

COVARIANCES ENTRE APPARENTÉS QUELCONQUES AVEC LINKAGE ET ÉPISTASIE

I. — EXPRESSION GÉNÉRALE

A. GALLAIS

*Station d'Amélioration des Plantes fourragères, 86-Lusignan
Institut national de la Recherche agronomique*

RÉSUMÉ

Pour le cas de deux loci, des coefficients de parenté généralisés entre deux individus I et J , sont définis comme la probabilité de tirer à l'un et à l'autre des loci concernés, un ensemble de gènes chez I et un ensemble de gènes chez J , en tenant compte des relations d'identité entre les gènes homologues. En supposant l'épistasie limitée à deux loci, à partir de ces coefficients, des expressions générales de la moyenne, de la variance et de la covariance entre apparentés quelconques sont établies pour une population consanguine, dérivant sans sélection d'une population panmictique, illimitée, avec équilibre des phases de linkage. L'effet combiné du linkage et de l'épistasie sur les statistiques caractéristiques d'une population peut alors être précisé. Des cas particuliers déjà connus sont envisagés, à partir de l'expression générale.

INTRODUCTION

En amélioration des plantes ou des animaux, toute sélection pour un caractère quantitatif donné a d'autant plus de chance d'être efficace que la variabilité génétique est importante. Une méthode de sélection, pour avoir le progrès génétique maximum, doit donc agir à la fois sur la moyenne et la variance du caractère quantitatif, par l'intermédiaire du choix des unités génétiques et des systèmes de reproduction. Pour utiliser au maximum les effets génétiques et leur variabilité, il faut donc connaître l'influence des systèmes de reproduction sur la valeur et la variabilité d'un caractère quantitatif.

Le choix d'une méthode de sélection doit aussi être fondé sur la nature de la variabilité génétique. Une expérience est donc nécessaire pour apprécier les paramètres de cette variabilité avec différents systèmes de reproduction. Mais, l'analyse statistique de l'hérédité des caractères quantitatifs repose sur des modèles mathématiques construits à l'aide de certains concepts. Avant toute expérience,

destinée à orienter le choix d'une méthode de sélection, il importe donc que le modèle logique et mathématique soit développé, avec le minimum d'hypothèses pour avoir une application biologique efficace.

Ces développements mathématiques permettent d'établir pour un caractère quantitatif des expressions théoriques de la moyenne, de la variance et de la covariance entre apparentés, en fonction de paramètres relatifs à l'action des gènes (additivité, dominance, épistasie), à leur liaison (linkage) et au mode de reproduction utilisé. L'étude des covariances entre individus apparentés a pour but essentiel de fournir un modèle d'analyse des composantes de la variabilité génétique, de permettre l'étude de l'influence d'un système de reproduction sur cette variabilité et donc d'arriver à prévoir le progrès génétique propre à une méthode de sélection.

C'est FISHER (1918) qui le premier applique les lois de l'hérédité mendélienne aux caractères quantitatifs; il a montré qu'il était possible de décomposer la variance génétique totale en variance génétique additive et variance due à la dominance et même variance due à l'épistasie, mais, dans des populations panmictiques illimitées; il a aussi donné pour cette situation l'expression du coefficient de corrélation théorique entre deux individus dans des cas particuliers d'apparentement. WRIGHT (1921-1922) introduit par la méthode des « coefficients de piste », le coefficient de consanguinité F , défini comme un coefficient de corrélation gamétique. Cette notion lui permet de généraliser les expressions de la corrélation théorique entre zygotes apparentés et consanguins mais en ne considérant que l'additivité.

L'originalité de l'apport de MALECOT (1948) est dans la définition probabiliste du coefficient de parenté φ et du coefficient de consanguinité f . Il généralise ainsi de façon élégante les résultats de FISHER sur l'expression des covariances entre apparentés non consanguins d'une population panmictique. KEMPTHORNE (1954, 1955, 1957) et COCKERHAM (1954, 1963) reprenant les résultats de MALECOT les généralisent en exprimant la covariance entre apparentés non consanguins dans une population panmictique avec épistasie, mais en l'absence de linkage. Ils donnent ainsi une expression de la covariance entre certains types d'apparentés issus de croisement d'individus consanguins mais non parents.

GILLOIS (1964) en introduisant le concept de situation d'identité restreinte et de coefficient d'identité, est le premier à donner une généralisation de la covariance entre apparentés, quelles que soient les relations de parenté et la consanguinité, mais, en l'absence d'épistasie et de linkage. Ces résultats ont été généralisés par BOUFFETTE (1966) aux autotétraploïdes. HARRIS (1964), indépendamment de GILLOIS, a donné une expression des covariances entre apparentés en régime de reproduction en consanguinité, et généralisé au cas d'épistasie, mais avec l'hypothèse d'absence de linkage.

Une étude générale de l'influence du linkage sur les covariances entre apparentés quelconques n'a pas encore été faite. SCHNELL (1963) a cependant résolu le problème de l'expression des covariances entre apparentés dans une population illimitée panmictique en équilibre, avec linkage et épistasie. Dans ce cas, le linkage sans épistasie ne change rien à l'expression déjà donnée par MALECOT en 1948.

De nombreux auteurs ont abordé le problème de l'expression des covariances

entre apparentés, en régime d'autofécondation, en tenant compte du linkage [BATEMAN et MATHER (1951), MATHER (1949), NELDER (1952), HAYMAN and MATHER (1955)]. L'étude la plus complète, mais en l'absence d'épistasie et avec biallélisme, a été celle de GATES, COMSTOCK, ROBINSON (1957).

Les études faites jusqu'ici ne sont que partielles. Or, pour l'analyse de l'hérédité d'un caractère quantitatif, il faut un modèle mathématique faisant intervenir le minimum d'hypothèses, inévitable dans l'état actuel de nos connaissances en génétique mathématique. Aussi, en utilisant les concepts développés par GILLOIS (1964), le but de cette étude, dans un premier article, est d'établir une expression générale de la covariance entre apparentés quelconques consanguins ou non, issus d'une population illimitée, panmictique, à l'équilibre, avec épistasie et linkage. Pour faciliter la compréhension des développements et limiter leur extension, l'étude se limite au cas de deux loci, où à tous les ensembles, considérés indépendants, de deux loci. La généralisation à l loci serait très longue avec les concepts utilisés. Dans un deuxième article, l'évolution des coefficients introduits, sera étudiée en régime d'autofécondation.

Les hypothèses de l'étude

1. — Diploïdie ou comportement diploïde à la méiose,
2. — Absence d'effet extra-chromosomique,
3. — Absence d'effet de position sur le chromosome,
4. — Absence de migration, mutation et sélection,
5. — Population de départ illimitée, panmictique, en équilibre de linkage.

La première hypothèse limite l'étude aux espèces à comportement diploïde à la méiose. L'extension des résultats aux espèces autotétraploïdes est possible à envisager, avec certaines difficultés, dues au nombre de paramètres en utilisant les concepts développés par BOUFFETTE (1966).

Le manque de connaissance approfondie sur les mécanismes de l'hérédité cytoplasmique rend actuellement difficile l'introduction de ce paramètre dans les modèles de mathématique génétique.

L'effet de position est distingué ici de l'interaction entre gènes non allèles. Il serait possible d'inclure dans l'étude un paramètre de position « Cis » ou « Trans ». Mais, dans cette première approche du problème, nous ne l'avons pas fait.

L'absence de sélection, naturelle ou artificielle revient à attribuer une participation égale de chaque génotype, à chaque génération, pour la formation de la génération suivante. Dans une expérience, il est certes difficile d'éliminer la sélection au niveau gamétique, mais une certaine sélection au niveau zygotique peut être évitée. Les deux hypothèses classiques d'absence de migration et de sélection ne semblent pas graves et pourraient d'ailleurs être levées.

Enfin la cinquième hypothèse limite l'étude aux populations illimitées panmictique, en équilibre du point de vue du linkage. L'hypothèse d'effectif illimité et de panmixie en l'absence de sélection permet de considérer que des gènes non identiques, au sens de GILLOIS (1964), sont indépendants. Nous excluons ainsi

toute relation de dépendance entre gènes allèles. L'hypothèse d'équilibre de linkage, en l'absence de sélection, permettra de considérer qu'à tout moment les « phases » de deux gènes non allèles sont en fréquence égale : $P = Q = 1/2$ (voir figure).



FIG. 1.

- (a) gènes i et k , j et l en position « Cis » de fréquence P .
 (b) gènes i et k , j et l en position « Trans » de fréquence Q .

Nous introduisons ici les termes « Cis » et « Trans » au lieu des termes classiques de « couplage » et « répulsion » introduit par MATHER, pour ne pas faire d'hypothèse sur la dominance et le nombre d'allèles à un locus. Dans une population panmixtique, l'équilibre des deux phases est atteint de façon asymptotique, et d'autant plus lentement que le taux de recombinaison (p) est faible,

$$\gamma_{(ik)_m} - p_i p_k = (1 - p)^m [\gamma_{(ik)_0} - p_i p_k] \quad (\text{Malécot } (*))$$

m représentant le nombre de générations de multiplication en panmixie, γ_{ik} la fréquence des gamètes portant les gènes i et k , de fréquences respectives p_i et p_k .

Par suite de cette hypothèse, les expressions établies ne permettront pas de retrouver les expressions correspondant à la liaison absolue ($p = 0$), équivalentes des expressions correspondant à un locus, et qui supposent :

$$P = 1, \quad Q = 0$$

Le raisonnement plus général qui permettrait d'englober ce cas n'a pas été envisagé. Nous supposons donc dans ce qui suit que le paramètre de recombinaison p est tel que :

$$0 < p \leq 1/2$$

I. — LES CONCEPTS UTILISÉS GÉNÉRALISATION DU COEFFICIENT DE PARENTÉ

I. — La relation d'identité pour une classe d'homologues

« La relation d'identité induit sur l'ensemble des gènes d'une classe d'homologues une partition en sous-ensembles disjoints appelée *situation d'identité restreinte*. Toute probabilité attachée à une *situation d'identité restreinte* est un *coefficient d'identité restreinte* » [GILLOIS (1964)].

L'ordre d'une situation d'identité restreinte étant le nombre de loci homologues considérés, et son rang le nombre de zygotes entre lesquels se distribuent ses loci, nous allons considérer les situations d'identité restreinte d'ordre 4 et de rang 2. En tenant compte de l'origine paternelle ou maternelle des gènes,

(*) Cours de Génétique mathématique, Faculté des Sciences de Paris, en 1963.

