# Inbreeding depression and heterosis :
# Expected means and variances among inbred lines
# and their crosses

C. CHEVALET et M. GILLOIS

*Laboratoire de Génétique cellulaire,*
*Centre de Recherches de Toulouse, I.N.R.A.,*
*B.P. 12, 31320 Castanet-Tolosan, France*

## Summary

Complementary phenomena of heterosis and inbreeding depression are investigated from a theoretical point of view in an experiment involving several inbred lines originating from the same population. It is assumed that dominance interactions exist between gene effects, contributing to a quantitative character. A general theory on the evolution of genetic variances within and between subdivisions of the population is developed. We base ourselves on the theory and on the calculation of identity coefficients. In the absence of selection, the evolution of variances is determined only through the knowledge of genetic components typical of the quantitative character, and of the genetic effective number of inbred lines. The given results generalize and consolidate the former studies. The main possible evolutions are discussed by means of numerical illustrations. The population structure differs from the one described by WRIGHT in the case without dominance, depending on the importance of the dominance interaction and on the initial frequencies of alleles.

Two main phenomena are expected:

(*i*) Preservation of temporary increase in the mean genetic variance within lines is possible in cases more general than those that are already known; the immediate and constant reduction of this variance in inbred lines does not constitute a widespread phenomenon.

(*ii*) The genetic variance in populations $F_1$ obtained by interbreeding inbred lines may be considerably lower than the mean genetic variance in parental populations. Thus the reduction in variance experimentally observed in $F_1$ must not always be attributed entirely to a reduction in the environmental variance.

Furthermore, the general results are illustrated by investigating crossings of pure lines of autogamous vegetable species.

## I. — Introduction :
## Heterosis and inbreeding depression phenomena

Many experiments led to the definition of the relations hip between the increase in the inbreeding coefficient and the quantitative performance. FALCONER (1960) described their essential characteristics. This phenotypic mean depression

distinguishes itself by its linearity as a function of the inbreeding coefficient F, in the (0-0.60) interval. In mammals, the essential characteristic of which is the period of gestation, performances such as the size of the litter, the viability of the offspring and their weight before weaning are dependent both on the mother's and their own inbreeding coefficient (DICKERSON *et al.*, 1954).

Heterosis corresponds to an observation initially made on the plant Kingdom: this expression conveys the superior value of the phenotypic vigour of plants born of parents of very distinct lines. Genetically distinct lines can be obtained through genetic drift from the same initial stock. ROBERTS, whose works have been quoted by FALCONER (1960), has reared thirty inbred lines of mice that showed an inbreeding depression in the size of the litter. Next, he proceeded to random matings between lines. In these matings the size of the litter reached the initial level of the original population. Presented along these lines, heterosis appears as the restoration of the loss caused by inbreeding, within one generation. Similar results have been reported in pigs by SELLIER (1970). Revealed in terms of means, this phenomenon is also accompanied by an evolution in the phenotypic and genotypic variances of populations. Generally it is agreed that inbreeding in many lines born of the same population increases the genetic homogeneity or uniformity within lines as well as the genetic variance or differentiation between lines. A very important aspect from the biological and the breeding point of view is the increase in the environmental variance in absolute and relative value, with the increase in inbreeding and the reduction in this variance with the rise of heterosis (FALCONER, 1960).

To interpret these phenomena the simplest classical quantitative genetics model is necessary; the gene effect involved in the expression of a character shows dominance interactions but no epistasis. The evolution of means will be reported, and possible evolutions of different genetic components of the variance will be described, but with this simple model it will not be possible to explain the increase of the environmental variance value in inbred lines, nor its reduction in hybrids.

# II. — Genetic model of interpretation

## A. — *The ideal population*

The above mentioned experimental results have been obtained from populations the structures of which can be idealized. It is assumed that a large population in which mating is panmictic, is available initially. This panmictic population bred a large number of subpopulations or " lines " panmictic themselves but not very large. This subdivision comprises situations such as geographic or ecological isolation or situations controlled by laboratory or domestic animal breeders.

The large population will be referred to in terms of " original population ", the small ones will be referred to in terms of " lines ". In each one of these small populations or lines, the genic frequencies are submitted to a drift mechanism, to which is attributed the essential part in revealing heterosis experimentally.

The simplest specific conditions governing this ideal population are:

1. Matings are achieved exclusively between individuals of the same line. Consequently, they are isolated. Migration is excluded from the model.

2. Generations are distinct and do not overlap.

3. The number of animals kept for breeding purposes in each line is the same for all lines and generations.

4. Within each line matings are panmictic.

5. There is no selection. Mutations are not taken into account.

6. For a pair of parents the number of offspring kept for breeding purposes for the following generation is determined by the law of Poisson.

## B. — Gene Effect Model

To explain the heterosis phenomenon, we shall restrict ourvelves to a particular model in which the genes present only dominance interactions. According to the breaking down introduced by FISHER (1918) a genotypic random variable $Z_\alpha$ is attached to each locus ($\alpha$) contributing to the character and made up of the sum of additive effects $X'_\alpha$ and $X''_\alpha$ of genes, and of the residue of dominance $D_\alpha$. Without epistasis, the total random genotypic variable Z is the sum:

$$Z = \sum_\alpha X'_\alpha + \sum_\alpha X''_\alpha + \sum_\alpha D_\alpha$$

Then, the genetic variance in the original population comprises one part of " additive genetic variance ", $2\,VAR_0(X) = V_A$, and one part of " dominance genetic variance ", $VAR_0(D) = V_D$. Whenever, we find it necessary to describe a particular line with greater precision we consider a character governed by a single biallelic locus, the numerical contributions of the three genotypes $(aa)$, $(aA)$ and $(AA)$ being written respectively:

$$2u, \ u + d \ \text{and} \ 0$$

$q$ referring to the allele $(a)$ frequency, $p$ to that of $(A)$.

The first approach in which the character is described in general terms by means of its statistical parameters, enables us to obtain straightaway results on an average, through identity, inbreeding and kinship coefficients. The second approach which is less general, enables us to understand what is taking place " locally " when two given lines are crossed; its development is based on the consideration of moments of orders 2, 3 and 4 of the genic frequency $q$ within a line. On this occasion, we must recall the very close relationship that exists between these genic frequency moments and identity coefficients.

## C. — Summary and notations

In finite populations, the statistical description of a quantitative character refers not only to the initial components $V_A$ and $V_D$ (GILLOIS, 1964, 1965, 1966a; HARRIS, 1964), but also to five genetic parameters and to the identity coefficients

defined by Gillois (1964, 1965). Thus, the genotypic expectation, variance and covariance in an inbred population are written:

$$(1) \qquad E(Z_I) = (1 - f_I)E_0(Z) + f_I E_I(Z)$$
$$= E_0(Z) + f_I \mathfrak{D}$$

$$(2) \qquad VAR(Z_I) = (1 + f_I)V_A + (1 - f_I)V_D$$
$$+ f_I V_D + f_I C'_{AD} + f_I(1 - f_I)D^2$$

$$(3) \qquad COV(Z_I Z_J) = 2\phi V_A + (\delta_9 + \delta_{12})V_D$$
$$+ \delta_1 V_D + \left(\delta_1 + \frac{1}{4}(\delta_2 + \delta_3 + \delta_4 + \delta_5)\right)C'_{AD}$$
$$+ (\delta_1 + \delta_6 - f_I f_J)D^2$$

$f_I$ stands for the inbreeding coefficient of the individual I, $\phi$ and $\delta_i$ being respectively the kinship and the $i^{th}$ identity coefficients of the pair of individuals (I, J). The significance of variance components introduced is defined in Table 1. Ep and Ec, Gillois's original notations, have been replaced here by $E_0$ and $E_1$; in more general terms $E_f$ stands for an expectation taken under the condition in which the inbreeding coefficient equals $f$.

Mathematical expectations referred to by symbol E are taken with respect to the distribution of frequencies $q$ due to the phenomenon of genetic drift. The means represented by the symbol M stand, strictly speaking, for mathematical expectations E conditioned by the value of the genic frequency $q$: if for a given line $(i)$ and an instant $(t)$, $q_{it}$ refers to the random variable which under these conditions takes the frequency of allele $(a)$ as a value and $Z_{it}$ refers to the random variable which takes as value the genotypic value of an individual taken at random from this line $(i)$ at the instant $(t)$, the expression $E(Z_{it}/q_{it})$ is replaced by the symbol $M_{it}(Z)$. This notation can be justified since this conditional expectation depends only on the frequency $q_{it}$ which characterizes line $(i)$ at the instant $(t)$.

