

GÉNÉTIQUE DES POPULATIONS NATURELLES DANS LE CAS D'UN SEUL LOCUS

II. — ÉTUDE DU COEFFICIENT DE PARENTÉ

G. MALÉCOT

*Mathématiques appliquées (Université de Lyon) B. P. n° 37,
69 - Villeurbanne-Charpennes*

RÉSUMÉ

Les notations étant les mêmes que dans la partie I, les variances et covariances *a priori* des fréquences q d'un même gène a dans des groupes panmictiques d'effectif fini N communiquant par migration sont maintenant ramenées [chap. II, 3°] aux « coefficients de parenté ».

Le coefficient de parenté φ_{xy} entre le groupe d'emplacement x et le groupe d'emplacement y n'est autre que la probabilité d'« identité » (par descendance mendélienne à partir d'un ancêtre commun) de deux loci tirés au sort *indépendamment* dans les emplacements x et y ; la notion de « bétamutation » (passage à l'indépendance, avec probabilité k dans chaque transmission de locus à la génération suivante ⁽¹⁾) permet de prendre, comme alternative à l'identité sans bétamutation, l'indépendance stochastique qui a donc la probabilité *a priori* $1 - \varphi_{xy}$.

Le calcul de φ_{xx} — qui est aussi le coefficient de consanguinité d'un diploïde appartenant à un groupe panmictique d'emplacement x — permet (chap. III) de retrouver, après multiplication par $C - C^2$, les résultats obtenus dans la partie I pour la variance intergroupe.

La décroissance, en fonction de la distance $y - x = d$, du coefficient de parenté $\varphi_{xy} = \varphi(d)$, ainsi que de la covariance et de la corrélation qui lui sont proportionnelles, est, dans le cas unidimensionnel, [chap. III, II, 1°] une décroissance exponentielle qui ne dépend que de k et de la variance σ^2 de la migration (résultat déjà établi dans la partie I). Dans le cas de migration bidimensionnelle discontinue [II, 2°] la différence numérique avec le cas unidimensionnel est faible sauf si k/σ^2 est très petit, alors la tendance vers zéro aux grandes distances est un peu plus rapide (ce que confirme le modèle de la migration continue isotrope).

⁽¹⁾ k coïncide, lorsque la sélection est absente ou lorsqu'elle est linéarisée, avec le coefficient noté également k dans la partie I.

CHAPITRE I

COEFFICIENT DE PARENTÉ (ENTRE GROUPES PANMICTIQUES)
(EN L'ABSENCE DE SÉLECTION)

Nous avons vu comment ⁽¹⁾ lorsqu'on connaît la génération initiale F_0 , on peut évaluer des probabilités *a priori* pour les générations successives F_n , des espérances et covariances *a priori* (ainsi que leurs limites quand n augmente indéfiniment).

Une méthode très différente peut être utilisée *dans le cas où il n'y a pas sélection* : elle consiste à utiliser ce que nous appellerons les *coefficients de parenté dans F_n* , c'est-à-dire les probabilités *a priori* pour que deux loci tirés au hasard chez deux individus de F_n (ou deux gamètes tirés au hasard dans le réservoir gamétique pour F_n) dérivent par descendance mendélienne et sans aucune mutation d'un même locus d'un ancêtre commun appartenant à une génération F_p ($0 \leq p < n$) : on qualifiera de *loci identiques* deux tels loci, qui portent forcément le même gène. Cette probabilité *a priori* sera notée $\varphi_{xy}(n)$ pour deux loci pris au hasard, l'un dans l'emplacement x , l'autre dans l'emplacement y . Que signifiera la non identité [événement de probabilité *a priori* $1 - \varphi_{xy}(n)$] ? Deux cas sont à distinguer :

1° *Aucune mutation* n'a jamais lieu (nous avons vu dans l'introduction que, dans ce cas, la probabilité *a priori* \bar{q}_{nx} pour un locus tiré au hasard dans un groupe isolé de porter le gène a est égale à la fréquence initiale q_0 de ce gène (que nous supposons, pour simplifier, indépendante de l'emplacement). Dans ce cas $1 - \varphi$ est la probabilité *a priori* pour que les deux loci tirés descendent de deux loci différents de la génération initiale ; nous les supposons alors indépendants en probabilité ⁽²⁾).

