

ÉTUDE DE L'ÉVOLUTION DE PETITES POPULATIONS HOMOGAMES PAR LA THÉORIE DE LA CONTRAINTE; EXPÉRIENCES ANALOGIQUES DE CONTRÔLE ⁽¹⁾

J. CROIZÉ-POURCELET

*Station centrale de Génétique animale,
Centre national de Recherches zootechniques, 78-Jouy-en-Josas
Institut national de la Recherche agronomique*

RÉSUMÉ

Les relations d'Identité et de Contrainte, introduites par GILLOIS en 1964-1965, sont aujourd'hui bien connues. Cependant, avant de mettre en application ces théories nouvelles et relativement complexes, il faut s'assurer de leur « cohérence interne », c'est-à-dire vérifier qu'aucune erreur ou contradiction ne se sont glissées dans les raisonnements et les définitions, et que les résultats calculés sont conformes aux hypothèses. Lorsqu'est acquise la quasi-certitude de cette cohérence interne, il devient possible et nécessaire de tester la « cohérence externe » de la théorie grâce à des expériences concrètes, réalisées avec des populations animales ou végétales. Une méthode permettant de tester la cohérence interne de la théorie utilisée est décrite dans ce texte. En l'appliquant à un cas très schématique d'homogamie génotypique partielle dans des lignées frère-sœur, il a été possible de montrer que la définition même de ce mode d'accouplement comportait une contradiction, de préciser après une nouvelle étude logique quelles conditions générales doivent vérifier les lois d'accouplement dans les petites populations pour être cohérentes. Les équations stochastiques rigoureuses vérifiées par les coefficients de Contrainte de certaines populations homogames et leurs solutions ont alors pu être calculées.

INTRODUCTION

Tandis que le coefficient de consanguinité (coefficient of inbreeding) de WRIGHT (1925) est une corrélation linéaire entre zygotes non consanguins, les coefficients de consanguinité et de parenté introduits par MALÉCOT (1948) sont les probabilités que deux loci homologues appartenant respectivement à un zygote et à deux zygotes distincts soient identiques, parce qu'issus sans mutation d'un

⁽¹⁾ Cette étude reprend une partie des résultats exposés à plusieurs reprises, en particulier le 13 décembre 1968 à la soutenance de notre thèse (CROIZÉ-POURCELET, 1968).

même locus ancêtre. Prolongeant l'œuvre de MALECOT, GILLOIS (1964) définit les coefficients d'identité qui sont les probabilités qu'un certain nombre de gènes homologues appartenant à un certain nombre de zygotes se trouvent dans une situation d'identité *S*. Constatant enfin que deux gènes peuvent être non identiques, sans pour autant être indépendants, il introduisit la notion de contrainte (GILLOIS, 1965, *a-b*). Les théories de l'Identité et de la Contrainte permettent d'une part de calculer les corrélations génétiques entre apparentés consanguins ou non, d'autre part de prévoir l'évolution de populations d'effectifs limités soumises à des lois d'accouplement précises (GILLOIS, 1966, *a-b*). Le calcul des coefficients d'identité et de contrainte permet une étude systématique des populations d'effectifs limités dont ils décrivent logiquement les états successifs, si on choisit comme « contraire » de l'indépendance, non la dépendance en probabilité (ce qui est l'acception générale), mais la dépendance absolue (qui est un cas particulier). Avant de mettre en application cette théorie à la fois nouvelle et complexe, il est toutefois nécessaire de vérifier sa « cohérence interne », c'est-à-dire de montrer que les hypothèses ne sont pas contradictoires entre elles, que les notions utilisées le sont sans ambiguïté, et que les résultats calculés sont effectivement conformes aux hypothèses. C'est seulement quand des preuves de la cohérence interne de la théorie sont fournies, qu'on peut tester sa « cohérence externe », en montrant qu'elle décrit avec précision les phénomènes biologiques considérés.

Dans cet article sera exposée une méthode expérimentale faisant ressortir exclusivement les erreurs logiques éventuelles de la théorie. Appliquée à un cas très schématique d'homogamie génotypique partielle, cette méthode a montré qu'une telle erreur existait. Une nouvelle analyse logique a permis de la déterminer et de reconstruire une nouvelle théorie de l'homogamie génotypique partielle rigoureuse (CROIZÉ-POURCELET, 1968).

I. — VÉRIFICATION DE LA COHÉRENCE INTERNE D'UNE THÉORIE : EXPÉRIENCES ANALOGIQUES

Dans sa forme actuelle, la théorie de la Contrainte permet d'étudier l'évolution de petites populations à condition qu'elles vérifient deux hypothèses :

1) l'absence de sélection naturelle ou artificielle, par conséquent tous les types de couples doivent être également féconds en probabilité, et la ségrégation des caractères doit être strictement mendélienne;

2) l'absence de formes d'homogamie différentes de celles éventuellement considérées, et de barrières génétiques partielles ou totales. Cette hypothèse entraîne donc que les conjoints se choisissent effectivement au hasard quand les accouplements ne sont pas dirigés. Il faut, pour vérifier expérimentalement la cohérence interne de notre théorie, utiliser des populations qui vérifient parfaitement ces hypothèses. De plus, les marqueurs doivent être semi-dominants pour permettre la reconnaissance directe des génotypes d'après les phénotypes. Le seul animal actuellement bien connu génétiquement est la *Drosophile*. Or, même chez cette mouche, on ne connaît pratiquement aucun gène semi-dominant qui

ne soit fortement contre-sélectionné quand il est porté à l'état pur. Si on réalise par conséquent les expériences à l'aide de matériel vivant, et que les résultats expérimentaux diffèrent notablement des prévisions théoriques, on ne peut savoir si ces différences sont dues à une incohérence logique de la théorie ou à la non-conformité des populations choisies, aux hypothèses citées plus haut. La cohérence interne de la théorie de la Contrainte ne peut donc être testée en utilisant des populations d'animaux réels. Pour y parvenir, on est conduit à employer des « populations analogiques » qui sont composées de pseudo-zygotes, réduits à leur génotype, leur sexe et une variable aléatoire propre qui les désigne individuellement. Ces zygotes sont supposés observer les hypothèses précédentes et les lois de la génétique formelle. Les événements aléatoires survenant dans une telle population sont entièrement décidés par un processus aléatoire contrôlé et reproductible, par exemple un itinéraire sur les tables de nombres au hasard de FISHER et YATES (1963). Les expériences analogiques réalisées avec de telles populations sont donc des modèles mathématiques dépendant d'une loi du hasard. Si un écart apparaît alors entre les prévisions théoriques et les résultats de telles expériences, il ne peut être dû qu'à une erreur logique de notre étude puisque les hypothèses initiales sont effectivement observées par les populations analogiques. Si au contraire prévisions et résultats concordent, ces derniers seront une preuve valable de la cohérence interne de l'étude théorique.