When crossing two lines $(i)$ and $(j)$, heterosis is defined as the difference between the genotypic level in the interbred population $(i \times j$ subscript) and the mean of genotypic levels of parental lines. We may write:

$$H_{i \times j} = M_{i \times j}(Z) - \frac{1}{2}(M_i(Z) + M_j(Z))$$

Therefore, it is also an expectation conditioned by the knowledge of frequencies $q_i$ and $q_j$ in parental lines.

Finally, the time subscript has not been mentioned: it is implied therefore, that all quantities are relative to the same instant, determined by the rank held by the generation with respect to the original population, or in a similar fashion, by the value of the inbreeding coefficient F calculated with regard to this original population.

## III. — Evolution of means

A. — *On an Average Calculation of the Inbreeding Depression and of heterosis*

$A_1$. — *Inbreeding depression*

For a line, the inbreeding coefficient of which is $f$, the genotypic mean expectation is (formula 1):

$$E_f(Z) = E_0(Z) + f \mathfrak{D}$$

With respect to the original population the mean of which is $E_0(Z)$, the mean inbreeding depression for lines having the same inbreeding coefficient $f$ is written as follows.

$$\text{(4)} \qquad\qquad E_f(Z) - E_0(Z) = f \, \mathfrak{D}$$

In the particular biallelic model, this depression is written:

$$\text{(5)} \qquad\qquad E_f(Z) - E_0(Z) = -2d p_0 q_0 f$$

The inbreeding depression depends on the dominance, and on the inbreeding level. Between this depression and the inbreeding coefficient $f$, the relation is linear as long as it can be assumed that the average of genic frequencies of all lines represents the average of the original population (no allele loss for the population on the whole, composed of all lines, and expressed by the term $p_0 q_0$ in formula (5)).

$A_2$. — *Heterosis*

If two lines, with inbreeding coefficients $f_i$ and $f_j$ are interbred to obtain $F_1$ individuals, the gametes that unite in these crossed zygotes, represent independent drawings from the original population. Then, we can write straightaway that the expectation of the random genotypic variable of an individual $F_1$ is equal to that of an individual of the original population, that is $E_0(Z)$. The expectation of heterosis $H_{ij}$ associated with the crossing of two lines is therefore:

$$E(H_{ij}) = E_0(Z) - \frac{1}{2}\left(E_{f_i}(Z) + E_{f_j}(Z)\right)$$

$$\text{(6)} \qquad\qquad E(H_{ij}) = -\frac{f_i + f_j}{2} \, \mathfrak{D}$$

When two lines having the same inbreeding coefficient $f$, are interbred, heterosis corresponds exactly to the recovery of the initial level from which we had departed because of the genetic drift in isolated lines. It is important to emphasize that this result holds only on an average, on the whole for all $F_1$ matings that are possible with the available lines.

B. — *Calculation of the Inbreeding Depression
and of Heterosis in Terms of Genic Frequencies*

$B_1$. — *Inbreeding depression*

Let a line $(i)$ be described by the frequency $q_i$ of allele $(a)$, and the frequencies of genotypes $(aa)$, $(aA)$ and $(AA)$ respectively defined by:

$$P = q_i^2; \qquad Q = 2 q_i p_i; \qquad R = p_i^2$$

In doing so, we consider in fact a large population showing the same frequency $q_i$ as in the line $(i)$, and being in Hardy-Weinberg equilibrium, which could hold only approximately in the line of finite size. Then, the character mean of line $(i)$ is written:

$$M_i(Z) = 2u q_i + 2d q_i(1 - q_i)$$

The inbreeding depression is therefore:

(7) $$M_i(Z) - E_0(Z) = 2(u + d) (q_i - q_0) - 2d(q^2_i - q^2_0)$$

Without selection between lines, the expectation of $q_i$ is still $q_0$, so that the inbreeding depression has an expectation of:

(8) $$E(M_i(Z)) - E_0(Z) = - 2d \, VAR \, (q_i)$$

## $B_2$. — *Heterosis*

The individuals $F_1$ obtained by crossing individuals from lines $(i)$ and $(j)$, in which the frequencies of $(a)$ are $q_i$ and $q_j$ have the following approximate genotypic constitution:

$$P = q_i p_j; \qquad Q = q_i(1 - q_j) + q_j(1 - q_i); \qquad etc...$$

The genotypic mean of population $F_1$ is therefore written:

$$M_{i \times j}(Z) = 2u(q_i + q_j)/2 + 2d((q_i + q_j)/2 - q_i q_j)$$

and heterosis, calculated with respect to the average of parental lines, is:

(9) $$H_{ij} = d(q_i - q_j)^2$$

Heterosis appears therefore as soon as the lines show different genic frequencies, at the locus under consideration. It will be maximum, in absolute value, if the difference $|q_i - q_j|$ is maximum. The expectation of the previous quantity can be written:

$$E(H_{ij}) = d \, (VAR \, (q_i) + VAR \, (q_j)),$$

and as in paragraph A, we find that the heterosis expectation is equal to the contrary of the inbreeding depression.

## $B_3$. — *Conclusion*

Thanks to a simple genetic model, the basic conclusions revealed in the analysis of the experimental results recalled in section I, have been rediscovered. The analysis of this particular model shows that heterosis is the more important as the two interbred lines are different, the extreme case being represented here by the interbreeding of two lines one of which had drifted towards the homozygous state $(aa)$ and the other towards the homozygous state (AA).

A comparison between the average results directly obtained (paragraph A) and the average results inferred from a " local " analysis (paragraph B) lay stress on basic fact: a comparison especially between formulas (5) and (8), leads to the relation:

$$VAR \, (q_i) = p_0 q_0 / i,$$

that holds for any fixed instant $(t)$ and that establishes a link between the variance of the genic frequency on the one hand, and the inbreeding coefficient, on the other. As a rule, for a panmictic population, we can write:

(10) $$VAR \, (q) = p_0 q_0 F$$

where F is function of kinship and inbreeding coefficients. This formula proven in a classical manner for a hermaphroditic population (see CROW and KIMURA, 1970, chapter 7) may be generalized (COCKERHAM, 1969).

This equivalence between the concepts of genic frequency variance and that of inbreeding coefficient is a general property that can be stated as follows: any moment of order $k$ of the genic frequency $q$ can be expressed in terms of the initial frequency $q_0$ and of identity coefficients defined for $k$ homologous genes taken at random from a population in which the frequency $q$ is calculated (MALECOT, 1969a). It is precisely this property that makes it possible to explain the genotypic variance within inbred lines, for non-additive characters, with respect to a certain number of parameters depending on $q_0$ and on identity coefficient of " orders " 2, 3, and 4. Indeed, for a character with dominance, the expression of the genetic variance in a line relies on the third and fourth powers of the frequency $q$, and the computation of its expectation requires the use of identity coefficients of order 2, 3 and 4.

## IV. — Evolution of variances

Except in the particular case of an additive character—which presents no interest in heterosis—the question of variance distribution within and between lines has not a simple, nor a unique answer. The problem is set with the biallelic model; next, the general method based on statistical considerations is presented in the additive case. Despite the fact that formulas have been simplified (a simplification that holds for lines with a sufficiently large genetic effective number ($Ne > 16$)), they remain abstract and we think it useful to illustrate them by designing imaginary quantitative characters in which genes show dominance and overdominance interactions.

### A. — *Inbreeding phase*

$A_1$. — *Biallelic model*

Let a line ($i$) be characterized at an instant ($t$) by the frequency $q_i$ of allele ($a$). Assuming the law of Hardy-Weinberg, we define:

— the genotypic mean:
$$M_i(Z) = 2uq_i + 2dq_ip_i,$$

— the mean of the genotypic squares:
$$M_i(Z^2) = 4u^2q^2_i + (u + d)^2 2p_iq_i,$$

— and the genotypic variance in this line:
$$V_{wi} = M_i(Z^2) — (M_i(Z))^2.$$

We intend to study $V_{wi}$, or rather its *expectation* $V_w$, that is, roughly the observable mean of quantities $V_{wi}$, if many equivalent lines are available. Meanwhile, we are led to introduce the overall variance of genotypes, calculated on all individuals of all lines, that is $V_t$. Finally, $V_b = V_t — V_w$ will refer to the variance of the means $M_i(Z)$.

