

SELECTION IN REFERENCE TO BIOLOGICAL GROUPS

III.* GENERALIZED RESULTS OF INDIVIDUAL AND GROUP SELECTION IN TERMS OF PARENT-OFFSPRING COVARIANCES†

By B. GRIFFING‡

[*Manuscript received June 24, 1968*]

Summary

Transference of the model-building unit from that of the gene to that of the entire genotype permits the consequences of individual and group selection to be given in terms of parent-offspring covariances. This, in turn, solves two basic problems: (1) extending the genetic analysis to any arbitrary level of complexity; and (2) formulating the prediction equations of genetic advance in terms of parameters which are directly estimable.

I. INTRODUCTION

In this series of studies, the genetic model usually used in selection theory is extended to accommodate any form of interaction (cooperative or competitive) between genotypes within small groups. When selection theory is applied to this more complex genetic model, a consideration of two populations of groups is necessary: (1) the parent population of groups in which selection operates, and (2) the progeny population of groups in which the effects of selection are measured.

In the first paper of this series (Griffing 1967), the consequences of individual and group selection were examined for the case in which group sizes for parent and progeny populations are the same. In the second paper (Griffing 1968) the more generalized situation was considered in which the group sizes in the two populations are different.

In both papers the consequences of selection were examined for genotypes generated by an arbitrary number of alleles at a single locus. These results were illuminating in that they showed the nature of response to selection in terms of direct and associate additive genetic effects. However, two further extensions of the theory are now necessary before the total response pattern can be seen clearly and before the theory can be used practically. The first extension has to do with generalizing the gene model to include more than one locus, in order to evaluate the total response to selection. The other extension has to do with the practical problem of stating the consequences of selection in terms of parameters that can be estimated directly.

* Part II, *Aust. J. biol. Sci.*, 1968, **21**, 1163-70.

† This work was supported by a grant from the United States Atomic Energy Commission [contract number AT(11-1)-1761].

‡ Faculty of Genetics, Ohio State University, Columbus, Ohio 43210.

Both of these problems can be solved by transferring the model-building unit from that of the gene to that of the individual genotype. This permits the consequences of selection to be formulated in terms of covariances among relatives. These covariances are directly estimable and they can be interpreted in terms of gene models of any degree of complexity.

In the following presentation, parent and progeny populations are constructed, necessary parameters defined, and consequences of individual and group selection derived.

II. CONSEQUENCES OF SELECTION IN POPULATIONS OF GROUPS OF SIZE n

In this section the groups in the parent and progeny populations are assumed to be of the same size (n).

(a) Parent Population Parameters

The parent base population is assumed to be in equilibrium under random mating. In this population let

$$\begin{aligned} H_i &= \textit{i} \text{th parental genotype,} \\ f_i &= \text{relative frequency of } H_i, \text{ and, then,} \\ \sum_i f_i H_i &= \text{genotypic array of the parent base population.} \end{aligned}$$

The parent population of groups, each with n randomly associated genotypes, is obtained from the n -way combinatorial product of the base population as follows:

$$[\sum f_i H_i] \times [\sum f_i H_i] \times \dots \times [\sum f_i H_i] = \sum f_{i_1} f_{i_2} \dots f_{i_n} (H_{i_1}, H_{i_2}, \dots, H_{i_n}).$$

In this representation a group of n randomly associated genotypes is given by the n -tuple

$$(H_{i_1}, H_{i_2}, \dots, H_{i_n}).$$

The genotypic value of H_{i_1} in such a group is denoted as

$${}_{i_1} h_{i_2, \dots, i_n},$$

and coded so that

$$\sum f_{i_1} f_{i_2} \dots f_{i_n} ({}_{i_1} h_{i_2, \dots, i_n}) = 0.$$

The subscript in front of the symbol h identifies the genotype under consideration, and the subscripts following h identify the remaining $(n-1)$ associated genotypes in the group.