2° *Des mutations* sont, à chaque génération, possibles, avec probabilité u , pour chaque gène a , de muter en A , et probabilité v , pour chaque gène A , de muter en a (u et v sont, dans la pratique, extrêmement faibles). Nous savons qu'alors la probabilité *a priori* \bar{q}_{nx} évolue et tend vers la limite $\bar{q} = \frac{v}{u+v}$ ⁽³⁾ ; alors, la non identité implique trois éventualités possibles.

a) Les deux loci descendent, avec ou sans mutation, de deux loci différents de la génération initiale ; ils sont alors *indépendants*, dans l'hypothèse formulée au 1°.

b) Les deux loci descendent d'un même locus d'un ancêtre commun qui a subi en tout un nombre pair (> 0) de mutations ; alors, quoique non identiques, ils sont certainement porteurs du même gène ⁽⁴⁾. Nous dirons qu'ils sont *absolument dépendants isoactifs* ⁽⁵⁾.

⁽¹⁾ *Ann. Génét. Sél. anim.*, 1971, 3. (3), 255-280.

⁽²⁾ Ceci exige que la génération initiale soit panmictique sur toute l'aire occupée par l'espèce ; cette hypothèse trop restrictive ne pourra être abandonnée que quand nous aurons montré que, si la génération initiale F_0 est suffisamment reculée dans le temps, ou, ce qui revient au même, si l'on suppose le numéro n de la génération F_n suffisamment grand, les relations de parenté et de consanguinité, donc d'identité ou d'indépendance, qui peuvent exister dans F_0 n'ont plus d'influence sur la valeur numérique $\varphi_{xy}(n)$ du coefficient de parenté.

⁽³⁾ Il s'agit là de la quantité précédemment appelée C ; dans le présent chapitre, où elle est interprétée comme probabilité *a priori*, nous la noterons \bar{q} .

⁽⁴⁾ Cette remarque n'est valable que dans le cas de diallélisme.

⁽⁵⁾ GILLOIS (1966), auteur de cette expression, réunit à ce cas celui de l'identité (nombre nul de mutations).

c) Les deux loci descendent d'un même locus qui a subi en tout un nombre impair de mutations ; alors ils sont porteurs de deux gènes différents : si l'un porte A , l'autre a ⁽¹⁾ ; on dit, qu'ils sont *absolument dépendants hétéroactifs*.

L'existence, pour 2 loci dialléliques, de quatre cas possibles, crée des complications ⁽²⁾. Il est possible, *dans le cas de diallélisme* de se ramener à une seule alternative : identité ou indépendance ; il suffit d'imaginer que, sur chaque chaînon parental, soit possible, une fictive « bêtamutation » affectant le locus avec une probabilité *a priori* k indépendante du gène qui porte le locus, et transformant ce gène en un gène indépendant (précisons : en un gène tiré au sort avec les probabilités \bar{q} et \bar{p} d'être a ou A). Pour qu'un gène (quelconque) subisse (sur un chaînon fixe) une bêtamutation le transformant en un gène a (resp. A), les probabilités sont alors $k\bar{q}$ et $k\bar{p}$.

Or les taux u et v des mutations observables ou « alphas mutations » (GILLOIS cité plus haut) doivent coïncider avec ces probabilités.

D'où les conditions $u = k\bar{p}$ et $v = k\bar{q}$, qui fournissent

$$k = u + v$$

en plus de la condition d'équilibre $\bar{q} = \frac{v}{u + v}$.

3° Nous pourrions donc dorénavant (dans le cas stationnaire du diallélisme) considérer que, sur chaque chaînon, il y a probabilité $1 - k$ de « persistance par non-bêtamutation », et probabilité $k = u + v$ de « bêtamutation faisant passer à l'indépendance » (avec probabilités \bar{q} et \bar{p} pour a et A).

