

# THE GENOTYPIC COMPOSITION AND VARIABILITY IN PLANT POPULATIONS UNDER MIXED SELF-FERTILIZATION AND RANDOM MATING

BY G. L. GHAI

*Institute of Agricultural Research Statistics, New Delhi*

## INTRODUCTION

THE genetical studies have generally concerned themselves with the Mendelian characters and quantitative characters of populations which are subjected to only one system of mating—such as complete cross-fertilization or inbreeding such as complete self-fertilization.

In plant populations of interest to the breeder, complete self-fertilization or complete cross-fertilization is seldom found. There are crops like rice and wheat which are highly but not completely self-fertilized, crops like cotton which are moderately self-fertilized and crops like various 'brassica' species or maize which are largely cross-fertilized. The study of plant populations in which both self and cross-fertilization at random obtain in varying degrees, in regard to their genotypic composition, variability and mean values in successive generations is, therefore, of very considerable theoretical as well as practical interest in relation to breeding.

A beginning was made in this direction concerning the genotypic composition when Garber (1951) dealt with the approach to equilibrium with varying percentages of cross and self-fertilization in the case of one autosomal gene pair for the initial population having the genotypic composition as  $p^2.AA + 2pq.Aa + q^2.aa$ , where  $p$  and  $q$  are the gene frequencies of the alleles 'A' and 'a' respectively.

Bennet and Binet (1956) considered the association between Mendelian factors for two pairs of genes with mixed selfing and cross-fertilization in equilibrium. They concluded that when equilibrium is reached the gametic frequency is equal to the product of the proportions of the corresponding genes taken separately, whereas the genotypic frequency is not equal to the product of the frequencies of genotypes at separate loci.

The present investigation was undertaken to study the plant populations subjected to self and cross-fertilization at random in varying degrees, in respect of their genotypic composition in successive generations, their approach to equilibrium and association between factors for the general case of  $k$  loci, and also to study the mean values and genotypic variability in successive generations.

In the present study it has been assumed that at each stage of mating a proportion 'x' of the whole population is obtained by self-fertilization and the remaining portion 'y' ( $x + y = 1$ ) cross-fertilized at random. It has also been assumed that the factors involved segregate independently and fertility and viability disturbances are absent.

#### GENOTYPIC COMPOSITION AND ASSOCIATION OF FACTORS

An attempt has been made here to obtain frequencies of the various genotypes in the  $n$ -th generation of mixed selfing and random mating in an arbitrary initial population for different number of independently segregating factors or loci.

Let  $f_{ijk\dots}^{(n)}$  be the frequencies of the various genotypes in the  $n$ -th generation of mixed breeding, where the suffixes correspond to the various gene pairs  $A-a$ ,  $B-b$ ,  $C-c$ , etc., involved—each taking the values 0, 1, or 2 depending upon the number of dominant alleles of the factor present in the genotype, e.g., in the case of three factors  $f_{012}^{(n)}$  is the frequency of the genotype 'aaBbCC' in the  $n$ -th generation.  $f_{ijk\dots}^{(0)}$  are the frequencies of the corresponding genotypes in the initial population.

Also, let  $R_{ijk\dots}^{(n)}$  be the resulting proportions in  $n$ -th generation of the corresponding genotypes when the population in  $(n-1)$ -th generation is subjected to random mating alone.

*Single factor : A-a.*—The frequencies of the three genotypes  $AA$ ,  $Aa$ ,  $aa$ , in the  $n$ -th generation are given by:

$$f_2^{(n)} = yR_2^{(n)} + x(f_2^{(n-1)} + \frac{1}{4}f_1^{(n-1)}).$$

$$f_1^{(n)} = yR_1^{(n)} + \frac{x}{2}f_1^{(n-1)}.$$

$$f_0^{(n)} = yR_0^{(n)} + x(f_0^{(n-1)} + \frac{1}{4}f_1^{(n-1)}).$$

By substituting the successive values of  $f$ 's, these equations can be written as

$$f_2^{(n)} = y \sum_{r=0}^{n-1} x^r R_2^{(n-r)} + x^n f_2^{(0)} + \frac{1}{2} \cdot \left\{ x^n - \left(\frac{x}{2}\right)^n \right\} f_1^{(0)}$$

$$+ \left(\frac{y}{2}\right) \sum_{r=1}^{n-1} \left\{ x^r - \left(\frac{x}{2}\right)^r \right\} R_1^{(n-r)}, \text{ etc.,}$$

where

$$R_3^{(n)} = p^2(A), \text{ for all } n,$$

$$R_1^{(n)} = 2p(A)p(a), \text{ for all } n,$$

$$p(A) = f_2^{(0)} + \frac{1}{2}f_1^{(0)}; \quad p(a) = f_0^{(0)} + \frac{1}{2}f_1^{(0)}.$$

hence,

$$f_2^{(n)} = (1 - x^n) p^2(A) + x^n f_2^{(0)} + \frac{1}{2} \cdot \left\{ x^n - \left(\frac{x}{2}\right)^n \right\} f_1^{(0)}$$

$$+ \left\{ \frac{x}{4 - 2x} - \frac{1}{2} x^n + \frac{y \left(\frac{x}{2}\right)^n}{2 - x} \right\} 2p(A)p(a)$$

$$f_0^{(n)} = (1 - x^n) p^2(a) + x^n f_0^{(0)} + \frac{1}{2} \cdot \left\{ x^n - \left(\frac{x}{2}\right)^n \right\} f_1^{(0)}$$

$$+ \left\{ \frac{x}{4 - 2x} - \frac{x^n}{2} + \frac{y \left(\frac{x}{2}\right)^n}{2 - x} \right\} 2p(A)p(a) \tag{1}$$

and

$$f_1^{(n)} = 4y \left\{ \frac{1 - \left(\frac{x}{2}\right)^n}{2 - x} \right\} p(A)p(a) + \left(\frac{x}{2}\right)^n f_1^{(0)}.$$

In the limiting case when  $n \rightarrow \infty$

$$\left. \begin{aligned} f_3^{(\infty)} &= p^2(A) + \frac{x}{2-x} p(A)p(a) \\ f_1^{(\infty)} &= \frac{4(1-x)}{2-x} p(A)p(a) \\ f_0^{(\infty)} &= p^2(a) + \frac{x}{2-x} p(A)p(a). \end{aligned} \right\} \tag{2}$$

It is interesting to observe that genotypic frequencies in the limiting case are independent of the initial genotypic frequencies and are

functions of gene frequencies and proportion of selfing, and are in equilibrium. These genotypic frequencies in equilibrium are the same as obtained by Bennet and Binet (1956).

*Two factors: A-a, B-b.*—There will be nine genotypes in all for two factors with two alleles per locus. The genotypic frequency of a double heterozygote 'AaBb' is given by

$$f_{11}^{(n)} = yR_{11}^{(n)} + \left(\frac{x}{4}\right)f_{11}^{(n-1)},$$

putting successive values of  $f$ 's,

$$f_{11}^{(n)} = y \sum_{r=0}^{n-1} \left(\frac{x}{4}\right)^r R_{11}^{(n-r)} + \left(\frac{x}{4}\right)^n f_{11}^{(0)} \quad (3)$$

To obtain the values of  $R$ 's we require gametic frequencies which are given below.

Let  $p_n(AB)$ , etc., be the proportions of gametes ( $AB$ ), etc., in the  $n$ -th generation. The recurrence relation for gametic frequencies is

$$\begin{aligned} p_{n+1}(AB) &= \left(\frac{y}{2}\right)\{p_n(AB) + p(A)p(B)\} + xp_n(AB) \\ &= \left(1 - \frac{y}{2}\right)p_n(AB) + \frac{y}{2} \cdot p(A)p(B). \end{aligned}$$

Hence, if

$$\begin{aligned} L_n(AB) &= p_n(AB) - p(A)p(B) \\ &= \left(1 - \frac{y}{2}\right)L_{n-1}(AB) = u^n L_0(AB), \end{aligned}$$

where  $u = 1 - y/2$ , then

$$\begin{aligned} p_{n+1}(AB) &= p_n(AB) - \left(\frac{y}{2}\right)L_n(AB) \\ &= p_0(AB) - (1 - u^{n+1})L_0(AB). \end{aligned}$$

Therefore,

$$\begin{aligned} R_{11}^{(n)} &= 2[p_{n-1}(AB)p_{n-1}(ab) + p_{n-1}(Ab)p_{n-1}(aB)] \\ &= R_{11}^{(1)} - 2(1 - u^{n-1})[p_0(AB)L_0(ab) + p_0(ab)L_0(AB) \\ &\quad + p_0(Ab)L_0(aB) + p_0(aB)L_0(AB)] \\ &\quad + 2(1 - u^{n-1})^2[L_0(AB)L_0(ab) + L_0(Ab)L_0(aB)]. \end{aligned}$$