(b) Progeny Population Parameters

The progeny population is derived in terms of a full-sib structure. To do this let

$$\sum q_{ij k} (H_{ij k}) = \text{full-sib array from the mating } H_i \times H_j,$$

Then the progeny population of groups of size n are obtained by forming the n -way combinatorial product using the array of full-sib families,

$$\sum f_i f_j q_{ijk}(H_{ijk}),$$

as a base population in the following manner:

$$[\sum f_i f_j q_{ijk}(H_{ijk})] \times \dots \times [\sum f_i f_j q_{ijk}(H_{ijk})] \\ = \sum (f_{i_1} f_{j_1} q_{i_1 j_1 k_1}) \dots (f_{i_n} f_{j_n} q_{i_n j_n k_n})(H_{i_1 j_1 k_1}, \dots, H_{i_n j_n k_n}).$$

The genotypic value of $H_{i_1 j_1 k_1}$ in the group $(H_{i_1 j_1 k_1}, \dots, H_{i_n j_n k_n})$ is

$${}_{i_1 j_1 k_1} h_{i_2 j_2 k_2}, \dots, {}_{i_n j_n k_n}$$

and coded so that

$$\sum (f_{i_1} f_{j_1} q_{i_1 j_1 k_1}) \dots (f_{i_n} f_{j_n} q_{i_n j_n k_n})({}_{i_1 j_1 k_1} h_{i_2 j_2 k_2}, \dots, {}_{i_n j_n k_n}) = 0.$$

(c) *Definitions of Parent-Offspring Parameters*

When interaction between genotypes within groups is permitted, it is necessary to distinguish between the direct genotypic effect of an individual and the associated genotypic effects contributed by the other group members. Hence four different kinds of covariances between parents and offspring need to be distinguished. These are

$$\text{Cov}(P_d O_d) = \sum f_{i_r} ({}_{i_r} h_{\dots, \dots}) ({}_{i_r} \dots h_{\dots, \dots})$$

or

$$= \sum f_{i_r} ({}_{i_r} h_{\dots, \dots}) ({}_{i_r} \dots h_{\dots, \dots})$$

$$\text{Cov}(P_d O_a) = \sum f_{i_r} ({}_{i_r} h_{\dots, \dots}) (\dots h_{i_r} \dots, \dots) \\ = \text{etc.},$$

$$\text{Cov}(P_a O_d) = \sum f_{i_r} (\dots h_{i_r}, \dots) ({}_{i_r} \dots h_{\dots, \dots}) \\ = \text{etc.}, \text{ and}$$

$$\text{Cov}(P_a O_a) = \sum f_{i_r} (\dots h_{i_r}, \dots) (\dots h_{i_r} \dots, \dots) \\ = \text{etc.},$$

where ${}_{i_r} h_{\dots, \dots}$ = average direct genotypic value of H_{i_r} as expressed in the parent population,

$${}_{i_r} h_{\dots, \dots} = {}_{i_r} h_{\dots, \dots} = \text{etc.}$$

= average associate genotypic value of H_{i_r} as expressed in the parent population,

$${}_{i_r} \dots h_{\dots, \dots} = {}_{i_r} \dots h_{\dots, \dots}$$

= average direct genotypic value of the half-sib progeny of H_{i_r} as expressed in the progeny group population, and

$$\dots h_{i_r} \dots, \dots = \dots h_{i_r, \dots}, \dots = \dots h_{\dots, i_r, \dots}, \dots = \text{etc.}$$

= average associate genotypic value of the half-sib progeny of H_{i_r} as expressed in the progeny group population.

For greater elaboration of the notions of "direct" and "associate" effects as interpreted in terms of a gene model, see Griffing (1967).

(d) *Consequences of Individual Selection*

The selection value for the parent genotype, H_{ir} , is

$$w_{ir} = 1 + (\bar{i}/\sigma)_{\text{ind.}}(i_r \bar{h} \dots, \dots),$$

where \bar{i} = standardized selection differential, σ = phenotypic standard deviation, and the subscript "ind." indicates that \bar{i} and σ relate to individual observations.

