance effects.

Conclusion

The simulations presented show that there is still no formal argument to invalidate the algorithm proposed by Foulley and Chevalet (1981), in a situation of selection and within the limits of the additive infinitesimal model. This equation can be used to improve predictions of the evolution of total genetic variance in selected populations of limited size. The evolution of the average inbreeding coefficient is then predicted as a function of the number of selected parents and the selection rules. These issues are currently under examination. The final objective is to derive a general algorithm for selection schemes to optimize these with respect to cumulated genetic response in the middle- and long-terms. Moreover, this study highlights some of the joint effects of drift and selection on the genetic structure of finite populations.

Acknowledgements. The authors wish to thank D. Gianola for useful comments and an anonymous referee for the review of the manuscript.

Appendix

Derivation of E (Chw)

Using the superscript [t] to designate generation number, the expression for Chw in (3b) can be written more generally as:

$$\mathrm{Chw}^{[t]}=2\sum_i \alpha_i^2\,[f_i^{[t]}-(p_i^{[t]})^2]=2\sum_i \alpha_i^2\,\Delta_i^{[t]} \qquad \text{(A1)}$$

where α_i is the effect of the favorable allele at locus i, assuming that the other one has a null effect, and $\Delta_i^{[t]}$ is a measure at this locus of the departure from Hardy-Weinberg equilibrium in offspring born at generation t. Ignoring superscripts, note that

$$\Delta=f-p^2 \qquad \text{(A2)}$$

Δ can also be written as a function of genotypic frequencies (Bulmer 1976) and, under this form, is closely related to the criterion proposed by Haldane (1954) for testing randomness of mating. Let N_m and N_f be the numbers of male and female parents, respectively. In these two groups we can find x_m and x_f favorable alleles, respectively, and a total number x. The frequency of the favorable allele is then:

$$p_m=x_m/2N_m \qquad \text{(A3a)}$$

$$p_f=x_f/2N_f \qquad \text{(A3b)}$$

$$p=x/2(N_m+N_f) \qquad \text{(A3c)}$$

Using (A3) in (A2), we have:

$$\Delta=-[N_m^2 p_m^2+N_f^2 p_f^2-(N_m^2+N_f^2)\,p_m p_f]/[N_m+N_f]^2 \qquad \text{(A4)}$$

For the sake of simplicity, we will consider in detail only the simple case where the numbers of male and female parents are equal, and show how the results can be extended to the general case. If $N_m = N_f = N/2$, Eq. (A4) becomes:

$$\Delta = -1/4(p_m - p_f)^2 \quad \text{(A5)}$$

Clearly, Δ is negative and is due to the difference in frequency of the favorable allele in the two parental groups. Alternatively, using Eqs. (A3) in (A5),

$$\Delta = -(p_m - p)^2$$

Using superscripts, Eq. (A1) becomes:

$$\text{Chw}^{[t]} = -2\sum_i \alpha_i^2 [p_{i,m}^{*[t-1]} - p_i^{[t]}]^2 \quad \text{(A6)}$$

where the term in brackets represents the difference between the frequency of the favorable allele at locus i in the male parents chosen at generation $t-1$, and the probability that an offspring born at generation t will receive this allele.

Consider the conditional expectation of (A6) as if male parents were sampled from an infinite genetic pool at $t-1$:

$$E[[p_{i,m}^{*[t-1]} - p_i^{[t]}]^2 \mid p_i^{[t-1]}]$$

Due to the infinite size of the genetic pool, this term is equivalent to:

$$E[[p_{i,m}^{*[t-1]} - p_i^{*[t-1]}]^2 \mid p_i^{*[t-2]}] \quad \text{(A7)}$$

(A7) can be expanded as the sum of three terms which are functions of conditional variances, covariances, and expectations of gene frequencies. In a pure drift situation, we have:

$$E[p_{i,m}^{*[t-1]} \mid p_i^{*[t-2]}] = E[p_i^{*[t-1]} \mid p_i^{*[t-2]}] = p_i^{*[t-2]}$$

$$\text{Var}[p_{i,m}^{*[t-1]} \mid p_i^{*[t-2]}] = p_i^{*[t-2]}(1 - p_i^{*[t-2]})/N$$

$$\text{Var}[p_i^{*[t-1]} \mid p_i^{*[t-2]}] = p_i^{*[t-2]}(1 - p_i^{*[t-2]})/2N$$

$$\text{Cov}[p_{i,m}^{*[t-1]}, p_i^{*[t-1]} \mid p_i^{*[t-2]}] = p_i^{*[t-2]}(1 - p_i^{*[t-2]})/2N$$