The proportion of a single heterozygote, say 'AABb', in the  $n$ -th generation is given by

$$\begin{aligned}
 f_{21}^{(n)} &= yR_{21}^{(n)} + \left(\frac{x}{2}\right)f_{21}^{(n-1)} + \left(\frac{x}{8}\right)f_{11}^{(n-1)} \\
 &= y \sum_{r=0}^{n-1} \left(\frac{x}{2}\right)^r R_{21}^{(n-r)} + \left(\frac{x}{2}\right)^n f_{21}^{(0)} + \frac{\left(\frac{x}{2}\right)^n - \left(\frac{x}{4}\right)^n}{2} f_{11}^{(0)} \\
 &\quad + \left(\frac{y}{2}\right) \sum_{s=1}^{n-1} \left\{ \left(\frac{x}{2}\right)^s - \left(\frac{x}{4}\right)^s \right\} R_{11}^{(n-s)} \tag{4}
 \end{aligned}$$

where

$$\begin{aligned}
 R_{21}^{(n)} &= 2p_{n-1}(AB)p_{n-1}(Ab) \\
 &= R_{21}^{(1)} - 2(1-u^{n-1})[p_0(AB)L_0(Ab) + p_0(Ab)L_0(AB)] \\
 &\quad + 2(1-u^{n-1})^2 L_0(AB)L_0(Ab)
 \end{aligned}$$

and  $R_{11}^{(n)}$  as obtained earlier.

The proportion of a double homozygote, say 'AABB', in the  $n$ -th generation is given by

$$f_{22}^{(n)} = yR_{22}^{(n)} + xf_{22}^{(n-1)} + \frac{x}{4}(f_{21}^{(n-1)} + f_{12}^{(n-1)}) + \frac{x}{16}f_{11}^{(n-1)}$$

putting the successive values of  $f$ 's on the R.H.S., we have

$$\begin{aligned}
 f_{22}^{(n)} &= y \sum_{r=0}^{n-1} x^r R_{22}^{(n-r)} + x^n f_{22}^{(0)} + \left(\frac{1}{2}\right) \left\{ x^n - \left(\frac{x}{2}\right)^n \right\} (f_{21}^{(0)} + f_{12}^{(0)}) \\
 &\quad + \left(\frac{y}{2}\right) \sum_{r=1}^{n-1} \left\{ x^r - \left(\frac{x}{2}\right)^r \right\} (R_{21}^{(n-r)} + R_{12}^{(n-r)}) \\
 &\quad + \left[ \left(\frac{1}{12}\right) \left\{ x^n - \left(\frac{x}{4}\right)^n \right\} \right. \\
 &\quad \left. + \frac{1}{2} \sum_{r=0}^{n-2} \left(\frac{x}{4}\right)^{r+1} \left\{ x^{n-r-1} - \left(\frac{x}{2}\right)^{n-r-1} \right\} \right] f_{11}^{(0)}
 \end{aligned}$$

$$\begin{aligned}
 & + \left(\frac{y}{12}\right) \sum_{r=1}^{n-1} \left\{x^r - \left(\frac{x}{4}\right)^r\right\} R_{11}^{(n-r)} \\
 & + \left(\frac{y}{2}\right) \sum_{r=1}^{n-2} R_{11}^{(n-r-1)} \sum_{s=1}^r \left(\frac{x}{4}\right)^{(r-s+1)} \left\{x^s - \left(\frac{x}{2}\right)^s\right\}
 \end{aligned} \tag{5}$$

where

$$\begin{aligned}
 R_{22}^{(n)} & = p_{n-1}^2(AB) = R_{22}^{(1)} - 2(1 - u^{n-1}) p_0(AB) L_0(AB) \\
 & + (1 - u^{n-1})^2 L_0^2(AB)
 \end{aligned}$$

and

$$R_{21}^{(n)}, R_{12}^{(n)} \text{ and } R_{11}^{(n)}$$

have been obtained earlier.

The frequencies of other genotypes can be easily obtained by proper substitution of suffixes.

The limiting genotypic frequencies are given by

$$\left. \begin{aligned}
 f_{11}^{(\infty)} & = \frac{16(1-x)}{4-x} p(A) p(a) p(B) p(b) \\
 f_{21}^{(\infty)} & = \frac{4(1-x)}{2-x} p^2(A) p(B) p(b) \\
 & + \frac{4x(1-x)}{(2-x)(4-x)} p(A) p(a) p(B) p(b) \\
 f_{22}^{(\infty)} & = p^2(A) p^2(B) + \frac{x}{2-x} \{p(A) p(b) + p(a) p(B)\} \\
 & \times p(A) p(B) + \frac{x(2+x)}{(2-x)(4-x)} \\
 & \times p(A) p(a) p(B) p(b).
 \end{aligned} \right\} \tag{6}$$

This shows that limiting genotypic frequencies are independent of the frequencies in the initial population and are in equilibrium.

*Three factors: A-a, B-b, C-c.*—The frequencies of occurrence of genotypes, say ‘AABBCC’ (triple homozygote), ‘AABBcc’ (single heterozygote), ‘AABbCc’ (double heterozygote) and ‘AaBbCc’ (triple heterozygote) in the  $n$ -th generation are given by

$$f_{222}^{(n)} = yR_{222}^{(n)} + xf_{222}^{(n-1)} + \frac{x}{4} (f_{221}^{(n-1)} + f_{122}^{(n-1)} + f_{212}^{(n-1)})$$

$$+ \frac{x}{16} (f_{211}^{(n-1)} + f_{121}^{(n-1)} + f_{112}^{(n-1)}) + \frac{x}{64} f_{111}^{(n-1)}$$

$$f_{221}^{(n)} = yR_{221}^{(n)} + \frac{x}{2} f_{221}^{(n-1)} + \frac{x}{8} (f_{211}^{(n-1)} + f_{121}^{(n-1)}) + \frac{x}{32} f_{111}^{(n-1)}$$

$$f_{211}^{(n)} = yR_{211}^{(n)} + \frac{x}{4} f_{211}^{(n-1)} + \frac{x}{16} f_{111}^{(n-1)}$$

and

$$f_{111}^{(n)} = yR_{111}^{(n)} + \frac{x}{8} f_{111}^{(n-1)}$$

where the values of  $R$ 's are given as follows:

$$\left. \begin{aligned} R_{222}^{(n)} &= p_{n-1}^2(ABC) \\ R_{221}^{(n)} &= 2p_{n-1}(ABC)p_{n-1}(ABc) \\ R_{211}^{(n)} &= 2[p_{n-1}(ABC)p_{n-1}(Abc) + p_{n-1}(ABc)p_{n-1}(AbC)] \\ R_{111}^{(n)} &= 2[p_{n-1}(ABC)p_{n-1}(abc) + p_{n-1}(ABc)p_{n-1}(abc) \\ &\quad + p_{n-1}(ABc)p_{n-1}(aBc) + p_{n-1}(aBC)p_{n-1}(Abc)] \end{aligned} \right\} (7)$$

where  $p_n(ABC)$  are the gametic frequencies of the gametes  $(ABC)$ , etc., in the  $n$ -th generation and are given by

$$p_{n+1}(ABC) = \left(x + \frac{y}{4}\right) p_n(ABC) + \left(\frac{y}{4}\right) \sum_{\text{factors}}^3 p(A) p_n(BC), \text{ etc.}$$

Let

$$L_n(ABC) = p_n(ABC) - \Sigma p(A) L_n(BC) - p(A) p(B) p(C)$$

then,

$$L_n(ABC) = \left(x + \frac{y}{4}\right) p_{n-1}(ABC)$$

$$+ \frac{y}{4} \sum p(A) \{L_{n-1}(BC) + p(B) p(C)\}$$

$$- \sum p(A) \left(x + \frac{y}{2}\right) L_{n-1}(BC) - p(A) p(B) p(C)$$

$$= vL_{n-1}(ABC) = v^n \cdot L_0(ABC)$$

where

$$v = x + y/4.$$

Therefore,

$$\begin{aligned} p_{n+1}(ABC) &= v \cdot p_n(ABC) + \left(\frac{y}{4}\right) \sum p(A)\{L_n(BC) + p(B)p(C)\} \\ &= up_n(ABC) + \frac{y}{2} p(A)p(B)p(C) - \frac{y}{4} \cdot L_n(ABC), \end{aligned}$$

which in terms of the initial values is

$$\begin{aligned} p_{n+1}(ABC) &= u^{n+1} p_0(ABC) + (1 - u^{n+1}) p(A)p(B)p(C) \\ &\quad - (u^{n+1} - v^{n+1}) L_0(ABC) \end{aligned}$$

where

$$u = x + y/2.$$

$$v = x + y/4.$$

In the limiting case, when  $n \rightarrow \infty$ ,

$$p_\infty(ABC) = p(A)p(B)p(C)$$

which shows that the gametic frequency is equal to the product of the proportions of the corresponding genes taken separately.