The mean of the progeny population of groups which is derived from full-sib matings among selected parents can be given as follows:

$$\begin{aligned} \mu_1 &= \sum[(f_{i_1} w_{i_1})(f_{j_1} w_{j_1})(q_{i_1 j_1 k_1})] \dots [(f_{i_n} w_{i_n})(f_{j_n} w_{j_n})(q_{i_n j_n k_n})] \times (i_{1j_1 k_1} \bar{h}_{i_1 j_1 k_1}, \dots, i_{nj_n k_n}) \\ &\simeq \sum(f_{i_1} f_{j_1}) \dots (f_{i_n} f_{j_n}) \{1 + (\bar{i}/\sigma)_{\text{ind.}} [(i_1 \bar{h} \dots, \dots + j_1 \bar{h} \dots, \dots) + \dots \\ &\quad + (i_n \bar{h} \dots, \dots + j_n \bar{h} \dots, \dots)]\} \times (i_{1j_1} \bar{h}_{i_1 j_1}, \dots, i_{nj_n}) \\ &= (\bar{i}/\sigma)_{\text{ind.}} \sum f_{i_r}(i_r \bar{h} \dots, \dots) \{i_r \bar{h} \dots, \dots + \dots + i_r \bar{h} \dots, \dots + \dots + \dots + i_r \bar{h} \dots, \dots \\ &\quad + \dots + \dots + \bar{h} \dots, \dots, i_r\} \\ &= 2(\bar{i}/\sigma)_{\text{ind.}} \{ \text{Cov}(P_d O_d) + (n-1) \text{Cov}(P_d O_a) \}. \end{aligned}$$

(e) *Consequences of Group Selection*

Let the selection value of the parent group ($H_{i_1}, H_{i_2}, \dots, H_{i_n}$) be

$$w_{i_1 i_2 \dots i_n} = 1 + (\bar{i}/\sigma)_{\text{gr.}} (1/n) (i_1 \bar{h}_{i_2}, \dots, i_n + \dots + i_n \bar{h}_{i_1}, \dots, i_{n-1}),$$

where the subscript "gr." indicates that \bar{i} and σ are group parameters.

It is clear then that the selection value of H_{i_r} following group selection is

$$w_{i_r} = 1 + (\bar{i}/\sigma)_{\text{gr.}} (1/n) (i_r \bar{h} \dots, \dots + \bar{h}_{i_r}, \dots, \dots + \dots + \bar{h} \dots, i_r).$$

The progeny mean in terms of groups of size n is then

$$\begin{aligned} \mu_1 &= \sum(f_{i_1} w_{i_1}) \dots (f_{j_n} w_{j_n}) (i_{1j_1} \bar{h}_{i_1 j_1}, \dots, i_{nj_n}) \\ &= \sum(f_{i_1} f_{j_1}) \dots (f_{i_n} f_{j_n}) \{1 + (\bar{i}/\sigma)_{\text{gr.}} (1/n) [(i_1 \bar{h} \dots, \dots + \dots + \bar{h} \dots, i_1) + \dots \\ &\quad + (j_n \bar{h} \dots, \dots + \dots + \bar{h} \dots, j_n)]\} \times (i_{1j_1} \bar{h}_{i_1 j_1}, \dots, i_{nj_n}). \end{aligned}$$

This is equivalent to

$$\begin{aligned} &(\bar{i}/\sigma)_{\text{gr.}} (1/n) \{ \sum f_{i_r}(i_r \bar{h} \dots, \dots + (n-1) \bar{h}_{i_r}, \dots, \dots) [i_r \bar{h} \dots, \dots + \dots + i_r \bar{h} \dots, \dots, \dots \\ &\quad + (n-1)(\dots \bar{h}_{i_r}, \dots, \dots + \dots + \dots \bar{h}_{i_r}, \dots, \dots)] \} \\ &= (\bar{i}/\sigma)_{\text{gr.}} (2/n) \{ \text{Cov}(P_d O_d) + (n-1) \text{Cov}(P_d O_a) + (n-1) \text{Cov}(P_a O_d) + (n-1)^2 \text{Cov}(P_a O_a) \}. \end{aligned}$$

Since the progeny population not subject to selection is coded to zero, the above quantity represents the increment change in mean due to selection.

III. CONSEQUENCES OF SELECTION IN GROUPS OF ONE SIZE WHEN EVALUATED IN GROUPS OF A DIFFERENT SIZE

Griffing (1968) develops the necessary argument for the single-locus case in which selection operates with regard to groups of size n_1 and the effects of selection are measured with regard to groups of size n_2 . The generalized results in terms of parent-offspring covariances are given below.