TABLEAU I

Les 15 situations d'identité restreinte à un locus entre 2 zygotes I et J
(≡, signe d'identité; ○, signe d'indépendance)

(D'après GILLOIS, 1964)

	Situation d'identité restreinte	Probabilités
S ₁	[L _I ≡ L _I [*] ≡ L _J ≡ L _J [*]]	δ ₁
S ₂	[L _I ≡ L _I [*] ≡ L _J ○ L _J [*]]	δ ₂
S ₃	[L _I ≡ L _I [*] ≡ L _J [*] ○ L _J]	δ ₃
S ₄	[L _I ≡ L _J ≡ L _J [*] ○ L _I [*]]	δ ₄
S ₅	[L _I [*] ≡ L _J ≡ L _J [*] ○ L _I]	δ ₅
S ₆	[L _I [*] ≡ L _I ○ L _J ≡ L _J [*]]	δ ₆
S ₇	[L _I ≡ L _I [*] ○ L _J ○ L _J [*]]	δ ₇
S ₈	[L _I ○ L _I [*] ○ L _J ≡ L _J [*]]	δ ₈
S ₉	[L _I ≡ L _J ○ L _I [*] ≡ L _J [*]]	δ ₉
S ₁₀	[L _I ≡ L _J ○ L _I [*] ○ L _J [*]]	δ ₁₀
S ₁₁	[L _I ○ L _J ○ L _I [*] ≡ L _J [*]]	δ ₁₁
S ₁₂	[L _I ≡ L _J [*] ○ L _I [*] ≡ L _J]	δ ₁₂
S ₁₃	[L _I ≡ L _J [*] ○ L _I [*] ○ L _J]	δ ₁₃
S ₁₄	[L _I ○ L _J [*] ○ L _I [*] ≡ L _J]	δ ₁₄
S ₁₅	[L _I ○ L _I [*] ○ L _J ○ L _J [*]]	δ ₁₅

elles sont au nombre de 15 (tabl. I, GILLOIS, 1964). Sans tenir compte de l'origine des gènes, sachant que deux sont pris dans le zygote I, deux dans le zygote J, on distingue seulement 7 situations d'identité restreinte (tabl. 2).

Nous noterons ces situations d'identité en écrivant la répartition entre les deux zygotes (trait vertical) des gènes identiques de chaque classe d'identité renfermant au moins deux gènes quelle que soit leur origine, la situation S₀0 signifiant l'absence de relation d'identité dans l'ensemble des quatre gènes homologues. Cette notation a déjà été utilisée par BOUFFETTE (1966) pour les tétraploïdes. Nous pouvons remarquer que certaines des situations décrites sont des sous-ensembles de situations d'identité restreinte introduites par GILLOIS (1964). A chaque situation d'identité, nous pouvons attacher un *coefficient d'identité restreinte*.

D'autres situations et coefficients d'identité restreinte peuvent être définis, dont l'importance apparaît clairement lors de l'étude des covariances entre apparentés. Ces situations représentent la répartition entre 2 zygotes des gènes identiques pour des tirages de 2, 3 et 4 gènes quel que soit l'état des autres gènes non tirés (tabl. 3).

TABLEAU 2

Les 7 situations d'identité restreinte de rang 2, d'ordre 4, sans attacher d'importance à la provenance des gènes.

Situations d'identité restreinte			Probabilités	
$S \binom{4}{42}$	S_1	$S_2 2$	δ_1	π_1
$S \binom{22}{42}$	S_6	$S_0 \begin{smallmatrix} 2 \\ 0 \end{smallmatrix} \begin{smallmatrix} 0 \\ 2 \end{smallmatrix}$	δ_6	π_2
	$S_9 \cup S_{12}$	$S_1 \begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}$	$\delta_9 + \delta_{12}$	π_3
$S \binom{3}{42}$	$S_2 \cup S_3 \cup S_4 \cup S_5$	$S_2 1$	$\delta_2 + \delta_3 + \delta_4 + \delta_5$	π_4
$S \binom{211}{42}$	$S_7 \cup S_8$	$S_2 0$	$\delta_7 + \delta_8$	π_5
	$S_{10} \cup S_{11} \cup S_{13} \cup S_{14}$	$S_1 \begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}$	$\delta_{10} + \delta_{11} + \delta_{13} + \delta_{14}$	π_6
$S \binom{0}{42}$	S_{15}	$S_0 0$	δ_{15}	π_7

TABLEAU 3

Les coefficients de parenté généralisés à un locus
(le point d'interrogation signifie « quels que soient les gènes à ce locus »)

Nombre de gènes tirés	Situation d'identité restreinte	Probabilités
2	$\frac{1}{2} \begin{smallmatrix} J \\ ? \end{smallmatrix}$ ou $\frac{1}{2} \begin{smallmatrix} J \\ ? \end{smallmatrix} 2$	$\varphi 2 ?$
	$1 1$	$\varphi 1 1$
3	$\frac{1}{2} \begin{smallmatrix} J \\ 1 \end{smallmatrix}$ ou $\frac{1}{2} \begin{smallmatrix} J \\ 1 \end{smallmatrix} 2$	$\varphi 2 1$
	4	$1 1$
$1 1$		$\varphi 1 1$
$2 0$ $0 2$		$2 0$ $\varphi 0 2$
	$2 2$	$\varphi 2 2$

Les coefficients d'identité restreinte attachés à chaque situation d'identité décrite dans le tableau 3 sont les *coefficients de parentés généralisés*. Le coefficient de parenté φ_{IJ} de deux individus I et J définis par MALECOT comme « la probabilité pour que deux gènes tirés l'un chez I , l'autre chez J soient identiques par descendance mendélienne, en l'absence de mutation » est un coefficient d'identité restreinte particulier.

$$\varphi_{IJ} = \varphi^{1|1}$$

Le coefficient d'identité restreinte $\varphi^{2|2}$ attaché à la situation d'identité restreinte $\frac{I J}{2|2}$ n'est autre que le coefficient de consanguinité de l'individu I :

$$\begin{aligned}\varphi_{2|2}^{I J} &= \varphi_I^{2|} = f_I \\ \varphi_{2|2}^{I J} &= \varphi_J^{2|} = f_J\end{aligned}$$

Ainsi le coefficient de parenté de MALECOT φ_{IJ} , et le coefficient de consanguinité f , sont des coefficients d'identité particuliers. Les coefficients d'identité restreinte définis comme des coefficients de parenté généralisés interviennent directement dans l'expression de la covariance donnée par GILLOIS (1964) :

$$\begin{aligned}\text{Cov } Z_I Z_J &= 4\varphi^{1|1} E(X_i^2) + 4\varphi^{2|1} E(X_i D_{ii}) + \varphi^{1|1} E(D_{ij}^2) + \varphi^{2|2} E(D_{ii}^2) \\ &+ \left[\begin{array}{cc} 2|0 & \frac{I J}{\varphi^{2|2}} \frac{I J}{\varphi^{2|2}} \end{array} \right] [E(D_{ii})]^2\end{aligned}$$

avec

$$\begin{aligned}\varphi^{1|1} &= \pi_1 + I/2\pi_3 + I/2\pi_4 + I/4\pi_6 \quad (*) \\ \varphi^{2|1} &= \pi_1 + I/4\pi_4 \\ \varphi^{1|1} &= \pi_3 \\ \varphi^{2|2} &= \pi_1 \\ \varphi^{2|0} &= \pi_2 \\ \varphi^{2|2} \varphi^{2|2} &= f_I f_J\end{aligned}$$

2. — Généralisation à plusieurs classes d'homologues

a) Situations d'identité

Pour l classe d'homologues, le nombre de situations d'identité restreinte devient très grand, même sans tenir compte de l'origine des gènes. Nous limiterons notre étude au cas de $l = 2$. Dans ce cas, sans tenir compte de l'origine des gènes nous aurons 53 situations d'identité restreinte de rang 2 et d'ordre 4-4 (2 zygotes et 4 gènes tirés à chaque locus), en admettant que les probabilités attachées aux situations symétriques ne soient pas égales. Avec les hypothèses faites, il y a

(*) Les probabilités π_i sont celles définies au tableau 2.

TABLEAU 4

Les situations d'identité restreinte de rang 2, d'ordre 4-4, pour 2 classes d'homologues. La notation d'une situation est un tableau indiquant la répartition entre les deux zygotes, trait vertical, et entre les 2 loci, trait horizontal, des 8 gènes tirés, selon les classes d'identité renfermant au moins 2 gènes. Une ligne du tableau correspond à une classe d'identité qui peut donc renfermer de 2 à 4 gènes identiques.

		Situations d'identité restreinte au locus n° 1						
		o o	(2 o)ou(o 2)	1 1	(2 1)ou(1 2)	$\frac{1}{1} \frac{1}{1}$	$\frac{2}{0} \frac{0}{2}$	2 2
Situations d'identité restreinte au locus n° 2	o o	$\frac{o o}{o o}$ P ₁	$\frac{2 o}{o o}$ P ₂	$\frac{1 1}{o o}$ P ₃	$\frac{2 1}{o o}$ P ₄	$\frac{1}{1} \frac{1}{1}$ P ₅	$\frac{2}{0} \frac{0}{2}$ P ₆	$\frac{2}{2} \frac{2}{2}$ P ₇
	(2 o)ou(o 2)	$\frac{o o}{2 o}$ P' ₂	$\frac{2 o}{2 o} \frac{2 o}{o 2}$ P ₈ P ₉	$\frac{1 1}{2 o}$ P ₁₀	$\frac{2 1}{2 o} \frac{2 1}{o 2}$ P ₁₁ P ₁₂	$\frac{1}{1} \frac{1}{1}$ P ₁₃	$\frac{2}{0} \frac{0}{2}$ P ₁₄	$\frac{2}{2} \frac{2}{2}$ P ₁₅
	1 1	$\frac{o o}{1 1}$ P' ₃	$\frac{2 o}{1 1}$ P' ₁₀	$\frac{1 1}{1 1}$ P ₁₆	$\frac{2 1}{1 1}$ P ₁₇	$\frac{1}{1} \frac{1}{1}$ P ₁₈	$\frac{2}{0} \frac{0}{2}$ P ₁₉	$\frac{2}{2} \frac{1}{1}$ P ₂₀
	(2 1)ou(1 2)	$\frac{o o}{2 1}$ P' ₄	$\frac{2 o}{2 1} \frac{2 o}{1 2}$ P' ₁₁ P' ₁₂	$\frac{1 1}{2 1}$ P' ₁₇	$\frac{2 1}{2 1} \frac{2 1}{1 2}$ P ₂₁ P ₂₂	$\frac{1}{1} \frac{1}{1}$ P ₂₃	$\frac{2}{0} \frac{0}{2}$ P ₂₄	$\frac{2}{2} \frac{2}{1}$ P ₂₅
	$\frac{1}{1} \frac{1}{1}$	$\frac{o o}{1 1}$ P' ₅	$\frac{2 o}{1 1}$ P' ₁₃	$\frac{1 1}{1 1}$ P' ₁₈	$\frac{2 1}{1 1}$ P' ₂₃	$\frac{1}{1} \frac{1}{1}$ P ₂₆	$\frac{2}{0} \frac{0}{2}$ P ₂₇	$\frac{2}{2} \frac{1}{1}$ P ₂₈
	$\frac{2}{0} \frac{0}{2}$	$\frac{o o}{2 0}$ P' ₆	$\frac{2 o}{o 2}$ P' ₁₄	$\frac{1 1}{o 2}$ P' ₁₉	$\frac{2 1}{o 2}$ P' ₂₄	$\frac{1}{1} \frac{1}{1}$ P ₂₇	$\frac{2}{0} \frac{0}{2}$ P ₂₉	$\frac{2}{2} \frac{0}{2}$ P ₃₀
	2 2	$\frac{o o}{2 2}$ P' ₇	$\frac{2 o}{2 2}$ P' ₁₅	$\frac{1 1}{2 2}$ P' ₂₀	$\frac{2 1}{2 2}$ P' ₂₅	$\frac{1}{1} \frac{1}{1}$ P ₂₈	$\frac{2}{0} \frac{0}{2}$ P' ₃₀	$\frac{2}{2} \frac{2}{2}$ P ₃₁

symétrie, et il n'y a que 3I situations d'identité restreinte de rang 2 et d'ordre 4-4 comme le montre le tableau de leur formation (tabl. 4) avec P_i ($i = 1$ à 3I) les coefficients d'identité restreinte attachés à ces probabilités.

