Theoretical and Applied Genetics 46, 157-163 (1975)

A Model for the Determination of the Variance in Genetic Relationship among Offspring from Open-pollinated Plant Populations

Hans-Rolf Gregorius
Lehrstuhl für Forstgenetik und Forstpflanzenzüchtung der Universität Göttingen (BRD)

Summary. Continuing a study of the average coefficient of kinship and inbreeding among offspring from specified mother plants belonging to a population of monoecious, diploid seed-plants, the variances of these coefficients have been computed. The variance in kinship was considered among offspring from a single mother plant and between offspring from two different mother plants. Special interest has been paid to the role played by the rate of self-fertilization and the effective size of neighbourhood and common neighbourhood. The graphical representation of some numerical examples indicates that it is impossible to predict general tendencies which hold true for the behaviour of all three types of variance if they are regarded as functions of the rate of self-fertilization; only the coefficient of inbreeding and kinship among offspring from the same parent showed similar tendencies. The influence of the effective size of neighbourhood and common neighbourhood on the respective variances proved to be of minor importance.

Introduction

Although the coefficient of kinship has been defined on a probabilistic basis by Malécot (1948), it can be conceived as a deterministic quantity in so far as its definition refers to a fixed pair of individuals. However, in many cases this pair of individuals results from a random sample taken from specified populations, whose members are genetically related in different ways. In order to extend the concept of the coefficient of kinship for a given pair of individuals to this situation, usually the average for all possible pairs is taken. More precisely, this means that the coefficient of kinship is regarded as a realization of a random variable, and the average is the expectation of this variable. The information about genetic relationship within a population or between populations, which is contained in the average, may be very misleading, because it makes no statements about the size and frequency of existing deviations. Thus it is necessary to determine the variance of the coefficient of kinship in order to get an idea of the degree of heterogeneity of the genetic relationship within or between populations. Despite the fact that considerable deviations in the coefficient of kinship from expectation can lead to severe consequences in the field of applied breeding theory when merely considering averages, to my knowledge it is hard to find any publication, except the paper of Leviandier and Jacquard (1974), treating variances of coefficients of kinship. Because of this the aim of the present paper was to derive these variances for the special situations arising in monoecious seed plant populations. The basic assumptions underlying the model employed are the same as those chosen in a preceding paper (Gregorius 1975), which the present paper completes. This model serves to determine the variance of the coefficient of kinship of two seeds randomly taken from the seed production of one mother plant or two different mother plants and the variance of the coefficient of inbreeding within the seed production of one mother plant.

The Model

The basic assumptions taken from the paper published recently (Gregorius 1975) will be repeated briefly:

The considerations are based on monoecious, diploid seed plant populations which are distributed over their habitats according to a specified population density. The seed production of each plant is broken down into one part resulting from self-fertilization and a remaining part resulting from cross-fertilization. All members of the population flower at the same time and any form of gametic selection, mutation and immigration of pollen is regarded as negligible. The mating probabilities between single individuals are assumed to contain the influence due to the mode of pollen-dispersal and differential pollen production.

The following notation will be used:
$\mu(x)$:= population density at place x.
$q(x)$:= rate of self-fertilization of a plant located at place x.
$p(x|y)$:= probability-density that pollen which came to fertilization at a plant located at place y originates from a plant located at place x.

$K(x)$:= probability that the two male gametes (pollen) contained in two seeds belonging to the cross-fertilized part of the seed production of a plant located at place x originate from the same father plant. Obviously

$$K(x) = \int \frac{p(y|x)^2}{\mu(y)}\,dy\,.$$

$K(x,y)$:= probability that the two male gametes (pollen) contained in two seeds belonging to the cross-fertilized parts of the seed production of two plants located at places x and y, respectively, originate from the same father plant. Obviously $K(x,y) = \int \frac{p(z|x)p(z|y)}{\mu(z)}\,dz\,.$

$f(x)$:= coefficient of inbreeding of a plant located at place x. $f(x)$ is conceived as a random variable.

$\Phi(x,y)$:= coefficient of kinship between two plants located at places x and y, respectively. $\Phi(x,y)$ is conceived as a random variable.

Note that $\frac{p(x|y)}{\mu(x)}$ can be regarded as an approximation for the probability that pollen which came to fertilization at a plant located at place y originates from just one different plant located at place x.