For individual selection the change in the progeny mean is

$$n_1(\Delta\mu)_{n_2} = 2[\frac{i}{\sigma}]_{\text{ind.}} \{ \text{Cov}(P_{d_{n_1}} O_{d_{n_2}}) + (n_2 - 1) \text{Cov}(P_{d_{n_1}} O_{a_{n_2}}) \}.$$

For group selection the change in the progeny mean is

$$n_1(\Delta\mu)_{n_2} = [\frac{i}{\sigma}]_{\text{gr.}} (2/n_1) \{ \text{Cov}(P_{d_{n_1}} O_{d_{n_2}}) + (n_2 - 1) \text{Cov}(P_{d_{n_1}} O_{a_{n_2}}) \\ + (n_1 - 1) \text{Cov}(P_{a_{n_1}} O_{d_{n_2}}) + (n_1 - 1)(n_2 - 1) \text{Cov}(P_{a_{n_1}} O_{a_{n_2}}) \}.$$

IV. INTERPRETATION OF THE PARENT-OFFSPRING COVARIANCES IN TERMS OF THE GENE MODEL FOR TWO LINKED LOCI

The obvious extension of the single-locus model used in the earlier studies is to that of a two-locus model with the following characteristics. Let

- $\sum p_i^1(A_i^1)$ = array of an arbitrary number of alleles at the first locus,
- $\sum p_k^2(A_k^2)$ = array of an arbitrary number of alleles at the second locus, and
- y = recombination value exhibited by the two loci.

Assume any system of dominance and epistatic parameters, and also assume that the population is in equilibrium with random mating. Then the genotypic value for $(A_i^1 A_k^2)(A_j^1 A_t^2)$ in the group

$$\{(A_i^1 A_k^2)(A_j^1 A_t^2), (A_r^1 A_t^2)(A_s^1 A_u^2), \dots\}$$

is

$$\begin{aligned} (ik)(jl)d_{(rt)(su)} \dots &= [d\alpha_i^1 + d\alpha_j^1 + d\delta_{ij}^1 + d\alpha_k^2 + d\alpha_t^2 + d\delta_{kl}^2 + da(\alpha\alpha)_{ik} + da(\alpha\alpha)_{il} + da(\alpha\alpha)_{jk} \\ &+ da(\alpha\alpha)_{jl} + da(\alpha\delta)_{ikl} + da(\alpha\delta)_{jkl} + da(\delta\alpha)_{ijk} + da(\delta\alpha)_{ijl} + da(\delta\delta)_{ijkl}] \\ &+ [a\alpha_r^1 + a\alpha_s^1 + a\delta_{rs}^1 + a\alpha_t^2 + a\alpha_u^2 + a\delta_{tu}^2 + aa(\alpha\alpha)_{rt} + aa(\alpha\alpha)_{ru} + aa(\alpha\alpha)_{st} \\ &+ aa(\alpha\alpha)_{su} + aa(\alpha\delta)_{rtu} + aa(\alpha\delta)_{stu} + aa(\delta\alpha)_{rst} + aa(\delta\alpha)_{rsu} + aa(\delta\delta)_{rstu}] \\ &+ da(\alpha\alpha)_{ir} + \text{other direct-associate terms} + \dots \end{aligned}$$

In this representation the usual gene model is used. To fully characterize all main effects and interactions of the two genotypes listed above, a four-locus model is needed. For a group of n genotypes, each of two loci, a $2n$ -locus model is required. The subscript d in front of the elements in the model indicates that the elements are "direct" effects; the subscript a indicates "associate" effects, etc.

Interpretation of the various parent-offspring covariances is now given in terms of variances and covariances associated with the above model.