La notation des situations d'identité restreinte exprime la répartition des gènes identiques dans chaque zygote I et J ; pour chacun des deux loci, chaque situation non symétrique en I et J est l'union de deux situations, par exemple :

$$\left\{ \begin{array}{l} 2|0 \\ 0|0 \end{array} \right\} = \left\{ \begin{array}{l} I J \\ 2|0 \\ 0|0 \end{array} \right\} \cup \left\{ \begin{array}{l} I J \\ 0|2 \\ 0|0 \end{array} \right\}$$

donc

$$P_2 = Pr \left\{ \begin{array}{l} I J \\ 2|0 \\ 0|0 \end{array} \right\} + Pr \left\{ \begin{array}{l} I J \\ 0|2 \\ 0|0 \end{array} \right\}$$

Le trait horizontal sépare les deux classes d'homologues, ainsi on écrira :

$$P'_2 = Pr \left\{ \begin{array}{l} I J \\ 0|0 \\ 2|0 \end{array} \right\} + Pr \left\{ \begin{array}{l} I J \\ 0|0 \\ 0|2 \end{array} \right\}$$

b) *Coefficients de parenté généralisés*

Ces situations d'identité restreintes étant définies, il est possible d'introduire, comme dans le cas d'une classe d'homologues, des situations d'identité particulières, auxquelles nous attacherons des *coefficients d'identité restreinte, généralisation des coefficients de parenté*. Par homogénéité avec ce qui précède, nous

les noterons $\left\{ \begin{array}{l} I J \\ x|x \\ x|x \end{array} \right\}$ sans écrire les situations symétriques (tabl. 5) et ne retenant que ceux qui seront utiles par la suite.

Les tirages ne portant que sur un locus $\left\{ \left(\frac{2}{?} \right), \left(\frac{3}{?} \right) \text{ et } \left(\frac{4}{?} \right) \right\}$ quelle que soit l'origine des gènes, conduisent évidemment aux coefficients de parenté définis précédemment pour une classe d'homologues.

Le raisonnement présenté ici dépasse celui de SHIKATA (1965) : SHIKATA reconnaissait les limites de sa généralisation. En effet, les seuls coefficients de parenté définis par SHIKATA peuvent s'écrire avec une notation analogue à celle introduite, mais faisant figurer toutes les classes d'identité :

$$\begin{aligned} R(I,I) &= \varphi_{1|1}^{1|1} \\ R(I,0) &= \varphi_{1|0}^{0|1} \\ R(0,0) &= \varphi_{0|1}^{0|1} \end{aligned}$$

TABLEAU 5

Les coefficients de parenté généralisés à deux loci (φ).
Le nombre de gènes tiré à chaque locus étant fixé, la notation est équivalente à celle des tableaux 3 et 4.

Nombre de gènes tirés		Situation d'identité restreinte	Probabilités
Locus n° 1	Locus n° 2		
2	0	$\frac{2 ?}{? ?}$ ou $\frac{? 2}{? ?}$ $\frac{1 1}{? ?}$	$\varphi \frac{? 2}{? ?}$ $\varphi \frac{1 1}{? ?}$
3	0	$\frac{2 1}{? ?}$ ou $\frac{1 2}{? ?}$	$\varphi \frac{2 1}{? ?}$
4	0	$\frac{1 1}{1 1}$ $\frac{1 1}{? ?}$ $\frac{2 0}{0 2}$ $\frac{2 ?}{? ?}$ $\frac{2 2}{? ?}$	$\varphi \frac{1 1}{1 1}$ $\varphi \frac{2 0}{0 2}$ $\varphi \frac{2 2}{? ?}$
2	2	$\frac{2 ?}{? 2}$ ou $\frac{? 2}{2 ?}$ $\frac{2 ?}{2 ?}$ ou $\frac{? 2}{? 2}$ $\frac{1 1}{? 2}$ ou $\frac{1 1}{2 ?}$ $\frac{1 1}{1 1}$	$\varphi \frac{2 ?}{? 2}$ $\varphi \frac{2 ?}{2 ?}$ $\varphi \frac{1 1}{? 2}$ $\varphi \frac{1 1}{1 1}$
3	2	$\frac{2 1}{2 ?}$ ou $\frac{1 2}{? 2}$ $\frac{2 1}{? 2}$ ou $\frac{1 2}{2 ?}$ $\frac{2 1}{1 1}$ ou $\frac{1 2}{1 1}$	$\varphi \frac{2 1}{2 ?}$ $\varphi \frac{1 2}{? 2}$ $\varphi \frac{2 1}{1 1}$
4	2	$\frac{1 1}{1 1}$ $\frac{1 1}{2 ?}$ ou $\frac{1 1}{? 2}$	$\varphi \frac{1 1}{1 1}$ $\varphi \frac{1 1}{2 ?}$

TABLEAU 5 (suite)

		$\frac{1 1}{1 1}$ $\frac{1 1}{1 1}$ $\frac{2 0}{0 2}$ $\frac{2 ?}{2 ?}$	$\frac{2 0}{0 2}$ $\frac{2 ?}{? 2}$	$\frac{1 1}{\phi 1 1}$ $\frac{1 1}{1 1}$ $\frac{2 0}{\phi 0 2}$ $\frac{2 ?}{2 ?}$
		$\frac{2 0}{0 2}$ $\frac{1 1}{1 1}$ $\frac{2 2}{2 ?}$ $\frac{2 2}{1 1}$	$\frac{2 2}{? 2}$	$\frac{2 0}{\phi 0 2}$ $\frac{1 1}{1 1}$ $\frac{2 2}{2 ?}$ $\frac{2 2}{\phi 1 1}$
3	3	$\frac{2 1}{2 1}$ $\frac{2 1}{1 2}$	$\frac{1 2}{1 2}$ $\frac{1 2}{2 1}$	$\frac{2 1}{\phi 2 1}$ $\frac{2 1}{\phi 1 2}$
4	3	$\frac{1 1}{1 1}$ $\frac{2 1}{2 1}$ $\frac{2 1}{2 1}$ $\frac{2 2}{2 1}$	$\frac{1 1}{1 1}$ $\frac{1 2}{1 2}$ $\frac{2 0}{0 2}$ $\frac{1 2}{1 2}$ $\frac{2 2}{1 2}$	$\frac{1 1}{\phi 1 1}$ $\frac{2 1}{2 1}$ $\frac{2 0}{\phi 0 2}$ $\frac{2 1}{2 1}$ $\frac{2 2}{\phi 2 1}$
4	4	$\frac{1 1}{1 1}$ $\frac{1 1}{1 1}$ $\frac{2 0}{0 2}$ $\frac{1 1}{1 1}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 2}{2 2}$ $\frac{2 2}{2 2}$ $\frac{2 2}{2 2}$		$\frac{1 1}{\phi 1 1}$ $\frac{1 1}{1 1}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 0}{0 2}$ $\frac{2 2}{2 2}$ $\frac{2 2}{2 2}$ $\frac{2 2}{2 2}$

c) Coefficients de consanguinité généralisés

Les coefficients de consanguinité ne sont qu'un cas particulier des coefficients de parenté, le tirage ne portant que dans un seul individu. Par analogie avec la définition du coefficient de consanguinité à un locus, il est possible de définir le coefficient de consanguinité généralisé pour 2 loci comme la probabilité d'avoir 2 gènes identiques à chacun des 2 loci; ainsi nous noterons :

$$\varphi_{2|2}^{2|2} = \varphi_{2|}^{2|} = f_{(2,2)}$$

et

$$\varphi_{2|1}^{2|1} = \varphi_{2|}^{2|} = f_{(2)}$$

$f_{(2)}$ représentant le coefficient de consanguinité à un locus défini par MALECOT. Mais ces coefficients ne donnent pas une description complète des états d'identité possible à chacun des 2 loci : pour cela, il faut introduire les situations d'identité de rang 1, faisant intervenir les 4 gènes, et schématisées dans la figure 2.

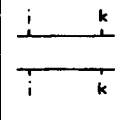
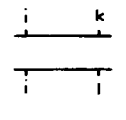
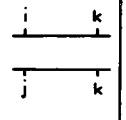
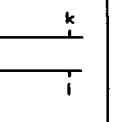
Situations				
Probabilités	Δ_1	Δ_2	Δ'_2	Δ_3

FIG. 2. — Les situations d'identité de rang 1 et d'ordre 2-2 pour 2 loci, sans tenir compte des « phases » des gènes non allèles.

Il est immédiat que :

$$f_{(2,2)} = \Delta_1$$

et Δ_2 , Δ'_2 , Δ_3 pourraient s'écrire

$$\Delta_2 = \varphi_{1|}^{2|}$$

$$\Delta'_2 = \varphi_{2|}^{1|}$$

$$\Delta_3 = \varphi_{1|}^{1|}$$

Les coefficients introduits par SHIKATA sont alors tels que :

$$\Delta_1 = F(1,1)$$

$$\Delta_2 = F(0,1)$$

$$\Delta'_2 = F(1,0)$$

$$\Delta_3 = F(0,0)$$

d) Expression des coefficients de parenté généralisés en fonction des probabilités attachées aux situations d'identité

Chaque coefficient de parenté peut être exprimé sous forme d'une combinaison linéaire des probabilités P_i attachées à chaque situation d'identité restreinte définies dans le tableau 4. Les résultats sont donnés en annexe.

II. — LES VARIABLES ALÉATOIRES GÉNÉTIQUES

I. — Structure d'une population consanguine

Soit G la variable aléatoire génotypique et $G_{(ij)(kl)}$ la mesure du caractère étudié sur un zygote de génotype $A_iA_jB_kB_l$,

i et j pour les allèles présents au locus A

k et l pour les allèles présents au locus B

G prend autant de valeurs qu'il y a de génotypes possibles.

Le tableau 6 suivant indique les différentes catégories de génotypes, leur valeur et leurs probabilités dans la population.

