Theor. Appl. Genet. 58, 177-180 (1980)

Is Fisher's Model Necessary for the Theory of Population Improvement ?

A. Gallais
Station d'Amélioration des Plantes Fourragères, INRA, Lusignan (France)

Summary. It is shown here that genetic advance in one cycle of recurrent selection can be formulated directly in terms of covariances between relatives by application of the general statistical principle of linear prediction. For practical use of such formulae it is necessary to estimate the corresponding covariance between relatives from the mating design used. With General Combining Ability selection such estimation is direct. For other types of selection, it is necessary to derive associated covariances from other types of covariances but it is not necessary to use classical results of covariances between relatives in terms of genetic effects. Indeed, covariances can be derived without factorial decomposition of the genetic effects at one locus, i.e., without the concept of additivity and dominance. This approach allows a simple derivation of the genetic advance after n cycles of selection, followed by m generations of intercrossing, with a minimum of assumptions.

Key words: Breeding theory – Quantitative genetics – Population improvement – Recurrent selection

Introduction

By Fisher's model I mean the factorial decomposition in additive and dominance effects of the value of a genotype reduced to one locus. I intend to show that this model is not necessary: (1) for formulating a general expression of the genetic advance whatever the population improvement procedure, (2) for estimating the covariances between relatives associated with the breeding method – i.e. – the covariance between the value of the parents according to any particular system of testing and the value of their offspring after intercrossing. Demonstration of the theory will be in terms of population improvement by recurrent selection.

A General Formulation of Genetic Advance in Recurrent Selection

Recurrent selection is a method of population improvement where the selected units (individuals or families) are intercrossed once or several times to produce the next generation. The main issue addressed is the prediction of the value of the offspring after intercrossing, knowing the value of the parents according to a particular system of test (Gallais 1977). To simplify the notation I denote by T the value of the parents according to a particular system of testing and by M the value of their offspring after one generation of intercrossing.

Suppose (assumption 1) that T and M are distributed according to a bivariate normal distribution with means respectively E(T) and E(M) and with variances and covariance var T, var M, cov TM. Then, whatever the genetic effects, level of ploïdy, and criterion of test, according to the linear principle of prediction, the value of M, knowing T, will be, with selection on one sex:

$$\hat{M} = E(M) + \frac{\text{cov TM}}{\text{var T}} (T - \bar{T}) \qquad (1)$$

This is equivalent to predicting the general combining ability of the individuals tested.

From (1) the general expression of genetic advance from generation n to generation n + 1 gives:

$$\mu_{n+1} - \mu_n = (\mu_{n,1} - \mu_n) + i\frac{\text{cov TM}}{\sqrt{\text{var T}}}, \qquad (2)$$

in which μ_n is the value of the population at generation n, $\mu_{n,1}$ is its value after intercrossing without selection in n and i is the selection intensity in standard units.

In establishing Eq. (1) and (2), no direct genetic assumptions are involved. However, to use these formulae to predict genetic advance, it is necessary to know the difference $(\mu_{n,1} - \mu_n)$ and cov TM (var T being estimated directly).

0040-5752/80/0058/0177/$ 1.00

It is often assumed that $(\mu_{n1} - \mu_n)$ is zero. For diploïds, this is true only in the absence of epistasis (assumption 2). For autopolyploïds it is true only in the absence of interactions between alleles (Gallais 1975). In what follows I will only consider the situation of diploïdy in the absence of epistasis; consequently it is only necessary to estimate cov TM. The case of epistasis will be considered in the appendix.

Estimation of the Covariances Between Relatives Associated with the Breeding Method

Cov TM is the covariance associated with the breeding method. It is defined at the population level. This is a covariance between relatives. For example, with individual phenotypic (mass) selection, this is the covariance between parents and their offspring (cov PO) and with general combining ability selection (GCA) it is a covariance between half-sibs (cov HS). The problem is then to estimate cov TM.

Note that with GCA selection, the value according to T is equivalent to the value according to M; hence without assumption (except no. 1) cov TM will be estimated by the variance among half sib families.

For other breeding methods the principle is to derive cov TM from estimable covariances between relatives of the mating design. According to known results from the application of the Fisher's model, it is known that without epistasis (assumption 2):

$$\begin{aligned} &\text{cov PO} = 1/2\,\sigma_A^2, \text{cov HS} = 1/4\,\sigma_A^2, \\ &\text{cov FS} = 1/2\,\sigma_A^2 + 1/4\,\sigma_D^2 \\ &\text{cov HUN} = 1/8\,\sigma_A^2 \text{ (half uncle-nephew).} \end{aligned} \tag{3}$$

cov HS and cov FS can be estimated in a two-factor mating design, and cov HS alone from the variance among half-sib families from a within-population top cross. Hence,

$$\begin{aligned} &\text{cov PO} = 2 \text{ cov Hs, cov HUN} = 1/2 \text{ cov HS} \\ &\sigma_G^2 = 4(\text{cov FS} - \text{cov HS}) \text{ etc. } \ldots \end{aligned} \tag{4}$$

It is possible to establish such relationships without relying on the concepts of additivity and dominance introduced by Fisher (1918). Let G_{ij} be the expected value of a genotype A_iA_j, i.e., the mean of all individuals in the population with genotype A_iA_j at the considered locus. In a random-mating population alleles in pairs of zygotes exist in three states of identity by descent (a letter representing a class of identity by descent two different letters represent independent homologous genes), as follows:

(ij | ij), (ij | ik), (ij | kl).