Remarquons qu'il existe une différence de principe entre les expériences analogiques décrites plus haut et la méthode de Monte-Carlo. Cette dernière est en effet utilisée pour permettre d'éviter une étude théorique préalable jugée trop complexe.

II. — L'HOMOGAMIE GÉNOTYPIQUE

A. — *Étude théorique.*

L'étude de l'homogamie génotypique a déjà été réalisée par GILLOIS (1966, *b*). Soit une petite population, dans laquelle à chaque génération on isole au hasard pour la reproduction N_1 mâles et N_2 femelles qui constituent la génération zygotique utile de leur génération. On dit que cette population est soumise à un taux d'homogamie λ , ($0 \leq \lambda \leq 1$) si deux conjoints se choisissent systématiquement de même génotype pour un caractère donné, en vue de l'accouplement, avec la probabilité λ , tandis que les accouplements s'effectuent au hasard avec la probabilité complémentaire $1 - \lambda$.

Dans ce cas, les coefficients de contrainte vérifient les équations suivantes (GILLOIS, 1966-*b*) :

$$\begin{aligned}
 \text{(I)} \quad F_1(n+1) &= \lambda(1 - Q(n)) + (1 - \lambda) \Phi_1(n); \\
 \text{(Ib)} \quad \Phi_1(n+1) &= (1 + F_1(n))/2N + (1 - 1/N) \Phi_1(n). \\
 \text{(II)} \quad F_2(n+1) &= \lambda \cdot Q(n) + (1 - \lambda) \Phi_2(n). \\
 \text{(IIb)} \quad \Phi_2(n+1) &= F_2(n)/2N + (1 - 1/N) \Phi_2(n). \\
 \text{(III)} \quad F_3(n+1) &= \lambda \cdot O + (1 - \lambda) \Phi_3(n). \\
 \text{(IIIb)} \quad \Phi_3(n+1) &= F_3(n)/2N + (1 - 1/N) \Phi_3(n).
 \end{aligned}$$

avec $\mathbf{1}/N = (\mathbf{1}/N_1 + \mathbf{1}/N_2)/4$. En cas de biallélisme, si on appelle A et Z les deux allèles, et si on appelle p la probabilité de tirer A de la population initiale, $p(n)$ celle de le tirer à la n^e génération, on a à toute génération $p(n) = p(0) = p$. Pour tester les raisonnements qui ont conduit à ces résultats, il suffit donc de réaliser une série d'expériences analogiques sur des populations indépendantes. Afin d'être reproduites un grand nombre de fois rapidement, elles sont schématisées à l'extrême : on choisit ainsi une population ne comptant qu'un mâle et qu'une femelle à chaque génération zygotique utile ($N_1 = N_2 = \mathbf{1}$). Afin d'avoir une évolution rapide mais non instantanée, on donne au taux d'homogamie λ la valeur 0,7. Enfin, on pose $p = q = 0,5$ pour rendre symétriques les rôles de A et de Z. On peut alors écrire :

$$\begin{aligned} P(n) &= F_1(n) \cdot p + F_3(n) \cdot p^2 = 0,5 \cdot F_1(n) + 0,25 \cdot F_3(n) \\ 2 \cdot Q(n) &= F_2(n) + 2 \cdot F_3(n) \cdot p \cdot q = F_2(n) + 0,5 \cdot F_3(n) \\ R(n) &= F_1(n) \cdot q + F_3(n) \cdot q^2 = 0,5 \cdot F_1(n) + 0,25 \cdot F_3(n) \\ &\text{donc } P(n) = R(n) \end{aligned}$$

Les équations de récurrence sont alors linéaires et s'écrivent en notations vectorielles :

$$\begin{aligned} \left\| \begin{array}{c} F_1(n + \mathbf{1}) \\ \Phi_1(n + \mathbf{1}) \end{array} \right\| &= \left\| \begin{array}{cc} \lambda & \mathbf{1} - \lambda \\ \mathbf{1}/2N & \mathbf{1} - \mathbf{1}/N \end{array} \right\| \cdot \left\| \begin{array}{c} F_1(n) \\ \Phi_1(n) \end{array} \right\| + \left\| \begin{array}{c} \lambda F_2(n)/2 + 3F_3(n)/4 \\ \mathbf{1}/2N \end{array} \right\| \\ \left\| \begin{array}{c} F_2(n + \mathbf{1}) \\ \Phi_2(n + \mathbf{1}) \end{array} \right\| &= \left\| \begin{array}{cc} \lambda/2 & \mathbf{1} - \lambda \\ \mathbf{1}/2N & \mathbf{1} - \mathbf{1}/N \end{array} \right\| \cdot \left\| \begin{array}{c} F_2(n) \\ \Phi_2(n) \end{array} \right\| + \left\| \begin{array}{c} \lambda \cdot F_3(n)/4 \\ \mathbf{0} \end{array} \right\| \\ \left\| \begin{array}{c} F_3(n + \mathbf{1}) \\ \Phi_3(n + \mathbf{1}) \end{array} \right\| &= \left\| \begin{array}{cc} \mathbf{0} & \mathbf{1} - \lambda \\ \mathbf{1}/2N & \mathbf{1} - \mathbf{1}/N \end{array} \right\| \cdot \left\| \begin{array}{c} F_3(n) \\ \Phi_3(n) \end{array} \right\| + \left\| \begin{array}{c} \mathbf{0} \\ \mathbf{0} \end{array} \right\| \end{aligned}$$

La résolution de (III, IIIb) et de (II, IIb) est classique; (I, Ib) est calculé comme complément à $\{\mathbf{1}, \mathbf{1}\}$ des deux premières solutions $\{F_3(n), \Phi_3(n)\}$ et $\{F_2(n), \Phi_2(n)\}$ en appelant $\{F_3(n), \Phi_3(n)\}$ la solution du système (III, IIIb) et $\{F_2(n), \Phi_2(n)\}$ celle de (II, IIb).