Substituting the values of  $p_n(ABC)$ , etc., in (7) we can obtain the values of  $R$ 's in terms of the values in the initial population.

Let  $\phi_i^{(n)}$  be defined as the frequency of any  $i$ -factor heterozygote, i.e., a genotype in which  $i$ -factors are in heterozygous state and the remaining homozygous ( $i=0, 1, \dots, k$ ), in the  $n$ -th generation of a mixed selfing and random mating population and  $\psi_i^{(n)}$  as the resulting proportion of the corresponding genotype in  $n$ -th generation when there is only random mating in the  $(n-1)$ -th generation.

Further, let  $\phi_{i,j}^{(n)} (i \geq j)$  be defined as the sum of the proportions of those  $i$ -factor heterozygotes in the  $n$ -th generation which would give rise to the desired  $j$ -factor heterozygote in the next generation when subjected to self-fertilization, e.g.,  $f_{211}^{(n-1)}$  and  $f_{121}^{(n-1)}$  are the proportions of those two factors heterozygotes  $AABbCc$  and  $AaBBcC$  in  $(n-1)$ -th generation which can give rise to the single heterozygote 'AABbCc' in the next generation when selfed. Therefore in this case

$$\phi_{2,1}^{(n-1)} = f_{211}^{(n-1)} + f_{121}^{(n-1)},$$



The frequency of the occurrence of genotype 'AABBCC' in  $n$ -th generation is given by

$$f_{222}^{(n)} = \gamma R_{222}^{(n)} + c_{00} f_{222}^{(n-1)} + c_{10} (f_{221}^{(n-1)} + f_{212}^{(n-1)} + f_{122}^{(n-1)}) \\ + c_{20} (f_{211}^{(n-1)} + f_{121}^{(n-1)} + f_{112}^{(n-1)}) + c_{30} f_{111}^{(n-1)}$$

where

$$c_{ij} = x \cdot \left(\frac{1}{2}\right)^i \left(\frac{1}{4}\right)^{i-j}; \quad i \geq j = 0, 1, 2 \dots k, \quad k = \text{number of factors.}$$

In the new notation this can be written as

$$\phi_0^{(n)} = \gamma \psi_{0,0}^{(n)} + c_{00} \phi_{0,0}^{(n-1)} + c_{10} \phi_{1,0}^{(n-1)} + c_{20} \phi_{2,0}^{(n-1)} \\ + c_{30} \phi_{3,0}^{(n-1)} \quad (8)$$

where  $\psi_{i,j}$ 's are similar to  $\phi_{i,j}$ 's with the only difference that  $f$ 's in  $\phi_{i,j}$ 's are to be replaced by corresponding  $R$ 's. The values of  $\phi_{i,j}$ 's and  $\psi_{i,j}$ 's will be different for different genotypes. For particular homozygote 'AABBCC',

$$\phi_{0,0}^{(n-1)} = f_{222}^{(n-1)} = \phi_0^{(n-1)} \\ \phi_{1,0}^{(n-1)} = f_{221}^{(n-1)} + f_{122}^{(n-1)} + f_{212}^{(n-1)} \\ \phi_{2,0}^{(n-1)} = f_{211}^{(n-1)} + f_{121}^{(n-1)} + f_{112}^{(n-1)} \\ \phi_{3,0}^{(n-1)} = f_{111}^{(n-1)} = \phi_3^{(n-1)}$$

Now

$$\phi_{1,0}^{(n-1)} = f_{221}^{(n-1)} + f_{122}^{(n-1)} + f_{212}^{(n-1)},$$

putting the values of  $f$ 's

$$\phi_{1,0}^{(n-1)} = \gamma \psi_{1,0}^{(n-1)} + \delta_{11,0} \phi_{1,0}^{(n-2)} + \delta_{21,0} \phi_{2,0}^{(n-2)} + \delta_{31,0} \phi_{3,0}^{(n-2)} \\ = \gamma \psi_{1,0}^{(n-1)} + c_{11} \phi_{1,0}^{(n-2)} + \delta_{21,0} \phi_{2,0}^{(n-2)} + \delta_{31,0} \phi_{3,0}^{(n-2)}$$

where

$$\delta_{rs,t} = \binom{r-t}{s-t} \cdot c_{rs}; \quad r \geq s \geq t = 0, 1, 2 \dots k, \quad k = \text{number} \\ \text{of factors.}$$

$$\delta_{rr,t} = c_{rr},$$

similarly

$$\phi_{2,0}^{(n-1)} = \gamma \psi_{2,0}^{(n-1)} + \delta_{22,0} \phi_{2,0}^{(n-2)} + \delta_{32,0} \phi_{3,0}^{(n-2)} \\ = \gamma \psi_{2,0}^{(n-1)} + c_{22} \phi_{2,0}^{(n-2)} + \delta_{32,0} \phi_{3,0}^{(n-2)},$$

Putting the successive values of  $\phi$ 's in (8), we have

$$\begin{aligned}
 \phi_0^{(n)} = & y \sum_{r=0}^{n-1} (c_{00})^r \psi_{0,0}^{(n-r)} + (c_{00})^n \phi_{0,0}^{(0)} + c_{10} \omega_{01}^{(n)} \phi_{1,0}^{(0)} \\
 & + y \sum_{r=1}^{n-1} c_{10} \omega_{01}^{(r)} \psi_{1,0}^{(n-r)} \\
 & + \left\{ c_{20} \omega_{02}^{(n)} + \delta_{21,0} c_{10} \sum_{r=1}^{n-1} \omega_{01}^{(r)} (c_{22})^{n-r-1} \right\} \phi_{2,0}^{(0)} \\
 & + y \left\{ c_{20} \sum_{r=1}^{n-1} \omega_{02}^{(r)} \psi_{2,0}^{(n-r)} \right. \\
 & \left. + \delta_{21,0} c_{10} \sum_{r=1}^{n-2} \psi_{2,0}^{(n-r-1)} \sum_{s=1}^r (c_{22})^{r-s} \omega_{01}^{(s)} \right\} \\
 & + \left\{ c_{30} \omega_{03}^{(n)} + \sum_{i=1,2} \delta_{3i,0} c_{i0} \sum_{r=1}^{n-1} (c_{33})^{n-r-1} \omega_{0i}^{(r)} \right. \\
 & \left. + \delta_{32,0} \delta_{21,0} c_{10} \sum_{r=1}^{n-2} \sum_{s=1}^r (c_{33})^{n-r-2} (c_{22})^{r-s} \omega_{01}^{(s)} \right\} \phi_{3,0}^{(10)} \\
 & + y \left[ c_{30} \sum_{r=1}^{n-1} \omega_{03}^{(r)} \psi_{3,0}^{(n-r)} \right. \\
 & + \sum_{i=1,2} \delta_{3i,0} c_{i0} \sum_{r=1}^{n-2} \psi_{3,0}^{(n-r-1)} \sum_{s=1}^r (c_{33})^{r-s} \omega_{0i}^{(s)} \\
 & \left. + \delta_{32,0} \delta_{21,0} c_{10} \sum_{r=1}^{n-3} \psi_{3,0}^{(n-r-2)} \sum_{s=1}^r \sum_{t=1}^s (c_{33})^{r-s} (c_{22})^{s-t} \omega_{01}^{(t)} \right] \tag{9}
 \end{aligned}$$

where

$$\omega_{ij}^{(n)} = \frac{(c_{ii})^n - (c_{jj})^n}{c_{ii} - c_{jj}}, \quad i \neq j.$$

This will give us the genotypic frequency of any homozygote in the  $n$ -th generation. The values of  $\phi_{i,j}$ 's and  $\psi_{i,j}$ 's to be put in the R.H.S. of (9) will depend upon the homozygote whose frequency is required.