For the situation in which the group size is n for both parent and progeny populations, the parent-offspring covariances are as follows:

$$\begin{aligned} \text{Cov}(P_a O_d) &= \frac{1}{2}[da\sigma_A^2] + \frac{1}{4}[da, da\sigma_{AA}^2], \\ \text{Cov}(P_d O_a) &= \frac{1}{2}[(da)\sigma_A] + \frac{1}{4}[(da, da)\sigma_{AA}], \\ \text{Cov}(P_a O_a) &= \frac{1}{2}[(aa)\sigma_A] + \frac{1}{4}[(aa, da)\sigma_{AA}], \\ \text{Cov}(P_d O_d) &= \frac{1}{2}[aa\sigma_A^2] + \frac{1}{4}[aa, aa\sigma_{AA}^2], \end{aligned}$$

where

$$\begin{aligned} da\sigma_A^2 &= 2\{\sum p_i^1 (d\alpha_i^1)^2 + \sum p_k^2 (d\alpha_k^2)^2\}, \\ aa\sigma_A^2 &= 2\{\sum p_i^1 (a\alpha_i^1)^2 + \sum p_k^2 (a\alpha_k^2)^2\}, \\ (da)\sigma_A &= 2\{\sum p_i^1 (d\alpha_i^1)(a\alpha_i^1) + \sum p_k^2 (d\alpha_k^2)(a\alpha_k^2)\}, \\ (aa)\sigma_A &= (da)\sigma_A, \\ da, da\sigma_{AA}^2 &= 4\{\sum p_i^1 p_k^2 [da(\alpha\alpha)]_{ik}\}^2, \\ aa, aa\sigma_{AA}^2 &= 4\{\sum p_i^1 p_k^2 [aa(\alpha\alpha)]_{ik}\}^2, \\ (da, da)\sigma_{AA} &= 4\{\sum p_i^1 p_k^2 [da(\alpha\alpha)]_{ik}[aa(\alpha\alpha)]_{ik}\}, \\ (aa, da)\sigma_{AA} &= (da, aa)\sigma_{AA}. \end{aligned}$$

Since

$$(aa)\sigma_A = (da)\sigma_A,$$

and

$$(aa, da)\sigma_{AA} = (da, aa)\sigma_{AA},$$

then

$$\text{Cov}(P_d O_a) = \text{Cov}(P_a O_d).$$

Hence the estimation procedure is simplified in terms of the number of different covariances needed to predict group selection.

For the situation in which selection occurs with regard to groups of size n_1 in the parent population, and the effects of selection are measured with regard to groups of size n_2 in the progeny population, the parent-offspring covariances are as follows:

$$\begin{aligned} \text{Cov}(P_{d_{n_1}} O_{d_{n_2}}) &= \frac{1}{2}[(d_{n_1} d_{n_2})\sigma_A] + \frac{1}{4}[(d_{n_1} d_{n_1}, d_{n_2} d_{n_2})\sigma_{AA}], \\ \text{Cov}(P_{d_{n_1}} O_{a_{n_2}}) &= \frac{1}{2}[(d_{n_1} a_{n_2})\sigma_A] + \frac{1}{4}[(d_{n_1} d_{n_1}, a_{n_2} a_{n_2})\sigma_{AA}], \\ \text{Cov}(P_{a_{n_1}} O_{d_{n_2}}) &= \frac{1}{2}[(a_{n_1} d_{n_2})\sigma_A] + \frac{1}{4}[(a_{n_1} a_{n_1}, d_{n_2} d_{n_2})\sigma_{AA}], \\ \text{Cov}(P_{a_{n_1}} O_{a_{n_2}}) &= \frac{1}{2}[(a_{n_1} a_{n_2})\sigma_A] + \frac{1}{4}[(a_{n_1} a_{n_1}, a_{n_2} a_{n_2})\sigma_{AA}], \end{aligned}$$