Pour calculer la solution du système (II, IIb) on en cherche d'une part une solution particulière sous la forme :

$$F'_2(n) = K \cdot F_3(n); \quad \Phi'_2(n) = L \cdot \Phi_3(n)$$

d'autre part la solution générale du système sans second membre associé à (II, IIb), que l'on appelle $\{F''_2(n), \Phi''_2(n)\}$. La solution générale du système (II, IIb) s'écrit :

$$\begin{cases} F_2(n) = F'_2(n) + F''_2(n) \\ \Phi_2(n) = \Phi'_2(n) + \Phi''_2(n) \end{cases}$$

Des considérations simples permettent d'éviter de nombreux calculs : en effet, le système (II, IIb) ne diffère de son homologue (III, IIIb) que par la présence en (II) du terme : $\lambda \cdot F_2(n)/2 + \lambda \cdot F_3(n)/4$. Considérons alors le couple de fonctions $\{f(n), \varphi(n)\}$ vérifiant :

$$\begin{cases} f(n) + F_3(n)/2 = \mathbf{0} \\ \varphi(n) + \Phi_3(n)/2 = \mathbf{0} \end{cases} \quad \text{soit} \quad \begin{cases} f(n) = -F_3(n)/2 \\ \varphi(n) = -\Phi_3(n)/2 \end{cases}$$

$\{f(n), \varphi(n)\}$ étant proportionnel à $\{F_3(n), \Phi_3(n)\}$ vérifie (III, III b). De plus, il vérifie aussi (II, II b) puisque $f(n) + F_3(n)/2 = 0$, donc $\{f(n), \varphi(n)\}$ est solution particulière de (II, II b). On peut donc écrire :

$$F_2(n) = f(n) + F''_2(n) \text{ et } \Phi_2(n) = \varphi(n) + \Phi''_2(n) \text{ donc}$$

$$2Q(n) = f(n) + F''_2(n) + F_3(n)/2 = F''_2(n) \text{ soit}$$

$$2Q(n) = F''_2(n) \text{ et } P(n) = R(n) = (1 - F''_2(n))/2$$

Il est donc inutile, dans le cas particulier où $p = q = 0,5$, de calculer $F_1(n)$, $F_2(n)$ et $F_3(n)$, il suffit de résoudre le système sans second membre associé à (II, II b) :

$$\begin{vmatrix} F''_2(n + 1) \\ \Phi''_2(n + 1) \end{vmatrix} = \begin{vmatrix} \lambda/2 & 1 - \lambda \\ 1/2N & 1 - 1/N \end{vmatrix} \cdot \begin{vmatrix} F''_2(n) \\ \Phi''_2(n) \end{vmatrix}$$

Cette résolution est très classique : soit $\delta = \sqrt{(1 - \lambda/2)^2 - (\lambda N - 1)}$; les valeurs propres sont :

$$s_1 = (1 + \lambda/2 - 1/N + \delta)/2; \quad s_2 = (1 + \lambda/2 - 1/N - \delta)/2$$

$F''_2(n)$, qui est la seule fonction utile au calcul de la structure des générations de la population, est de la forme $F''_2(n) = A \cdot s_1^n + B \cdot s_2^n$, A et B dépendant des conditions initiales. Si celles-ci sont la panmixie dans une population infinie, la structure initiale est donnée par la loi de Hardy-Weinberg et alors :

$$A = (1 - \lambda/2 - s_2)/2 \cdot \delta; \quad B = -(1 - \lambda/2 - s_1)/2 \cdot \delta$$

Numériquement si $N_1 = N_2 = 1$ et si $\lambda = 0,7$; $N = 2$

$$\delta^2 = (1 - 0,35)^2 - 0,5(0,7 - 0,5) = 0,3225; \quad \delta = 0,5679$$

$$\begin{cases} s_1 = 0,5(0,85 + 0,5679) = 0,7089 \\ s_2 = 0,5(0,85 - 0,5679) = 0,1411 \\ A = 0,5(0,65 - 0,1411)/0,5679 = 0,4481 \\ B = -0,5(0,65 - 0,7089)/0,5679 = 0,0519 \\ 2Q(n) = 0,4481 \cdot 0,7089^n + 0,0519 \cdot 0,1411^n \\ P(n) = R(n) = 0,5 - 0,2240 \cdot 0,7089^n - 0,0259 \cdot 0,1411^n \end{cases}$$

L'évolution moyenne probable de la structure d'une telle population homogame est entièrement définie par ces deux fonctions.

B. — *Expériences analogiques de contrôle.*

Une série d'expériences analogiques va permettre de tester la précision des résultats calculés et la cohérence de la théorie développée ci-dessus.

Ces expériences analogiques consistent à mesurer l'évolution moyenne de 100, puis de 400 populations pratiquant dans 70 p. 100 des cas une homogamie génotypique, et dans 30 p. 100 la panmixie, et à comparer ces moyennes aux valeurs

probables calculées. Afin d'avoir la certitude que les hypothèses de base sont effectivement observées, les zygotes sont remplacés par des « zygotes analogiques » notés :

A quand leur génotype est A/A
 H quand leur génotype est A/Z
 Z quand leur génotype est Z/Z

Puisque les lois de Mendel sont exactement observées, un zygote analogique donné a la probabilité :

1,00 d'être A , 0,00 d'être H , 0,00 d'être Z s'il provient d'un couple $A \times A$
 0,50 d'être A , 0,50 d'être H , 0,00 d'être Z s'il provient d'un couple $A \times H$
 0,00 d'être A , 1,00 d'être H , 0,00 d'être Z s'il provient d'un couple $A \times Z$
 0,25 d'être A , 0,50 d'être H , 0,25 d'être Z s'il provient d'un couple $H \times H$
 0,00 d'être A , 0,50 d'être H , 0,50 d'être Z s'il provient d'un couple $H \times Z$
 0,00 d'être A , 0,00 d'être H , 1,00 d'être Z s'il provient d'un couple $Z \times Z$
 (Il n'y a pas lieu de distinguer ici le rôle génétique du mâle et de la femelle.)

Procédé de simulation analogique. Cas des 100 populations : considérons les 100 lignes des tables V et VI de nombres au hasard de FISHER. Affectons une ligne à chaque population la n^e ligne affectée à la n^e population. La population initiale est infinie panmictique; la fréquence du gène A est $p = 0,5$, celle du gène Z est $q = 0,5$. D'après la loi de Hardy-Weinberg, la structure de cette population est :

$$P(o) = 0,25; \quad 2Q(o) = 0,50; \quad R(o) = 0,25$$

Constituons la première génération zygotique utile d'une population donnée en lisant le nombre formé par les deux premiers chiffres de la ligne.

Si ce nombre est compris entre 01 et 25 le 1^{er} conjoint est de type A

Si ce nombre est compris entre 26 et 75 le 1^{er} conjoint est de type H

Si ce nombre est compris entre 76 et 100 le 1^{er} conjoint est de type Z
 (le nombre 00 est assimilé à 100).

Soient alors les 2 chiffres suivants de la même ligne (le 3^e et le 4^e) qui forment un nombre.

Si ce nombre est compris entre 01 et 70 le couple est homogame (70 %).

Si ce nombre est compris entre 71 et 100 le couple est panmictique.

Si le couple est homogame, le 2^e conjoint a le même génotype que le premier, il n'y a donc pas lieu de se servir de la table pour définir son type.

Si le couple est panmictique, il faut choisir le génotype du 2^e conjoint. On utilise alors le nombre formé par le 5^e et le 6^e chiffre de la ligne : comme plus haut, pour le choix du 1^{er} conjoint :

Si ce nombre est compris entre 01 et 25 le 2^e conjoint est de type A

Si ce nombre est compris entre 26 et 75 le 2^e conjoint est de type H

Si ce nombre est compris entre 76 et 100 le 2^e conjoint est de type Z

Il faut donc 6 chiffres pour constituer un couple panmictique, 4 pour constituer un couple homogame.