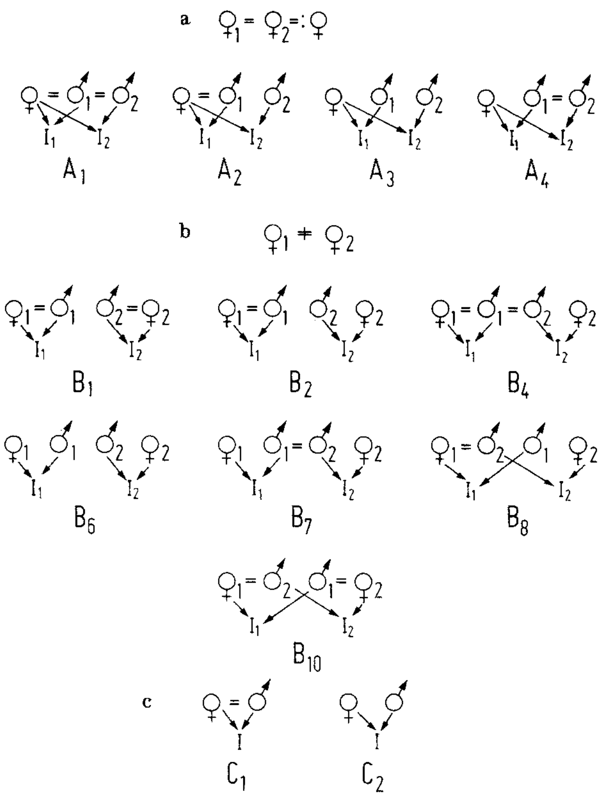
The parental population is assumed to result from random distribution of the individuals over the habitat. This implies that the random variables $\Phi(x,y)$ have identical probability-distributions for all places x and y, the same being true for the variables $f(x)$ for all places x, and that the variables $\Phi(x,y)$ and $f(x)$ are mutually stochastically independent for all x and y.

Based on these assumptions, the variance of the coefficient of kinship among the seed production of a plant located at a specified place and between the seed production of two different plants located at specified places and the variance of the coefficient of inbreeding within the seed production of a plant located at a specified place will be computed.

The coefficient of kinship of two specified individuals (and of inbreeding in one specified individual) is completely determined by the genetic relationship between their (or its) parents. In our case parents are identified by the places at which they are located, which, for the present consideration, requires distinguishing between three situations:

a) two seeds are produced from the same plant located at place x; in this case the variable $L_a(x)$ shall designate the two places from which the male gamete (pollen) present in each seed originates, and the variable $\Phi_a(x)$ gives the coefficient of kinship between these seeds.

b) two seeds produced from two different plants located at places x and y each; in this case the variable $L_b(x,y)$ shall designate the two places from which the male gametes present in each seed originate, and the variable $\Phi_b(x,y)$ gives the coefficient of kinship between these seeds.



The remaining three cases B_3, B_5, B_9 can be obtained from B_2, B_4, B_8 using the same order, by changing subscripts 1 into 2 and 2 into 1.

Fig. 1. Types of identity between parents (♀ = female parent, ♂ = male parent): a) The two offspring (seed) I_1 and I_2 originate from the same mother plant (♀$_1$ = ♀$_2$). b) The two offspring (seed) I_1 and I_2 originate from two different mother plants (♀$_1$ ≠ ♀$_2$). c) Types of identity between the two parents of one offspring (seed) I

c) a seed is produced from a plant located at place x; in this case the variable $L_c(x)$ shall designate the place from which the male gamete present in this seed originates, and the variable $f_c(x)$ gives the coefficient of inbreeding of this seed.

Since the fact that some of the parents (e.g. in situation a) up to all four) may happen to be the same individual introduces an effect on the genetic relationship among offspring which is not just due to the relatedness of their parents, this influence should be taken into account by distinguishing the different types of identity between parents for each of the above three situations (as in Fig. 1).

It is now possible to give a compilation of all the values that the variables $\Phi_a(x)$, $\Phi_b(x,y)$ and $f_c(x)$ can adopt:

a) $\Phi_a(x)$ in case $A_i (i=1,\dots,4)$ and $L_a(x)$ is given:

$A_1,\ L_a(x) = (x,x) : \frac{1}{2}(1+f(x))$;

$A_2,\ L_a(x) = (x,y) : \frac{1}{2}\left[\frac{1}{2}(1+f(x)) + \Phi(x,y)\right]$

$A_3,\ L_a(x) = (z,y) : \frac{1}{4}\left[\frac{1}{2}(1+f(x)) + \Phi(x,y) + \Phi(z,x) + \Phi(z,y)\right]$

$A_4,\ L_a(x) = (y,y) : \frac{1}{4}\left[\frac{1}{2}(1+f(x)) + 2\Phi(x,y) + \frac{1}{2}(1+f(y))\right]$.

b) $\Phi_b(x,y)$ in case $B_i (i=1,\dots,10)$ and $L_b(x,y)$ is given:

$B_1,\ L_b(x,y) = (x,y) : \Phi(x,y)$

$B_2,\ L_b(x,y) = (x,z) : \frac{1}{2}[\Phi(x,y) + \Phi(x,z)]$

$B_3,\ L_b(x,y) = (z,y) : \frac{1}{2}[\Phi(x,y) + \Phi(y,z)]$

$B_4,\ L_b(x,y) = (x,x) : \frac{1}{2}[\Phi(x,y) + \frac{1}{2}(1+f(x))]$

$B_5,\ L_b(x,y) = (y,y) : \frac{1}{2}[\Phi(x,y) + \frac{1}{2}(1+f(y))]$

$B_6,\ L_b(x,y) = (u,v) : \frac{1}{4}[\Phi(x,y) + \Phi(x,v) + \Phi(u,y) + \Phi(u,v)]$

$B_7,\ L_b(x,y) = (u,u) : \frac{1}{4}[\Phi(x,y) + \Phi(x,u) + \Phi(u,y) + \frac{1}{2}(1+f(u))]$

$B_8,\ L_b(x,y) = (u,x) : \frac{1}{4}[\Phi(x,y) + \frac{1}{2}(1+f(x)) + \Phi(u,y) + \Phi(u,x)]$

$B_9,\ L_b(x,y) = (y,u) : \frac{1}{4}[\Phi(x,y) + \frac{1}{2}(1+f(y)) + \Phi(u,x) + \Phi(u,y)]$

$B_{10},\ L_b(y,x) = (y,x) : \frac{1}{4}[2\Phi(x,y) + \frac{1}{2}(1+f(x)) + \frac{1}{2}(1+f(y))]$.

c) $f_c(x)$ in case $C_i (i=1,2)$ and $L_c(x)$ is given:

$C_1,\ L_c(x) = x : \frac{1}{2}(1+f(x))$; $C_2,\ L_c(x) = y : \Phi(x,y)$

Returning to our actual intention, namely the computation of the variances $V(\dots)$ of $\Phi_a(x)$, $\Phi_b(x,y)$ and $f_c(x)$, which in terms of expectations $E(\dots)$ may be given the representation $V(\Phi_a(x)) = E(\Phi_a(x)^2) - E(\Phi_a(x))^2$ etc., we are now in a position to solve the problem, applying some elementary properties for conditional expectations. The conditional expectation of e.g. $\Phi_a(x)$ given $L_a(x)$ will be written as $E(\Phi_a(x)|L_a(x))$. With this notation, and using $P(A)$ for the probability of any event A, we obtain

$$E(\Phi_a(x)) = \sum_{i=1}^{4} P(A_i)\cdot E(E(\Phi_a(x)|L_a(x))|A_i) \qquad (1)$$

$$E(\Phi_b(x,y)) = \sum_{i=1}^{10} P(B_i)\cdot E(E(\Phi_b(x,y)|L_b(x,y))|B_i) \qquad (2)$$

$$E(f_c(x)) = \sum_{i=1}^{2} P(C_i)\cdot E(E(f_c(x)|L_c(x))|C_i), \qquad (3)$$

the same representations holding true if Φ_a^2, Φ_b^2, f_c^2 are substituted for Φ_a, Φ_b, f_c respectively. Thus it remains to determine the probabilities $P(\dots)$ and the conditional expectations.

The probabilities $P(A_i)$, $P(B_i)$ and $P(C_i)$ can be derived directly from the statements contained in Fig. 1, recalling the meaning of the quantities

$q(x)$, $\frac{p(x|y)}{\mu(x)}$, $K(x)$ and $K(x,y)$:

$P(A_1) = q(x)^2$, $P(A_2) = 2q(x)(1-q(x))$,

$P(A_3) = (1-q(x))^2(1-K(x))$, $P(A_4) = (1-q(x))^2\cdot K(x)$;

$P(B_1) = q(x)q(y)$, $P(B_2) = q(x)(1-q(y))\left(1-\frac{p(x|y)}{\mu(x)}\right)$,

$P(B_3) = (1-q(x))q(y)\left(1-\frac{p(y|x)}{\mu(y)}\right)$, $P(B_4) = q(x)(1-q(y))\cdot\frac{p(x|y)}{\mu(x)}$, $P(B_5) = (1-q(x))q(y)\cdot\frac{p(y|x)}{\mu(y)}$,

$P(B_6) = (1-q(x))(1-q(y))\cdot\left[\left(1-\frac{p(x|y)}{\mu(x)}\right)\cdot\left(1-\frac{p(y|x)}{\mu(y)}\right) - K(x,y)\right]$, $P(B_7) = (1-q(x))(1-q(y))\cdot K(x,y)$,

$P(B_8) = (1-q(x))(1-q(y))\left(1-\frac{p(y|x)}{\mu(y)}\right)\cdot\frac{p(x|y)}{\mu(x)}$,

$P(B_9) = (1-q(x))(1-q(y))\left(1-\frac{p(x|y)}{\mu(x)}\right)\cdot\frac{p(y|x)}{\mu(y)}$,

$P(B_{10}) = (1-q(x))(1-q(y))\cdot\frac{p(y|x)\cdot p(x|y)}{\mu(y)\cdot\mu(x)}$;

$P(C_1) = q(x)$, $P(C_2) = 1-q(x)$.