Cette nouvelle formulation statistique de l'incidence des mutations (même dans le cas de multiallélisme, avec égalité de tous les taux de mutation vers chaque allèle) valable rigoureusement dans le cas de diallélisme, reste approximativement valable aussi lorsque chaque mutation donne naissance à l'un quelconque parmi une nombreuse série d'allèles, comme c'est le cas pour les substitutions d'amino-acides dans les chaînes protéiniques, chaque substitution pouvant probablement être regardée, suivant la suggestion de KIMURA (1968 a), comme donnant naissance à un allèle nouveau.

Nous supposerons dorénavant que, pour *chaque transmission d'un locus le long d'un chaînon, le maintien de l'identité a la probabilité $1 - k$ et le passage à l'indépendance la probabilité k* ⁽³⁾.

4° Pour deux individus I et J dont on connaîtrait les chaînes de parenté qui les relie (chaque chaîne étant la réunion de deux chaînes d'ascendance remontant, par n_i et p_i chaînons, à un premier ancêtre commun A_i) deux loci tirés au hasard, l'un chez I , l'autre chez J , auraient la probabilité $\left(\frac{1}{2}\right)^{n_i + p_i}$ de provenir tous deux de A_i , auquel cas ils ont la probabilité $\frac{1}{2}$ de provenir du même locus de A_i (ce qui entraîne

⁽¹⁾ Cette remarque n'est valable que dans le cas de diallélisme.

⁽²⁾ Notons que, si la dépendance résulte d'une homogamie basée sur la ressemblance des gènes, des génotypes, ou des phénotypes, sa probabilité ne peut pas être calculée indépendamment de la fréquence des gènes (alors que cette simplification se produira lorsqu'on calculera seulement les probabilités d'identité ou de « bêtamutation » définies plus loin).

⁽³⁾ Dans l'application aux mutations biochimiques, k est le taux total de survenue, dans le locus considéré, de tous les changements biochimiques possibles ; l'identité doit être, *a contrario*, entendue au sens biochimique le plus fin, et non seulement au sens des manifestations génotypiques habituelles ; k sera alors plus grand, pour un locus que pour un site biochimique et le coefficient de parenté sera plus faible pour un locus donné que pour un site donné.

leur identité) et la probabilité $\frac{1}{2} f_{A_i}$ de provenir des deux loci homologues *et* d'être identiques, (f_{A_i} désignant la probabilité *a priori* pour que les deux loci homologues de A_i soient identiques, probabilité qu'on appelle le coefficient de consanguinité de A_i). Comme la probabilité de maintien de l'identité (non-bêtamutation) sur les $n_i + p_i$ chaînons est $(1 - k)^{n_i + p_i}$ la probabilité pour que les deux loci de I et J soient identiques *par transmission sans bêtamutation depuis A_i le long de la chaîne de parenté considérée* est :

$$\left(\frac{1 - k}{2}\right)^{n_i + p_i} \frac{1 + f_{A_i}}{2}.$$

Il peut exister entre I et J plusieurs chaînes de parenté distinctes par au moins un chaînon (correspondant à des ancêtres communs distincts ou non) (fig. 1). Comme la transmission le long d'une chaîne exclut la transmission le long de toute autre chaîne, le coefficient de parenté φ_{IJ} de I et de J sera :

$$\varphi_{IJ} = \sum \left(\frac{1 - k}{2}\right)^{n_i + p_i} \frac{1 + f_{A_i}}{2}.$$

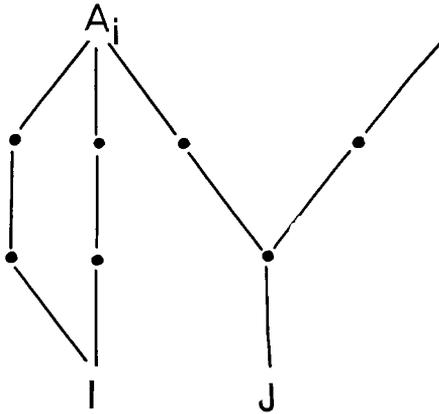


FIG. 1. — Exemple de liaison de parenté entre les individus I et J par l'intermédiaire de l'ancêtre commun A_i

La sommation étant étendue sur toutes les chaînes distinctes. Quant à f_{A_i} , coefficient de consanguinité de A_i , il est le produit par $(1 - k)^2$ du coefficient de parenté de ses parents.