1^{er} exemple. — Population 1 : 1^{re} ligne de la table V.

1 ^{er} nombre : 28. Le premier conjoint est de type H	}	couple $H \times H$
2 ^e nombre : 89. Le couple est panmictique		
3 ^e nombre : 65. Le second conjoint est de type H		

2^e exemple. — Population 2 : 2^e ligne de la table V.

1 ^{er} nombre : 30. Le premier conjoint est de type H	}	couple $H \times H$
2 ^e nombre : 29. Le couple est homogame :		
Le second conjoint est de type H		

La génération zygotique utile d'une population donnée engendre ses descendants qui forment la première génération zygotique $G(I)$ de cette population. On en tire alors la génération zygotique utile $G'(I)$ selon le procédé indiqué plus haut, avec les conditions suivantes :

Cas 1. Si le premier couple géniteur était $A \times A$, la population est fixée, la structure de cette population devient définitivement $P = 1,00$; $2Q = R = 0,00$ à chaque génération.

Cas 2. Si le premier couple géniteur était $A \times H$, et

si sort un nombre compris entre 1 et 50, le descendant choisi est de type A
 si sort un nombre compris entre 51 et 100, le descendant choisi est de type H

Cas 3. Si le premier couple géniteur était $A \times Z$, tous les descendants sont H

Cas 4. Si le premier couple géniteur était $H \times H$,

si sort un nombre compris entre 1 et 25, le descendant choisi est de type A
 si sort un nombre compris entre 26 et 75, le descendant choisi est de type H
 si sort un nombre compris entre 76 et 100, le descendant choisi est de type Z

Cas 5. Si le premier couple géniteur était $H \times Z$,

si sort un nombre compris entre 1 et 50, le descendant choisi est de type H
 si sort un nombre compris entre 51 et 100, le descendant choisi est de type Z

Cas 6. Si le premier couple géniteur était $Z \times Z$, la population est fixée, la structure de cette population devient définitivement $P = 2Q = 0,00$; $R = 1,00$ à chaque génération.

1^{er} exemple. — Population 1 (1^{re} ligne de la table V). Le couple géniteur de la génération zygotique utile zéro était $H \times H$ (*cas 4*). La structure probable de sa descendance est $P = 0,25$; $2Q = 0,50$; $R = 0,25$. Le choix du couple géniteur formant la génération zygotique utile *un* est le suivant :

1^{er} nombre (5^e et 6^e chiffres de la 1^{re} ligne) : 87, le premier est de type Z .
 2^e nombre (7^e et 8^e chiffres de la 1^{re} ligne) : 08, accouplement homogame : le deuxième conjoint est de même type que le premier, donc le couple est de type $Z \times Z$: *cas 6*. La population est fixée, la structure de cette population va rester : $P = 2Q = 0,00$; $R = 1,00$ à chaque génération.

2^e exemple. — Population 5 (5^e ligne de la table V). Le couple géniteur de la génération zygotique utile zéro était $A \times H$ (cas 2). La structure probable de sa descendance est $P = 0,50$; $2Q = 0,50$; $R = 0,00$. Le choix du couple géniteur de la génération zygotique utile *un* est le suivant :

1^{er} nombre (7^e et 8^e chiffres de la 5^e ligne) : 96 le premier zygote est de type H

2^e nombre (9^e et 10^e chiffres de la 5^e ligne) : 86 accouplement panmictique

3^e nombre (11^e et 12^e chiffres de la 5^e ligne) : 19 le deuxième zygote est de type A

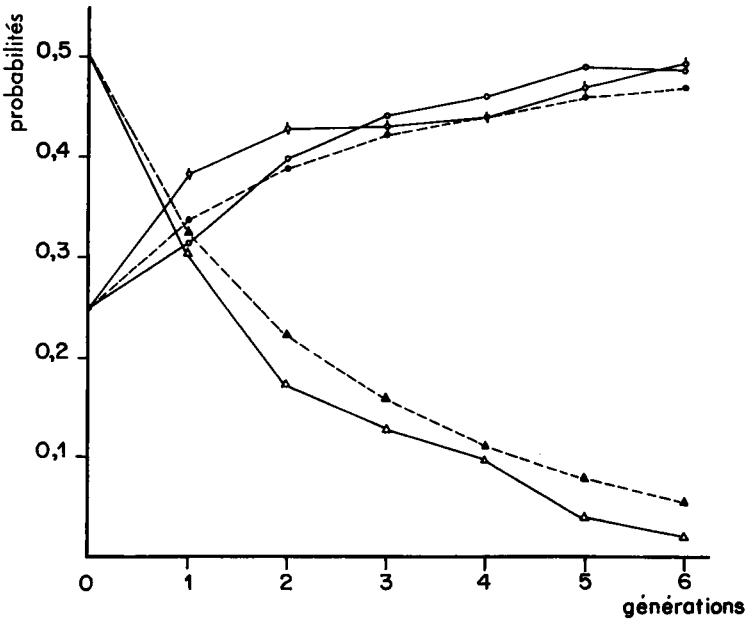


FIG. 1. — Comparaison de l'évolution moyenne de 100 populations homogames indépendantes (taux d'homogamie : $\lambda = 0,7$), aux courbes théoriques initiales.

- évolutions théoriques de $P(n)$ et de $R(n)$
- ▲----- évolution théorique de $2Q(n)$
- évolution moyenne de $P(n)$
- évolution moyenne de $R(n)$
- △----- évolution moyenne de $2Q(n)$

Le couple est de type $H \times A$ (ou $A \times H$) : cas 2. La structure probable de sa descendance est $P = 0,50$; $2Q = 0,50$; $R = 0,00$. Pour choisir le couple géniteur formant la génération zygotique utile *deux* il suffit de répéter l'opération avec les chiffres suivants de la cinquième ligne, etc.

Ce procédé est donc un moyen sûr et simple de simuler l'évolution de populations comportant à chaque génération un mâle et une femelle reproducteurs, issues d'une seule population infinie panmictique, et pratiquant toutes l'homogamie génotypique partielle avec la probabilité 0,7.

La moyenne des structures de ces 100 populations a été superposée à chaque génération au graphique où était représentée la structure moyenne probable calculée (fig. 1). Il apparaît sur cette figure que les évolutions théorique et expé-

rimentale, sont voisines sans cependant se superposer vraiment. Pour savoir si l'écart observé ici est dû au hasard ou à une imperfection logique, il suffit de réaliser une nouvelle série d'expériences. En reprenant la même expérience, avec cette fois 400 populations indépendantes (et en suivant un itinéraire plus compact sur la table de FISHER et YATES, 1963), on doit obtenir des résultats expérimentaux deux fois plus précis. Or, les tendances observées sur la figure 1 se confirment (fig. 2). Il existe un écart faible, mais certain, entre les courbes expérimentales et les prévisions théoriques.