Initially, $V_b = 0$, there is only one variance parameter, $V_0$:

$$V_0 = 2u^2 p_0 q_0 + 4ud p_0 q_0 (1 - 2q_0) + 2d^2 p_0 q_0 (1 - 2p_0 q_0),$$

(table 1). $V_{wi}$ has the same expression, substituting $q_i$, the present frequency, for $q_0$, and the expectations $V_w$ and $V_b$ can be written:

$$V_w = 2u^2 \mathrm{E}(p_i q_i) + 4ud\mathrm{E}\{p_i q_i (1 - 2q_i)\}$$
$$\quad + 2d^2 \mathrm{E}\{p_i q_i (1 - 2p_i q_i)\}$$
$$V_b = 4u^2 \mathrm{E}(q_i^2) + 4d^2 \mathrm{E}\{q_i^2 (1 - q_i)^2\} + 8ud\mathrm{E}\{q_i^2 (1 - q_i)\}$$
$$\quad - \{2uq_0 + 2d p_0 q_0 (1 - \mathrm{F})\}^2$$

according to the definition of F (formula (10), paragraph III B$_3$). Therefore, $V_w$ and $V_b$ require the computation of moments of order 3 and 4 of frequency $q_i$. Their behaviour as time goes on depends also on the ratio $d/u$. The sum $V_t$, on the other hand, is expressed directly:

$$V_t = \mathrm{E}\{M_i(Z^2)\} - \{\mathrm{E}(M_i(Z))\}^2$$
$$\quad = (1 - \mathrm{F})V_0 + \mathrm{F}4u^2 p_0 q_0 + \mathrm{F}(1 - \mathrm{F})4d^2 p_0^2 q_0^2$$

$4u^2 p_0 q_0$ stands for the total variance of Z, when all the lines have reached fixation, without selection;
$4d^2 p_0^2 q_0^2 = (-2d p_0 q_0)^2$ stands for the square of the parameter which characterizes the inbreeding depression. This formula is formula (2), written as follows:

(2')          $\mathrm{VAR}(Z) = (1 - f)(V_A + V_D) + f(2V_A + V_D + C'_{AD}) + f(1 - f)D^2$

where $V_A + V_D = \mathrm{VAR}_0(Z)$ stands for the initial variance,
and $2V_A + V'_D + C'_{AD} = \mathrm{VAR}_1(Z)$ the total asymptotic variance.


A$_2$. — *A particular case: the character is additive*

Then we can calculate $V_b$ directly, for instance: having $d = 0$ in the expression of the former paragraph, it becomes:

$$V_b = 4u^2 \mathrm{E}(q_i^2) - 4u^2 q_0^2$$
$$\quad = 4u^2 p_0 q_0 \mathrm{F}$$

And since $V_0 = 2u^2 p_0 q_0$, we get:

(11)                    $V_w = (1 - \mathrm{F})V_0$
$$V_b = 2\mathrm{F}V_0$$
$$V_t = (1 + \mathrm{F})V_0$$

(WRIGHT, 1951, 1952.)
This same breaking down can be obtained starting out from a different point of view: given an inbred population (of two sexes), characterized by the mean inbreeding coefficient $f$ and the mean kinship coefficient $\phi$, and the initial variance $V_0$ of an additive character,

$(1 + f)V_0$    represents the a-priori variance of the genotype $Z_I$ of a random individual I;

$2\phi V_0$    represents the a-priori covariance between the random variable $Z_I$ and $Z_J$ of two distinct individuals taken from the same line.

In an experiment involving several equivalent lines, this covariance will also represent the variance of line means and we shall be able to write:

(11′)
$$V_t = (1 + f)V_0$$
$$V_b = 2\phi V_0$$
$$V_w = (1 + f - 2\phi)V_0$$

Thus, we get the same beaking as WRIGHT got (formulas 11) noting that $f$ and $\phi$ are not very different from the same value F. The difference is related to the fact that the break-down is obtained here directly from genotypic frequencies. On the contrary, expressions introducing a single parameter F are based on frequencies and probabilities attached to gametes, from where we proceed to zygotic frequencies by means of the law of Hardy-Weinberg. This statistical approach of the interpretation of a priori variance and covariance formulas, leads to a natural extension to the case of any character due to many loci, with any number of alleles per locus.

$A_3$. — *A general case with dominance*

Taking into account the previous consideration and expressions (2) and (3) the break-down of the total variance $V_t$ into intra-line variance $V_w$ and inter-line variance $V_b$, can be expressed:

(12)
$$V_t = (1 + f)V_A + (1 - f)V_D + fV'_D + fC'_{AD} + f(1 - f)D^2$$
$$+ \sigma_t^2(f)(\mathfrak{D}^2 - D^2)$$
$$V_w = (1 + f - 2\phi)V_A + (1 - f - \delta_9 - \delta_{12})V_D + (f - \delta_1)V'_D$$
$$+ \{f - \delta_1 - (\delta_2 + \delta_3 + \delta_4 + \delta_5)/4\}C'_{AD} + (f - \delta_1 - \delta_6)D^2$$
$$+ \sigma_w^2(f)(\mathfrak{D}^2 - D^2)$$
$$V_b = 2\phi V_A + (\delta_9 + \delta_{12})V_D + \delta_1 V'_D$$
$$+ \{\delta_1 + (\delta_2 + \delta_3 + \delta_4 + \delta_5)/4\}C'_{AD} + (\delta_1 + \delta_6 - f^2)D^2$$
$$+ \sigma_b^2(f)(\mathfrak{D}^2 - D^2)$$

In the foregoing expressions, coefficients assume their mean values within a line. Symbols $\sigma_t^2(f)$, $\sigma_w^2(f)$ and $\sigma_b^2(f)$ hold for the variances of individual inbreeding coefficients, over the whole population, within line, and between lines, respectively. It should be noted that the corresponding component $(\mathfrak{D}^2 - D^2)$ arises only when more than one locus contribute to the character. The hypotheses proposed in section 2 concerning the structure of lines make it possible to simplify these expressions in so far as the genetic effective size Ne is sufficiently large, larger than say 16. If this is the case, identity coefficients $\delta_i$ assume approximate expressions, functions only of parameter F, that under the same hypothesis is equal to $f$ and $\phi$. Four probabilities will be introduced here.

F : the probability that at the same locus, two gametes taken at random carry identical genes:

$$F \simeq 1 - \exp(-t/2Ne)$$

G : the probability that three gametes taken at random carry identical genes:

$$G \simeq 1 - \frac{3}{2}(1 - F) + \frac{1}{2}(1 - F)^3$$

H : the probability that four gametes taken at random carry four identical genes:

$$H \simeq 1 - \frac{9}{5}(1 - F) + (1 - F)^3 - \frac{1}{5}(1 - F)^6$$

K : the probability that, four distinct gametes being drawn at random, the four genes of the same locus are distributed into two independent groups, each made up of two identical genes:

$$K = 2(G - H)$$

(according to CHEVALET, GILLOIS, NASSAR, 1977).

Also, from the work of WEIR and COCKERHAM (1969), it turns out that the variances of inbreeding coefficients may be approximated as follow:

$$\sigma_t^2(f) \simeq \sigma_w^2(f) \simeq \frac{(1 - F)^2}{12 . Ne} \qquad \text{and} \qquad \sigma_b^2(f) \simeq 0$$

It will be found in the numerical illustrations (section 6) that the corresponding component may generally be neglected. Formulas (12) become:

$$(12')\quad V_t = (1 + F)V_A + (1 - F)V_D + FV'_D + FC'_{AD} + F(1 - F)D^2$$
$$+ \frac{(1 - F)^2}{12 . Ne}(\mathfrak{D}^2 - D^2)$$

$$V_b = 2FV_A + \frac{2}{3}KV_D + HV'_D + GC'_{AD} + \left(H + \frac{1}{3}K - F^2\right)D^2$$

$$V_w = (1 - F)V_A + \left(1 - F - \frac{2}{3}K\right)V_D + (F - H)V'_D + (F - G)C'_{AD}$$
$$+ \left(F - H - \frac{1}{3}K\right)D^2 + \frac{(1 - F)^2}{12 . Ne}(\mathfrak{D}^2 - D^2)$$

The behaviour of these variances with time can therefore be studied as soon as relative values for five genetic variance components are given; these results make it possible to generalize the work carried out by ROBERTSON (1952) in the case of one character governed by a very rare recessive gene.

To complete this break-down of the variance, it may be interesting to consider the " additive " part, in so far as this additive genetic fraction of the variance can be estimated in inbred lines. This break-down has been achieved by CROW and KIMURA in the particular case studied here (1970, formulas 7-5-20, 21, 22, page 343). Their extension to a general character with dominance characterized by parameters $V_A'$, $V_D'$, $V'_D$, $C'_{AD}$ and $D^2$ can be written as follows:

● Intra-line additive genetic variance:

$$(13)\quad V_{wA} = (1 - F)V_A + \left(2F - 2G - \frac{4}{3}K\right)V_D$$
$$+ 2(G - H)V'_D + (F - G)C'_{AD} + \left(2G - 2H - \frac{2}{3}K\right)D^2$$

● Genetic variance of intra-line dominance:

$$V_{wD} = \left(1 - 3F + 2G + \frac{2}{3}K\right)V_D + (F - 2G + H)V'_D$$
$$+ \left(F - 2G + H + \frac{1}{3}K\right)D^2 + \frac{(1-F)^2}{12.Ne}(\mathfrak{D}^2 - D^2)$$

A peculiar phenomenon has been revealed by ROBERTSON if, in our biallelic model, we have $(d = -u)$ and $(q \simeq 0)$: at the beginning of the inbreeding process, the intra-line variance increases at first, before finally decreasing towards 0. For a recessive gene this property persits if: $(6q^2_0 < 1)$ (see CROW, KIMURA, 1970).