where

$$\begin{aligned}
 (d_{n_1 d_{n_2}})\sigma_A &= 2\{\sum p_i^1(d_{n_1}\alpha_i^1)(d_{n_2}\alpha_i^1) + \sum p_k^2(d_{n_1}\alpha_k^2)(d_{n_2}\alpha_k^2)\}, \\
 (a_{n_1 a_{n_2}})\sigma_A &= 2\{\sum p_i^1(a_{n_1}\alpha_i^1)(a_{n_2}\alpha_i^1) + \sum p_k^2(a_{n_1}\alpha_k^2)(a_{n_2}\alpha_k^2)\}, \\
 (d_{n_1 a_{n_2}})\sigma_A &= 2\{\sum p_i^1(d_{n_1}\alpha_i^1)(a_{n_2}\alpha_i^1) + \sum p_k^2(d_{n_1}\alpha_k^2)(a_{n_2}\alpha_k^2)\}, \\
 (a_{n_1 d_{n_2}})\sigma_A &= 2\{\sum p_i^1(a_{n_1}\alpha_i^1)(d_{n_2}\alpha_i^1) + \sum p_k^2(a_{n_1}\alpha_k^2)(d_{n_2}\alpha_k^2)\}, \\
 (d_{n_1 d_{n_1}, d_{n_2 d_{n_2}}})\sigma_{AA} &= 4\{\sum p_i^1 p_k^2 [d_{n_1 d_{n_1}}(\alpha\alpha)_{ik}] [d_{n_2 d_{n_2}}(\alpha\alpha)_{ik}]\}, \\
 (d_{n_1 d_{n_1}, a_{n_2 a_{n_2}}})\sigma_{AA} &= 4\{\sum p_i^1 p_k^2 [d_{n_1 d_{n_1}}(\alpha\alpha)_{ik}] [a_{n_2 a_{n_2}}(\alpha\alpha)_{ik}]\}, \\
 (a_{n_1 a_{n_1}, d_{n_2 d_{n_2}}})\sigma_{AA} &= 4\{\sum p_i^1 p_k^2 [a_{n_1 a_{n_1}}(\alpha\alpha)_{ik}] [d_{n_2 d_{n_2}}(\alpha\alpha)_{ik}]\}, \\
 (a_{n_1 a_{n_1}, a_{n_2 a_{n_2}}})\sigma_{AA} &= 4\{\sum p_i^1 p_k^2 [a_{n_1 a_{n_1}}(\alpha\alpha)_{ik}] [a_{n_2 a_{n_2}}(\alpha\alpha)_{ik}]\}.
 \end{aligned}$$

It should be noted that when $n_1 \neq n_2$, $(a_{n_1 a_{n_2}})\sigma_A$ and $(a_{n_1 d_{n_2}})\sigma_A$ need not be equal. Also $(d_{n_1 d_{n_1}, a_{n_2 a_{n_2}}})\sigma_{AA}$ need not equal $(a_{n_1 a_{n_1}, d_{n_2 d_{n_2}}})\sigma_{AA}$. Hence the values for $\text{Cov}(P_{a_{n_1}} O_{a_{n_2}})$ and $\text{Cov}(P_{a_{n_1}} O_{d_{n_2}})$ may be different and, therefore, it is necessary to estimate each separately.

Finally, extensions to more complex gene models involving more loci are obvious and can be derived from the fact that in general the parent-offspring covariance for a random mating population in equilibrium has the following form for n loci:

$$\text{Cov}(PO) = \frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_{AA}^2 + \frac{1}{8}\sigma_{AAA}^2 + \dots + (\frac{1}{2})^n \sigma_{AA\dots A}^2.$$

It is apparent that the parent-offspring covariances, as developed in the above theory, are variations of this basic formula.

V. DISCUSSION

Transferring the basis of analyses from the gene to the individual has solved the two remaining basic problems concerning prediction of individual and group selection. The first problem is that of extending the theory to accommodate more complex gene models. The second problem is a practical one of estimating genetic advance. Both of these problems are solved by formulating the prediction equations in terms of parent-offspring covariances. The reasons that these covariances are useful in solving these two basic problems are that the genetic compositions of these covariances are known for any level of genetic complexity, and because the parent-offspring covariances are directly estimable from experimental data. A variety of genetic experimental designs can be used for this estimation procedure. These designs will be presented in a subsequent paper in this series.

One of the important features which the group theory analyses bring to light is the possible importance of associate effects in changing the population genotypic structure by selection. It is clear that as the effective number in the group increases, the associate effects may take on an increasingly dominant role in determining

the consequences of selection. This role is manifest through the parent-offspring covariances involving cross-products of direct and associate effects as well as covariances which are functions of only associate effects. Up to now these covariances have been completely, or at least largely, ignored.

VI. REFERENCES

- GRIFFING, B. (1967).—Selection in reference to biological groups. I. Individual and group selection applied to populations of unordered groups. *Aust. J. biol. Sci.* **20**, 127–39.
- GRIFFING, B. (1968).—Selection in reference to biological groups. II. Consequences of selection in groups of one size when evaluated in groups of a different size. *Aust. J. biol. Sci.* **21**, 1163–70.