TABLEAU 6

Génotypes, valeurs génotypiques et probabilités conditionnées correspondantes (Conditions C_1, C_2, C'_2, C_3 de la figure 2)

Génotypes $i \neq j \neq k \neq l$	G	Probabilités conditionnées par la situation			
		C_3	C_2	C'_2	C_1
$A_iA_jB_kB_l$	$G_{(ij)(kl)}$	$f_{(ij)(kl)}$			
$A_iA_iB_kB_l$	$G_{(ii)(kl)}$	$f_{(ii)(kl)}$	$f_{(i)(kl)}$		
$A_iA_jB_kB_k$	$G_{(ij)(kk)}$	$f_{(ij)(kk)}$		$f_{(ij)(k)}$	
$A_iA_iB_kB_k$	$G_{(ii)(kk)}$	$f_{(ii)(kk)}$	$f_{(i)(kk)}$	$f_{(i)(k)}$	$f_{(i)(k)}$

Avec équilibre de linkage au départ, la structure de la population consanguine s'écrit donc, avec $i \neq j \neq k \neq l$, et en tenant compte de l'ordre des allèles présents à un locus :

$$Pr(A_iA_jB_kB_l) = \Delta_3 \phi_i \phi_j \phi_k \phi_l$$

$$Pr(A_iA_iB_kB_l) = \Delta_3 \phi_i^2 \phi_k \phi_l + \Delta_2 \phi_i \phi_k \phi_l$$

$$Pr(A_iA_jB_kB_k) = \Delta_3 \phi_i \phi_j \phi_k^2 + \Delta'_2 \phi_i \phi_j \phi_k$$

$$Pr(A_iA_iB_kB_k) = \Delta_3 \phi_i^2 \phi_k^2 + \Delta_2 \phi_i \phi_k^2 + \Delta'_2 \phi_i^2 \phi_k + \Delta_1 \phi_i \phi_k$$

2. — Définitions et estimations des aléatoires génétiques

a) Variable aléatoire génique

L'effet d'un gène se manifeste par une contribution numérique à l'expression d'un caractère. Soit X_i la contribution de l'allèle A_i ; par définition la variable aléatoire X prend la valeur X_i avec la probabilité p_i .

L'estimation que nous pouvons avoir de X_i est donnée par la moyenne des valeurs des génotypes ayant leurs gènes homologues non identiques et portant l'allèle A_i au locus A :

$$X_i = E_c [G_{(ij)(kl)}] = \sum_{j,k,l} p_j p_k p_l G_{(ij)(kl)}$$

condition c , i fixé; c'est l'hypothèse d'équilibre de linkage dans la population, qui permet d'écrire le dernier membre de l'égalité.

b) Variable aléatoire de dominance D

Deux gènes allèles i et j présents dans le zygote considéré apportent à la réalisation du caractère une contribution numérique globale G_{ij} . L'interaction entre ces 2 gènes permet de définir la valeur D_{ij} que prend la variable aléatoire D avec la probabilité $p_i p_j$

$$D_{ij} = G_{ij} - (X_i + X_j)$$

L'estimation de G_{ij} est:

$$G_{ij} = \sum_{k,l} p_k p_l G_{(ij)(kl)}$$

c) Variable aléatoire d'épistasie I

Il faut distinguer 3 types de variables aléatoires d'épistasie:

1^o Celles qui font intervenir deux gènes non allèles:

Soit i et k deux tels gènes, par analogie avec la définition de l'aléatoire de dominance, nous aurons:

$$I_{(ik)} = \sum_{j,l} p_j p_l G_{(ij)(kl)} - (X_i + X_k)$$

nous posons

$$G_{ik} = \sum_{j,l} p_j p_l G_{(ij)(kl)}$$

2^o Celles qui font intervenir deux gènes allèles et un gène non allèle:

$$I_{(ijk)} = \sum_l p_l G_{(ij)(kl)} - I_{ik} - I_{jk} - D_{ij} - X_i - X_j - X_k$$

$$G_{ijk} = \sum_l p_l G_{(ij)(kl)}$$

3° Celles qui font intervenir les quatre gènes :

$$I_{ijkl} = G_{ijkl} - I_{ijk} - I_{ijl} - I_{ikl} - I_{jkl} - I_{ik} - I_{il} - I_{jk} - I_{jl} \\ - D_{ij} - D_{kl} - X_i - X_j - X_k - X_l$$

d) *Propriétés des aléatoires génétiques*

Définies à partir d'une population en équilibre de linkage, les aléatoires génétiques montrent certaines propriétés qui seront utilisées.

1° En centrant l'aléatoire zygotique G_{ijkl} , il résulte d'après le théorème des moyennes conditionnées et la définition des aléatoires :

$$\begin{aligned} E(X_i) &= E(X_k) = 0 \\ E(G_{ij}) &= E(D_{ij}) = E(D_{kl}) = 0 \\ E(G_{ik}) &= E(I_{ik}) = 0 \\ E(G_{ijk}) &= E(I_{ijk}) = E(I_{ikl}) = 0 \\ E(G_{jkl}) &= E(I_{ijkl}) = 0 \end{aligned}$$

2° Espérances d'aléatoires avec allèles identiques :

nous savons déjà que $E(D_{ii}) \neq 0$ (GILLOIS, 1964). De même, on montre que :

$$E(I_{iikk}) \neq 0$$

Lorsque dans la condition envisagée un gène n'est pas identique à l'un des autres gènes, l'espérance conditionnée correspondante est nulle. Ainsi :

$$E(I_{iik}) = E[E_c(I_{iik})]$$

condition c , i fixé

$$\text{or } E_c(I_{iik}) = E_c(G_{iik}) - 2E_c(I_{ik}) - D_{ii} - 2X_i - E(X_k)$$

$$E(X_k) = 0$$

$$E_c(I_{ik}) = E_c(G_{ik}) - X_i = 0$$

$$E_c(G_{iik}) = G_{ii} = D_{ii} + 2X_i$$

$$E_c(I_{iik}) = 0$$

$$\text{d'où } E(I_{iik}) = 0$$

De même :

$$E(I_{iikl}) = 0$$

3° Espérances du produit de deux aléatoires :

GILLOIS (1964) a montré que :

$$E(X_i D_{ii}) \neq 0$$

$$\text{et } E(X_i D_{ij}) = 0$$

De même, par l'application du théorème des moyennes conditionnées :

$$E(X_i I_{ik}) = E[E_c(X_i I_{ik})]$$

$c = i$ fixé

or
$$E_c(X_i I_{ik}) = X_i E_c(I_{ik}) = 0$$

d'où
$$E(X_i I_{ik}) = 0$$

Il en est de même pour tout produit de deux aléatoires dès que figure dans l'une d'elles un gène non identique à l'un quelconque des autres gènes correspondant à ces aléatoires.

Par exemple :

$$E(D_{ij} I_{ijk}) = 0$$

mais
$$E(D_{ij} I_{ijkk}) \neq 0$$

III. — EXPRESSION GÉNÉRALE DE LA MOYENNE, DE LA VARIANCE ET DE LA COVARIANCE ENTRE APPARENTÉS

I. — Expression de la moyenne d'une population consanguine

D'après les propriétés des aléatoires génétiques, et par application du théorème des moyennes conditionnées :

$$E(D_{ij}) = f_{(2)} E(D_{ii})$$

$$E(D_{kl}) = f_{(2)} E(D_{kk})$$

et
$$E(I_{ijkl}) = f_{(2,2)} E(I_{iikkk})$$

$f_{(2)}$ étant le coefficient de consanguinité classique et $f_{(2,2)} = \Delta_1$, la généralisation du coefficient de consanguinité à deux loci.

Il en résulte que la variable aléatoire génotypique a pour moyenne :

$$E(G) = f_{(2)} [E(D_{ii}) + E(D_{kk})] + f_{(2,2)} E(I_{iikkk})$$

G est une variable aléatoire centrée par rapport à la moyenne de la population panmictique. Dans le cas d'indépendance :

$$f_{(2,2)} = [f_{(2)}]^2$$

nous retrouvons là dans ce cas particulier un résultat déjà établi par KEMP-THORNE (1957). *En l'absence d'épistasie*, le linkage ne peut avoir aucun effet sur la moyenne. Ce résultat a déjà été donné par MATHER (1949). En faisant la *sommation sur tous les couples de loci*, l'expression de la moyenne devient en attribuant un $f_{(2,2)}$ à chaque couple :

$$E(G) = f_{(2)} \sum_i [E(D_{ii}^{(i)})] + \sum_{\substack{i, i' \\ i < i'}} f_{(2,2)}^{(i, i')} [E(I_{iikkk}^{(i, i')})]$$

L'expression qui résulte de la sommation est donc immédiate, à partir des résultats relatifs à deux loci. Ainsi, pour simplifier l'écriture, dans ce qui suit, la sommation sur tous les couples de loci ne sera pas faite.

2. — Expression de la covariance entre deux zygotes I et J

$$\text{cov } Z_I Z_J = E(Z_I Z_J) - [E(Z_I)][E(Z_J)]$$

En notant i, j, k, l les gènes du zygote I et par i', j', k', l' , les gènes du zygote J , l'espérance du produit des aléatoires $Z_I Z_J$ peut se décomposer comme suit :

$$\begin{aligned} E(Z_I Z_J) = & 4[E(X_i X_{i'}) + E(X_k X_{k'})] + 4[E(X_i X_{k'}) + E(X_k X_{i'})] \\ & + 4[E(X_i D_{i'j'}) + E(X_k D_{k'l'})] + 4[E(X_i D_{k'l'}) + E(X_k D_{i'j'})] \\ & + 8[E(X_i I_{i'k'}) + E(X_i I_{ik})] + 8[E(X_k I_{i'k'}) + E(X_k I_{ik})] \\ & + [E(D_{ij} D_{i'j'}) + E(D_{kl} D_{k'l'})] + [E(D_{ij} D_{k'l'}) + E(D_{i'j'} D_{kl})] \\ & + 16E(I_{ik} I_{i'k'}) \\ & + 4[E(D_{ij} I_{i'k'}) + E(D_{i'j'} I_{ik})] + 4[E(D_{kl} I_{i'k'}) + E(D_{k'l'} I_{ik})] \\ & + 4[E(I_{ijk} X_{i'}) + E(I_{i'j'k'} X_i)] + 4[E(I_{ikl} X_{k'}) + E(I_{i'k'l'} X_k)] \\ & + 4[E(I_{ijk} X_{k'}) + E(I_{i'j'k'} X_k)] + 4[E(I_{ikl} X_{i'}) + E(I_{i'k'l'} X_i)] \\ & + 2[E(I_{ijk} D_{i'j'}) + E(I_{i'j'k'} D_{ij})] + 2[E(I_{ikl} D_{k'l'}) + E(I_{i'k'l'} D_{kl})] \\ & + 2[E(I_{ijk} D_{k'l'}) + E(I_{i'j'k'} D_{kl})] + 2[E(I_{ikl} D_{i'j'}) + E(I_{i'k'l'} D_{ij})] \\ & + 8[E(I_{ijk} I_{i'k'}) + E(I_{i'j'k'} I_{ik})] + 8[E(I_{ikl} I_{i'k'}) + E(I_{i'k'l'} I_{ik})] \\ & + 4[E(I_{ijk} I_{i'j'k'}) + E(I_{ikl} I_{i'k'l'})] + 4[E(I_{i'k'l'} I_{ijk}) + E(I_{i'j'k'} I_{ikl})] \\ & + 2[E(I_{ijk} X_{i'}) + E(I_{i'j'k'l'} X_i)] + 2[E(I_{ijk} X_{k'}) + E(I_{i'j'k'l'} X_k)] \\ & + [E(I_{ijk} D_{i'j'}) + E(I_{i'j'k'l'} D_{ij})] + [E(I_{ijk} D_{k'l'}) + E(I_{i'j'k'l'} D_{kl})] \\ & + 4[E(I_{ijk} I_{i'k'}) + E(I_{i'j'k'l'} I_{ik})] \\ & + 2[E(I_{ijk} I_{i'j'k'}) + E(I_{i'j'k'l'} I_{ijk})] + 2[E(I_{ijk} I_{i'k'l'}) + E(I_{i'j'k'l'} I_{ikl})] \\ & + E(I_{ijk} I_{i'j'k'l'}) \end{aligned}$$