As a general rule, it can be shown that the phenomenon appears as soon as the following relation holds:

(14)                    $$V_A + V_D < V'_D + C'_{AD} + D^2$$

basing ourselves on approached expressions of identity coefficients.

*Note;* Formulas (12) can be extended to any quantitative character, governed by genes in epistatic interactions, by introducing generalized identity coefficients (see GALLAIS, 1970).

## B. — *Mating Phase; Population $F_1$*

$B_1$. — *Biallelic model*

Two inbred lines are interbred. These lines are characterized by:

— the frequencies $q_i$ and $q_j$,
— the means $M_i(Z)$ and $M_j(Z)$,
— the variances $V_{wi}$ and $V_{wj}$.

The subscript $(i \times j)$ stands for the quantities related to the population $F_1$, in which the genotypes have the following frequencies:

$$P = q_i q_j; \qquad Q = q_i p_j + q_j p_i; \qquad R = p_i q_j$$

We have therefore:

$$M_{i \times j}(Z) = (u + d)(q_i + q_j) - 2dq_i q_j$$
$$M_{i \times j}(Z^2) = 4u^2 q_i q_j + (u + d)^2 (q_i p_j + p_i q_j)$$

and the genetic variance of population $F_1$ can be written:

$$V_{i \times j} = 4u^2 q_i q_j + (u + d)^2 (q_i p_j + p_i q_j)$$
$$- (u + d)^2 (q_i + q_j)^2 - 4d^2 q_i^2 q_j^2$$
$$+ 4d(u + d)(q_i + q_j)q_i q_j$$

Consequently, the expression of this quantity, refers only to the moments of order two of the frequency, because of the independence between $q_i$ and $q_j$. Using relation (10), we get:

$$E(V_{i \times j}) = (1 - F)V_0 + F(1 - F)4d^2 p_0^2 q_0^2$$

In the particular case of an additive character, the average of variances of populations $F_1$ does not differ from that of parental lines. This property holds " locally " since:

$$V_{i \times j} - \frac{1}{2}(V_{wi} + V_{wj}) = 2d(q_i - q_j)^2 (q_i + q_j - 1)^2 \{2u - d(q_i + q_j - 1)\}$$

cancels itself if $d = 0$. If $d$ is not equal to zero, this difference can take any signs, but it can be noted that if heterosis is large in the crossing between lines $(i)$ and $(j)$, $|q_i - q_j|$ is large; and this implies that $|q_i + q_j - 1|$ is small (since: $|q_i - q_j|$ $+ |q_i + q_j - 1| < 1$). The difference between former variances is close to zero. This fact shows that the reduction in the variance observed in $F_1$ is not of genetic nature. This observation that applies to the interbreeding of two very differentiated lines, does not on average and has no reason to hold if many lines and all crossings are available.

### $B_2$. — *A general Model*

Just like in the study in which the means were investigated (paragraph III, $A_2$) we note that two gametes, one of which has been drawn from a line $(i)$ and the other from a distinct line $(j)$ are surely " non-identical ". This is the same situation as the one studied by MALÉCOT in 1939, in which only one parameter is useful: coefficient F, and we get:

— a priori variance $V_{tF_1}$ of the genotype of a zygote taken at random from any crossing $F_1$:

$$(15.1) \qquad\qquad V_{tF_1} = V_A + V_D = V_0$$

— variance $V_{bF_1}$ of the means of all possible population $F_1$:

$$(15.2) \qquad\qquad V_{bF_1} = FV_A + F^2 V_D \qquad \text{(FALCONER, 1960, chapter 16)}$$

— genetic variance expectation in a population $F_1$, $V_{wF_1}$:

$$(15.3) \qquad\qquad V_{wF_1} = (1 - F)(V_A + V_D) + F(1 - F)V_D$$

It is not easy to compare these expressions with formulas (12) or (12') especially because of the components $V'_D$, $C'_{AD}$ and $D^2$ the relative values of which vary with the gene effect and with the initial frequencies of different alleles.

## V. — Lines derived from interbred populations

If several lines obtained through systematic experiments, or made up of animal samples of isolated races are available, many types of interbreeding can be considered. The first are of types $F_1$, and can reveal heterosis. Complex three or more way crossings, with several levels, etc., can be attempted to show new heterosis reactions. It is difficult to give a useful theoretical description of such phenomena. We shall only recall here the result concerning a line $F_2$ derived from interbreeding individuals belonging to the same interbred population $F_1$.

Then two essential phenomena appear: loss of half of the heterosis (for a

model with dominance without maternal or paternal type effects), and the reappearance of a part of the genetic variance.

If on an average, the loss of heterosis can be easily described, and if broadly speaking the components of the mean of different types of crossings can be analysed in complex situations (HAYMAN, MATHER, 1955; DICKERSON, 1969) the genetic variance reapparance mode is not easy to understand as soon as we depart from the purely additive case.

Again we consider two parental populations ($i$) and ($j$) having undergone identical evolutions in as far as probability is concerned, and having interbred at a moment when the frequencies of an allele ($a$) were respectively $q_i$ and $q_j$. To simplify, we assume that all gametes produced by the zygotes of the population $F_1$ have the same probability of carrying the allele a, equal to $(q_i + q_j)/2$. These gametes unite at random to form the zygotes $F_2$, then the closed line evolution proceeds normally from this population $F_2$ on. The forming of $F_2$ zygotes allows us to consider this $F_2$ generation as the first generation of a line born of an original population in which the initial frequency of ($a$) would be $(q_i + q_j)/2$. This fictitious initial population does not identify itself with that of $F_1$, since a panmictic structure with genotypic frequencies deduced through the law of Hardy-Weinberg, corresponding to the frequencies $q'_0 = (q_i + q_j)/2$, must be attributed to it. To this imaginary structure correspond an " initial " expectation $E'_0(Z)$, and an " initial " variance $V_0$, just as fictitious as the population. Although attributed to the frequency of ($a$) in population $F_1$, these quantities have no meaning for this interbred population merely because these genotypic frequencies are different. In this approach, heterosis in $F_1$ gives evidence of a non-random association of gametes in zygotes, and coming back to random association of panmixia, from $F_2$ on, shows a loss of information together with the disappearing of a part of heterosis.

## A. — Evolution of the Mean

In the biallelic model, with notations of paragraph III, B, writing $M^2_{i \times j}(Z)$ for the genotypic mean in $F_2$, and $N'_e$ for the genetic effective number of the new closed line $F_1$, $F_2$, etc., we get according to formula (9) and to the value $q'_0 = (q_i + q_j)/2$:

$$(9') \qquad M_{i \times j} = \{M_i(Z) + M_j(Z)\}/2 + d(q_i - q_j)^2$$

$$M^2_{i \times j} = \{M_i(Z) + M_j(Z)\}/2 + \frac{1}{2}d(q_i - q_j)^2 + \frac{1}{2N'_e}q'_0 p'_0 2d$$

The first two terms correspond to the starting levels of the initial fictitious populations (frequency $q'_0$), and the last term to the new inbreeding depression, $1/2N'_e$ representing the inbreeding coefficient of $F_2$ *with respect to* $F_1$. In the absence of any other effects, the mean level of line $F_2$, $F_3$, etc., must therefore decrease in expectation, from level:

$$E'_0 = \frac{1}{2}(M_i + M_j) + \frac{1}{2}H_{ij}$$

As we pass from $F_1$ to $F_2$, we lose half of heterosis, then the mean level of the line continues to decrease, gradually if the effective number $N'_e$ is sufficiently large.