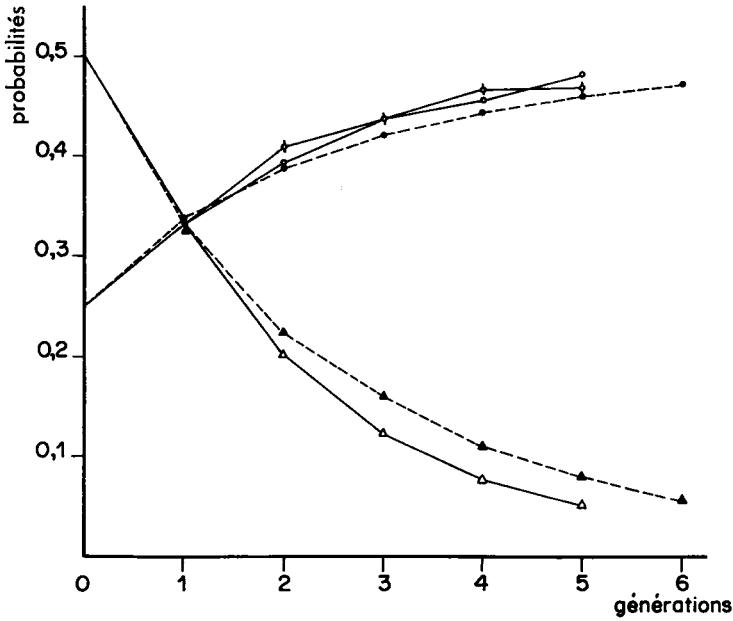


FIG. 2. — Comparaison de l'évolution moyenne de 400 populations homogames indépendantes (taux d'homogamie: $\lambda = 0,7$), aux courbes théoriques initiales.

C. — Les lois d'accouplement: cas des populations très restreintes.

L'écart observé entre les valeurs probables calculées et les résultats expérimentaux analogiques démontre l'existence d'une incohérence logique sans doute assez subtile dans la théorie de l'homogamie telle qu'elle est exposée plus haut. Une analyse poussée montre que c'est la définition de l'homogamie dans les petites populations qui porte en elle une contradiction. Si on choisit en effet les N_1 mâles et les N_2 femelles reproducteurs au hasard, il n'est pas certain que l'on pourra réaliser des accouplements homogames entre eux. Supposons par exemple que les générations zygotiques comptent 2 mâles et 2 femelles à chaque génération. Si par hasard les deux femelles sont différentes des deux mâles à une certaine génération, il ne sera pas possible de réaliser d'accouplement homogame à cette génération. Naturellement, une telle éventualité est d'autant moins probable que N_1 et N_2 sont grands. Il faut enfin noter que si $N_1 = N_2 = 1$, cette définition de

l'homogamie est absolument caduque, puisque le choix des conjoints devient purement aléatoire, donc indépendant de λ . Une autre façon de concevoir l'homogamie, où l'on tire différemment les générations zygotiques utiles, aboutit aux mêmes équations. Elle consiste à tirer un zygote quelconque de sa génération, puis selon les cas à lui choisir un conjoint, du même type ou au hasard, parmi les zygotes de l'autre sexe. L'on constitue ainsi un ensemble de zygotes. Mais il n'a aucune raison de compter à la fois N_1 mâles et N_2 femelles distincts exactement (N_1 et N_2 étant choisis avant les tirages).

Ces difficultés proviennent du fait que l'homogamie, comme toute loi d'appariement dans les petites populations, est définie par deux lois : d'une part celle qui préside au tirage des zygotes destinés à former la génération zygotique utile, d'autre part celle qui définit comment se réalisent les accouplements. Or, ces deux lois ne sont pas indépendantes : si l'on définit *a priori* l'une, alors l'autre ne doit pas être contradictoire avec elle; si le choix des zygotes de la génération zygotique utile est défini *a priori*, la probabilité de réaliser un accouplement homogame doit être conditionnée par la possibilité de réunir un couple homogame; si au contraire on définit *a priori* la loi d'accouplement, alors le choix des zygotes des générations zygotiques utiles en dépend, tant pour les nombres N_1 de mâles et N_2 de femelles qui les composent que pour le type de ces zygotes. En outre, dans ce cas, les couples ou groupes de couples formés ne peuvent être changés sans que soit changée l'épreuve qui a permis de les constituer.

Plus précisément, on doit alors distinguer deux sortes de zygotes : les « zygotes libres » dont le nombre et le choix sont arbitraires, et les « zygotes liés », dont le nombre et le type sont partiellement définis par les lois précédentes et le type des premiers. Le cas traité dans les expériences analogiques est de ce genre : la « loi *a priori* » est le mode d'accouplement : l'homogamie génotypique partielle de taux λ , ($0 \leq \lambda \leq 1$). La loi conditionnée est celle du choix des zygotes des générations zygotiques utiles : les zygotes libres sont N_1 mâles tirés au hasard de l'ensemble des mâles de leur génération. Les zygotes liés sont les femelles choisies par les mâles parmi l'infinité des femelles de la génération, selon le mode d'accouplement défini plus haut. La probabilité de tirer N_1 femelles distinctes est dans ce cas 1. On constitue donc N_1 couples disjoints. Ces couples sont évidemment stables, car des recombinaisons entre eux conduiraient à former des couples selon une autre catégorie d'épreuve.

D. — Équations stochastiques de l'homogamie dans le cas où la population est très petite

On est donc conduit, quand les populations sont très petites, à adopter une définition plus précise de l'homogamie. Considérons celle où la loi *a priori* est le mode d'accouplement, c'est-à-dire l'homogamie génotypique de taux λ , la loi conditionnée celle du choix des zygotes reproducteurs. Les zygotes libres sont les N_1 mâles à qui on choisit une femelle en suivant la loi d'accouplement, pour former N_1 couples disjoints et stables. Ces N_1 couples vont engendrer la génération suivante supposée très nombreuse, dont sera extraite selon la même loi une nouvelle génération zygotique utile de N_1 couples, etc.

Considérons alors la génération $G(n)$ ayant la structure : $\{P(n), 2Q(n), R(n)\}$. Tirons-en la génération zygotique utile à N_1 couples. Ces N_1 couples vont engendrer la génération zygotique $G(n + 1)$. Tirons au hasard un zygote de $G(n + 1)$.

Supposons qu'il provienne d'un couple homogame (probabilité λ). Ses parents avaient tous deux soit le génotype A/A (probabilité $P(n)$), dans ce cas les deux gènes du zygote considéré seront isocontraints, soit le génotype H/H (probabilité $2Q(n)$) dans ce cas les deux gènes de ce zygote auront une chance sur deux d'être isocontraints, et une chance sur deux d'être hétérocontraints, soit enfin le génotype Z/Z (probabilité $R(n)$), dans ce cas les deux gènes de ce zygote seront isocontraints. Dans aucun de ces cas ils ne seront non-contraints.