The frequencies of other genotypes can be derived from (9) itself by proper substitution of suffixes. Thus, the frequency,  $\phi_1(n)$ , of any single factor heterozygote will be

$$\begin{aligned}
\phi_1^{(n)} = & y \sum_{r=0}^{n-1} (c_{11})^r \psi_{1.1}^{(n-r)} + (c_{11})^n \phi_{1.1}^{(0)} + c_{21} \omega_{12}^{(n)} \phi_{2.1}^{(0)} \\
& + y c_{21} \sum_{r=1}^{n-1} \omega_{12}^{(r)} \psi_{2.1}^{(n-r)} + \left\{ c_{31} \omega_{13}^{(n)} + \delta_{32.1} c_{21} \right. \\
& \times \sum_{r=1}^{n-1} \omega_{12}^{(r)} (c_{33})^{n-r-1} \left. \right\} \phi_{3.1}^{(0)} \\
& + y \left[ c_{31} \sum_{r=1}^{n-1} \omega_{13}^{(r)} \psi_{3.1}^{(n-r)} + \delta_{32.1} c_{21} \right. \\
& \times \left. \sum_{r=1}^{n-2} \sum_{s=1}^r (c_{33})^{r-s} \omega_{12}^{(s)} \psi_{3.1}^{(n-r-1)} \right];
\end{aligned}$$

similarly,

$$\begin{aligned}
\phi_2^{(n)} = & y \sum_{r=0}^{n-1} (c_{22})^r \psi_{2.2}^{(n-r)} + (c_{22})^n \phi_{2.2}^{(0)} + c_{32} \omega_{23}^{(n)} \phi_{3.2}^{(0)} \\
& + y c_{32} \sum_{r=1}^{n-1} \omega_{23}^{(r)} \psi_{3.2}^{(n-r)}
\end{aligned}$$

and

$$\phi_3^{(n)} = y \sum_{r=0}^{n-1} (c_{33})^r \psi_{3.3}^{(n-r)} + (c_{33})^n \phi_{3.3}^{(0)}.$$

So far we have obtained the genotypic frequencies in the  $n$ -th generation for one, two and three independently segregating factors with arbitrary initial population. As the number of factors increases, the algebra becomes more cumbersome and the recurrence relations for gametic frequencies itself get more complicated even for random mating. As such it was not feasible to work out the genotypic frequencies in the general case of  $k$ -factors for arbitrary initial population. The results for the particular case of panmictic initial population are discussed in the next section.

### *Genotypic Frequencies and the Association of Factors for the Panmictic Initial Population 1*

The initial population considered is

$$\text{II} [p^2(A) \cdot AA + 2p(A)p(a) \cdot Aa + p^2(a) \cdot aa] \quad (10)$$

With this as the initial population the gametic frequencies under mixed selfing and random mating will not change from generation to genera-

tion and will be the same as in the initial population because  $L_0(AB)$ ,  $L_0(ABC)$ ; etc., will vanish and therefore,

$$R^{(n)}_{ijk} \dots = R^{(1)}_{ijk} \dots = f^{(0)}_{ijk} \dots$$

where  $f^{(0)}_{ijk} \dots$ 's are given by (10).

*Single factors,  $k = 1$ .*—The initial genotypic frequencies are given by  $p^2(A) \cdot AA + 2p(A)p(a) \cdot Aa + p^2(a) \cdot aa$ . Substituting these values in (1), we have

$$\left. \begin{aligned} f_2^{(n)} &= p^2(A) + \frac{x}{2-x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} p(A)p(a) \\ f_1^{(n)} &= 2 \left[ 1 - \frac{x}{2-x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} \right] p(A)p(a) \\ f_0^{(n)} &= p^2(a) + \frac{x}{2-x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} p(A)p(a) \end{aligned} \right\} (11)$$

The frequencies in equilibrium are already given in (2).

*Two factors,  $k = 2$ .*—Substituting the values of initial genotypic frequencies given by (10) for  $k = 2$ , in (3), (4) and (5) we have

$$f_{11}^{(n)} = \left[ 1 - \frac{3x}{4-x} \left\{ 1 - \left(\frac{x}{4}\right)^n \right\} \right] f_{11}^{(0)} \quad (12)$$

$$\begin{aligned} f_{21}^{(n)} &= \left[ 1 - \frac{x}{2-x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} \right] f_{21}^{(0)} \\ &+ \frac{1}{2} \left[ \frac{3x}{4-x} \left\{ 1 - \left(\frac{x}{4}\right)^n \right\} - \frac{x}{2-x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} \right] f_{11}^{(0)} \end{aligned} \quad (13)$$

$$\begin{aligned} f_{22}^{(n)} &= f_{22}^{(0)} + \frac{x}{4-2x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} (f_{21}^{(0)} + f_{12}^{(0)}) \\ &+ \left[ \frac{1}{2} \cdot \frac{x}{2-x} \left\{ 1 - \left(\frac{x}{2}\right)^n \right\} \right. \\ &\left. - \frac{3}{4} \cdot \frac{x}{4-x} \left\{ 1 - \left(\frac{x}{4}\right)^n \right\} \right] f_{11}^{(0)} \end{aligned} \quad (14)$$

The frequencies in the limiting case are independent of the initial frequencies and are already given in (6).

*Three factors,  $k = 3$ .*—Since the values of  $\psi$ 's in the case of panmictic initial population don't change with  $n$ , i.e., are equal to  $\phi^{(0)}$ 's

therefore, the equation (9) becomes,

$$\phi_0^{(n)} = \sum_{i=0}^3 \sum_{r=0}^i \phi_{i,0}^{(0)} \left(\frac{1}{2}\right)^{i-0} \binom{i-0}{r-0} (-1)^{r+0} \alpha_r^{(n)}$$

where

$$\alpha_r^{(n)} = \frac{y}{1 - c_{rr}} (1 - c_{rr}^n) + (c_{rr})^n,$$

similarly,

$$\phi_1^{(n)} = \sum_{i=1}^3 \sum_{r=1}^i \phi_{i,1}^{(0)} \left(\frac{1}{2}\right)^{i-1} \binom{i-1}{r-1} (-1)^{r+1} \alpha_r^{(n)}$$

$$\phi_2^{(n)} = \sum_{i=2}^3 \sum_{r=2}^i \phi_{i,2}^{(0)} \left(\frac{1}{2}\right)^{i-2} \binom{i-2}{r-2} (-1)^{r+2} \alpha_r^{(n)}$$

and

$$\phi_3^{(n)} = \alpha_3^{(n)} \phi_{3,3}^{(0)}.$$

From these equations it is evident that in the case of three independently segregating factors, the frequency of any  $j$ -factor heterozygote ( $j = 0, 1, 2, 3$ ) is given by

$$\phi_j^{(n)} = \sum_{i=j}^3 \sum_{r=j}^i \left(\frac{1}{2}\right)^{i-j} \binom{i-j}{r-j} (-1)^{r+j} \alpha_r^{(n)} \phi_{i,j}^{(0)} \quad (15)$$

and in the limiting case

$$\phi_j^{(\infty)} = \sum_{i=j}^3 \sum_{r=j}^i \phi_{i,j}^{(0)} (-1)^{r+j} \left(\frac{1}{2}\right)^{i-j} \binom{i-j}{r-j} \left(\frac{y}{1 - c_{rr}}\right). \quad (16)$$

The genotypic frequencies obtained earlier for one and two segregating factors can also be written in this form.

*k-Factors.*—From the genotypic frequencies obtained for one, two and three factors it can be deduced that the genotypic frequency of any  $j$ -factor heterozygote with mixed selfing and random mating in different generations, in the general case of ' $k$ ' independently segregating factors is given by

$$\begin{aligned} \phi_j^{(n)} &= \sum_{i=j}^k \sum_{r=j}^i \phi_{i,j}^{(0)} \left(\frac{1}{2}\right)^{i-j} \binom{i-j}{r-j} (-1)^{r+j} \alpha_r^{(n)} \\ &= \sum_{r=0}^{k-j} \phi_{r+j,j}^{(0)} \left(\frac{1}{2}\right)^r \sum_{s=0}^r \binom{r}{s} (-1)^s \alpha_{s+j}^{(n)} \end{aligned} \quad (17)$$

where

$\phi_{i,j}^{(0)}$  = sum of the genotypic frequencies in the initial population of those  $i$ -factor heterozygotes which would give rise to the desired  $j$ -factor heterozygote in the next generation when self-fertilized.

$x$  = proportion self-fertilized.

$y$  =  $(1 - x)$  = proportion cross-fertilized at random.

In the limiting case, the frequency of any  $j$ -factor heterozygote is given by

$$\phi_j^{(\infty)} = \sum_{r=0}^{k-j} \phi_{r+j,j}^{(0)} \left(\frac{1}{2}\right)^r \sum_{s=0}^r \binom{r}{s} (-1)^s \left(\frac{y}{1 - \frac{x}{2^{s+j}}}\right) \quad (18)$$

These frequencies are seen to be in equilibrium. The excess of the genotypic frequency in equilibrium of any  $j$ -factor heterozygote ( $j = 0, 1, 2 \dots, k$ ) over the product of equilibrium frequencies for genotypes at the separate loci is given by

$$\begin{aligned} p_{\infty}(Aa, Bb \dots Jj; II, \dots KK) - \prod p_{\infty}(Aa) \prod p_{\infty}(II) \\ = \sum_{r=0}^{k-j} \phi_{r+j,j}^{(0)} \left[ y \cdot \left(\frac{1}{2}\right)^r \sum_{s=0}^r \binom{r}{s} \frac{(-1)^s}{1 - \frac{x}{2^{s+j}}} \right. \\ \left. - \left(\frac{2y}{2-x}\right)^j \left(\frac{x}{4-2x}\right)^r \right] \end{aligned} \quad (19)$$

which shows that in equilibrium the genotypic frequency is not equal to the product of the genotypic frequencies at separate loci.