The same result is obtained directly in the general case with dominance, by

using the fact that the mean inbreeding coefficient of individuals in $F_2$ is equal to $F/2$, F being the inbreeding coefficient reached in parental lines. P, $F_1$, $F_2$, being used as subscripts for expectations defined in parental populations $F_1$, $F_2$, respectively, we have:

$$E_P(Z) = E_0(Z) + F.\mathfrak{D}$$
$$E_{F_1}(Z) = E_0(Z) \qquad\qquad = E_P(Z) - F.\mathfrak{D}$$
$$E_{F_2}(Z) = E_0(Z) + \frac{F}{2}.E_1D = E\ (Z) - \frac{F}{2}.\mathfrak{D}$$

### B. — Evolution of the Variance

Just like in the expression for the variance in parental lines and in populations $F_1$, we refer to the identity coefficients. The quantities F, G, H, K are the same as before (paragraph IV, $A_3$) and they still depend on parental lines. With the same approximations, we get after several calculations, and ignoring the contributions due to the variance in inbreeding coefficients:

- a priori variance of the genotype of a zygote from a line $F_2$:

$$(16.1)\quad V_{tF_2} = \left(1 + \frac{F}{2}\right)V_A + \left(1 - \frac{F}{2}\right)V_D + \frac{F}{2}V'_D + \frac{F}{2}C'_{AD} + \frac{F}{2}\left(1 - \frac{F}{2}\right)D^2$$

- variance of the means of all possible $F_2$ populations:

$$(16.2)\qquad V_{bF_2} = FV_A + \left(\frac{1}{4}F^2 + \frac{1}{12}K\right)V_D + \frac{1}{8}HV'_D + \frac{1}{4}GC'_{AD}$$
$$+ \left(\frac{1}{8}H + \frac{1}{24}K - \frac{1}{8}F^2\right)D^2$$

- genetic variance in a population $F_2$ (in expectation):

$$(16.3)\qquad V_{wF_2} = \left(1 - \frac{F}{2}\right)V_A + \left(1 - \frac{F}{2} - \frac{F^2}{4} - \frac{K}{12}\right)V_D + \left(\frac{F}{2} - \frac{H}{8}\right)V'_D$$
$$+ \left(\frac{F}{2} - \frac{G}{4}\right)C'_{AD} + \left(\frac{F}{2} - \frac{F^2}{8} - \frac{H}{8} - \frac{K}{24}\right)D^2$$

Except in the additive case, in which we can see that in $F_2$ half of the genetic variance lost during the inbreeding phase in recovered, these formulas do not give obvious information concerning the behaviour of the variance. Only in several examples do these formulas become helpful (section 6).

The complexity of these formulas gives an insight into the difficulty of relating a heritability estimated in an interbred population to heritability coefficients estimated in parental populations, even if the changes undergone by the component of the phenotypic variance attributed to the environment are not taken into account. Only the particular case in which the parental lines have all reached fixation can be further analysed.

## C. — *Application to the case in which parental lines are homozygous*

The particular case in which parental lines have reached fixation is often considered in plant genetics, in autogamous species in which pure lines can be easily obtained. The analysis of genetic variability can be explained by having:

$$F = G = H = 1 \quad \text{and} \quad K = 0$$

in the preceding formulas, and interpreted in simple statistical terms.

Let us assume that $n$ pure lines were born of a basic population in which the quantitative character considered is characterized by the parameter $\mathfrak{D}$ of inbreeding depression and by its variance components $V_A$, $V_D$, $V'_D$, $C'_{AD}$, and $D^2$: the intra-line variance, $_0\sigma^2_w$, represents the environmental effect variance, that is $_0V_e$ in pure line, and the inter line variance, that is $_0\sigma^2_b$, represents the genetic variance.

$V_b = V_t = 2V_A + V'_D + C'_{AD}$, since there is no longer an intraline genetic variance ($V_w = 0$).

Then, we can consider the experiment that consists in carrying out all inter-breeding of type $F_1$ between these lines, next from each $i \times j$ crossings, subpopulations of type $F_2$ are created (simply through self-fecundation if the species is autogamous). In subpopulations $F_{1i \times j}$, a variance analysis can be carried out, according to a random model:

$$Y_{ijk} = \mu + \alpha_i + \alpha_j + (\alpha\alpha)_{ij} + \varepsilon_{ijk}$$

in which $\alpha_i$ stands for the main effect of the $i^{th}$ line, $(\alpha\alpha)_{ij}$ the interaction between $i^{th}$ and $j^{th}$ lines, and $\varepsilon_{ijk}$, the residual effect. This variance analysis makes it possible to estimate:

$$_1\sigma^2_\alpha = \text{variance of } \alpha_i \text{ effects,}$$
$$_1\sigma^2_\alpha = \text{variance of interaction effects } (\alpha\alpha)_{ij},$$
$$_1\sigma^2_\alpha = \text{residual variance.}$$

The same approach, carried out on subpopulations $F_{2i \times j}$ leads to the estimation of similar quantities: $_2\sigma^2_\alpha$, $_2\sigma^2_{\alpha\alpha}$ and $_2\sigma^2_\varepsilon$. The genetic interpretation of quantities $\sigma^2$, that can be estimated, is then as follows, taking into account formulas (2), (15) and (16):

$$_0\sigma^2_b = 2V_A + V'_D + C'_{AD}.$$
$$_0\sigma^2_w = {}_0V_e$$
$$_1\sigma^2_\alpha = \frac{1}{2}V_A$$
$$_1\sigma^2_{\alpha\alpha} = V_D$$
$$_1\sigma^2_\varepsilon = {}_1V_e$$
$$_2\sigma^2_\alpha = \frac{1}{2}V_A + \frac{1}{8}C'_{AD} + \frac{1}{16}V'_D$$
$$_2\sigma^2_{\alpha\alpha} = \frac{1}{4}V_D$$
$$_2\sigma^2_\varepsilon = \left(\frac{1}{2}V_A + \frac{1}{4}V'_D + \frac{1}{4}C'_{AD}\right) + \left(\frac{1}{8}V'_D + \frac{1}{4}V_D + \frac{1}{4}D^2\right) + {}_2V_e$$

In the expression of $_2\sigma^2_\varepsilon$, the break-down enclosed in the parentheses corresponds to the break-down of the genetic variance in subpopulations $F_2$ into " additive

genetic variance " and " genetic variance of dominance ". $_2V_e$ stands for the variance of environmental effects in the generation $F_2$.

These formulas call for the following comments:

1) The additive genetic variance in subpopulation $F_2$ is equal to one quarter of the variance between parental lines. This is a case in which a " heritability " in crossing can be associated with the variability of parental lines. We can write:

$$h^2_{F_2} = \frac{1}{4} \frac{_0\sigma^2_b}{_2\sigma^2_\varepsilon}$$

as for an additive character.

We must note, however, that such a " heritability " presents only a limited practical interest for at least two reasons: the genetic variance in the subpopulations $F_2$ is calculated on an average, and therefore it has no significance for a particular subpopulation $F_{2ij}$; the experiment described, that assumes the existence of pure lines can be achieved in practice only on autogamous plants in which heritability has no predictive value in so far as selection is concerned.

2) If the environmental variance $_2V_e$ in $F_2$ can be correctly evaluated from environmental variances $_1V_e$ and $_0V_e$ or from other experiments, the experimental diagram described allows the evaluation of the genetic components introduced to explain the estimated variances $\sigma^2$. At least for the characters which are assumed not to be subject to epistatic interactions, this diagram provides a method to estimate the variance components $V'_{D'}$, $C'_{AD'}$, and $D^2$ related to inbreeding.

3) The introduction of dominance effects differentiates the genetic interpretations of quantities $_1\sigma^2_\alpha$ and $_2\sigma^2_\alpha$ that can be estimated. The existence of negative $C'_{AD}$ quantities particularly, can account for a reduction in variances $\sigma^2_\alpha$ between $F_1$ and $F_2$ generations.


## VI. — Numerical illustrations


Even though within the additive model the evolution of variances follows a simple rule (formulas (11) developed by WRIGHT), it is difficult to get a general idea of the phenomenon as soon as dominance is introduced. Here we present several curves in order to illustrate formulas (12) and (13), and a corresponding table, of the redistribution of the variance in interbred populations $F_1$ and $F_2$ (formulas (15) and (16)).