Méthode de calcul des différentes espérances :

En appliquant la propriété du produit de deux aléatoires où figure dans l'une d'elles un gène non identique à l'un quelconque des autres gènes correspondant à ces aléatoires, il est des espérances qui sont immédiatement nulles, par exemple :

$$\begin{aligned} E(X_i X_{k'}) &= E(X_{i'} X_k) = 0 \\ E(X_i D_{k'l'}) &= E(X_k D_{i'j'}) = 0 \\ E(X_i I_{i'k'}) &= E(X_k I_{i'k'}) = 0 \\ E(D_{ij} I_{i'k'}) &= E(D_{kl} I_{i'k'}) = 0 \\ E(X_i I_{i'j'k'}) &= E(X_k I_{i'k'l'}) = 0 \end{aligned}$$

Les autres espérances se décomposent par application du théorème des moyennes conditionnées et se simplifient par application de la propriété du produit de deux aléatoires.

$$E(X_i X_{i'}) = [Pr(i \equiv i')]E(X_i^2) + [1 - Pr(i \equiv i')]E(X_i X_j)$$

or
$$E(X_i X_j) = 0$$

$$i \neq j'$$

et
$$Pr(i \equiv i') = \frac{1|1}{\varphi^{2|2}} = \varphi^{1|1}$$

d'où
$$E(X_i X_{i'}) = \varphi^{1|1} E(X_i^2)$$

De même on montre que :

$$E(X_i D_{i'j'}) = \frac{1 \ J}{\varphi^{2|2}} E(X_i D_{ii}) = \frac{1 \ J}{\varphi^{1|2}} E(X_i D_{ii})$$

$$E(D_{ij} D_{i'j'}) = \varphi^{2|2} E(D_{ii}^2) + \varphi^{0|2} E(D_{ii} D_{jj}) + \varphi^{1|1} E(D_{ij}^2)$$

Ces résultats ont déjà été donnés par GILLOIS (1964). Pour les espérances faisant intervenir une ou deux aléatoires d'épistasie, le même raisonnement s'applique.

$$E(D_{ij} D_{k'l'}) = \frac{1 \ J}{\varphi^{2|2}} E(D_{ii} D_{kk})$$

$$E(I_{ik} I_{i'k'}) = \frac{1|1}{\varphi^{1|1}} E(I_{ik}^2)$$

$$E(I_{ijk} X_{k'}) = \frac{1 \ J}{\varphi^{1|1}} E(I_{iik} X_k)$$

$$E(I_{ijk} D_{k'l'}) = \frac{1 \ J}{\varphi^{1|2}} E(I_{iik} D_{kk})$$

— etc —

$$\begin{aligned} E(I_{ijkl} I_{i'j'k'l'}) &= \frac{2|2}{\varphi^{2|2}} E(I_{iikk}^2) + \frac{0|2}{\varphi^{2|2}} E(I_{iikk} I_{jjkk}) \\ &+ \varphi^{2|0} E(I_{iikk} I_{iill}) + \frac{0|2}{\varphi^{2|0}} E(I_{iikk} I_{jjll}) \\ &+ \frac{2|2}{\varphi^{1|1}} E(I_{iikl}^2) + \frac{1|1}{\varphi^{2|2}} E(I_{ijkk}^2) \\ &+ \frac{0|2}{\varphi^{1|1}} E(I_{iikl} I_{jjkl}) + \frac{1|1}{\varphi^{0|2}} E(I_{ijkk} I_{ijll}) \\ &+ \frac{1|1}{\varphi^{1|1}} E(I_{ijkl}^2) \end{aligned}$$

Les tirages non symétriques en I et J peuvent être regroupés; par exemple, on a :

$$E(I_{IJk}I_{I'k'}) = \frac{1J}{\varphi_{11}^2} E(I_{IIk}I_{Ik})$$

et
$$E(I_{I'j'k}I_{Ik}) = \frac{1J}{\varphi_{11}^2} E(I_{IIk}I_{Ik})$$

or
$$\frac{1J}{\varphi_{11}^2} + \frac{1J}{\varphi_{11}^2} = 2 \frac{21}{\varphi_{11}^2}$$

d'où
$$E(I_{IJk}I_{I'k'}) + E(I_{I'j'k}I_{Ik}) = 2 \frac{21}{\varphi_{11}^2} E(I_{IIk}I_{Ik}) \tag{1}$$

et par symétrie :

$$E(I_{Ikk}I_{I'k'}) + E(I_{I'k'l}I_{I'k'}) = 2 \frac{11}{\varphi_{21}^2} E(I_{Ikk}I_{Ik}) \tag{2}$$

or, avec nos hypothèses :

$$\frac{21}{\varphi_{11}^2} = \frac{11}{\varphi_{21}^2}$$

la somme (1) + (2) s'écrira donc :

$$(1) + (2) = 2 \frac{21}{\varphi_{11}^2} [E(I_{IIk}I_{Ik}) + E(I_{Ikk}I_{Ik})]$$

L'écriture finale de l'espérance du produit des deux aléatoires génotypiques est donnée ci-dessous, en attachant toujours les gènes i et j au locus n° 1, et k, l au locus n° 2. Pour l'homogénéité d'écriture, l'espérance du produit de deux *aléatoires orthogonales* est gardée sous sa forme initiale; par exemple :

$E(D_{ii}D_{jj})$ pourrait s'écrire :

$i \neq j$

$$E(D_{ii})E(D_{jj}) = [E(D_{ii})]^2$$

et $E(D_{ii}D_{kk}) = E(D_{ii})E(D_{kk})$ (i et k sont non allèles).

$$\begin{aligned} E(G_I G_J) &= 4 \varphi^{11} [E(X_i^2) + E(X_k^2)] + 4 \varphi^{21} [E(X_i D_{ii}) + E(X_k D_{kk})] \\ &+ \varphi^{22} [E(D_{ii}^2) + E(D_{kk}^2)] + \varphi^{02} [E(D_{ii} D_{jj}) + E(D_{kk} D_{ll})] \\ &+ 2 \varphi^{22} E(D_{ii} D_{kk}) + \varphi^{11} [E(D_{ij}^2) + E(D_{kl}^2)] \\ &+ 16 \varphi^{11} E(I_{ik}^2) + 8 \varphi^{22} [E(I_{ikk} X_i) + E(I_{iik} X_k)] \\ &+ 4 \varphi^{22} [E(I_{ikk} D_{ii}) + E(I_{iik} D_{kk})] + 16 \varphi^{11} [(E(I_{iik} I_{ik}) + E(I_{ikk} I_{ik})] \\ &+ 4 \varphi^{11} [E(I_{iik}^2) + E(I_{ikk}^2)] + 4 \varphi^{02} [E(I_{iik} I_{jjk}) + E(I_{ikk} I_{lll})] \\ &+ 4 \varphi^{11} [E(I_{ijl}^2) + E(I_{ikl}^2)] + 8 \varphi^{21} E(I_{ikk} I_{iik}) \end{aligned}$$

$$\begin{aligned}
 &+ 4 \varphi^{\frac{2|1}{2|?}} [\mathbb{E}(I_{iik}X_i) + \mathbb{E}(I_{iik}X_k)] + 2 \varphi^{\frac{2|2}{2|?}} [\mathbb{E}(I_{iik}D_{ii}) + \mathbb{E}(I_{iik}D_{kk})] \\
 &+ 2 \varphi^{\frac{1|1}{1|?}} [\mathbb{E}(I_{ijk}D_{ij}) + \mathbb{E}(I_{iik}D_{ki})] + 2 \varphi^{\frac{2|0}{0|2}} [\mathbb{E}(I_{iik}D_{jj}) + \mathbb{E}(I_{iik}D_{ii})] \\
 &+ 8 \varphi^{\frac{2|1}{2|1}} \mathbb{E}(I_{iik}I_{ik}) + 4 \varphi^{\frac{2|2}{2|1}} [\mathbb{E}(I_{iik}I_{ii}) + \mathbb{E}(I_{iik}I_{ikk})] \\
 &+ 4 \varphi^{\frac{1|1}{1|1}} [\mathbb{E}(I_{ijk}I_{ijk}) + \mathbb{E}(I_{iik}I_{iki})] + 4 \varphi^{\frac{2|0}{0|2}} [\mathbb{E}(I_{iik}I_{jjk}) + \mathbb{E}(I_{iik}I_{iii})] \\
 &+ \varphi^{\frac{2|2}{2|2}} \mathbb{E}(I_{iik}^2) + \varphi^{\frac{2|0}{0|2}} [\mathbb{E}(I_{iik}I_{jjk}) + \mathbb{E}(I_{iik}I_{iii})] \\
 &+ \varphi^{\frac{2|0}{0|2}} \mathbb{E}(I_{iik}I_{jii}) + \varphi^{\frac{1|1}{1|2}} [\mathbb{E}(I_{iik}^2) + \mathbb{E}(I_{ijk}^2)] \\
 &+ \varphi^{\frac{2|0}{1|1}} [\mathbb{E}(I_{iik}I_{jjk}) + \mathbb{E}(I_{ijk}I_{jii})] + \varphi^{\frac{1|1}{1|1}} \mathbb{E}(I_{ijk}^2)
 \end{aligned}$$

Le produit des espérances à pour expression :

$$\begin{aligned}
 \mathbb{E}(G_I) \mathbb{E}(G_J) &= f_{(2)_I} f_{(2)_J} [\mathbb{E}(D_{ii}D_{jj}) + \mathbb{E}(D_{kk}D_{ii}) + 2\mathbb{E}(D_{ii}D_{kk})] \\
 &+ [f_{(2)_I} f_{(2,2)_J} + f_{(2)_J} f_{(2,2)_I}] [\mathbb{E}(D_{jj}I_{iik}) + \mathbb{E}(D_{ii}I_{iik})] \\
 &+ f_{(2,2)_I} f_{(2,2)_J} \mathbb{E}(I_{iik}I_{jii})
 \end{aligned}$$

L'expression de la covariance est immédiate :

$$\text{Cov}(G_I G_J) = \mathbb{E}(G_I G_J) - [\mathbb{E}(G_I)][\mathbb{E}(G_J)]$$

3. — Expression de la variance

L'expression de la variance se déduit immédiatement de l'expression de la variance en faisant $I = J$. Dans ce cas, $f_{(2)I} = f_{(2)J}$ et $f_{(2,2)I} = f_{(2,2)J}$.