Then, (A7) amounts to:

$$p_i^{*[t-2]}(1 - p_i^{*[t-2]})/2N \quad \text{(A8)}$$

As $p_i^{*[t-2]} = p_i^{[t-1]}$, and substituting (A8) into (A6), one obtains:

$$E[\text{Chw}^{[t]} \mid p_i^{[t-1]}] = -2\sum_i \alpha_i^2 p_i^{[t-1]}(1 - p_i^{[t-1]})/2N = -\text{Va}^{[t-1]}/2N$$

where $\text{Va}^{[t-1]}$ is the observed genic variance at generation $t-1$. Finally, the expectation of Chw over replicates is:

$$E[\text{Chw}^{[t]}] = -E[\text{Va}^{[t-1]}]/2N \quad \text{(A9)}$$

In the general case, where numbers of parents may be unequal, Eq. (A4) can be written as follows:

$$\Delta = -[N_m/N_f]p_m^2 + [p(N_m + N_f)/N_f]p_m - p^2 \quad \text{(A10)}$$

Determining the sign of (A10) shows that Δ is not always negative or null. However, using previous arguments, one obtains an expression analogous to (A9):

$$E[\text{Chw}^{[t]}] = -E[\text{Va}^{[t-1]}]/2(N_m + N_f) \quad \text{(A11)}$$

References

Bulmer MG (1971) The effect of selection on genetic variability. Am Nat 105:201–211

Bulmer MG (1976) The effect of selection on genetic variability: a simulation study. Genet Res 28:101–117

Bulmer MG (1980) The mathematical theory of quantitative genetics. Clarendon Press, Oxford

Chevalet C (1987) Control of genetic drift in selected populations. Proc 2nd Int Conf Quant Genet. Raleigh NC, pp 379–394

Crow JF, Kimura M (1970) An introduction to population genetics. Harper and Row, New York

Dempfle L (1975) A note on increasing the limit of selection through selection within families. Genet Res 24:127–135

Fernando RL, Gianola D, Henderson CR (1985) Inferences on breeding values and on genetic and environmental parameters when selection and non-random mating occur. In: Annu meeting Am Soc Animal Sci. University of Georgia, Athens/GA. Abstract no. 31

Foulley JL, Chevalet C (1981) Méthode de prise en compte de la consanguinité dans un modèle simple de simulation des performances. Ann Génét Sél Anim 13:189–196

Goffinet B (1983) Selection on selected records. Génet Sél Evol 15:91–97

Haldane JBS (1954) An exact test for randomness of mating. J Genet 52:631–635

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423–449

Henderson CR (1976) A simple method for computing the inverse of a numerator matrix used in prediction of breeding values. Biometrics 32:69–84

Jacquard A (1974) The genetic structure of populations. Springer, Berlin Heidelberg New York

Kimura M (1957) Some problems of stochastic processes in genetics. Ann Math Stat: 28, 882–901

Lush JL (1945) Animal breeding plans. 3rd edn. Iowa State University Press, Ames/IA

Maki-Tanila A, Kennedy BW (1986) Mixed model methodology under genetic models with a small number of additive and non-additive loci. Proc 3rd World Cong Genet appl to Livestock Prod, vol. 12:443–448

Mueller JP, James JW (1983) Effect on linkage disequilibrium of selection for a quantitative character with epistasis. Theor Appl Genet 65:25–30

Pearson K (1903) Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. Phil Trans R Soc London Ser A 200:1–66

Quaas RL (1976) Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics 32:949–953

Quaas RL, Pollak EJ (1980) Mixed model methodology for farm and ranch beef cattle testing programs. J Anim Sci 51: 1277–1287

Robertson A (1960) A theory of limits in artificial selection. Proc R Soc London Ser B 153:234–249

Sirkkomaa S, Lindstrom UB (1981) Simulation of response to selection for body weight in rainbow trout. Acta Agric Scand 31:426–432

Sorensen DA, Kennedy BW (1984) Estimation of response to selection using least-squares and mixed model methodology. J Anim Sci 58:1097–1106

Thompson R (1977) The estimation of heritability with balanced data. II. Data available on more than two generations. Biometrics 33:497–504

Wright S (1931) Evolution in mendelian populations. Genetics 16:97–159

Ziehe M, Gregorius HR (1981) Deviations of genotypic structures from Hardy-Weinberg proportions under random mating and differential selection between the sexes. Genetics 98:215–230