Four quantitative characters are considered:

(a) Character governed by four triallelic loci in which alleles $A_1$, $A_2$, $A_3$ have the genotypic effects defined in the following table:

|        | $A_1$ | $A_2$ | $A_3$ |
|--------|-------|-------|-------|
| $A_1$ . . | 2 | 2 | 2 |
| $A_2$ . . | 2 | 1 | 1 |
| $A_3$ . . | 2 | 1 | 0 |

($A_1$ dominant over $A_2$ and $A_3$, $A_2$ dominant over $A_3$).    The initial frequencies of these different alleles at the four loci, are:

|         | $A_1$ | $A_2$ | $A_3$ |
|---------|-------|-------|-------|
| Locus 1 | 1/2   | 1/4   | 1/4   |
| Locus 2 | 1/4   | 1/2   | 1/4   |
| Locus 3 | 1/4   | 1/4   | 1/2   |
| Locus 4 | 1/3   | 1/3   | 1/3   |

After the initial variance of the original population is reduced to one, the components are written:

$$V_A = 0.7477 \qquad V_D = 0.2523 \qquad \mathfrak{D} = -1.256$$
$$V'_D = 0.0993 \qquad C'_{AD} = -0.1992 \qquad D^2 = 0.3959$$

(b) Character governed by only one biallelic locus, in which the contributions of three genotypes are 1, 0, 0 for (aa), (aA) and (AA) respectively, the initial frequency of (a) being equal to 0.1, as it has been presented by ROBERTSON in 1952. The values of components are:

$$V_A = 0.1818 \qquad V_D = 0.8182 \qquad \mathfrak{D} = -0.9045$$
$$V'_D = 5.8182 \qquad C'_{AD} = 2.9091 \qquad D^2 = 0.8182$$

(c) Character governed only by locus number 1 of the preceding character (a), with the same effects and the same initial frequencies.    The case has been considered because it is an extension of the previous case (b), in the sense that the relation (14) is also proven by its parameters, which are:

$$V_A = 0.6207 \qquad V_D = 0.3793 \qquad \mathfrak{D} = -0.7505$$
$$V'_D = 0.1379 \qquad C'_{AD} = 0.6437 \qquad D^2 = 0.5632$$

(d) Character governed by four triallelic loci the three alleles of which have the following contributions:

|         | $A_1$ | $A_2$ | $A_3$ |
|---------|-------|-------|-------|
| $A_1$ . . | 1   | 2   | 2   |
| $A_2$ . . | 2   | 0.5 | 2   |
| $A_3$ . . | 2   | 2   | 0   |

(advantage given to heterozygotes).    The initial frequencies at four loci are respectively identical with those of homonymous alleles of character (a).    After reducing the initial variance to one the components are:

$$V_A = 0.1601 \qquad V_D = 0.8399 \qquad \mathfrak{D} = -2.510$$
$$V'_D = 0.2143 \qquad C'_{AD} = -0.2592 \qquad D^2 = 1.5765$$

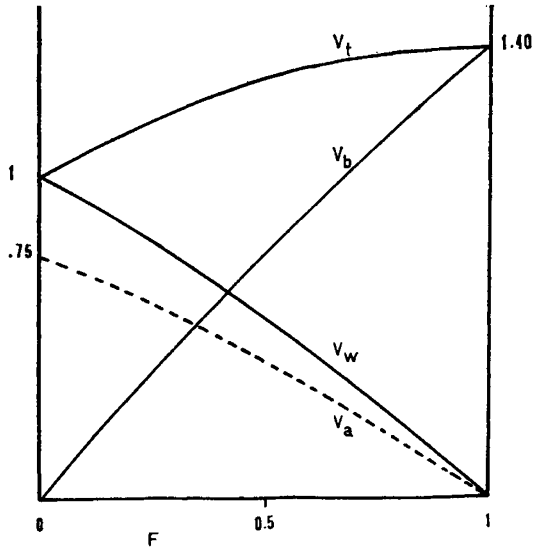The results are shown in figure 1 and in Table 2.
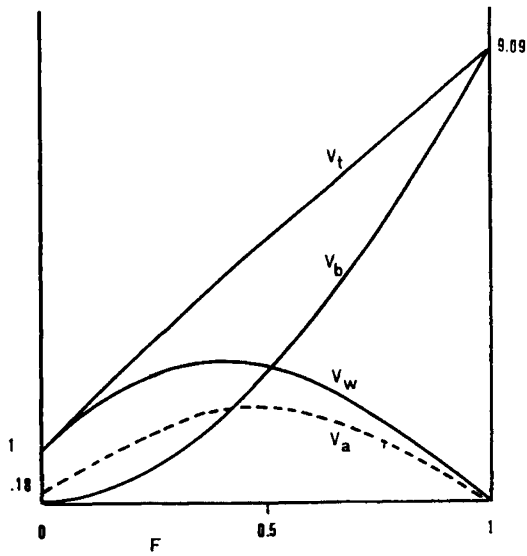
FIG. 1a



FIG. 1b

FIG. 1. — *Change with time of the expected genetic variances within and between subdivisions of a population. Figures 1a to 1d refer to the four characters defined in section 6.*

The meanings of symbols are:

$V_t$ : overall genetic variance,

$V_w$ : expectation of the within line genetic variance,

$V_a$ : expectation of the additive genetic variance within lines,

$V_b$ : genetic variance between line means,

F  : stands for the mean inbreeding coefficient of the lines, and is taken as the measure of time.
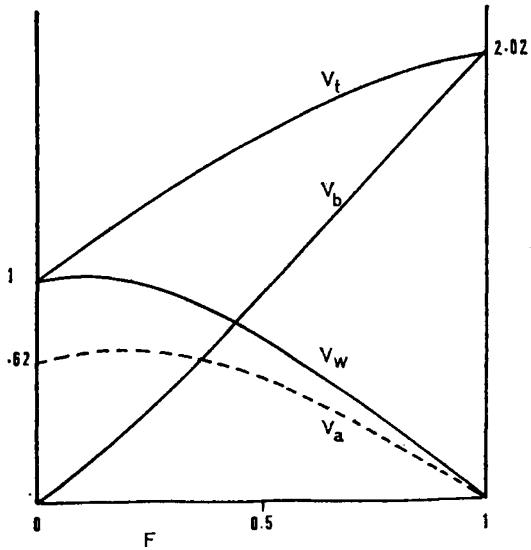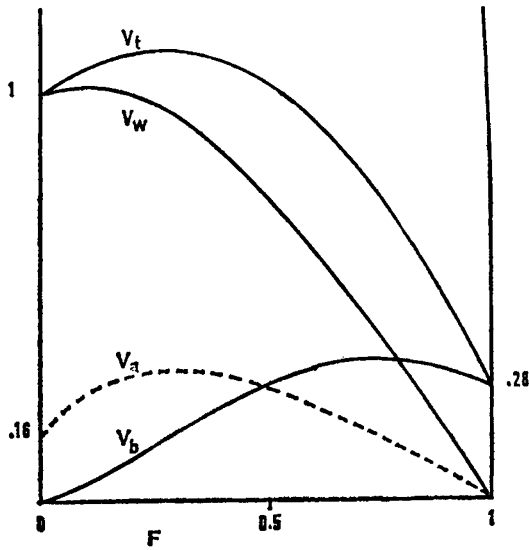
FIG. 1c

FIG. 1d

*Évolution de la variance génétique dans et entre les subdivisions d'une population.*
*Les figures 1a à 1d se réfèrent aux quatre caractères définis en section 6.*

$V_t$ : variance génétique totale,
$V_w$ : espérance de la variance génétique intra-lignée,
$V_a$ : espérance de la variance génétique additive intra-lignée,
$V_b$ : variance génétique entre lignées,
$F$ : coefficient moyen de consanguinité dans les lignées, à l'instant $t$ défini par :
$F = 1 - \exp(-t/2N_e)$.

## A. — *Break-down into intra- and inter-line variance* (fig. 1)

For character ($a$), in which only pure dominance interactions were introduced, and in which intermediate frequencies of alleles were considered, the general evolution of variances is very similar to that observed in the additive case, with only a marked reduction in the total asymptotic variance, which is $1.4 \, V_0$ here and not $2V_0$. The fraction of the variance attributed to additive effects $V_{wA}/V_{w'}$ remains relatively constant.

For the other three characters considered, an initial increase in the genetic variance of lines is observed, as well as an increase in the additive genetic variance.