Clearly $\sum_{i=1}^{4} P(A_i) = \sum_{i=1}^{10} P(B_i) = \sum_{i=1}^{2} P(C_i) = 1$.

Because the $\Phi(x,y)$ have been assumed to possess identical distributions for all places x, y, as do the $f(x)$ for all places x, the above conditional expectations of $\Phi_a(x)$, $\Phi_b(x,y)$ and $f_c(x)$ attain the following representations, putting $E(\Phi(x,y)) = \Phi$ and $E(f(x)) = f$:

$E(E(\Phi_a(x)|L_a(x))|A_i)$ in case

$A_1 : \frac{1}{2}(1+f)$; $A_2 : \frac{1}{4}(1+f) + \frac{1}{2}\cdot\Phi$; $A_3 : \frac{1}{8}(1+f) + \frac{3}{4}\cdot\Phi$; $A_4 : \frac{1}{4}(1+f) + \frac{1}{2}\cdot\Phi$.

$E(E(\Phi_b(x,y)|L_b(x,y))|B_i)$ in case

B_1, B_2, B_3 and $B_6 : \Phi$; B_4, B_5 and $B_{10} : \frac{1}{4}(1+f) + \frac{1}{2}\cdot\Phi$;

B_7, B_8 and $B_9 : \frac{1}{8}(1+f) + \frac{3}{4}\cdot\Phi$.

$E(E(f_c(x)|L_c(x))|C_i)$ in case

$C_1 : \frac{1}{2}(1+f)$; $C_2 : \Phi$.

Applying formulae (1), (2) und (3) we are now immediately led to

$$E(\Phi_a(x)) = \Phi + [\tfrac{1}{2}(1+f) - \Phi] \cdot \tfrac{1}{4}[(1+q(x))^2 + (1-q(x))^2 \cdot K(x)]$$

$$E(\Phi_b(x,y)) = \Phi + [\tfrac{1}{2}(1+f) - \Phi] \cdot \tfrac{1}{4}[\frac{p(x|y)}{\mu(x)}(1+q(x))(1-q(y)) + \\ + \frac{p(y|x)}{\mu(y)}(1-q(x))(1+q(y)) + (1-q(x))(1-q(y)) \cdot K(x,y)]$$

$$E(f_c(x)) = \Phi + [\tfrac{1}{2}(1+f) - \Phi] \cdot q(x),$$

which are exactly the same results I derived earlier (Gregorius 1975, equ. (1a), (2a), (3a)) using a different approach.

The assumption of mutual independence among the variables $\Phi(x,y)$ and $f(x)$ for all places x and y implies the following results for the conditional expectations of Φ_a^2, Φ_b^2 and f_c^2, putting $V(\Phi(x,y)) = V_\Phi$ and $V(f(x)) = V_f$:

$E(E(\Phi_a(x)^2 | L_a(x)) | A_i)$ in case

$$A_1 : \tfrac{1}{4}(1+f)^2 + \tfrac{1}{4} V_f ; \quad A_2 : \tfrac{1}{4}(\tfrac{1}{4}(1+f)^2 + \tfrac{1}{4} V_f + V_\Phi + \Phi^2 + \\ + (1+f)\Phi) ;$$

$$A_3 : \tfrac{1}{16} \tfrac{1}{4}(1+f)^2 + \tfrac{1}{4} V_f + 3V_\Phi + 3(1+f)\Phi + 9\Phi^2) ;$$

$$A_4 : \tfrac{1}{4}(\tfrac{1}{4}(1+f)^2 + \tfrac{1}{8} V_f + V_\Phi + \Phi^2 + (1+f)\Phi) .$$

$E(E(\Phi_b(x,y)^2 | L_b(x,y)) | B_i)$ in case

$$B_1 : V_\Phi + \Phi^2 ; \quad B_2 \text{ and } B_3 : \tfrac{1}{2} V_\Phi + \Phi^2 ;$$

$$B_4 \text{ and } B_5 : \tfrac{1}{4}(V_\Phi + \Phi^2 + \tfrac{1}{4} V_f + \tfrac{1}{4}(1+f)^2 + (1+f)\Phi) ;$$

$$B_6 : \tfrac{1}{4} V_\Phi + \Phi^2 ; \quad B_7, B_8 \text{ and } B_9 : \tfrac{1}{16}(\tfrac{1}{4} V_f + 3 V_\Phi + \\ + \tfrac{1}{4}(1+f)^2 + 3(1+f)\Phi + 9\Phi^2) ;$$

$$B_{10} : \tfrac{1}{4}(V_\Phi + \tfrac{1}{8} V_f + \Phi^2 + \tfrac{1}{4}(1+f)^2 + (1+f)\Phi) .$$