En définitive, pour deux loci pris chez deux individus I et J (resp. pour les deux loci homologues de A_i) : φ_{IJ} (resp. f_{A_i}) représente la probabilité d'identité ; $1 - \varphi_{IJ}$ (resp. $1 - f_{A_i}$) représente la probabilité d'indépendance.

5° Mais cette formule ne permettrait l'évaluation *a priori* du coefficient de parenté dans F_n que si tous les individus de F_n avaient deux à deux la même parenté connue d'avance (« système de croisement régulier » de SEWALL-WRIGHT (1931). Or, dans un ou plusieurs groupes panmictiques où les croisements sont aléatoires, les parentés aussi sont aléatoires. Dans ce cas, pour deux individus I et J pris au

hasard, on peut seulement définir les probabilités d'avoir des ancêtres communs de divers ordres (ces probabilités sont supérieures à zéro dans un groupe panmictique d'effectif limité, ou dans plusieurs tels groupes communiquant par migration). On peut aussi obtenir une équation de récurrence pour la fonction $\varphi_{xy}(n)$ qui désigne (rappelons-le) la probabilité d'identité de deux loci de F_n tirés au hasard, l'un chez un individu I d'emplacement x , l'autre chez un individu J d'emplacement y [si les deux emplacements x et y coïncident, $\varphi_{xx}(n)$ désigne la probabilité d'identité de deux loci de deux individus *distincts* I et J tirés au hasard indépendamment dans le même emplacement x ; et aussi, *dans le cas de panmixie*, la probabilité d'identité des deux loci *homologues* d'un même individu en x , ou « coefficient de consanguinité »].

Le calcul, pour un groupe et pour plusieurs groupes panmictiques, sera fait au chapitre III. Mais, au préalable, nous allons établir les importantes relations entre coefficients de parenté, probabilités *a priori*, et variances *a priori* [ce qui expliquera parfaitement le parallélisme entre les résultats du chapitre III (plus loin) et du chapitre III de la partie I].

CHAPITRE II

APPLICATIONS DU COEFFICIENT DE PARENTÉ

I. — Corrélation gamétique *a priori*

Si, dans F_n nous tirons au hasard un gamète ou un locus dans l'emplacement x , et, indépendamment, un autre gamète ou locus dans l'emplacement y (emplacement coïncidant ou non avec x), et que nous attachions à ces deux gamètes, deux variables aléatoires X ou X' égales à 1 lorsque le locus porte a et à 0 lorsque le locus porte A (les probabilités *a priori* respectives étant \bar{q}_n et \bar{p}_n , que nous supposons indépendants du groupe, donc $E(X) = E(X') = \bar{q}_n$), le coefficient de corrélation *a priori* entre X et X' peut se calculer à partir de la probabilité *a priori* $\varphi_{xy}(n)$ pour que X et X' soient « égales par identité » (« identiques » dirons-nous en bref) et de la probabilité $1 - \varphi_{xy}(n)$ pour qu'elles soient indépendantes (avec probabilités marginales \bar{q}_n et \bar{p}_n que l'on peut dorénavant prendre égales à \bar{q} et \bar{p} si la génération initiale est suffisamment lointaine); en désignant par E_p ou E_c les espérances mathématiques conditionnées par l'hypothèse d'indépendance ou par l'hypothèse d'identité, nous avons :

$$E[(X - \bar{q})(X' - \bar{q})] = [1 - \varphi_{xy}(n)]E_p[(X - \bar{q})(X' - \bar{q})] + \varphi_{xy}(n)E_c[(X - \bar{q})(X' - \bar{q})].$$

Mais l'espérance E_p du produit est nulle en raison de l'indépendance, et l'espérance E_c n'est autre que :

$$E[(X - \bar{q})^2] = \bar{q} - \bar{q}^2.$$

Le coefficient $\frac{E[(X - \bar{q})(X' - \bar{q})]}{\sqrt{E[(X - \bar{q})^2]E[(X' - \bar{q})^2]}}$ n'est donc autre asymptotiquement

que le coefficient de parenté $\varphi_{xy}(n)$, qui mérite le nom de *corrélacion gamétique a priori* (entre les emplacements x et y) ⁽¹⁾.