In each state a probability noted $\varphi(\ |\ )$, can be associated with the state of identity between brackets. After summation over the set of loci, the general expression of covariance between relatives X and Y in a random mating population is:

$$\begin{aligned} &\text{cov XY} = \varphi_{Xy}\,(\text{ij} \mid \text{ij})\ \text{cov}\ G_{ij}G_{ij} + \\ &+ \varphi_{Xy}\,(\text{ij} \mid \text{ik})\ \text{cov}\ G_{ij}G_{ik}, \end{aligned} \tag{5}$$

$$\text{with cov}\ G_{ij}G_{ij} = \sum_s E(G_{ij}^s - \mu_s)^2 = \sigma_G^2,$$

$$\text{cov}\ G_{ij}G_{ik} = \sum_s E(G_{ij}^s - \mu_s)(G_{ik}^s - \mu_s) = \sigma_{GG'},$$

s denoting a particular locus.

Then, according to the values of the coefficients $\varphi(\ |\ )$:

$$\left.\begin{aligned} \text{cov FS} &= 1/4\,\sigma_G^2 + 1/2\,\sigma_{GG'} \\ \text{cov HS} &= 1/2\,\sigma_{GG'} \\ \text{cov HUN} &= 1/4\,\sigma_{GG'} \\ \text{cov PO} &= \sigma_{GG'}. \end{aligned}\right\} \tag{6}$$

Thus the relationships among variances and covariances are found without expressing covariances in terms of additive and dominance variance components. The methodology used is an application of Cotterman's approach (1948) with new and independent developments given by Gillois for inbred populations (1964) (see also Gallais 1970, 1976 for a general formulation). Clearly, using the factorial decomposition of the genotypic value, this leads to $\sigma_{GG'} = 1/2\,\sigma_A^2$ and to the classical expression of the covariances. With the proposed approach the concepts of additivity and dominance do not appear necessary. In the appendix the same approach is developed to take into account epistasis and to derive the general expression of genetic advance after n cycles of selection followed by m generations of intercrossing.

Conclusion

Only assumption 1, (bivariate normal distribution of the value of the parents and of the value of their offspring), is essential to formulate and to apply the theory of genetic advance under recurrent selection. Griffing (1960) has shown that such an assumption is fulfilled if the effects of a locus (or group of loci) are small in comparison to the phenotypic standard deviation. However the assumption 1 is not as strong; it can also be replaced by the assumption of linear relationship between the value of the parents according to the system of testing and the value of their offspring after one or several cycles of intercrossing.

This statistical assumption allows a general expression of genetic advance in one or several cycles of selection, or after relaxation of the selection, in terms of covariances between relatives. It does not appear necessary to use the

factorial decomposition of genotypic value in terms of additivity and dominance, i.e., it is not necessary to decompose the contribution of a locus. The covariances between relatives can be directly defined at the genotypic level for a locus or a group of loci of size related to the degree of epistasis considered. The results are extended to the set of involved loci by summing on all loci or possible groups of loci.

There is nothing new in the derived results. However the methodology used, illustrated in the appendix, allows a simple derivation of already known results, with the minimum of assumptions.

Appendix

Prediction of Genetic Advance in the Case of Epistasis without Direct Reference to Fisher's Model

To compute $\mu_{n,1}$, I consider the general case where selection is followed by m cycles of intercrossing. Applying the general principle of linear prediction (with assumption 1), the value of generation n+1 after m cycles of intercrossing will be (Gallais 1979):

$$\mu_{n+1,m-1} = \mu_{n,m} + i\theta\, 2^{m-1}\, \frac{\text{cov } TM_m}{\sqrt{\text{var } T}}\,, \qquad (7)$$

in which $\theta = 1$ or 2 if selection is n one sex or two. M_m is the value of offspring from one genotype in generation n+1 after m−1 cycles of intercrossing. $\theta\, 2^{m-1}$ represents the number of genotypes contributing to the value of the generation n+1 after (m−1) cycles of intercrossing (without limitation of the population size, assumption (3)).