$E(E(f_c(x)^2 | L_c(x)) | C_i)$ in case

$$C_1 : \tfrac{1}{4}(1+f)^2 + \tfrac{1}{4} V_f ; \quad C_2 : V_\Phi + \Phi^2 .$$

With that, all the quantities which are needed to compute the variances according to formulae (1), (2) and (3) are specified. In order to arrive at a representation which is as easy as possible to survey, and at the same time displays the role played by V_Φ, V_f, Φ and f clearly, it is necessary to perform some tedious rearrangements leading to the following final results:

$$V(\Phi_a(x)) = V_\Phi \cdot a_1(x) + \tfrac{1}{4} V_f \cdot a_2(x) + [\tfrac{1}{2}(1+f) - \Phi]^2 \cdot a_3(x) \tag{4}$$

$$V(\Phi_b(x,y)) = V_\Phi \cdot b_1(x,y) + \tfrac{1}{4} V_f \cdot b_2(x,y) + \\ + [\tfrac{1}{2}(1+f) - \Phi]^2 \cdot b_3(x,y) \tag{5}$$

$$V(f_c(x)) = V_\Phi \cdot (1-q(x)) + \tfrac{1}{4} V_f \cdot q(x) + \\ + [\tfrac{1}{2}(1+f) - \Phi]^2 \cdot (1-q(x))q(x) \tag{6}$$

$$a_1(x) = \tfrac{1}{16}(1-q(x))^2 \cdot K(x) + \tfrac{1}{2}(1-q(x))(\tfrac{3}{8} + \tfrac{5}{8} q(x))$$

$$a_2(x) = \tfrac{1}{16}(1-q(x))^2 \cdot K(x) + (\tfrac{1}{4} + \tfrac{3}{4} q(x))^2$$

$$a_3(x) = \tfrac{3}{16}(1-q(x))^2 \cdot K(x) + (\tfrac{1}{4} + \tfrac{3}{4} q(x))^2 - \\ - \tfrac{1}{16}[(1+q(x))^2 + (1-q(x))^2 \cdot K(x)]^2$$

$$b_1(x,y) = \tfrac{1}{4}(1+q(x))(1+q(y)) - \\ - \tfrac{1}{16}(1-q(x))(1-q(y))[K(x,y) - \\ - 2\frac{p(x|y)p(y|x)}{\mu(x)\mu(y)}] - \tfrac{1}{4}(1-q(y))(\tfrac{1}{4} + \tfrac{3}{4} q(x)) \cdot \\ \cdot \frac{p(x|y)}{\mu(x)} - \tfrac{1}{4}(1-q(x))(\tfrac{1}{4} + \tfrac{3}{4} q(y)) \cdot \frac{p(y|x)}{\mu(y)}$$

$$b_2(x,y) = \tfrac{1}{4}(1-q(y))(\tfrac{1}{4} + \tfrac{3}{4} q(x)) \cdot \frac{p(x|y)}{\mu(x)} + \\ + \tfrac{1}{4}(1-q(x))(\tfrac{1}{4} + \tfrac{3}{4} q(y)) \cdot \frac{p(y|x)}{\mu(y)} + \\ + \tfrac{1}{16}(1-q(x))(1-q(y)) \cdot K(x,y)$$

$$b_3(x,y) = \tfrac{1}{4}(1-q(y))(\tfrac{1}{4} + \tfrac{3}{4} q(x)) \cdot \frac{p(x|y)}{\mu(x)} + \\ + \tfrac{1}{4}(1-q(x))(\tfrac{1}{4} + \tfrac{3}{4} q(y)) \cdot \frac{p(y|x)}{\mu(y)} + \\ + \tfrac{1}{16}(1-q(x))(1-q(y))[K(x,y) + \\ + 2 \cdot \frac{p(x|y)p(y|x)}{\mu(x)\mu(y)}] - \\ - \tfrac{1}{16}[\frac{p(x|y)}{\mu(x)}(1+q(x))(1-q(y)) + \\ + \frac{p(y|x)}{\mu(y)}(1+q(y))(1-q(x)) + \\ + (1-q(x))(1-q(y)) \cdot K(x,y)]^2 .$$

The surprisingly high formal agreement between (4), (5) and (6) allows some statements to be made which are applicable to all of the three situations. The variances increase in a linear manner with V_Φ, V_f and the squared difference between the average coefficient of kinship of a plant with itself $\frac{1}{2}(1+f)$ and the average coefficient of kinship between two different plants Φ. Furthermore, provided we accept the case $\frac{1}{2}(1+f) < \Phi$ to be not realistic, the variances of f and Φ increase and decrease, respectively, with increasing values for f and Φ. This last finding deserves particular attention, since it tells us that an increasing average coefficient of kinship in the parental generation always implies increasing homogenization of kinship as well as inbreeding effects within the offspring generation, while an increasing coefficient of inbreeding has just the opposite consequence.