En particulier $\varphi_{xx}(n)$ est la corrélation gamétique *a priori* entre deux gamètes distincts tirés au sort indépendamment dans le groupe d'emplacement x , donc, s'il s'agit d'un *groupe panmictique*, le coefficient de corrélation *a priori* entre les gamètes qui s'unissent, et par suite le coefficient de corrélation *a priori* entre les aléatoires attachées aux deux loci homologues d'un même individu ⁽²⁾.

2. — Probabilités a priori des génotypes

Les probabilités *a priori* \overline{P}_n , $2\overline{Q}_n$, \overline{R}_n , pour qu'un individu I (diploïde) tiré au hasard dans F_n dans le groupe *panmictique* d'emplacement x présente chacun des trois génotypes AA , Aa , aa , sont les probabilités respectives de ces trois combinaisons sur deux gamètes tirés au sort indépendamment dans le groupe d'emplacement x , gamètes qui ont la probabilité $\varphi_{xx}(n)$ de porter deux loci identiques et la probabilité $1 - \varphi_{xx}(n)$ de porter deux loci indépendants.

On a donc :

$$\begin{aligned}\overline{P}_n &= \varphi_{xx}(n)\overline{p} + [1 - \varphi_{xx}(n)]\overline{p}^2 \\ 2\overline{Q}_n &= 2[1 - \varphi_{xx}(n)]\overline{p}\overline{q} \\ \overline{R}_n &= \varphi_{xx}(n)\overline{q} + [1 - \varphi_{xx}(n)]\overline{q}^2.\end{aligned}$$

Or, les probabilités *a priori* (pour un diploïde tiré au sort) sont aussi les espérances *a priori* des fréquences des génotypes dans F_n (même raisonnement que dans l'introduction de la partie I).

On peut de même calculer les probabilités *a priori* de croisement des différents types à partir des coefficients d'identité restreinte de GILLOIS (1964).

Les probabilités *a priori* expriment aussi, (MALÉCOT, 1959, p. 182) les fréquences moyennes des croisements des différents types évaluées sur un grand nombre de groupes (de même coefficient φ), ou sur un grand nombre de générations, ou sur un grand nombre de caractères.

3. — Variances et covariances a priori (entre groupes panmictiques)

Désignons par X_j ($1 \leq j \leq 2N$) les aléatoires attachées, dans la génération F_n , aux $2N$ loci du groupe panmictique occupant l'emplacement x .

Soient de même X'_k ($1 \leq k \leq 2N$) les aléatoires attachées au groupe panmictique (supposé de même effectif N) occupant l'emplacement y .

⁽¹⁾ Si les aléatoires X et X' prennent chacune (avec probabilités \overline{q} et \overline{p}) des valeurs numériques quelconques t et s (au lieu de 1 et 0), ces valeurs pourraient se déduire de 1 et 0 par une transformation linéaire ; donc le coefficient de corrélation resterait le même. Il en résulte que, si t et s sont les contributions numériques apportées par les gènes a et A , à un caractère héréditaire quantitatif, $\varphi_{xy}(n)$ est encore interprétable comme coefficient de corrélation entre les contributions numériques apportées par les deux loci considérés. Nous développerons dans le chapitre de génétique quantitative, les applications de cette remarque (qui devront tenir compte de ce que *chaque* individu a reçu *deux* loci).

⁽²⁾ Nous ne distinguons pas les sexes et admettons que l'autofécondation est possible (sans quoi une légère correction serait nécessaire) (MALÉCOT, 1951).