It has also been shown that for an arbitrary initial population with one, two and three independently segregating factors, the limiting

frequencies are independent of the initial genotypic frequencies and are in equilibrium. These depend only on the gene frequencies and the proportions of self-fertilization and thus are the same as obtained for the panmictic initial population having the same gene frequencies. While this study could not be extended beyond  $k = 3$  for arbitrary initial population, it would appear that the conclusion holds in the general case of  $k$ -factors.

#### MEAN AND THE GENOTYPIC VARIANCE

In any breeding programme it is equally important to have an idea of the variability that will be present in the population in successive stages of breeding as is the knowledge about genotypic composition. Work has already been done in this direction for populations subjected to only one system of breeding, *e.g.*, self-fertilization or cross-fertilization. But there does not seem to be any study made so far in the case of populations experiencing mixture of breeding systems. An attempt has, therefore, been made here to fill the gap to some extent by obtaining mean values and the genotypic variance in successive generations of a population undergoing mixed selfing and random mating.

In the discussion that follows, in addition to the assumptions made earlier, it has also been assumed that factors or loci are additive in action. This assumption is justified on the consideration that actions of genes can be made additive approximately by suitable transformation of the scale and even if there is small deviation from additivity it would show effects similar to those of environments (Fisher, 1913).

For the purpose of present study the initial population has been taken as

$$II [p^k(A) \cdot AA + 2p(A)p(a) \cdot Aa + p^2(a) \cdot aa].$$

*Single factor,  $k = 1$ .*—Let  $d_a$ ,  $h_a$ ,  $-d_a$  be the average effects of three genotypes  $AA$ ,  $Aa$ ,  $aa$  respectively for the gene pair  $A-a$ , measured from the mid-parental value. Then the mean  $\mu_n(A)$  in the  $n$ -th generation is given by

$$\mu_n(A) = (f_2^{(n)} - f_0^{(n)})d_a + f_1^{(n)}h_a,$$

where  $f_2^{(n)}$ ,  $f_1^{(n)}$ ,  $f_0^{(n)}$  have been obtained earlier in (11).

Therefore,

$$\begin{aligned} \mu_n(A) &= \overline{p(A) - p(a)} d_a + 2p(A)p(a)h_a(1 - 2z_n) \\ &= \mu_0(A) - 4z_n p(A)p(a)h_a \end{aligned} \quad (20)$$

where

$$z_n = \left(\frac{x}{4}\right) \frac{1 - \left(\frac{x}{2}\right)^n}{1 - \frac{x}{2}}, \text{ for all } n \text{ and } x$$

and the genotypic variance,  $V_n(A)$ , in the  $n$ -th generation is given by

$$\begin{aligned} V_n(A) &= (f_2^{(n)} + f_0^{(n)})d_a^2 + f_1^{(n)}h_a^2 - \{\mu_n(A)\}^2 \\ &= (1 - 2z_n)V_0(A) + 8p(A)p(a)z_n d_a^2 + 8p^2(A) \\ &\quad \times p^2(a)z_n(1 - 2z_n)h_a^2 \end{aligned} \quad (21)$$

where

$$\begin{aligned} V_0(A) &= 2p(A)p(a)\{d_a - \overline{p(A) - p(a)}h_a\}^2 \\ &\quad + 4p^2(A)p^2(a)h_a^2. \end{aligned}$$

*No dominance: ( $h = 0$ ).*—In the absence of dominance

$$\{\mu_n(A)\}_{ad} = \overline{p(A) - p(a)}d_a = \{\mu_0(A)\}_{ad} \quad (22)$$

where the suffix 'ad' stands for absence of dominance and

$$\begin{aligned} \{V_n(A)\}_{ad} &= \frac{1 - \left(\frac{x}{2}\right)^{n+1}}{1 - \frac{x}{2}} \cdot 2p(A)p(a)d_a^2 \\ &= \frac{1 - \left(\frac{x}{2}\right)^{n+1}}{1 - \frac{x}{2}} \{V_0(A)\}_{ad} \end{aligned} \quad (23)$$

In the limiting case

$$\{V_\infty(A)\}_{ad} = \frac{4}{2-x} p(A)p(a)d_a^2 = \frac{2}{2-x} \{V_0(A)\}_{ad} \quad (24)$$

In the absence of dominance, therefore, mean does not change and retains its initial value, whereas the genotypic variance increases



from generation to generation and also with the increase in the percentage of selfing.

*Complete dominance* ( $h_o = d_a$ ):

$$\begin{aligned}\{\mu_n(A)\}_{cd} &= \overline{\{p(A) - p(a) + 2p(A)p(a)(1 - 2z_n)\}} d_a \\ &= \{\mu_0(A)\}_{cd} - 4z_n p(A)p(a) d_a\end{aligned}\quad (25)$$

where suffix 'cd' stands for complete dominance.

$$\begin{aligned}\{V_n(A)\}_{cd} &= (1 - 2z_n) \{V_0(A)\}_{cd} + 8z_n p(A)p(a) d_a^2 \\ &\quad + 8p^2(A)p^2(a)z_n(1 - 2z_n) d_a^2 \\ \{V_0(A)\}_{cd} &= 4p^2(a) \{1 - p^2(a)\} d_a^2\end{aligned}$$

Hence,

$$\begin{aligned}\{V_n(A)\}_{cd} &= \{V_0(A)\}_{cd} + 8p(A)p(a)z_n d_a^2 \\ &\quad \times \{p(A) - p(a) + 2p(A)p(a)(1 - z_n)\}\end{aligned}\quad (26)$$

$$\begin{aligned}\{V_n(A) - V_{n-1}(A)\}_{cd} &= 8p(A)p(a) d_a^2 (z_n - z_{n-1}) \\ &\quad \times \{1 - 2p^2(a) - 2p(A)p(a)(z_n + z_{n-1})\} \\ &= 8p(A)p(a) d_a^2 (z_n - z_{n-1}) \\ &\quad \times \{p(A) - p(a) + 2p(A)p(a)(1 - z_n - z_{n-1})\}.\end{aligned}$$

It is evident from these equations that in the presence of dominance, the mean value decreases with the increase in the number of generations and also with the increasing percentage of selfing.

The genotypic variance decreases with the number of generations if  $p^2(a) \geq \frac{1}{2}$ , i.e., proportions of recessives in the initial population is equal to or greater than one half. It increases steadily in successive generations if  $p(a) \leq \frac{1}{2}$ .

The influence of gene frequency and proportion of selfing on genotypic variance for different generations ( $n = 1, 2, 3$  and  $\infty$ ) relative to the initial genotypic variance is given in Table I.

*Two factors,  $k = 2$ .*—The mean  $\mu_n(A, B)$  and genotypic variance  $V_n(A, B)$  in the  $n$ -th generation for two pairs of genes  $A-a$  and  $B-b$  is given by

$$\mu_n(A, B) = \Sigma \{f_{2.}^{(n)} - f_{0.}^{(n)}\} d_a + f_{1.}^{(n)} h_a;$$

where

$$f_2 = f_{22} + f_{21} + f_{20}, \text{ etc.}$$

$$\begin{aligned} \mu_n(A, B) &= \mu_n(A) + \mu_n(B) \\ &= \mu_0(A, B) - 4z_n \{p(A) p(a) h_a + p(B) p(b) h_b\} \end{aligned}$$

where  $d_b, h_b, -d_b$  are the average effects of genotypes  $BB, Bb, bb$  for the gene pair  $B-b$ , measured from the mid-parental value.

$$\begin{aligned} V_n(A, B) &= \Sigma [(f_{2.}^{(n)} + f_{0.}^{(n)} - (f_{2.}^{(n)} - f_{0.}^{(n)})^2) d_a^2 \\ &\quad + f_{1.}^{(n)} (1 - f_{1.}^{(n)}) h_a^2 - 2d_a h_a f_{1.}^{(n)} (f_{2.}^{(n)} - f_{0.}^{(n)}) \\ &\quad + 2h_o h_b (f_{11}^{(n)} - f_{1.}^{(n)} f_{.1}^{(n)})] \\ &= V_n(A) + V_n(B) + 8\lambda_n p(A) p(a) p(B) p(b) h_a h_b \end{aligned}$$

where

$$\lambda_n = \frac{\left\{ y + 3 \left( \frac{x}{4} \right)^{n+1} \right\}}{\left( 1 - \frac{x}{4} \right)} - \left\{ \frac{y + \left( \frac{x}{2} \right)^{n+1}}{1 - \frac{x}{2}} \right\}^2$$

and is depicted graphically in Fig. (1) for different values of  $n$  and percentage of selfing.  $\lambda_n$  measures the difference between the frequency

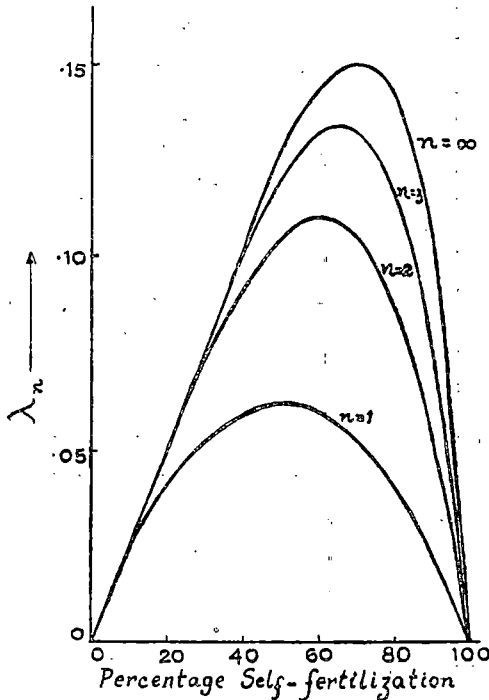


FIG. 1. Effect of proportion of selfing on ' $\lambda_n$ ' in different generations.