Considering a joint change in Φ and f homogenization is approached to the second order according to the decreasing difference $\frac{1}{2}(1+f) - \Phi$.

If a plant population reproduces by pure self-fertilization only, i.e. $q(x) \equiv 1$, then $a_1(x) = a_3(x) = b_2(x,y) = b_3(x,y) \equiv 0$ and $b_1(x,y) = a_2(x) \equiv 1$ and therefore $V(\Phi_a(x)) = V(f_c(x)) = \frac{1}{4} V_f$, $V(\Phi_b(x,y) \equiv V_\Phi$ as could have been expected. The impact of varying rates of self-fertilization on the variances can be demonstrated by Figs. 2, 3, 4, where for simplicity we assume $q(x)$ to be constant throughout the population, i.e. $q(x) \equiv q$. The fact that the expectations of Φ_a, Φ_b and f_c in succession, are increasing, decreasing and increasing function of q, should be regarded as the background to the following results (Figs. 2, 3, 4):

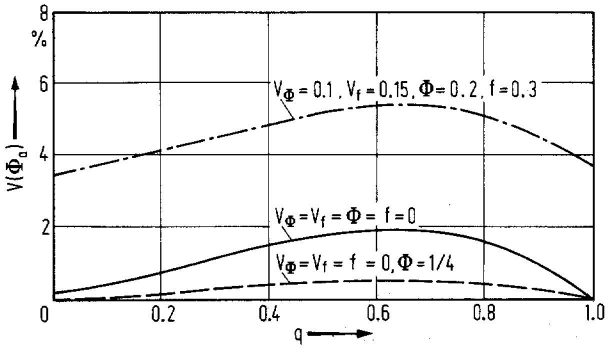


Fig. 2. The variance of the coefficient of kinship Φ_a between two seeds produced from the same mother plant as a function of the rate of self-fertilization q. The representations are based on $K(x) = 0.15$

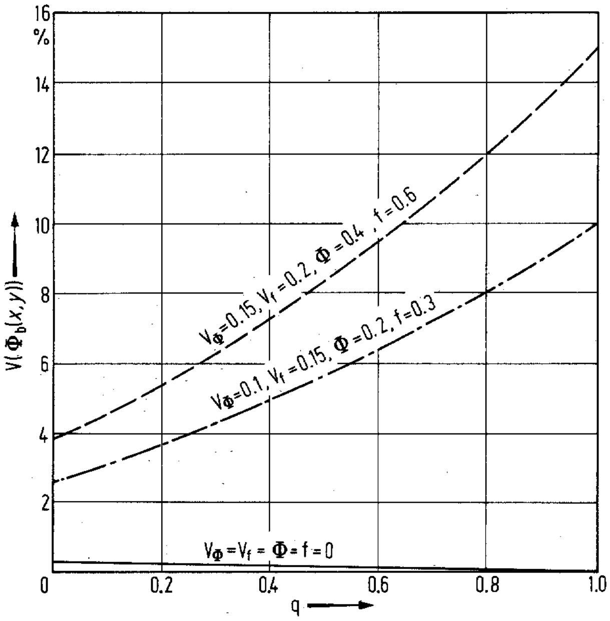


Fig. 3. The variance of the coefficient of kinship $\Phi_b(x,y)$ between two seeds produced from two different mother plants located at places x and y respectively as a function of the rate of self-fertilization q. The representations are based on

$K(x,y) = 0.13$ and $\frac{p(x|y)}{\mu(x)} = \frac{p(y|x)}{\mu(y)} = 0.04$

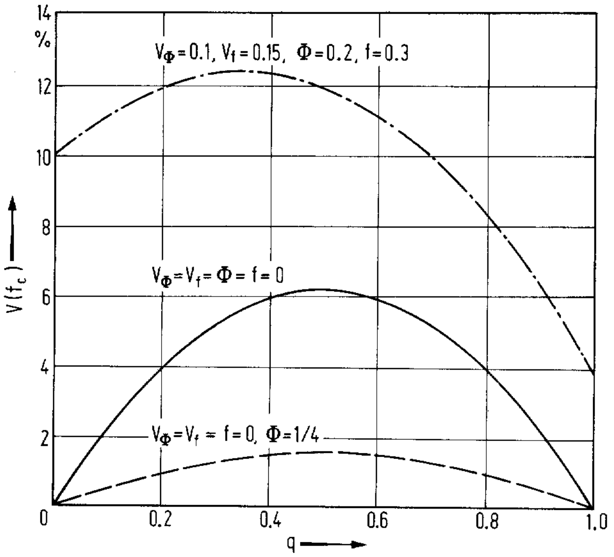


Fig. 4. The variance of the coefficient of inbreeding f_c of a seed as a function of the rate of self-fertilization q