Les situations d'identité restreinte se limitent à 4 :

TABLEAU 7

Les situations d'identité d'un individu avec lui-même

Situations d'identité	$\begin{pmatrix} I & I \\ I & I \\ I & I \\ I & I \end{pmatrix}$	$\begin{pmatrix} I & I \\ I & I \\ 2 & 2 \end{pmatrix}$	$\begin{pmatrix} 2 & 2 \\ I & I \\ I & I \end{pmatrix}$	$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$
Probabilités	P_{26}	P'_{28}	P_{28}	P_{31}

Ce qui annule 6 coefficients de parenté sur 28; il ne figurera dans la variance que 22 termes au lieu de 28 dans la covariance.

4. — Étude de cas particuliers

Covariance entre apparentés quelconques, consanguins

1° Absence d'épistasie et de linkage

$$\begin{aligned} \text{Cov} (G_I G_J) &= 4 \varphi^{11} [\mathbf{E}(X_I^2) + \mathbf{E}(X_K^2)] + 4 \varphi^{21} [\mathbf{E}(X_I D_{II}) + \mathbf{E}(X_K D_{KK})] \\ &+ \varphi^{22} [\mathbf{E}(D_{II}^2) + \mathbf{E}(D_{KK}^2)] + \varphi^{02} [\mathbf{E}(D_{II} D_{JJ}) + \mathbf{E}(D_{KK} D_{II})] \\ &+ \varphi^{11} [\mathbf{E}(D_{IJ}^2) + \mathbf{E}(D_{KI}^2)] - f_{(2)1} f_{(2)2} [\mathbf{E}(D_{II} D_{JJ}) + \mathbf{E}(D_{KK} D_{II})] \end{aligned}$$

C'est l'extension à deux loci indépendants de l'expression donnée par GILLOIS (1964). En utilisant la sommation sur l loci, pour les différentes composantes, la généralisation à un ensemble de l loci indépendants est immédiate.

2° Linkage avec absence d'épistasie

Dans ce cas, il s'ajoute à l'expression précédente une quantité caractérisant les covariances entre les aléatoires génotypiques des deux loci $G_I^{(1)}$ et $G_J^{(2)}$ d'une part et $G_I^{(2)}$ $G_J^{(1)}$ d'autre part, dont la somme a pour expression :

$$\text{Cov} G_I^{(1)} G_J^{(2)} + \text{Cov} G_J^{(2)} G_I^{(1)} = 2 [\varphi^{22} - f_{(2)1} f_{(2)2}] \mathbf{E}(D_{II} D_{KK})$$

dans le cas particulier d'indépendance :

$$\varphi^{22} = f_{(2)1} f_{(2)2}$$

la covariance entre les aléatoires des deux loci est nulle et nous retrouvons l'expression précédente.

3° Liaison absolue

Avec les hypothèses faites, il n'est théoriquement pas possible d'envisager ce cas. Certains résultats sont toutefois immédiats, ou particuliers :

$$f_{(2,2)} \text{ devient } f_{(2)}$$

et seules ont un sens les espérances rassemblées dans le tableau suivant avec leur interprétation et leur coefficient de parenté correspondant (tabl. 8).

Avec ces relations, l'expression générale reconduit donc à l'expression donnée par GILLOIS (1964).

TABLEAU 8

Relations entre paramètres de liaison quelconque et de liaison absolue

Espérances		Coefficients de parenté	
Liaison quelconque	Liaison absolue	Liaison quelconque	Liaison absolue
$E(I_{iik}I_{jji})$	$E(D_{ii}D_{jj})$	$\begin{array}{c} 2 0 \\ 0 2 \\ \varphi \frac{2}{0} \\ 0 2 \end{array}$	$\begin{array}{c} 2 0 \\ \varphi^0 2 \end{array}$
$E(I_{ijk}^2)$	$E(D_{ij}^2)$	$\begin{array}{c} 1 1 \\ 1 1 \\ \varphi \frac{1}{1} \\ 1 1 \end{array}$	$\begin{array}{c} 1 1 \\ \varphi^1 1 \end{array}$
$E(I_{iik}^2)$	$E(D_{ii}^2)$	$\begin{array}{c} 2 2 \\ \varphi \frac{2}{2} \\ 2 2 \end{array}$	$\varphi^2 2$
$1/2 E(I_{iik}I_{ik})$	$E(X_i D_{ii})$	$\begin{array}{c} 2 1 \\ \varphi \frac{2}{1} \\ 2 1 \end{array}$	$\varphi^2 1$
$1/4 E(I_{ik}^2)$	$E(X_i^2)$	$\begin{array}{c} 1 1 \\ \varphi \frac{1}{1} \\ 1 1 \end{array}$	$\varphi^1 1$

4° *Indépendance*

Chaque coefficient de parenté relatif à des tirages portant aux deux loci est alors égal au produit des coefficients de parenté attachés à chacun des deux loci considérés séparément. Ainsi, par exemple :

$$\begin{aligned} \varphi^{\frac{1|1}{1|1}} &= [\varphi^{1|1}]^2 \\ \varphi^{\frac{2|1}{1|1}} &= \varphi^{2|1} \varphi^{1|1} \text{ etc...} \end{aligned}$$

Notre expression conduit alors aux résultats de HARRIS (1964) pour 2 loci.

Covariance entre apparentés dans une population panmictique

Toutes les espérances renfermant des aléatoires avec des allèles identiques sont nulles :

$$\begin{aligned} \text{Cov } G_I G_J &= 4 \varphi^{1|1} [E(X_i^2) + E(X_k^2)] + \varphi^{\frac{1|1}{1|1}} [E(D_{ij}^2) + E(D_{kl}^2)] \\ &+ 16 \varphi^{\frac{1|1}{1|1}} E(I_{ik}^2) + 4 \varphi^{\frac{1|1}{1|1}} [E(I_{ijk}^2) + E(I_{ikl}^2)] + \varphi^{\frac{1|1}{1|1}} E(I_{ijk}^2) \end{aligned}$$

Dans le cas d'indépendance totale, nous retrouvons les résultats de KEMPTHORNE (1957) avec :

$$\begin{aligned}\varphi^{1|1} &= \frac{\varphi + \varphi'}{4} \\ \varphi_{1|1}^{1|1} &= \varphi\varphi' \\ \varphi_{1|1}^{1|1} &= [\varphi^{1|1}]^2 = \left(\frac{\varphi + \varphi'}{4}\right)^2 \\ \varphi_{1|1}^{1|1} &= \varphi^{1|1} \varphi_{1|1}^{1|1} \\ \varphi_{1|1}^{1|1} &= [\varphi_{1|1}^{1|1}]^2\end{aligned}$$

et en posant :

$$\begin{aligned}\mathbb{E}(X_i^2) + \mathbb{E}(X_k^2) &= 1/2 \sigma_A^2 \\ \mathbb{E}(D_{ij}^2) + \mathbb{E}(D_{kl}^2) &= \sigma_D^2 \\ \mathbb{E}(I_{ik}^2) &= 1/4 \sigma_{AA}^2 \\ \mathbb{E}(I_{ijk}^2) + \mathbb{E}(I_{ikl}^2) &= 1/2 \sigma_{AD}^2 \\ \mathbb{E}(I_{ijkl}^2) &= \sigma_{DD}^2\end{aligned}$$

d'où la formule classique de KEMPTHORNE (1957) et COCKERHAM (1954)

$$\text{Cov}(G_I G_J) = \frac{\varphi + \varphi'}{2} \sigma_A^2 + \varphi\varphi' \sigma_D^2 + \left(\frac{\varphi + \varphi'}{2}\right)^2 \sigma_{AA}^2 + \frac{\varphi + \varphi'}{2} \varphi\varphi' \sigma_{AD}^2 + (\varphi\varphi')^2 \sigma_{DD}^2$$

Dans le cas de liaison quelconque, nous retrouvons les résultats de SCHNELL (1961).

$$\text{Cov}(G_I G_J) = 2 \varphi^{1|1} \sigma_A^2 + \varphi_{1|1}^{1|1} \sigma_D^2 + 4 \varphi_{1|1}^{1|1} \sigma_{AA}^2 + 2 \varphi_{1|1}^{1|1} \sigma_{AD}^2 + \varphi_{1|1}^{1|1} \sigma_{DD}^2$$

$$\text{avec } \varphi^{1|1} = 1/4[\theta_i + \theta'_i]$$

$$\varphi_{1|1}^{1|1} = \theta_i \theta'_i$$

$$\varphi_{1|1}^{1|1} = 1/16 [\theta_{ij} + \theta'_{ij} + \theta'_i \theta_j + \theta_i \theta'_j]$$

$$\varphi_{1|1}^{1|1} = 1/4 [\theta_{ij} \theta'_j + \theta'_i \theta_j]$$

$$\varphi_{1|1}^{1|1} = \theta_{ij} \theta'_{ij}$$

5° Autre formulation de la covariance et de la variance avec un modèle global

Soit $G^{(1)(2)}$ la valeur du génotype attachée aux loci (1) et (2) et $G^{(1)}$ la valeur attachée au locus n° 1 et $G^{(2)}$ la valeur attachée au locus n° 2, nous pouvons écrire, en centrant les valeurs :

$$G^{(1)(2)} = G^{(1)} + G^{(2)} + H^{(1)(2)}$$

$H^{(1)(2)}$ désignant la variable aléatoire d'épistasie; ainsi par exemple pour un génotype $A_1^{(1)} A_j^{(1)} A_k^{(2)} A_l^{(2)}$:

$$\begin{aligned} G_{ij}^{(1)} &= X_i + X_j + D_{ij} \\ G_{kl}^{(2)} &= X_k + X_l + D_{kl} \\ H_{ijkl}^{(1,2)} &= I_{ik} + I_{il} + I_{jk} + I_{jl} + I_{ijk} + I_{ijl} + I_{ikl} + I_{jkl} + I_{ijkl} \end{aligned}$$

Nous allons considérer $G^{(1)}$, $G^{(2)}$ et $H^{(1,2)}$ comme des variables aléatoires simples. L'expression générale de la covariance entre deux individus I et J sera donc :

$$\begin{aligned} \text{Cov} [G_I^{(1,2)} G_J^{(1,2)}] &= E(G_I^{(1)} G_J^{(1)}) + E(G_I^{(2)} G_J^{(2)}) \\ &+ E(G_I^{(1)} G_J^{(2)}) + E(G_I^{(2)} G_J^{(1)}) \\ &+ E(G_I^{(1)} H_J^{(1,2)}) + E(G_J^{(1)} H_I^{(1,2)}) + E(G_I^{(1)} H_J^{(1,2)}) + E(G_J^{(2)} H_I^{(1,2)}) \\ &+ E(H_I^{(1,2)} H_J^{(1,2)}) \\ &- E(G_I^{(1,2)}) E(G_J^{(1,2)}) \end{aligned}$$

Les variables aléatoires utilisées dans ce modèle global possèdent des propriétés analogues à celles du premier modèle. En particulier :

$$E(G_{ij}) = f_{(2)} E(G_{ii})$$

de même $E(H_{ijkl}) = 0$

et $E(H_{iikl}) = E(H_{ijkk}) = 0$

Dans le tableau 9 nous donnons les résultats.