TABLE I

*Influence of the gene frequency and percentage selfing on the relative change in the genotypic variance  $(V_n - V_0)/V_0$  in different generations (n) from segregation of one factor subject to complete dominance ( $d = h = 1$ )*

$q = p(a)$	n	Percentage selfing									
		10	20	30	40	50	60	70	80	90	100
.1	1	0.443	0.883	1.318	1.749	2.176	2.599	3.018	3.433	3.843	4.250
	2	0.465	0.970	1.512	2.091	2.704	3.350	4.027	4.722	5.465	6.222
	3	0.466	0.979	1.542	2.159	2.835	3.572	4.374	5.240	6.172	7.169
	$\infty$	0.467	0.980	1.547	2.176	2.879	3.668	4.560	5.576	6.742	8.091
.2	1	0.190	0.377	0.560	0.740	0.917	1.090	1.260	1.427	1.590	1.750
	2	0.199	0.414	0.641	0.882	1.133	1.394	1.662	1.938	2.217	2.500
	3	0.200	0.417	0.654	0.910	1.187	1.482	1.798	2.132	2.482	2.844
	$\infty$	0.200	0.418	0.656	0.917	1.204	1.520	1.871	2.259	2.690	3.167
.3	1	0.104	0.205	0.303	0.399	0.492	0.582	0.676	0.755	0.837	0.917
	2	0.109	0.225	0.347	0.474	0.604	0.738	0.873	1.008	1.143	1.274
	3	0.109	0.227	0.353	0.488	0.632	0.783	0.940	1.102	1.266	1.427
	$\infty$	0.109	0.227	0.354	0.492	0.641	0.802	0.976	1.162	1.360	1.564
.4	1	0.060	0.117	0.172	0.226	0.277	0.326	0.372	0.417	0.460	0.500
	2	0.062	0.128	0.197	0.267	0.338	0.408	0.478	0.546	0.610	0.670
	3	0.063	0.130	0.200	0.275	0.352	0.432	0.512	0.591	0.666	0.734
	$\infty$	0.063	0.130	0.201	0.277	0.357	0.442	0.530	0.619	0.707	0.786

.5	1	0.032	0.063	0.092	0.120	0.146	0.170	0.192	0.213	0.232	0.250
	2	0.034	0.069	0.105	0.141	0.176	0.209	0.240	0.269	0.293	0.312
	3	0.034	0.070	0.107	0.145	0.183	0.220	0.255	0.286	0.311	0.328
∞	∞	0.034	0.070	0.107	0.146	0.185	0.224	0.262	0.296	0.342	0.333
.6	1	0.014	0.027	0.038	0.048	0.057	0.065	0.071	0.077	0.081	0.083
	2	0.015	0.029	0.043	0.056	0.077	0.076	0.082	0.085	0.084	0.078
	3	0.015	0.029	0.044	0.057	0.079	0.078	0.084	0.085	0.079	0.064
∞	∞	0.015	0.029	0.044	0.057	0.079	0.079	0.084	0.083	0.071	0.042
.7	1	0.000	0.000	0.001	0.004	0.007	0.001	0.016	0.022	0.029	0.036
	2	0.000	0.000	0.002	0.006	0.012	0.020	0.030	0.046	0.064	0.087
	3	0.000	0.000	0.002	0.007	0.013	0.024	0.038	0.058	0.085	0.120
∞	∞	0.000	0.000	0.002	0.007	0.014	0.025	0.042	0.067	0.104	0.160
.8	1	0.010	0.020	0.032	0.043	0.056	0.068	0.082	0.096	0.110	0.112
	2	0.010	0.023	0.037	0.053	0.072	0.093	0.117	0.144	0.174	0.208
	3	0.010	0.023	0.038	0.055	0.076	0.100	0.130	0.164	0.206	0.255
∞	∞	0.010	0.023	0.038	0.056	0.077	0.104	0.137	0.180	0.233	0.306
.9	1	0.018	0.037	0.056	0.075	0.094	0.114	0.133	0.153	0.174	0.194
	2	0.019	0.040	0.064	0.090	0.118	0.149	0.183	0.220	0.259	0.302
	3	0.019	0.041	0.065	0.093	0.125	0.160	0.201	0.247	0.299	0.358
∞	∞	0.019	0.041	0.066	0.094	0.127	0.165	0.205	0.265	0.332	0.415

of double heterozygote in  $n$ -th generation for two independent loci and the product of the frequencies of single heterozygotes at these two individual loci expressed as a multiple of  $4 p(A) p(a) p(B) p(b)$ .

$\lambda_n$  is a non-negative quantity which vanishes for  $x = 0, y = 1$ , i.e., complete cross-fertilization and for  $y = 0, x = 1$ , i.e., complete self-fertilization.

It is obvious from the expressions for mean and the genotypic variance that mean value due to two factors segregating independently in a population undergoing mixed self-fertilization and random mating is equal to the sum of the means due to each factor separately, whereas in the case of genotypic variance in addition to the sum of variances due to each factor separately, there appears another term which vanishes in the case of complete cross-fertilization, complete self-fertilization or in the absence of dominance.

Thus the usual assumption that variance due to independently segregating factors is equal to the sum of variances due to each factor separately is subject to limitations—viz., it holds good only when the population is experiencing only one system of breeding or when there is no dominance.

No dominance: ( $h_a = h_b = 0$ ) :—

$$\{\mu_n(A, B)\}_{ad} = \{\mu_0(A, B)\}_{ad} = \Sigma \overline{p(A) - p(a)} d_a$$

and

$$\begin{aligned} \{V_n(A, B)\}_{ad} &= 4 \frac{1 - \left(\frac{x}{2}\right)^{n+1}}{2 - x} \{\Sigma p(A) p(a) d_a^2\} \\ &= 2 \cdot \frac{1 - \left(\frac{x}{2}\right)^{n+1}}{2 - x} \{V_0(A, B)\}_{ad}. \end{aligned}$$

In the limiting case

$$\{V_\infty(A, B)\}_{ad} = \frac{2}{2 - x} \{V_0(A, B)\}_{ad}.$$

Therefore in the absence of dominance, the genotypic variance increases steadily as the proportion of self-fertilization increases. It also increases in successive generations for a given amount of self-fertilization.

*Complete dominance:* ( $h_a = d_a$ ;  $h_b = d_b$ ).—Let us assume that the genotypic values and gene frequencies of the two factors are equal, i.e.,

$$d_a = d_b = d, \text{ say}$$

$$p(a) = p(b) = q, \text{ say}$$

then,

$$\{\mu_n(A, B)\}_{ca} = 2 \{p - q + 2pq - 4pqz_n\} d, \text{ where } p = 1 - q,$$

and

$$\begin{aligned} \{V_n(A, B)\}_{ca} &= 8 [q^2(1 - q^2) + 2pqz_n \{p - q + 2pq(1 - z_n)\}] \\ &\quad \times d^2 + 8p^2q^2\lambda_n d^2 \end{aligned}$$

where  $8p^2q^2d^2\lambda_n$  is the additional increase due to mixture of breeding systems. This additional increase relative to the total genotypic variance in different generations ( $n = 1, 2, 3$  and  $\infty$ ) is given in Table II for different values of equal gene frequency 'q' and 'x' the fraction self-fertilized.