TABLE I

*The statistical parameters describing a quantitative character*
*(Components of the mean and of the variance)*
*Paramètres statistiques décrivant un caractère quantitatif*
*(Composantes de la moyenne et de la variance)*

| Significance for one locus | Notation for one locus ($\alpha$) | Notation of the sums of terms related to several loci | Biallelic model |
|---|---|---|---|
| Genotypic mean in the original population ($F = 0$) | $E_0(Z_\alpha)$ | $E_0(Z)$ | $2uq_0 + 2dp_0q_0$ |
| Genotypic mean calculated in identity condition ($F = 1$) . . . . . . . | $E_1(Z_\alpha)$ | $E_1(Z)$ | $2uq_0$ |
| Mean of the dominance deviation computed in identity condition, i.e. parameter of inbreeding depression . . . . . . . | $E_1(D_\alpha) = E_1(Z_\alpha)$ $- E_0(Z_\alpha) = \mathfrak{D}$ | $E_1(D) = \mathfrak{D}$ | $- 2dp_0q_0$ |
| Additive genetic variance in the original population. . | $2 \cdot VAR(X_\alpha)$ | $V_A$ | $2p_0q_0(u + (1 - 2q_0)d)^2$ |
| Dominance genetic variance in the original population | $VAR_0(D_\alpha)$ | $V_D$ | $4d^2p^2_0q_0{}^2$ |
| Dominance genetic variance computed in identity condition . . . . . . . | $VAR_1(D_\alpha)$ | $V'_D$ | $4d^2p_0q_0(1 - 2q_0)^2$ |
| Covariance between additive effects and dominance deviation, computed in identity conditions . . . . | $4 \cdot COV_1(X_\alpha D_\alpha)$ | $C'_{AD}$ | $- 8dp_0q_0(1 - 2q_0)$ $(u + d(1 - 2q_0))$ |
| Square of the parameter of inbreeding depression (equal to $VAR_0(D)$ in the biallelic case) . . . . . . | $(E_1(D_\alpha))^2$ | $D^2$ | $4d^2p_0{}^2q_0{}^2$ |

The phenomenon can last a long time (case $(b)$) or it can only be a practically temporary maintenance of the initial variance (case $(c)$). The most noteworthly of these increases seems to be that of the additive genetic variance. This increase in the genetic variance $V_w$ or $V_{wA}$, is not uniform in all lines. $V_w$ is indeed an expectation and its increase originates from the fact that in a few lines $(j)$ the frequency of some genes, scarce at first will increase and the genetic variance $V_{wj}$ will become very large with respect to the initial variance, whereas, in most lines $(i)$ these genes will disappear and variances $V_{wi}$ will become very small.

A remarkable thing happens to character $(d)$: the final reduction in the total variance $V_t$. This is due to the fact that in the initial population most of the variance is due to the difference between the performances of homozygotes on the one hand, and those of the heterozygotes on the other. At the end of the process, only homozygous populations are left which are not very different from each other. It can even said that a model which would attribute the same contribution to all homozygotes $A_iA_i$, would lead to an asymptotic total variance equal to O. According to formula (2) or (2') the total variance passes through a maximum if the following relation holds:

$$|\text{VAR}_0(Z) - \text{VAR}_1(Z)| < D^2$$

For characters $(a)$ et $(d)$ which involve several loci, it is found that the component attached to the variance of inbreeding coefficients does not contribute significantly to the expected variance. In the case $(d)$, where $\mathfrak{D}^2 - D^2$ is about 4.72 times the initial variance $V_A + V_D$, and with a small effective size ($N_e = 16$), the contribution does not exceed 2 p. 100, and is of the same order of magnitude as the approximation involved in the computation of identity coefficients. However, it might be questionable to ignore this component for a character contributed by a large number, L, of loci yielding equal contributions $\mathfrak{D}_\alpha$: then the component $\mathfrak{D}^2 - D^2 = (L - 1)D^2$ could be important. This may only happen for a character showing a very great inbreeding depression.

### B. — Redistribution of the Variance in interbred Population $F_1$ and $F_2$

In table 2, the results are summed up in condensed form, on the one hand for an additive character to be used as reference and for characters $(a)$, $(b)$, $(c)$ and $(d)$. The same distinction between character $(a)$ and the other characters must still be made, the first having a behaviour that is very similar to that of an additive character. The other characters $(b)$, $(c)$, $(d)$ have a common property which seems to present some interest, that of being able to recover from lines $F_2$ an important fraction of the genetic variance of the initial population, not only at the beginning when intraline variances are large but during the entire process, still the limit is reached. Another aspect common to character $(b)$, $(c)$ can be pointed out: variance $V_{wF_1}$ is systematically and clearly smaller than $V_w$, whereas the divergence is small for characters $(a)$ and $(d)$. In a way this questions the comparison made between the variance in inbred parental lines, and the variance in hybrid populations $F_1$ that descend from them. It would not be fair to characters such as $(b)$ and $(c)$ to attribute entirely the difference in phenotypic variances observed experimentally to an increase in the environmental variance.

The preceding results are only illustrations but they seem to include most of the possible cases. If we restrict ourselves to biallelic loci in which one of the alleles is dominant over the other, but considering various distributions of initial

**TABLE 2**

Redistribution of the genetic variance within the interbred populations $F_1(V_{wF1})$ and $F_2(V_{wF2})$ from inbred lines of the inbreeding coefficient F. Characters (a), (b), (c), (d) are defined in the text (section VI)

Distribution de la variance génétique dans les populations $F_1(V_{wF1})$ et $F_2(V_{wF2})$ obtenues par croisements entre lignées consanguines de même coefficient de parenté moyenne F. Les caractères (a), (b), (c), (d) sont définis en section VI

| F | Additive character | | |
|---|---|---|---|
| | $V_w$ | $V_{wF1}$ | $V_{wF2}$ |
| 0.0 | 1.0 | 1.0 | 1.0 |
| 0.2 | 0.8 | 0.8 | 0.9 |
| 0.4 | 0.6 | 0.6 | 0.8 |
| 0.6 | 0.4 | 0.4 | 0.7 |
| 0.8 | 0.2 | 0.2 | 0.6 |
| 1.0 | 0.0 | 0.0 | 0.5 |

| F | Character (a) | | | Character (b) | | | Character (c) | | | Character (d) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $V_w$ | $V_{wF1}$ | $V_{wF2}$ | $V_w$ | $V_{wF1}$ | $V_{wF2}$ | $V_w$ | $V_{wF1}$ | $V_{wF2}$ | $V_w$ | $V_{wF1}$ | $V_{wF2}$ |
| 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.2 | 0.839 | 0.840 | 0.924 | 2.357 | 0.931 | 1.778 | 0.987 | 0.861 | 1.013 | 1.007 | 0.934 | 1.026 |
| 0.4 | 0.648 | 0.661 | 0.838 | 2.839 | 0.769 | 2.387 | 0.844 | 0.691 | 0.989 | 0.863 | 0.802 | 1.004 |
| 0.6 | 0.440 | 0.461 | 0.742 | 2.448 | 0.596 | 2.836 | 0.610 | 0.491 | 0.935 | 0.624 | 0.602 | 0.940 |
| 0.8 | 0.223 | 0.240 | 0.637 | 1.396 | 0.331 | 3.161 | 0.319 | 0.261 | 0.856 | 0.326 | 0.334 | 0.838 |
| 1.0 | 0.0 | 0.0 | 0.523 | 0.0 | 0.0 | 3.409 | 0.0 | 0.0 | 0.759 | 0.0 | 0.0 | 0.700 |

frequencies between loci (the law of Pearson between o and 1) we find again the cases studied in this paper and in addition to them, a situation in which the intra-line mean variance $V_w$ decreases faster than in the case of an additive character, contrary to the cases illustrated.

# VII. — Conclusion

In this article heterosis was defined only with respect to interbreeding of inbred lines born of the same stock. No attempt was made to study the incidence of epistatic interactions. Interbreeding distinct races in which different alleles can segregate, the experimental results of plant genetics, may therefore appear to have been abusively neglected. In fact, the ideas justifying this are:
— Heterosis, in as far as its genesis is concerned, should not be dissociated from its complementary phenomenon, the inbreeding depression.
— The dominance model has the two fold advantage of giving a satisfactory theoretical explanation and of being supported by physiological and molecular genetics data, at least when only one locus is being considered.
— The overdominance is revealed at the level of allelic series with visible effects, but it is refuted at the molecular level and it is interpreted more as the result of an interaction between closely related locus effects.
— Despite the fact that epistasis has been shown to exist in the vegetable kingdom, it has no explicative value as long as it is introduced at the statistical level to make up for all deviations from a simple model.
Section III — evolution of means — surely does not bring any new results in recalling how the dominance model explains the existence and the linearity of the inbreeding depression as well as heterosis in interbreeding inbred lines. Its essential interests, in our opinion, is of didactic nature.
Two standard computation methods have been used in parallel, but the relationships are not always clearly understood. On the one hand, we have recalled the " on an average " computation in which MALECOT's inbreeding coefficient $f$ makes its appearance defined as a probability of identity between genes. On the other hand, we used a more standard description, by means of the frequency of an allele in a line, and obtained a formula in which the variance of the frequency of this allele VAR($q$) appears this time. A comparison between these two expressions leads to a formula connecting the two quantities F and VAR($q$) : VAR($q$) = $p_0 q_0$F.
Without being a general demonstration, this comparison between methods is a good introduction to the systematic use of identity coefficients, that takes place next to analyse the evolution of variances.
Section IV is the one that brings most innovations in so far as it shows original results concerning break-down of the variance between and within lines. These results are established after a simple statistical analysis, but the simplification of formulas and their expression in terms of only one parameter F, relies on the results recently obtained concerning the evolution of these identity coefficients in panmictic populations.
This field is still left to be explored, since we do not know how to formulate the evolution of variances, in general terms. This is due to a basic difficulty which lies in characterizing a quantitative character through five variance components. The difficulties in estimating them are known (CHEVALET, 1976) but the

knowledge of identity coefficients in panmictic lines may lead to new estimation methods, that are simpler than those based on intra-line analysis and on explicit calculation of coefficients depending on the pedigree.