Le terme correctif à pour expression :

$$\begin{aligned} E(G_I) E(G_J) &= \{ f_{(2)_I} [E(G_{ii}) + E(G_{kk})] + f_{(2,2)_I} E(H_{iikk}) \} \{ f_{(2)_J} [E(G_{jj}) + \\ &E(G_{ll})] + f_{(2,2)_J} E(H_{jjll}) \} \\ &= f_{(2)_I} f_{(2)_J} [E(G_{ii} G_{jj}) + E(G_{kk} G_{ll}) + 2E(G_{ii} G_{kk})] \\ &+ [f_{(2)_I} f_{(2,2)_J} + f_{(2)_J} f_{(2,2)_I}] [E(G_{ii} H_{jjkk}) + E(G_{kk} H_{jjll})] \\ &+ f_{(2,2)_I} f_{(2,2)_J} E(H_{iikk} H_{jjll}) \end{aligned}$$

TABEAU 9

i, j, m au locus n° 1, k, l, r au locus n° 2

	Espérance	Coefficient ¹	
$E(G^{(1)}G^{(1)}) + E(G^{(2)}G^{(2)})$	(1) $E(G_{ij}G_{im}) + E(G_{ki}G_{kr})$	$Pr \begin{matrix} 1 1 \\ 0 0 \end{matrix}$	$\begin{matrix} 1 1 \\ \varphi_1 0 \\ 0 1 \end{matrix}$
	(2) $E(G^2_{ij}) + E(G^2_{kl})$	$Pr \begin{matrix} 1 1 \\ 1 1 \end{matrix}$	$\begin{matrix} \varphi_1 1 \\ 1 1 \end{matrix}$
	$E(G_{ij}G_{il}) + E(G_{ki}G_{kk})$	$Pr \begin{matrix} 2 1 \\ 0 1 \end{matrix}$	$\begin{matrix} \varphi_2 1 \\ \varphi_0 1 \end{matrix}$
	(3) $E(G^2_{il}) + E(G^2_{kk})$	$Pr \begin{matrix} 2 2 \\ 0 2 \end{matrix}$	$\begin{matrix} \varphi_2 2 \\ \varphi_0 2 \end{matrix}$
	(4) $E(G_{il}G_{jj}) + E(G_{kk}G_{ll})$	$Pr \begin{matrix} 2 0 \\ 0 2 \end{matrix}$	$\begin{matrix} \varphi_2 0 \\ \varphi_0 2 \end{matrix}$
$E(G^{(1)}G^{(2)})$	(5) $E(G_{il})E(G_{kk})$	$Pr \begin{matrix} 2 ? \\ ? 2 \end{matrix}$	$\begin{matrix} \varphi_2 ? \\ ? 2 \end{matrix}$
$E(G^{(1)}H^{(1,2)}) + E(G^{(2)}H^{(1,2)})$	(6) $E(G_{ij}H_{imkk}) + E(G_{ki}H_{ilkr})$	$Pr \begin{matrix} 1 1 \\ ? 2 \end{matrix}$	$\begin{matrix} 1 1 \\ 1 0 \\ \varphi_0 1 \\ ? 2 \end{matrix}$
	(7) $E(G_{il}H_{jjkk}) + E(G_{kk}H_{llll})$	$Pr \begin{matrix} 2 0 \\ 0 2 \\ ? 2 \end{matrix}$	$\begin{matrix} 2 0 \\ \varphi_0 2 \\ \varphi_0 2 \end{matrix}$
	(8) $E(G_{ij}H_{ijkk}) + E(G_{ki}H_{ilkl})$	$Pr \begin{matrix} 1 1 \\ 1 1 \\ ? 2 \end{matrix}$	$\begin{matrix} 1 1 \\ 1 1 \\ \varphi_0 1 \\ ? 2 \end{matrix}$
	(9) $E(G_{il}H_{ijkk}) + E(G_{kk}H_{ilkl})$	$Pr \begin{matrix} 2 1 \\ ? 2 \end{matrix}$	$\begin{matrix} 2 1 \\ \varphi_0 1 \\ ? 2 \end{matrix}$
	(10) $E(G_{ij}H_{ilkk}) + E(G_{ki}H_{ilkk})$	$Pr \begin{matrix} 1 2 \\ ? 2 \end{matrix}$	$\begin{matrix} 1 2 \\ 1 0 \\ \varphi_0 1 \\ ? 2 \end{matrix}$
	(11) $E(G_{il}H_{ilkk}) + E(G_{kk}H_{ilkk})$	$Pr \begin{matrix} 2 2 \\ ? 2 \end{matrix}$	$\begin{matrix} 2 2 \\ \varphi_0 2 \\ ? 2 \end{matrix}$
$E(H^{(1,2)}H^{(1,2)})$	(12) $E(H_{ijki}H_{imkr})$	$Pr \begin{matrix} 1 1 \\ 1 1 \end{matrix}$	P_{16}
	(13) $E(H_{ijkk}H_{imkl}) + E(H_{ilkl}H_{ijkr})$	$Pr \begin{matrix} 1 1 \\ 2 1 \end{matrix}$	P_{17}
	(14) $E(H_{ijki}H_{imkl}) + E(H_{ijki}H_{ijkr})$	$Pr \begin{matrix} 1 1 \\ 1 1 \\ 1 1 \end{matrix}$	P_{18}
	(15) $E(H_{ijkk}H_{imll}) + E(H_{ilkl}H_{jjkr})$	$Pr \begin{matrix} 1 1 \\ 2 0 \\ 0 2 \end{matrix}$	P_{19}
	(16) $E(H_{ijkk}H_{imkk}) + E(H_{ilkl}H_{ilkr})$	$Pr \begin{matrix} 1 1 \\ 2 2 \end{matrix}$	P_0

TABLEAU 9 (suite)

	Espérance	Coefficient ¹
(17)	$E(H_{iikkk}H_{ijkl})$	$Pr \frac{2 I}{2 I} \quad P_{21}$
(18)	$E(H_{iikkk}H_{ijkk})$	$Pr \frac{2 I}{I 2} \quad P_{22}$
(19)	$E(H_{ijkk}H_{ijkl}) + E(H_{iikkk}H_{ijkl})$	$Pr \frac{I I}{I I} \quad P_{23}$
(20)	$E(H_{iikkk}H_{jjkl}) + E(H_{iikkk}H_{ijll})$	$Pr \frac{2 I}{2 0} \quad P_{24}$
(21)	$E(H_{iikkk}H_{ijkk}) + E(H_{iikkk}H_{iikl})$	$Pr \frac{2 I}{2 2} \quad P_{25}$
(22)	$E(H^2_{ijkl})$	$Pr \frac{I I}{I I} \quad P_{26}$
(23)	$E(H_{iikkk}H_{jjkl}) + E(H_{ijkk}H_{ijll})$	$Pr \frac{2 0}{0 2} \quad P_{27}$
(24)	$E(H^2_{ijkk}) + E(H^2_{iikl})$	$Pr \frac{I I}{I I} \quad P_{28}$
(25)	$E(H_{iikkk}H_{jjll})$	$Pr \frac{2 0}{0 2} \quad P_{29}$
(26)	$E(H_{iikkk}H_{jjkk}) + E(H_{iikkk}H_{iill})$	$Pr \frac{2 0}{0 2} \quad P_{30}$
(27)	$E(H^2_{iikkk})$	$Pr \frac{2 2}{2 2} \quad P_{31}$

(¹) Tous les tirages sont figurés.

Il y a donc 28 termes soit autant que dans le premier modèle. Au niveau de la variance il n'y a que 10 termes au lieu de 22 : ce modèle peut donc être utilisé dans certaines conditions. En considérant comme seule variable aléatoire la variable aléatoire génotypique, il est même possible d'exprimer la variance seulement en fonction de 6 paramètres :

$$E(G^2_{ijkl})$$

$$E(G^2_{iikl}) + E(G^2_{ijkk})$$

$$E(G^2_{iikkk})$$

$$[E(G_{iikl})]^2, [E(G_{iikl})E(G_{jjrr})], [E(G_{iikkk})]^2$$

De telles décompositions, en limitant le nombre de paramètres à estimer, peuvent être d'intérêt dans certaines situations, selon l'objectif poursuivi.

CONCLUSIONS

La contribution de cette étude peut se situer à deux niveaux. Au *niveau des moyennes*, elle permettra une interprétation de résultats concernant l'influence de la consanguinité avec *épistasie* et *linkage*, et, bien qu'elle se limite à l'ensemble de tous les couples de loci, considérés indépendants, elle doit permettre une première approche de l'influence de l'épistasie avec *linkage*. Au *niveau de l'étude de la variabilité génétique, de sa nature et de son évolution* au cours de la reproduction selon un système consanguin particulier, elle apporte un modèle mathématique indispensable, d'une part pour l'étude de l'efficacité d'une méthode de sélection, et d'autre part, pour l'interprétation d'expériences destinées à orienter le choix de la méthode la plus efficace dans le matériel dont on dispose.

L'hypothèse d'absence de sélection qui a été faite limite certes l'utilisation de cette étude pour la théorie de la sélection artificielle dans les populations consanguines à un seul cycle de sélection, c'est-à-dire à une à plusieurs générations de reproduction en consanguinité, avant la sélection. Un seul cycle de sélection peut déjà apporter des informations très intéressantes sur l'efficacité d'un schéma de sélection. L'influence d'une sélection pendant la phase de consanguinité, sur l'intensité de la sélection possible au niveau de la phase de croisement est encore mal connue, tant du point de vue théorique qu'expérimental.

Avec les concepts utilisés, le nombre de paramètres nécessaires pour décrire la variabilité génétique d'un caractère quantitatif devient vite très grand dès que l'on ne veut pas faire d'hypothèses trop restrictives sur l'importance des relations d'épistasie. C'est un obstacle très important pour l'estimation des composantes de la variabilité génétique. A partir de l'expression générale établie, des situations particulières entraînant une diminution du nombre de paramètres, pourront être recherchées (par exemple, populations bialléliques; de telles populations sont possibles à réaliser chez le maïs). Des plans d'expériences conduisant non pas à l'estimation de tous les paramètres, mais à un nombre limité de fonctions de ces paramètres devront être établis. Certaines méthodes de l'analyse multivariée pourraient aussi être employées dans ce but.

Reçu pour publication en mai 1970.