It obviously is impossible to predict general tendencies which hold true for the behaviour of all three variances if they are regarded as functions of the rate of self-fertilization q. Furthermore, this statement can even be maintained considering just one of the variances, namely $V(\Phi_b(x,y))$, whose type of dependence on q, to a high degree is determined by the 'initial conditions' for V_Φ, V_f, Φ and f. On the other hand, comparing figures 4 and 2, inbreeding and the relationship between offspring from the same parent lead to the same tendencies for their respective variances, but differ clearly with respect to magnitude in favour of $V(f_c)$.

The Influence of Neighbourhood-Size

We shall use Wright's (1946) concept of neighbourhood in a sense which is directly applicable to the conditions realized in plant populations (see Gregorius 1975). In the latter paper the notations $N_e(x)$ and $N_e(x,y)$ were used respectively to describe the effective size of neighbourhood of a plant located at place x and of the effective size of common neighbourhood of two plants located at places x and y, and it could have been shown that $K(x) = \frac{1}{N_e(x)}$ and $K(x,y) = \frac{N_e(x,y)}{N_e(x) \cdot N_e(y)}$. These quantities, including $\frac{p(x|y)}{\mu(x)}$, comprise all effects on the variances which are due to the pollen-dispersal; consequently pollen dispersal does not affect the variance nor the expectation of f_c, as can be seen from the respective representations. On the other hand, expectations and variances for Φ_a and Φ_b are governed by pollen dispersal, but in two entirely different ways. $E(\Phi_a(x))$ and $V(\Phi(x))$

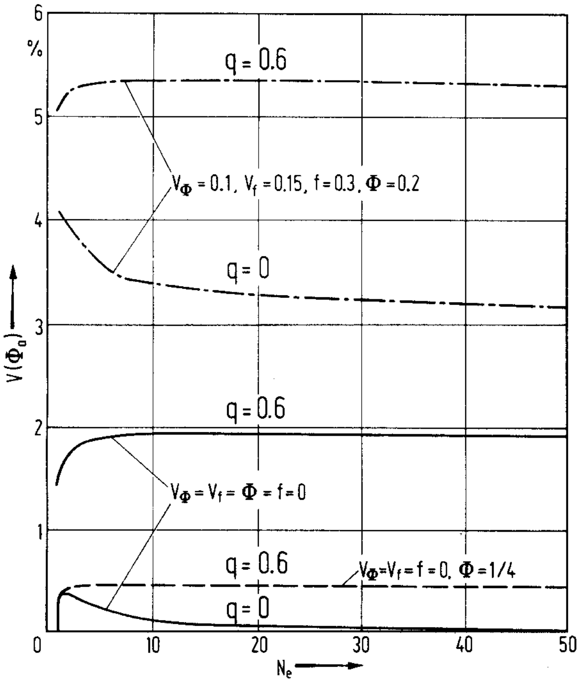


Fig. 5. The variance of the coefficient of kinship Φ_a of two seeds produced from the same mother plant as a function of the effective size of neighbourhood N_e