Nevertheless, the behaviour observed in the given illustrations and which are also found when we study loci where initial allelic frequencies follow various distributions, must retain our attention since they show the possibility of an evolution, which does not follow WRIGHT's diagram even at the cost of a transformation of data.

(*i*) The maintenance of the mean genetic variance in the lines, or its temporary increase is possible in cases that are more general than those already known; the immediate and continuous reduction in this variance of inbred lines is not a general phenomenon.

(*ii*) The genetic variance of populations $F_1$ obtained from interbreeding inbred lines can be considerably smaller than the mean genetic variance in parental populations; the reduction in variance, experimentally observed in $F_1$ should not always be attributed to a reduction in the environmental variance.

(*iii*) Considering a model with several loci shows that a component of variance attached to the variability in individual inbreeding coefficients must be considered. Although it seems to be generally small, it should be taken into account if experimental evidence shows a great inbreeding depression or a clear contribution of the variance of inbreeding coefficients to the phenotypic variance.

The use of the results obtained from this study, to interpret the results of an experiment, or to devise a method for the estimation of variance components is illustrated in the case in which pure lines are available. This application which holds for autogamous vegetable species, may be extended to any inbred lines by means of formulas developed in this paper and by means of approached expressions of identity coefficients. It might also be worth generalizing such results to cases where the founder individuals of the lines are not unrelated.

Finally, this study of variances is based explicitly on the hypothesis of lack of selection, mutation and migration. Mutation and migration phenomena can be treated by the theory of identity between genes (MALECOT, 1948, 1969*b*, 1971, 1972; GILLOIS, 1964, 1966*b*). The evolution of identity coefficients with mutation is well known (GILLOIS, 1964; CHEVALET, GILLOIS, NASSAR, 1977) and it is possible to consider extending MALECOT's methods to identity coefficients. The study of the influence of selection can rely on the results of KIMURA (1964) on the moments of genic frequencies. For sufficiently low selection pressures, it is possible to assume that the formulae we have given are good explanations for at least a certain period of time. The characterization of this field of validity will constitute the subject of a forth-coming study.

## Résumé

*Hétérosis et dépression dues à la consanguinité.*
*Évolution des variances génétiques dans les lignées consanguines*
*et dans leurs croisements*

Les phénomènes complémentaires d'hétérosis et de dépression de consanguinité sont définis par rapport à une expérience où une population est subdivisée en lignées isolées et d'effectifs limités. Pour un caractère quantitatif dû à l'action de plusieurs locus où se manifestent des interactions de dominance, on donne un traitement général de l'évolution des moyennes et des varian-

ces génétiques dans et entre les lignées. En absence de sélection, les variances génétiques dans les lignées consanguines, et dans les croisements entre ces lignées, s'expriment au moyen des valeurs initiales des variances génétiques, et en fonction des coefficients d'identité entre 2, 3 et 4 gènes. On donne des formules approchées de ces variances, où les coefficients d'identité s'expriment en fonction de puissances du coefficient moyen de consanguinité dans les lignées. Ces résultats généralisent notamment les travaux de WRIGHT, ROBERTSON, CROW et KIMURA, qui sont limités ou bien à des caractères additifs, ou bien à un seul locus. Les principaux types d'évolution des variances génétiques, et de leur redistribution dans des populations croisées $F_1$ et $F_2$, sont discutés au moyen d'illustrations numériques. L'interprétation statistique des résultats est esquissée dans le cas particulier du croisement entre des lignées homozygotes.

Deux résultats qualitatifs méritent d'être soulignés :

(i) Le maintien, ou l'accroissement temporaire, de l'espérance de la variance génétique intralignée semble possible dans des cas plus généraux que ceux déjà connus.

(ii) La variance génétique dans un croisement $F_1$ entre deux lignées consanguines peut être nettement plus petite que la moyenne des variances génétiques dans les lignées parentales. La réduction de la variance phénotypique, souvent observée dans une telle expérience, ne devrait donc pas être toujours attribuée exclusivement à une réduction de la variance des effets du milieu.

Deux types d'application des résultats obtenus sont envisagés : l'estimation des composantes génétiques de la variance dans les populations consanguines, et l'obtention de prédictions approchées de la réponse à la sélection dans une petite population.

# Références

CHEVALET C., 1976. L'estimation des composantes de la variance phénotypique dans une population consanguine. *Ann. Génét. Sél. anim.*, **8** (2), 181-232.

CHEVALET C., GILLOIS M., NASSAR R. F., 1977. Identity coefficients in finite populations. I: Evolution of identity coefficients in a random mating diploid dioecious population. *Genetics*, **86**, 697-713.

CROW J. F., KIPUMA M., 1970. An introduction to population genetics theory. Harper, New York.

COCKERHAM C. C., 1969. Variance of gene frequencies. *Evolution*, **23**, 72-84.

DICKERSON G. E., BLUNN C. T., CHAPMAN A. B., KOTTMAN R. M., KRIDER J. L., WARWICK E. J., WATTLEY J. A. (Jr.), 1954. Evaluation of selection in developing inbred lines of swines. Res. Bull. Mo. agric. Exp. Sta., N° 551, 60 p. (*Anim. Breed. Abstr.*, **23**, N° 1283).

DICKERSON G. E., 1969. Experimental approaches in utilizing breed ressources. *Anim. Breed. Abstr.*, **37**, 191-202.

FALCONER D. S., 1960. Introduction to quantitative genetics. Oliver and Boyd, London.

FISHER R. A., 1918. The correlations between relatives on the supposition of mendelian inheritance. *Trans. Roy. Soc. Edinburgh*, **52**, 399-433.

GALLAIS A., 1970. Covariances entre apparentés quelconques avec linkage et épistasie. I: Expression générale. *Ann. Génét. anim.*, **2**, 281-310.

GILLOIS M., 1964. La relation d'identité en génétique. Unpublished thesis, Faculty of Sciences, Paris.

GILLOIS M., 1965. Relation d'identité en génétique. *Ann. Inst. Henri Poincaré*, B, **2**, 1-94.

GILLOIS M., 1966a. Note sur la variance et la covariance génotypiques entre apparentés. *Ann. Inst. Henri Poincaré*, B, **2**, 349-352.

GILLOIS M., 1966b. La relation de dépendance en génétique. *Ann. Inst. Henri Poincaré*, B, **2**, 261-278.

HARRIS D. L., 1964. Genotypic covariances between inbred relatives. *Genetics*, **50**, 1320-1348.

HAYMAN B. I., MATHER K., 1955. The description of genic interactions in continuous variation. *Biometrics*, **11**, 69-8z.

KIMURA M., 1964. Diffusion models in population genetics. *J. Appl. Prob.*, **1**, 177-232.

MALECOT G., 1939. Théorie mathématique de l'hérédité mendélienne généralisée. Guilhot, Paris, 100 p.

MALECOT G., 1948. Les mathématiques de l'hérédité. Masson et Cie, Paris, 60 p.

MALECOT G., 1969a. Consanguinité panmictique et consanguinité systématique (coefficients de WRIGHT et de MALECOT). *Ann. Génét. Sél. anim.*, **1**, 237-242.

MALECOT G., 1969b.   The mathematics of heredity. Freeman, San Francisco, 86 p.

MALECOT G., 1971.   Génétique des populations diploïdes naturelles dans le cas d'un seul locus.
   I : Évolution de la fréquence d'un gène. Étude des variances et covariances. *Ann. Génét.
   Sél. anim.*, **3**, 255-280.

MALECOT G., 1972.   Génétique des populations naturelles dans le cas d'un seul locus. II : Étude
   du coefficient de parenté. *Ann. Génét. Sél. anim.*, **4**, 385-409.

ROBERTSON A., 1952.   The effect of inbreeding on the variation due to recessive genes. *Genetics*,
   **37**, 189-207.

SELLIER P., 1970.   Hétérosis et croisements chez le porc (revue bibliographique). *Ann. Génét.
   Sél. anim.*, **2**, 145-207.

WEIR B. S., COCKERHAM C. C., 1969.   Group inbreeding with two linked loci. *Genetics*, **63**, 711-
   742.

WRIGHT S., 1951.   The genetical structure of populations. *Ann. Eugenics*, **15**, 323-354.

WRIGHT S., 1952.   The theoretical variance within and between subdivisions of a population
   that is in a steady state. *Genetics*, **37**, 312-321.