Supposons maintenant qu'il ne provienne pas d'un couple homogame (probabilité $1 - \lambda$); dans ce cas, les deux gènes de ce zygote seront dans la situation de contrainte de leurs gènes paternel et maternel. Or, le père et la mère sont deux zygotes quelconques de la génération $G(n)$, donc :

$$\begin{cases} F_1(n + 1) = \lambda(P(n) + Q(n) + R(n)) + (1 - \lambda) \Phi_1(n) \\ F_2(n + 1) = \lambda \cdot Q(n) + (1 - \lambda) \Phi_2(n) \\ F_3(n + 1) = (1 - \lambda) \Phi_3(n) \end{cases}$$

nous vérifions que $F_1(n + 1) + F_2(n + 1) + F_3(n + 1) = 1$.

Tirons au hasard maintenant deux zygotes de $G(n + 1)$.

Ces deux zygotes peuvent être pleins frères parce qu'issus du même couple (probabilité $1/N_1$). Considérons l'un des gènes du premier zygote et l'un des gènes du second. Ils peuvent être tous deux issus du père (probabilité $1/4$) ou tous deux de la mère (probabilité $1/4$), il ont donc une chance sur deux d'être issus du même parent. Dans ce cas, ils ont encore une chance sur deux d'être issus tous deux du même gène parental, ils sont alors identiques donc isocontraints. Ils ont une chance sur deux de n'être pas issus du même gène parental, alors ils seront dans la situation de contrainte des deux gènes de leur parent commun, donc isocontraints, hétérocontraints ou non-contraints avec les probabilités respectives $F_1(n)$, $F_2(n)$ et $F_3(n)$; ils ont aussi une chance sur deux d'être issus chacun d'un parent distinct de ce couple, c'est-à-dire que l'un des gènes est paternel, l'autre maternel. Dans ce cas, ils auront la même probabilité d'être dans une situation de contrainte donnée que les deux gènes d'un descendant de ce couple, qui reçoit de la même façon un gène paternel et un gène maternel, donc nous retrouvons les probabilités $F_1(n + 1)$, $F_2(n + 1)$, $F_3(n + 1)$.

Ces deux zygotes peuvent aussi être issus de deux couples distincts (probabilité : $1 - 1/N_1$). Considérons toujours l'un des gènes du premier et un des gènes du second. Ils proviennent de deux gènes tirés indépendamment de deux zygotes quelconques de la génération $G(n)$; la probabilité qu'ils soient dans une situation de contrainte donnée est celle que deux gènes quelconques tirés de deux zygotes quelconques de la génération $G(n)$ soient dans cette même situation de contrainte, donc :

$$\begin{cases} \Phi_1(n + 1) = (1/2N_1)[(1 + F_1(n))/2 + F_1(n + 1)/2] + (1 - 1/N_1) \Phi_1(n) \\ \Phi_2(n + 1) = (1/2N_1)[F_2(n)/2 + F_2(n + 1)/2] + (1 - 1/N_1) \Phi_2(n) \\ \Phi_3(n + 1) = (1/2N_1)[F_3(n)/2 + F_3(n + 1)/2] + (1 - 1/N_1) \Phi_3(n) \end{cases}$$

avec $1/N = (1/N_1 + 1/N_2)/4 = 1/2N_1$ si $N_1 = N_2$.

Nous avons donc en contractant :

$$\begin{aligned}
 \text{(I)} \quad & F_1(n+I) = \lambda(I - Q(n)) + (I - \lambda) \Phi_1(n) \\
 \text{(II)} \quad & F_2(n+I) = \lambda \cdot Q(n) + (I - \lambda) \Phi_2(n) \\
 \text{(III)} \quad & F_3(n+I) = (I - \lambda) \Phi_3(n) \\
 \text{(Ib)} \quad & \Phi_1(n+I) = (I + F_1(n)) / 2N + [I - (I + \lambda) / N] \Phi_1(n) + \lambda(I - Q(n)) / N \\
 \text{(IIb)} \quad & \Phi_2(n+I) = F_2(n) / 2N + [I - (I + \lambda) / N] \Phi_2(n) + \lambda \cdot Q(n) / N \\
 \text{(IIIb)} \quad & \Phi_3(n+I) = F_3(n) / 2N + [I - (I + \lambda) / N] \Phi_3(n) \text{ avec} \\
 & P(n) = F_1(n) \cdot p + F_3(n) \cdot p^2 \\
 & 2Q(n) = F_2(n) + 2F_3(n) \cdot p \cdot q \\
 & R(n) = F_1(n) \cdot q + F_3(n) \cdot q^2
 \end{aligned}$$

Il est évident qu'en posant $\lambda = 0$, on retrouve les équations de la panmixie. Cependant, ces équations ne recouvrent pas exactement le même phénomène; en effet dans la panmixie « normale », les N_1 mâles et N_2 femelles de la génération zygotique utile étaient pris au hasard *a priori* dans la génération zygotique, et la panmixie existait entre ces N_1 mâles et N_2 femelles, c'est-à-dire que l'on pouvait former jusqu'à $N_1 \cdot N_2$ couples distincts. Ici, au contraire, une femelle (par exemple) est choisie comme conjointe d'un mâle donné, donc nécessairement $N_2 = N_1$, il y a N_1 couples seulement, et en dehors de ces couples aucun accouplement ne se produit. La solution de ce système est tout à fait semblable à celle donnée au paragraphe II, à condition que $p = q = 0,5$. Dans ce cas en effet :

$$\begin{aligned}
 P(n) + Q(n) + R(n) &= I - Q(n) = F_1(n) + F_2(n) / 2 + 3F_3(n) / 4 \\
 Q(n) &= F_2(n) / 2 + F_3(n) / 4
 \end{aligned}$$

donc :

$$\begin{aligned}
 \text{(III)} \quad & \left\| F_3(n+I) \right\| = \left\| \begin{array}{cc} 0 & I - \lambda \\ I / 2N & I - (I + \lambda) / N \end{array} \right\| \cdot \left\| F_3(n) \right\| \\
 \text{(IIIb)} \quad & \left\| \Phi_3(n+I) \right\| = \left\| \begin{array}{cc} \lambda / 2 & I - \lambda \\ (I + \lambda) / 2N & I - (I + \lambda) / N \end{array} \right\| \cdot \left\| F_2(n) \right\| + \left\| \begin{array}{c} \lambda \cdot F_3(n) / 4 \\ \lambda \cdot F_3(n) / 4N \end{array} \right\|
 \end{aligned}$$