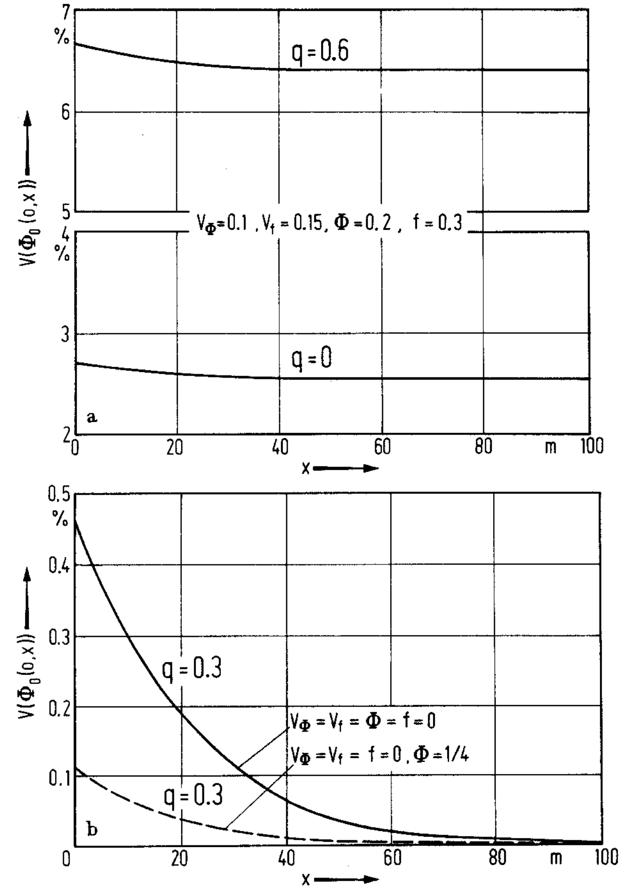


Fig. 6. The variance of the coefficient of kinship $\Phi_b(0,x)$ between two seeds produced from two different mother plants separated by a distance x as a function of this distance. The representations are based on a one-dimensional habitat, exponential pollen-distribution with range D = 50 m and population-density $\mu = 0.2\,m^{-1}$

depend solely on the effective neighbourhood-size $N_e(x)$, and $E(\Phi_a(x))$ is a decreasing function of this. An important aspect in this case is that the same size for $N_e(x)$ can be obtained from basically different conditions concerning e.g. dimensionality of the habitat, population-density or shape of pollen-dispersal. For this reason we shall give a graphical representation of $V(\Phi_a(x))$ as a function of $N_e(x)$ (see Fig. 5).

It can be seen from this figure that the variance of Φ_a very rapidly attains an approximately constant value with increasing effective size of neighbourhood. This tells us that the range of pollen-dispersal, as well as population density, is of minor importance for the heterogeneity in relationship between offspring from the same parent. On the other hand the graphs confirm the strong dependence of the variance on the rate of self-fertilization, as stated before. Proceeding from investigating expectations and variances for Φ_a to those for Φ_b, an entirely new aspect is introduced in so far as distances between plants have to be considered. This new aspect is taken into account with the help of the quantities $N_e(x,y)$, i.e. the effective size of common neighbourhood, and $\frac{p(x|y)}{\mu(x)}$, $\frac{p(y|x)}{\mu(y)}$, which are the probabilities that one individual mates with the other. Because, in general $N_e(x,y)$ does not determine the probabilities $\frac{p(x|y)}{\mu(x)}$ and $\frac{p(y|x)}{\mu(y)}$ completely, it is impossible to regard expectations and variances for Φ_b just as functions of effective size. Thus population-density, dimensionality of the habitat and shape of pollen-dispersal have to be considered explicitely. On the other hand, the great majority of real situations show a common tendency for $N_e(x,y)$, $\frac{p(x|y)}{\mu(x)}$ and $\frac{p(y|x)}{\mu(y)}$ to decrease with increasing distance between the places x and y, which then implies a decline in $E(\Phi_b)$. To demonstrate this situation, we shall give some examples based on the same assumptions which were made in the previous paper (Gregorius 1975), i.e. one-dimensional habitat, homogeneous population-density ($\mu(x) \equiv \mu$) and exponential probability-density for the pollen-distribution ($p(x|y) = \frac{3}{2D} \cdot \exp(-\frac{3}{D} \cdot |x-y|)$, where D can be conceived as a measure for the range of pollen-dispersal). The following Figs.6a and 6b show $V(\Phi_b(x,y))$ as a function of the distance $|x-y|$ between the locations x and y of the two mother-plants.

The conclusions we derived from Fig.5, concerning the extent to which the effective size of neighbourhood N_e influences the variance of Φ_a, obviously apply also to the influence of distance between mother plants (and thus the effective size of common neighbourhood) on the variance of Φ_b, when looking at the different cases presented in Fig.6. Again the dominating quantity is the rate of self-fertilization. In general, if there is no common neighbourhood for two plants located at places x and y, i.e. $N_e(x,y) = 0$ (and $p(x|y) = p(y|x) = 0$), the expectation and variance for $\Phi_b(x,y)$ is the same as in the case of a hypothetically infinite population, where cross-fertilization is at random, because in both cases $K(x,y) = p(x|y) = p(y|x) = 0$.

Literature

Gregorius, H.-R.: A model for genetic relationship among offspring from open-pollinated plant populations. Theor. Appl. Genet. (1975, in press)

Leviandier, Th.; Jacquard, A.: Homogénéisation de l'apparentement dans une population limitée. Ann. Génét. 17, (2), 99-103 (1974)

Malécot, G.: Les mathématiques de l'hérédité, Vol. 1. Paris: Masson et Cie. 1948

Wright, S.: Isolation by distance under diverse systems of mating. Genetics 31, 39-59 (1946)

Received January 23, 1975
Communicated by H. Stubbe

Dr. Hans-Rolf Gregorius
Lehrstuhl für Forstgenetik und Forstpflanzenzüchtung der Universität Göttingen
Bäsgenweg 2
D-34 Göttingen-Weende (Germany/BRD)