On a :

$$q_{nx} = \frac{\sum_j X_j}{2N}, \quad q_{ny} = \frac{\sum_k X_k}{2N}$$

Les X_j et X_k ayant toujours pour espérance mathématique la probabilité *a priori* \bar{q} (voir I°), \bar{q} est aussi l'espérance mathématique de q_{nx} et q_{ny} et l'on a :

$$\begin{aligned} \sigma_x^2(n) &= E[(q_{nx} - \bar{q})^2] = E \left\{ \left[\frac{1}{2N} \sum_{j=1}^{j=2N} (X_j - \bar{q}) \right]^2 \right\} \\ &= \frac{1}{4N^2} \sum_{j=1}^{j=2N} E[(X_j - \bar{q})^2] + 2 \sum_{(jk)} E[(X_j - \bar{q})(X_k - \bar{q})] \end{aligned}$$

La notation $\sum_{(jk)}$ désignant une sommation étendue aux $\frac{2N(2N-1)}{2}$ combinaisons de j et de k .

Or, nous avons vu que $E[(X_j - \bar{q})^2]$ de chaque aléatoire est asymptotiquement égale à $\bar{q} - \bar{q}^2$: la covariance *a priori* $E[(X_j - \bar{q})(X_k - \bar{q})]$ de deux aléatoires attachées à deux loci tirés au hasard indépendamment dans l'emplacement x est le produit de leur coefficient de corrélation $\varphi_{xx}(n)$ par la variance commune :

$$E[(X_j - \bar{q})^2] = \bar{q} - \bar{q}^2.$$

On a donc :

$$\begin{aligned} \sigma_x^2(n) &= \bar{q}(1 - \bar{q}) [2N + 2N(2N - 1) \varphi_{xx}(n)] / 4N^2 \\ &= \left(1 - \frac{1}{2N} \right) \varphi_{xx}(n) + \frac{1}{2N} \bar{q}(1 - \bar{q}). \end{aligned}$$

Cette formule reste valable si l'on ne mesure la fréquence locale que sur un échantillon au hasard de s individus prélevés parmi les N diploïdes : il suffit de remplacer N par s (voir chap. III, § IV).

On a de même, pour la covariance asymptotique entre q_{nx} et q_{ny} :

$$\begin{aligned} E[(q_{nx} - \bar{q})(q_{ny} - \bar{q})] &= \sum_j \sum_k E[(X_j - \bar{q})(X_k - \bar{q})] / 4N^2 \\ &= \varphi_{xy}(n) E[(X_j - \bar{q})^2] = \varphi_{xy}(n) \bar{q}(1 - \bar{q}) \end{aligned}$$

formule qui englobe comme cas particulier celle qui donne $\sigma_x^2(n)$ lorsque N ou s est assez grand [plus précisément : lorsque $1/2N$ ou $1/2s$ est plus petit par rapport à $\varphi_{xx}(n)$].

Ainsi se trouve établie une relation générale d'une extrême simplicité : la covariance asymptotique entre les fréquences de deux groupes (ou la variance de la fréquence d'un groupe suffisamment nombreux) est égale au produit de $\bar{q}(1 - \bar{q})$ par le coefficient de parenté (entre les deux groupes ou à l'intérieur du groupe).

Le problème, traité au chapitre IV (partie I), des covariances, est donc ramené au calcul du coefficient de parenté, qui fera l'objet du chapitre suivant.

CHAPITRE III

CALCUL DES COEFFICIENTS
DE CONSANGUINITÉ ET DE PARENTÉ

I. — Coefficients de consanguinité d'un groupe panmictique isolé

I. Cas d'une population monoïque.

Tous les individus (au nombre de N) produisant des gamètes des deux sexes et contribuant également au réservoir gamétique (voir note (1), p.).

Le coefficient de consanguinité f_n d'un individu pris au hasard dans F_n est la probabilité d'identité des deux loci homologues d'un gamète mâle et d'un gamète femelle pris au hasard et indépendamment dans le réservoir gamétique pour F_n . Or, la probabilité (d'ailleurs indépendante du sexe) pour que ces deux gamètes proviennent du même individu de F_{n-1} est $1/N$; la probabilité pour que leurs loci proviennent du même locus d'un individu de F_{n-1} est $1/2N$ et la probabilité conditionnée pour qu'ils soient identiques est $(1-k)^2$; s'ils proviennent de deux loci différents (du même individu ou d'individus différents avec probabilité totale $1 - 1/2N$), ces deux loci ont été apportés à F_{n-1} par deux gamètes différents résultant de deux tirages indépendants dans le réservoir gamétique pour F_{n-1} , leur probabilité d'identité est alors $(1-k)^2 f_{n-1}$.