$\{F_1(n), \Phi_1(n)\}$ se résolvant comme complément à $\{I, I\}$ de $\{F_2(n) + F_3(n), \Phi_2(n) + \Phi_3(n)\}$. Les simplifications introduites au paragraphe II sont valables ici, donc il suffit de calculer la solution du système sans second membre associé à (II, II bis) notée $\{F''_2(n), \Phi''_2(n)\}$ soit en notation vectorielle :

$$\left\| \begin{array}{cc} F''_2(n+I) \\ \Phi''_2(n+I) \end{array} \right\| = \left\| \begin{array}{cc} \lambda / 2 & I - \lambda \\ (I + \lambda) / 2N & I - (I + \lambda) / N \end{array} \right\| \cdot \left\| F''_2(n) \right\|$$

les valeurs propres notées s_1 et s_2 vérifient l'équation en s :

$$\left| \begin{array}{cc} \lambda / 2 - s & I - \lambda \\ (I + \lambda) / 2N & I - (I + \lambda) / N - s \end{array} \right| = 0$$

$s_1 = [\lambda(N - 2) + 2(N - I) + \delta] / 4N$; $s_2 = [\lambda(N - 2) + 2(N - I) - \delta] / 4N$
avec $\delta^2 = 4[\lambda^2(N_1 - I)^2 - 2\lambda(2N_1 - I)(N_1 + I) + (4N_1^2 + I)] \geq 0$ si $0 \leq \lambda \leq I$
 $F_2''(n) = A \cdot s_1^n + B \cdot s_2^n$, A et B étant définis par les conditions à l'origine.

$A + B = F_2''(0) = 2Q(0)$; $A \cdot s_1 + B \cdot s_2 = F_2''(I) = 2Q(I)$

$$A = 4N(Q(I) - s_2 \cdot Q(0)) / \delta; \quad B = -4N(Q(I) - s_1 \cdot Q(0)) / \delta$$

Exemples: Considérons une population initiale infinie panmictique. Tirons en 1 mâle et 1 femelle qui seront les deux premiers géniteurs d'une lignée frère sœur où $N_1 = N_2 = 1$ à chaque génération, donc où $N = 2$.

$$\delta_2 = 4 - 16(\lambda - 1) = 20 - 16\lambda = 4(5 - 4\lambda)$$

puisque $p = q = 0,5$, $P(0) = 0,25$; $2Q(0) = 0,50$; $R(0) = 0,25$

$$s_1 = (2 + 2\sqrt{5 - 4\lambda})/8 = (1 + \sqrt{5 - 4\lambda})/4; s_2 = (1 - \sqrt{5 - 4\lambda})/4$$

$$F_1(0) = F_2(0) = 0; F_3(0) = 1; \Phi_1(0) = \Phi_2(0) = 0; \Phi_3(0) = 1$$

$$2Q(0) = 0,5; 2Q(1) = 0,25\lambda + 0,5(1 - \lambda) = 0,5 - 0,25\lambda$$

$$A + B = 0,5 \quad A \cdot s_1 + B \cdot s_2 = 0,5 - 0,25\lambda$$

$$A = + (0,75 - 0,5\lambda + 0,25\sqrt{5 - 4\lambda}) / (5 - 4\lambda)$$

$$B = - (0,75 - 0,5\lambda - 0,25\sqrt{5 - 4\lambda}) / (5 - 4\lambda)$$

Le calcul automatique de l'évolution moyenne probable a été réalisé sur calculatrice électronique (Olivetti P.101), pour λ variant par dixièmes de zéro à un (Tableau 1).

TABLEAU 1

Évolution théorique de 2 Q(n) en fonction de n et de λ

n	λ=0,0	λ=0,1	λ=0,2	λ=0,3	λ=0,4	λ=0,5	λ=0,6	λ=0,7	λ=0,8	λ=0,9	λ=1,0
0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0
1	50,0	47,5	45,0	42,5	40,0	37,5	35,0	32,5	30,0	27,5	25,0
2	37,5	35,0	32,5	30,0	27,5	25,0	22,5	20,0	17,5	15,0	12,5
3	31,2	28,2	25,2	22,4	19,7	17,2	14,7	12,4	10,2	8,2	6,3
4	25,0	22,0	19,1	16,5	14,0	11,7	9,6	7,7	6,0	4,5	3,1
5	20,3	17,3	14,6	12,2	10,0	8,0	6,3	4,8	3,5	2,4	1,6
6	16,4	13,6	11,1	9,0	7,1	5,5	4,1	3,0	2,1	1,3	
7	13,3	10,7	8,5	6,6	5,0	3,7	2,7	1,8	1,2		
8	10,7	8,4	6,5	4,9	3,6	2,6	1,8				
9	8,7	6,6	4,9	3,6	2,5	1,7					
10	7,0	5,2	3,8	2,6	1,8						
11	5,7	4,1	2,9	2,0							
12	4,6	3,2	2,2								

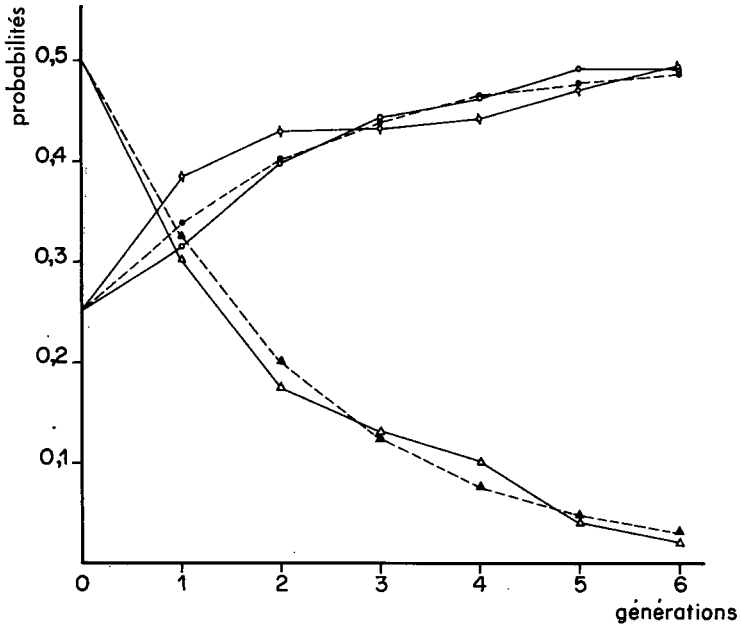


FIG. 3. — Comparaison de l'évolution moyenne de 100 populations homogames indépendantes (taux d'homogamie: $\lambda = 0,7$), aux courbes théoriques corrigées.