So far we have discussed the mean and genotypic variance in successive generations for one and two independently segregating factors both in the absence of dominance and complete dominance. The mean and genotypic variance in the  $n$ -th generation in the general case of 'k' independently segregating factors is given by

$$\mu_n(A, B, \dots, K)$$

$$= \sum^k \{p(A) - p(a)\} d_a + 2p(A) p(a) (1 - 2z_n) h_a$$

$$= \sum^k \mu_n(A) = \mu_0(A, B, \dots, K) - 4z_n \sum^k p(A) p(a) h_a$$

and

$$V_n(A, B, \dots, K) = \sum^k V_n(A) + 8\lambda_n \sum^{k \times 2} p(A) p(a) p(B) p(b) h_a h_b$$

Assuming the genotypic values and gene frequencies of different factors involved to be equal such that:

$$d_a = d_b = \dots = d_k = d, \text{ say,}$$

$$h_a = h_b = \dots = h_k = h, \text{ say,}$$

$$p(a) = p(b) = \dots = p(k) = q, \text{ say, then,}$$

TABLE II

*Additional increment in genotypic variance due to the mixture of breeding systems relative to the total genotypic variance in different generations (n) in the case of complete dominance (d = h = 1), for two independently segregating loci (k = 2)*

q	n	Percentage selfing								
		10	20	30	40	50	60	70	80	90
.1	1	0.012	0.017	0.018	0.017	0.016	0.013	0.010	0.007	0.004
	2	0.014	0.020	0.023	0.024	0.023	0.020	0.017	0.012	0.007
	3	0.014	0.021	0.024	0.026	0.025	0.023	0.020	0.015	0.008
	∞	0.014	0.021	0.025	0.026	0.026	0.025	0.022	0.017	0.010
.2	1	0.012	0.019	0.022	0.022	0.021	0.019	0.015	0.011	0.006
	2	0.014	0.023	0.029	0.032	0.032	0.030	0.025	0.019	0.011
	3	0.014	0.024	0.030	0.034	0.035	0.034	0.030	0.024	0.014
	∞	0.014	0.024	0.031	0.035	0.037	0.037	0.034	0.028	0.017
.3	1	0.011	0.017	0.021	0.022	0.022	0.020	0.017	0.012	0.006
	2	0.012	0.022	0.028	0.033	0.034	0.033	0.029	0.023	0.013
	3	0.012	0.022	0.030	0.035	0.038	0.038	0.035	0.028	0.017
	∞	0.012	0.022	0.030	0.036	0.040	0.041	0.039	0.033	0.021
.4	1	0.009	0.015	0.019	0.020	0.020	0.019	0.016	0.012	0.006
	2	0.010	0.019	0.026	0.030	0.033	0.032	0.029	0.023	0.014
	3	0.010	0.019	0.027	0.033	0.034	0.038	0.036	0.029	0.018
	∞	0.010	0.019	0.027	0.034	0.038	0.041	0.040	0.035	0.023

.5	1	0.007	0.012	0.016	0.017	0.018	0.017	0.014	0.011	0.006
	2	0.008	0.015	0.022	0.026	0.029	0.029	0.027	0.022	0.013
	3	0.008	0.016	0.023	0.029	0.033	0.034	0.034	0.029	0.018
.6	∞	0.008	0.016	0.023	0.029	0.034	0.038	0.038	0.034	0.024
	1	0.005	0.009	0.012	0.014	0.014	0.014	0.012	0.010	0.005
	2	0.006	0.012	0.017	0.021	0.024	0.025	0.024	0.019	0.012
.7	3	0.006	0.012	0.018	0.023	0.027	0.030	0.029	0.026	0.017
	∞	0.006	0.012	0.018	0.024	0.029	0.032	0.033	0.031	0.022
	1	0.004	0.007	0.009	0.010	0.011	0.010	0.009	0.007	0.004
.8	2	0.004	0.009	0.013	0.016	0.018	0.019	0.019	0.016	0.009
	3	0.004	0.009	0.013	0.018	0.021	0.023	0.024	0.021	0.014
	∞	0.004	0.009	0.014	0.018	0.022	0.025	0.030	0.026	0.018
.9	1	0.002	0.004	0.006	0.007	0.007	0.007	0.006	0.005	0.003
	2	0.003	0.005	0.008	0.011	0.012	0.013	0.013	0.011	0.007
	3	0.003	0.006	0.009	0.012	0.014	0.016	0.016	0.015	0.010
∞	∞	0.003	0.006	0.009	0.012	0.015	0.017	0.019	0.018	0.014
	1	0.001	0.002	0.003	0.003	0.004	0.003	0.003	0.002	0.001
	2	0.001	0.003	0.004	0.005	0.006	0.007	0.007	0.008	0.005
∞	3	0.001	0.003	0.004	0.006	0.007	0.009	0.009	0.008	0.005
	∞	0.001	0.003	0.004	0.006	0.007	0.009	0.010	0.010	0.007



TABLE III

*Additional increment in genotypic variance due to the mixture of breeding systems relative to the total genotypic variance in different generations (n) in the case of complete dominance (d = h = 1) for four independently segregating loci (k = 4)*

q	n	Percentage selfing								
		10	20	30	40	50	60	70	80	90
.1	1	0.037	0.049	0.053	0.051	0.046	0.039	0.031	0.022	0.011
	2	0.041	0.059	0.067	0.068	0.065	0.059	0.049	0.036	0.019
	3	0.041	0.060	0.070	0.073	0.072	0.066	0.057	0.043	0.024
	∞	0.041	0.060	0.070	0.074	0.074	0.070	0.062	0.049	0.029
.2	1	0.036	0.055	0.065	0.064	0.061	0.054	0.045	0.032	0.017
	2	0.040	0.066	0.082	0.089	0.090	0.084	0.073	0.055	0.031
	3	0.041	0.068	0.086	0.096	0.099	0.096	0.086	0.068	0.040
	∞	0.041	0.068	0.087	0.098	0.103	0.102	0.095	0.078	0.049
.3	1	0.032	0.051	0.061	0.065	0.063	0.058	0.048	0.035	0.019
	2	0.036	0.062	0.081	0.092	0.096	0.093	0.083	0.065	0.038
	3	0.036	0.064	0.085	0.099	0.107	0.107	0.099	0.081	0.050
	∞	0.036	0.064	0.086	0.101	0.111	0.114	0.110	0.094	0.062
.4	1	0.026	0.044	0.056	0.059	0.059	0.055	0.047	0.035	0.019
	2	0.030	0.054	0.073	0.086	0.092	0.091	0.083	0.067	0.040
	3	0.030	0.056	0.077	0.093	0.103	0.106	0.101	0.085	0.053
	∞	0.030	0.056	0.077	0.095	0.107	0.114	0.112	0.099	0.067

.5	1	0.021	0.036	0.046	0.051	0.052	0.049	0.042	0.032	0.018
	2	0.024	0.045	0.062	0.075	0.082	0.084	0.078	0.064	0.039
	3	0.024	0.046	0.065	0.081	0.093	0.098	0.095	0.082	0.053
$\infty$	$\infty$	0.024	0.046	0.066	0.083	0.097	0.105	0.106	0.097	0.068
.6	1	0.016	0.028	0.036	0.041	0.042	0.040	0.035	0.027	0.015
	2	0.018	0.035	0.050	0.062	0.069	0.071	0.068	0.056	0.035
	3	0.018	0.036	0.053	0.067	0.078	0.084	0.084	0.073	0.048
$\infty$	$\infty$	0.018	0.036	0.053	0.069	0.082	0.091	0.094	0.088	0.063
.7	1	0.012	0.021	0.027	0.031	0.032	0.031	0.027	0.021	0.012
	2	0.013	0.026	0.037	0.047	0.053	0.056	0.054	0.046	0.029
	3	0.013	0.027	0.039	0.051	0.061	0.067	0.068	0.060	0.041
$\infty$	$\infty$	0.013	0.027	0.040	0.053	0.064	0.072	0.077	0.073	0.054
.8	1	0.007	0.013	0.018	0.020	0.021	0.021	0.019	0.014	0.008
	2	0.008	0.017	0.025	0.031	0.036	0.039	0.038	0.032	0.021
	3	0.008	0.017	0.026	0.034	0.042	0.047	0.048	0.044	0.030
$\infty$	$\infty$	0.008	0.017	0.026	0.035	0.044	0.051	0.055	0.053	0.040
.9	1	0.004	0.006	0.009	0.010	0.011	0.010	0.009	0.007	0.004
	2	0.004	0.008	0.012	0.016	0.013	0.020	0.020	0.017	0.011
	3	0.004	0.008	0.013	0.017	0.021	0.024	0.025	0.023	0.016
$\infty$	$\infty$	0.004	0.008	0.013	0.018	0.022	0.026	0.029	0.029	0.022

(i) In the absence of dominance ( $h = 0$ ) mean does not change and retains its initial value, and the genotypic variance is given by

$$\{V_n(A, B, \dots K)\}_{ca} = \frac{1 - \left(\frac{x}{2}\right)^{n+1}}{2 - x} \cdot 4kpqd^2.$$

This increases from generation to generation and also with the increasing amount of self-fertilization in any given generation.