Applying the principle of conditional expectations, cov TM_m can be written, with epistasis restricted to pairs of loci:

$$\text{cov } TM_m = \varphi_m(1,0)\,\text{cov}(1,0) + \varphi_m(1,1)\,\text{cov}(1,1), \qquad (8)$$

in which cov(1,0) and cov(1,1) represent covariances between two-locus genotypes having respectively one or two non homologous genes in common. $\varphi(1,0)$ and $\varphi(1\mid 1)$ are the probabilities (coefficients of kinship) associated to such situations. (in the situation (0,0) genotypes having no gene in common, the conditional covariance is zero). For example, in the absence of linkage:

$$\begin{aligned}&-\text{ covariance parent offspring} \quad \text{cov PO} = \text{cov}(1,1)\\ &-\text{ covariance between half-sibs}\\ &\quad \text{cov HS} = 1/4\,\text{cov}(1,0) + 1/4\,\text{cov}(1,1)\\ &-\text{ covariance half uncle-nephew}\\ &\quad \text{cov HUN} = 3/16\,\text{cov}(1,0) + 1/16\,\text{cov}(1,1),\end{aligned} \qquad (9)$$

It is clear that $\text{cov}(1,0) = 1/2\,\sigma^2_A$,

$$\text{cov}(1,1) = 1/2\,\sigma^2_A + 1/4\,\sigma^2_{AA},$$

so $\text{cov}(1,1) - \text{cov}(1,0) = 1/4\,\sigma^2_{AA}$

In the general situation putting $m' = m-1$

$$\varphi_{m'}(1,0) = (1/2)^{m'}\varphi_0(1,0) + (1/2)^{m'}(1-(1/2)^{m'})\varphi_0(1,1)$$

and

$$\varphi_{m'}(1,1) = (1/2)^{2m'}\varphi_0(1,1)$$

Hence,

$$\mu_{n+1,m'} = \mu_{n,m} + i\theta\,\{K_1\,\text{cov}(1,0) + (1/2)^{m'}K_{11}(\text{cov}(1,1) - \text{cov}(1,0))\}/\sqrt{\text{var } T},$$

with $K_1 = \varphi_0(1,0) + \varphi_0(1,1)$ and $K_{11} = \varphi_0(1,1)$. According to (9), $K_1 = K_{11} = 1$ for selection on the phenotype and $K_1 = 1/2$, $K_{11} = 1/4$ for G.C.A. selection. Solving the recurrence in r and m:

$$\mu_{n,m'} = \mu_0 + i\theta\,\Big\{nK_1\,\text{cov}(1,0) + (1/2)^{m'}K_{11}\sum_{r=1}^{r=n}(1/2)^{r-1}\quad(\text{cov}(1,1) - \text{cov}(1,0))\Big\}/\sqrt{\text{var } T}.$$

Noting that $\sum_{r=1}^{r=n}(1/2)^{r-1} = 2(1-(1/2)^n)$

$$\mu_{n,m-1} - \mu_n = -2\,i\theta\,K_{11}(1-(1/2)^{m-1}) - (1-(1/2)^n)(\text{cov}(1,1) - \text{cov}(1,0))/\sqrt{\text{var } T}.$$

With $m > 1$, this difference will be zero if $\text{cov}(1,1) - \text{cov}(1,0) = 0$, i.e., in the absence of epistasis.

To estimate $(\mu_{n,m-1} - \mu_n)$ or $(\mu_{n,1} - \mu_n)$ with $m = 2$, it is then necessary to estimate cov(1,1) and cov(1,0) from covariances between relatives such cov PO, cov HS or cov HUN.

Literature

Cotterman, C.W. (1940): Calculus for Statistical Genetics. Unpublished Thesis, Columbus, Ohio: Ohio St. University

Fisher, R.A. (1918): The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinburgh **52**, 399-433

Gallais, A. (1970): Covariances entre apparentés quelconques avec linkage et épistasie. 1: Expression générale. Ann. Genet. Sel. Anim. **2**, 281-310

Gallais, A. (1975): Selection with truncation in autotetraploids. Comparison with diploids. Theor. Appl. Genet. **46**, 387-395

Gallais, A. (1976): A general approach to dependance relationships among genes with some applications, (eds.: Kempthorne, O.; Pollak, E.), pp. 829-836. Proc. Int. Conf. Quantitative Genetics, Ames

Gallais, A. (1977): Amélioration des populations, méthodes de sélection et création de variétiés. 1: Ann. Amélior. Plantes **27**, 281-329

Gallais, A. (1979): Application du concept de valeur variétale à la théorie de la sélection de variétiés hybrides. Ann. Amélior. Plantes **29**, 23-24

Gillois, M. (1964): La relation d'identité en Génétique. Thèse Doctorat d'Etat. Fac. Sci. Paris

Griffing, B. (1960): Theoretical consequences of truncation selection based on the individual phenotype. Aust. J. Biol. Sci. **13**, 307-343

Received April 7, 1980
Communicated by R.W. Allard

Dr. A. Gallais
Station d'Amélioration des Plantes
Fourragères, INRA
86600 Lusignan, Vienne (France)