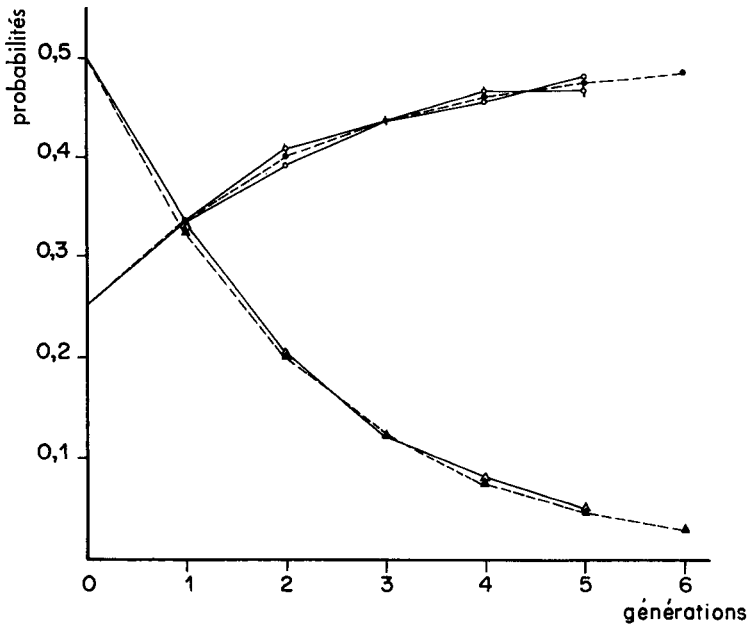


FIG. 4. — Comparaison de l'évolution moyenne de 400 populations homogames indépendantes (taux d'homogamie: $\lambda = 0,7$), aux courbes théoriques corrigées.

Lorsqu'on compare maintenant les courbes théoriques pour $\lambda = 0,7$ aux courbes moyennes données par les expériences analogiques précédentes, on constate aussitôt leur remarquable superposition (fig. 3 et 4).

DISCUSSION

Ce premier travail ne comporte aucune étude de la variabilité de l'évolution moyenne réelle de ces populations autour de sa valeur moyenne probable, car le calcul des moments du deuxième ordre est ardu. Il serait cependant possible de s'en faire une idée qualitative grâce aux expériences analogiques, en les considérant maintenant comme un processus de Monte Carlo, mais les valeurs (numériques) calculées ne permettraient en rien de faire progresser la théorie; la question demeure donc non résolue. Cependant, ces expériences analogiques qui étaient très schématiques puisque destinées à tester le modèle choisi, auront déjà permis :

1° d'affiner la théorie de l'homogamie en général, et à partir de cette nouvelle réflexion de préciser quelles conditions doivent observer les définitions des lois d'accouplements dans les petites populations pour être cohérentes;

2° d'acquérir la quasi-certitude que la théorie de l'homogamie, telle qu'elle est exposée ici (c'est-à-dire avec $p = q = 0,5$) est maintenant achevée en tant que modèle mathématique.

Remarquons enfin que les équations de l'homogamie dans les petites populations où la loi *a priori* est le tirage des zygotes géniteurs, et la loi conditionnée celle des modes d'accouplements, sont très lourdes à établir et ne sont pas linéaires. Elles sont d'autant plus complexes que N_1 et N_2 sont grands car ils interviennent alors comme exposants, mais ne diffèrent des équations de GILLOIS (1968) que par des termes correctifs, négligeables dès que N_1 et N_2 sont supérieurs à une dizaine. Il y a donc tout avantage à utiliser ces équations approchées de GILLOIS dans la plupart des cas.

Reçu pour publication en décembre 1969.

REMERCIEMENTS

Nous tenons à remercier ici M. MALECOT et M. GILLOIS pour leurs conseils et pour les corrections qu'ils nous ont permis d'apporter à ce texte.

SUMMARY

STUDY OF THE EVOLUTION OF HOMOGAMIC LITTLE POPULATIONS BY THE THEORY OF CONSTRAINT: ANALOGIC CONTROL EXPERIMENTS

The theories of Identity and of Constraint introduced by GILLOIS are well known. The second is relatively compound and must be tested at least at two levels, before it can be applied to animal or vegetal populations. We must verificate first its internal coherence—which is the logic of the reasonings, the notions used and of the hypothesis between themselves—, and then its external coherence—that is to say the precision with which this mathematical model

describes the studied phenomena. The "analogical experiments" described here, which deal with populations, of pseudo-zygotes, defined by their sex, their genotype and an aleatory variable, are a partially experimental and quick method which allows testing the internal coherence of the theory, because it is sure that these populations respect the fundamental and restrictive hypothesis, particularly the fail of selection on the considered character. Applied to the case of partial genotypic homogamy in small populations, this method showed that the usual definition of this type of matings contained a contradiction. Indeed the even uniformly aleatory choice of the breeding animals and their type of matings between themselves are non independent laws, which all cannot be defined *a priori*. Consequently, it is necessary to precise in the definition of the processes of matings in small populations, which is the *a priori* law, and how the other laws are conditioned by it, and between themselves, and to verify that they are not contradictory. The theory of the partially genotypic homogamy in small populations, where the *a priori* law is the type of matings, and the conditioned law the choice of the breeding animals, which is here treated by the theory of the absolute Constraint, is a logically coherent model and allows to calculate before hand the probable mean evolution of homogamous populations with a great precision, as it is shown by the analogical experiments. It leads to distinguish two classes of zygotes, the "free zygotes", which are chosen independently of the type of mating, and the "bound zygotes" whose choice depends upon it and upon his chosen partner.

RÉFÉRENCES BIBLIOGRAPHIQUES

- CROIZÉ-POURCELET J., 1968. Étude théorique et expérimentale de l'évolution de petites populations par la notion de Contrainte. *Thèse Sc. appl. Orsay*. Série A, n° 419, 127 p. + annexes.
- FISHER R.A., YATES F., 1963. Statistical tables for biological, agricultural and medical research. Oliver and Boyd. London, 146 p.
- GILLOIS M., 1964. La relation d'identité en génétique. *Thèse Doct. État*, Fac. Sci. Paris, Série A, 294 p.
- GILLOIS M., 1965 (a). La relation de dépendance en génétique. *Journée d'Études Féd. europ. Zootech.*, La Haye, juin 1965, 29 p.
- GILLOIS M., 1965 (b). Entropie de la description d'une population. *Journée d'Études Féd. europ. Zootech.*, La Haye, juin 1965, 4 p.
- GILLOIS M., 1966 (a). La relation de dépendance en génétique. *Ann. Inst. Henri Poincaré*. Section B., 2, 261-278.
- GILLOIS M., 1966 (b). L'homogamie dans une population d'effectif limité. *Ann. Inst. Henri Poincaré*. Section B., 2, 299-347.
- GILLOIS M., 1968. Logique et mathématique de l'hérédité, *I.N.R.A.* Paris (à paraître).
- MALECOT G., 1948. Les mathématiques de l'hérédité, republiées dans : MALECOT G., 1966. Probabilité et hérédité. *I.N.E.D.* cahier n° 47. P.U.F.
- WRIGHT S., 1925. An approximate method of calculating coefficient of inbreeding and relationship from livestock pedigrees. *J. Agric. Res.*, 13, 377-384.