(ii) In the presence of complete dominance ( $h = d$ ), the mean and genotypic variance is given by

$$\{\mu_n(A, B, \dots, K)\}_{ca} = \{\mu_0(A, B, \dots, K)\}_{ca} - 4pqkd \cdot z_n$$

which decreases in successive generations and also for increase in the proportion of self-fertilization, and

$$\begin{aligned} \{V_n(A, B, \dots, K)\}_{ca} \\ = 4kd^2 [q^2(1 - q^2) + 2pqz_n \{p - q + 2pq(1 - z_n)\}] \\ + 4k(k - 1)p^2q^2d^2\lambda_n. \end{aligned}$$

The additional increment due to mixture of breeding systems, relative to the total genotypic variance is

$$\frac{(k - 1)p^2q^2\lambda_n}{q^2(1 - q^2) + 2pqz_n(p - q + 2pq - 2pqz_n) + (k - 1)p^2q^2\lambda_n}.$$

The relative additional increments in different generations for  $k = 4$  and  $k = 6$  have been tabulated in Tables III and IV, whereas for  $k = 2$ , these increments are given in Table II.

It is evident from these tables that:

- (i) the additional relative increments increase with factors;
- (ii) are higher for intermediate gene frequencies than for the extreme values; and
- (iii) are higher for intermediate proportions of self-fertilization than for extreme values.

The increment for a given amount of self-fertilization reaches its maximum very rapidly and there is relatively small increase thereafter.

The maximum values of these additional increments up to six factors are given in Table V. The maximum additional increase for  $k = 6$  is of the order of 18% of the genotypic variance.

TABLE IV

*Additional increment in genotypic variance due to the mixture of breeding systems relative to the total genotypic variance in different generations (n) in the case of complete dominance (d = h = 1), for six independently segregating loci (k = 6)*

g	n	Percentage selfing								
		10	20	30	40	50	60	70	80	90
·1	1	0·060	0·080	0·085	0·082	0·074	0·064	0·051	0·036	0·019
	2	0·066	0·095	0·107	0·109	0·104	0·094	0·079	0·058	0·032
	3	0·066	0·097	0·111	0·116	0·114	0·106	0·091	0·070	0·040
	∞	0·066	0·097	0·112	0·118	0·118	0·112	0·100	0·079	0·048
·2	1	0·059	0·088	0·101	0·103	0·098	0·087	0·072	0·052	0·028
	2	0·066	0·106	0·130	0·141	0·142	0·133	0·116	0·089	0·051
	3	0·066	0·109	0·135	0·150	0·156	0·151	0·136	0·109	0·065
	∞	0·066	0·109	0·136	0·153	0·161	0·160	0·149	0·124	0·080
·3	1	0·052	0·082	0·098	0·104	0·101	0·093	0·078	0·058	0·032
	2	0·058	0·100	0·128	0·144	0·150	0·146	0·131	0·104	0·062
	3	0·058	0·102	0·134	0·155	0·166	0·167	0·155	0·128	0·080
	∞	0·058	0·102	0·135	0·158	0·172	0·177	0·170	0·147	0·099
·4	1	0·043	0·071	0·087	0·095	0·095	0·088	0·076	0·057	0·032
	2	0·049	0·087	0·116	0·135	0·144	0·144	0·132	0·107	0·065
	3	0·049	0·089	0·122	0·146	0·161	0·165	0·158	0·134	0·086
	∞	0·049	0·090	0·123	0·149	0·167	0·176	0·174	0·155	0·107

TABLE IV (Contd.)

		Percentage selfing									
		10	20	30	40	50	60	70	80	90	
.5	n	1	0.059	0.074	0.082	0.083	0.079	0.068	0.052	0.029	
	1	0.035	0.059	0.074	0.082	0.083	0.079	0.068	0.052	0.029	
	2	0.039	0.073	0.099	0.119	0.130	0.132	0.123	0.102	0.063	
	3	0.039	0.075	0.105	0.129	0.146	0.153	0.149	0.129	0.085	
.6	n	∞	0.039	0.106	0.132	0.152	0.164	0.166	0.152	0.108	
	1	0.027	0.046	0.059	0.067	0.069	0.066	0.057	0.044	0.025	
	2	0.030	0.058	0.081	0.099	0.110	0.114	0.108	0.090	0.057	
	3	0.030	0.059	0.085	0.107	0.124	0.133	0.132	0.117	0.078	
.7	n	∞	0.030	0.059	0.110	0.129	0.143	0.148	0.138	0.101	
	1	0.019	0.034	0.044	0.050	0.053	0.051	0.045	0.035	0.020	
	2	0.022	0.042	0.061	0.066	0.086	0.090	0.087	0.074	0.047	
	3	0.022	0.044	0.064	0.083	0.098	0.107	0.108	0.097	0.066	
.8	n	∞	0.022	0.044	0.065	0.085	0.102	0.115	0.122	0.086	
	1	0.012	0.022	0.029	0.034	0.035	0.034	0.031	0.024	0.014	
	2	0.014	0.028	0.041	0.051	0.059	0.063	0.062	0.053	0.034	
	3	0.014	0.029	0.043	0.056	0.068	0.076	0.078	0.071	0.049	
.9	n	∞	0.014	0.029	0.043	0.058	0.078	0.082	0.088	0.065	
	1	0.006	0.011	0.014	0.017	0.018	0.017	0.016	0.012	0.007	
	2	0.007	0.014	0.020	0.026	0.030	0.033	0.033	0.028	0.018	
	3	0.007	0.014	0.021	0.029	0.035	0.040	0.042	0.038	0.027	
	n	∞	0.007	0.014	0.022	0.029	0.037	0.043	0.048	0.036	
	1	0.007	0.014	0.022	0.029	0.037	0.043	0.048	0.047	0.036	
	2	0.007	0.014	0.022	0.029	0.037	0.043	0.048	0.047	0.036	
	3	0.007	0.014	0.022	0.029	0.037	0.043	0.048	0.047	0.036	

TABLE V

*Maximum values of the additional increment upto six factors relative to the genotypic variance in different generations (n)*

<i>k</i>	2				4				6			
<i>n</i>	1	2	3	$\infty$	1	2	3	$\infty$	1	2	3	$\infty$
<i>q</i>	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Percentage selfing	40	50	55	60	40	50	55	60	40	50	55	60
Maximum values %	2.25	3.41	3.85	4.12	6.47	9.59	10.72	11.43	10.35	15.03	16.68	17.70

## SUMMARY

In the present paper a study has been made of the plant populations experiencing self and cross-fertilization at random, in respect of their genotypic composition, their approach to equilibrium, mean values and genotypic variability in successive generations in the general case of several independently segregating factors.

Regarding the genotypic composition, the study has been made upto three factors for the arbitrary initial population which has been extended to the general case of  $k$ -factors for the panmictic initial population. In the former case it has been observed that the limiting genotypic frequencies are independent of the initial frequencies and depend only on gene frequencies and amount of self-fertilization and are in equilibrium. The equilibrium frequencies in the general case of  $k$ -factors in an arbitrary initial population would, therefore, seem to be the same as obtained for the panmictic initial population with the same gene frequencies.

The expressions for the mean and genotypic variance in the  $n$ th generation have been obtained for the panmictic initial population in the general case of  $k$ -factors. It is observed that usual assumption that the variance due to several independently segregating factors is equal to the sum of the variances due to each factors separately does not hold good in the case of mixed selfing and random-mating population and is true only when there is either one breeding system in operation or when there is no dominance. It is, therefore, evident that trend of changes in the genotypic variance due to several factors cannot be predicted from the behaviour of a single factor because of the appearance of an additional term in the expression for the genotypic variance.

## ACKNOWLEDGEMENT

My grateful thanks are due to Dr. V. G. Panse, Statistical Adviser, Institute of Agricultural Research Statistics (I.C.A.R.), for his keen interest in the present investigation and encouragement given.

## REFERENCES

1. Bennet, J. H. ... "On the theory of random mating," *Ann. Eugen.*, London, 1954, 18, 311-17.
- . ——— and Binet, F. E. ... "Association ° between Mendelian factors with mixed selfing and random mating," *Heredity*, 1956, 10, 51-55.

3. Fisher, R. A. .. *The Theory of Inbreeding*, Oliver and Boyd, Edinburgh, 1949.
4. Garber, M. J. .. "Approach to genotypic equilibrium with varying percentages of selfing and cross-fertilization," *J. Heredity*, 1951, 42, 299-300.
5. Kempthorne, O. .. *An Introduction to Genetic Statistics*, John Wiley & Sons, New York, 1957.
6. Li, C. C. .. *Population Genetics*, The University Press, Chicago, 1955.
7. Mather, K. .. *Biometrical Genetics*, Methuen, London, 1949.