REMERCIEMENTS

Nous tenons à adresser de très vifs remerciements à M. LEFORT, Professeur de Mathématiques à l'Institut National Agronomique qui a bien voulu critiquer et vérifier tous nos développements mathématiques et M. CHEVALET, Chargé de Recherches à la Station de Génétique animale du C.N.R.Z. qui a accepté de relire le texte. Mais, c'est grâce à M. GILLOIS, Maître de Recherches à la Station de Génétique animale du C.N.R.Z., que cette étude a pu voir le jour; nous lui en sommes très reconnaissants. Enfin, dans nos remerciements, nous n'oublions pas Monsieur le Professeur MALÉCOT qui a bien voulu apporter sa contribution par la lecture critique de cette étude.

SUMMARY

COVARIANCES BETWEEN ARBITRARY RELATIVES WITH LINKAGE AND EPISTASIS

I. — GENERAL EXPRESSION

Generalized coefficients of relationship between two individuals, I and J, are defined for two loci as the probability of drawing from one and the other of the loci a group of I genes and a group of J genes, taking into account the identity relationships between homologous genes. Supposing that epistasis is limited to two loci, general expressions of the mean, variance and covariance between relatives are established from these coefficients for an inbred population derived without selection from an unlimited random-mating population with linkage-phase balance. The combined effects of linkage and epistasis on the characteristic statistics of a population may then be studied. From the general expression already known cases are considered.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BATEMAN A. J., MATHER K., 1951. The progress of inbreeding in barley. *Heredity*, **8**, 321-348.
- BOUFFETTE J., 1966. *Expression de la covariance génotypique chez les tétraploïdes*. Thèse doctorat 3^e cycle, Fac. Sci. Lyon.
- COCKERHAM C. C., 1954. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*, **39**, 859-882.
- COCKERHAM C. C., 1963. Estimation of genetic variances. In W. D. HANSON, H. F. ROBINSON. *Statistical Genetics and Plant Breeding* Pub. Watl. Acad. Sci. Natl. pres. Council 982, 53-93.
- COMSTOCK R. E., ROBINSON H. F., 1952. Estimation of average dominance of genes. In J. GOWEN, *Heterosis*, Iowa State College Press, 494-516.
- FISHER R. A., 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh*, **52**, 399-433.
- GATES C. E., COMSTOCK R. E., ROBINSON H. F., 1957. Generalized genetic variance and covariance formulate for self-fertilized crops assuming linkage. *Genetics*, **42**, 749-763.
- GALLAIS A., 1968. Interactions between alleles and their variability in autotetraploid cross-fertilized plants. Consequences for selection. *V Congress of Eucarpia*, Milan 1-2 october.
- GILLOIS M., 1964. *La relation d'identité en génétique*. Thèse Fac. Sc. Paris.
- GILLOIS M., 1967. La notion de génotype. *Annales de Génétique*, **10**, 201-202.
- HAYMAN, MATHER, 1955. The description of genic interactions in continuous variation. *Biometrics*, **11**, 69-82.
- HARRIS D. L., 1964. Biometrical parameters of self fertilized diploid populations. *Genetics*, **50**, 931-956.
- HARRIS D. L., 1964. Genotypic covariances between inbred relatives. *Genetics*, **50**, 1319-1348.
- KEMP THORNE O., 1954. The correlation between relatives in a random mating population. *Proc. Royal. Soc. London, B*, **143**, 103-113.
- KEMP THORNE O., 1957. *An introduction to genetic statistics*. J. Wiley, New-York, 545 p.
- MALECOT G., 1948. *Les mathématiques de l'hérédité*. Masson et Cie, Paris, 63 p.
- MATHER K., 1949. *Biometrical genetics*. Methuen and Company Ltd, London.
- NELDER S. A., 1952. Some genotypic frequencies and variance components occurring in biometrical genetics. *Heredity*, **6**, 387-394.
- SCHNELL F. W., 1961. Some general formulations of linkage effects in inbreeding. *Genetics*, **46**, 947-957.
- SCHNELL F. W., 1963. The covariance between relatives in the presence of linkage. In W. D. HANSON, H. F. ROBINSON; *Statistical genetics and Plant Breeding*. Natl. Acad. Sci. Natl. Res. Council Pub. (982), 468-483.
- SHIKATA M., 1962. The generalized inbreeding coefficient in a Markov process. *Rep. Stat. Appl. Res. Juse*, **9**, 127-136.
- SHIKATA M., 1963. Representation and calculation of selfed populations by group ring. *J. Theoret. Biol.*, **5**, 142-160.
- SHIKATA M., 1965. A generalization of the inbreeding coefficient. *Biometrics*, **21**, 665-681.
- WRIGHT S., 1921. Systems of mating. II. The effects of inbreeding on the genetics composition of a population. *Genetics*, **6**, 124-143.
- WRIGHT, 1922. Coefficients of inbreeding and relationship. *Amer. Natur.*, **56**, 330-338.

ANNEXE

*Relations entre les coefficients de parenté
et les probabilités des situations d'identité*

$$\begin{aligned} \varphi_{\text{ind}}^{1|1} &= \text{I}/4[P_3 + P_{10} + P_{16} + P'_{17} + P'_{18} + P'_{19} + P'_{20}] \\ &+ \text{I}/2[P_5 + P_{13} + P_{18} + P_{23} + P_{26} + P'_{27} + P'_{28}] \\ &+ \text{I}/2[P_4 + P_{11} + P_{12} + P_{17} + P_{21} + P_{22} + P'_{23} + P'_{24} + P'_{25}] \\ &+ [P_7 + P_{15} + P_{20} + P_{25} + P_{28} + P_{30} + P_{31}] \end{aligned}$$

$= \varphi_{\text{ind}}^{1|1}$ l'indice « ind » signifiant indépendance de linkage, puisqu'il n'est pas tenu compte de l'état des gènes à l'autre locus.

$$\varphi_{\text{ind}}^{2|1} = \varphi_{\text{ind}}^{2|1}$$

$$\varphi_{\text{ind}}^{1|1} = \varphi_{\text{ind}}^{1|1}$$

$$\varphi_{\text{ind}}^{2|0} = \varphi_{\text{ind}}^{2|0}$$

$$\varphi_{\text{ind}}^{2|2} = \varphi_{\text{ind}}^{2|2}$$

$$\begin{aligned} \varphi_{\text{ind}}^{2|?} &= \text{I}/2P_9 + \text{I}/2(P_{12} + P'_{12}) + \text{I}/2(P_{14} + P'_{14}) + \text{I}/2(P_{15} + P'_{15}) + \text{I}/2P_{22} \\ &+ \text{I}/2(P_{24} + P'_{24}) + \text{I}/2(P_{25} + P'_{25}) + P_{29} + (P_{30} + P'_{30}) + P_{31} \end{aligned}$$

$$\begin{aligned} \varphi_{\text{ind}}^{1|1} &= \text{I}/16P_{16} + \text{I}/8(P_{17} + P'_{17}) + \text{I}/8(P_{18} + P'_{18}) + \text{I}/4(P_{20} + P'_{20}) \\ &+ \text{I}/4(P_{21} + P_{22}) + \text{I}/4(P_{23} + P'_{23}) + \text{I}/2(P_{25} + P'_{25}) + \text{I}/4P_{26} \\ &+ \text{I}/2(P_{28} + P'_{28}) + P_{31} \end{aligned}$$

$$\begin{aligned} \varphi_{\text{ind}}^{1|1} &= \text{I}/8P_{10} + \text{I}/8P'_{17} + \text{I}/4P'_{19} + \text{I}/4P'_{20} + \text{I}/4(P_{11} + P_{12}) \\ &+ \text{I}/4(P_{21} + P_{22}) + \text{I}/2P'_{24} + \text{I}/2P'_{25} + \text{I}/4P_{13} + \text{I}/4P_{23} \\ &+ \text{I}/2P'_{27} + \text{I}/2P'_{28} + \text{I}/2P_{15} + \text{I}/2P_{25} + P_{30} + P_{31} \end{aligned}$$

$$\varphi_{\text{ind}}^{2|1} = \text{I}/4P_{12} + \text{I}/4P_{22} + \text{I}/4P'_{24} + \text{I}/4P'_{25} + \text{I}/2P_{15} + \text{I}/2P_{25} + P_{30} + P_{31}$$

$$\begin{aligned} \varphi_{\text{ind}}^{2|1} &= \text{I}/16P_{17} + \text{I}/8(P_{21} + P_{22}) + \text{I}/8P'_{23} + \text{I}/4P'_{25} + \text{I}/4P_{20} + \text{I}/2P_{25} \\ &+ \text{I}/2P_{28} + P_{31} \end{aligned}$$

$$\varphi_{\text{ind}}^{2|2} = \text{I}/4P_{20} + \text{I}/2P_{25} + \text{I}/2P_{28} + P_{31}$$

$$\varphi_{\text{ind}}^{2|0} = \text{I}/4P_{19} + \text{I}/2P_{24} + \text{I}/2P_{27} + P'_{30}$$

$$\varphi_{\text{ind}}^{1|1} = \text{I}/4P_{18} + \text{I}/2P_{23} + \text{I}/2P_{26} + P'_{28}$$

$$\varphi_{\text{ind}}^{1|2} = \text{I}/8P_{22} + \text{I}/4P'_{25} + \text{I}/4P_{25} + P_{31}$$

$$\varphi_{\text{ind}}^{2|1} = \text{I}/4P_{11} + \text{I}/4P_{21} + \text{I}/4P'_{24} + \text{I}/4P'_{25} + \text{I}/2P_{15} + \text{I}/2P_{25} + P_{30} + P_{31}$$

$$\varphi_{\text{ind}}^{2|2} = \text{I}/2P_{15} + \text{I}/2P_{25} + P_{30} + P_{31}$$

$$\varphi_{2|1}^{\frac{1|1}{1|1}} = \mathbb{I}/2P_{13} + \mathbb{I}/2P_{23} + P'_{27} + P'_{28}$$

$$\varphi_{2|1}^{\frac{2|0}{0|2}} = \mathbb{I}/2P_{14} + \mathbb{I}/2P_{24} + P_{29} + P'_{30}$$

$$\varphi_{2|1}^{\frac{2|1}{2|1}} = \mathbb{I}/8P_{21} + \mathbb{I}/4P'_{25} + \mathbb{I}/4P_{25} + P_{31}$$

$$\varphi_{2|1}^{\frac{2|2}{2|1}} = \mathbb{I}/4P_{25} + P_{31}$$

$$\varphi_{2|1}^{\frac{1|1}{1|1}} = \mathbb{I}/4P_{23} + P'_{28}$$

$$\varphi_{2|1}^{\frac{2|0}{0|2}} = \mathbb{I}/4P_{24} + P'_{30}$$

$$\varphi_{2|2}^{\frac{2|2}{2|2}} = P_{31}$$

$$\varphi_{2|2}^{\frac{2|0}{0|2}} = P'_{30}$$

$$\varphi_{2|0}^{\frac{2|0}{0|2}} = P_{29}$$

$$\varphi_{2|2}^{\frac{1|1}{1|1}} = P'_{28}$$

$$\varphi_{2|0}^{\frac{1|1}{0|2}} = P'_{27}$$

$$\varphi_{1|1}^{\frac{1|1}{1|1}} = P_{26}$$