On obtient ainsi la récurrence :

$$f_n = \left(1 - \frac{1}{2N}\right) (1-k)^2 f_{n-1} + \frac{1}{2N} (1-k)^2$$

qui se résout par :

$$f_n = \frac{\frac{1}{2N} (1-k)^2}{1 - \left(1 - \frac{1}{2N}\right) (1-k)^2} + \left(1 - \frac{1}{2N}\right)^n (1-k)^{2n} \times \text{cte.}$$

f_n tend donc, quand $n \rightarrow \infty$, vers une limite qui, en négligeant k^2 , s'écrit :

$$\sim f = \frac{1/2N}{1/2N + 2k} = \frac{1}{1 + 4Nk}.$$

On approche de cette limite lorsque l'ordre de grandeur de n dépasse l'ordre de la plus petite des quantités $2N$ et $\frac{1}{2k}$, c'est-à-dire $2N$ en général.

Le cas bien connu de l'autofécondation indéfinie est le cas où $N = 1$: f_n tend alors vers la limite $\frac{1}{1+4k}$ qui est très voisine de 1. La probabilité asymptotique de

non-identité de deux loci pris au hasard est $\frac{4k}{1+4k} \sim 4k \sim 4(u+v)$.

2. Cas d'une population à sexes séparés.

N_1 mâles (contribuant également au réservoir gamétique mâle) et N_2 femelles (contribuant également au réservoir gamétique femelle).

Nous définissons toujours le coefficient de consanguinité f_n en prenant au hasard, dans le réservoir gamétique pour F_n , un gamète mâle Γ et un gamète femelle Γ' qui sont issus de deux individus *différents* P et M respectivement ; ces deux individus ont la probabilité $\frac{1}{N_2}$ d'avoir le même père — et la probabilité pour que Γ et Γ' proviennent (pour les loci considérés) d'un père commun à P et M est $\frac{1}{4N_1}$; P et M ont la probabilité $\frac{1}{N_2}$ d'avoir la même mère — et la probabilité pour que Γ et Γ' proviennent (pour les loci considérés) d'une mère commune à P et M est $\frac{1}{4N_2}$.

La probabilité pour que les deux loci considérés proviennent d'un *grand* parent commun est $\frac{1}{4N_1} + \frac{1}{4N_2}$, que nous noterons $\frac{1}{N}$ (N étant le double de la moyenne harmonique des deux effectifs) et dans ce cas la probabilité *conditionnée* d'identité est $(1 - k)^4 (1 + f_{n-2})/2$.

La probabilité pour qu'ils proviennent de grand-parents différents est $1 - \frac{1}{N}$, auquel cas leur probabilité conditionnée d'identité est $(1 - k)^4 \varphi_{n-2}$, φ_{n-2} désignant le coefficient de *parenté* de deux individus différents pris au hasard dans F_{n-2} ; mais dans le cas de sexes séparés $f_{n-1} = (1 - k)^2 \varphi_{n-2}$, d'où :

$$f_n = \frac{1}{N} (1 - k)^4 (1 + f_{n-2})/2 + \left(1 - \frac{1}{N}\right) (1 - k)^2 f_{n-1}$$

qui donne la même limite $f = 1/(1 + 4Nk)$ qu'au 1° (1) ; cette limite est atteinte si k peut être regardé comme constant (en particulier si les taux de mutation sont nuls) avec la vitesse de tendance vers zéro de $\left(1 - 2k - \frac{1}{2N}\right)^n$. C'est en général (en particulier si $k = 0$) $2N$ qui fixe cette vitesse : alors, pour atteindre l'équilibre, le nombre de générations nécessaire est de l'ordre de l'inverse de l'effectif du groupe isolé.

Lorsque $k = 0$ (pas de mutations), la limite est 1 : une population d'effectif limité N tend plus ou moins rapidement (suivant que N est plus ou moins petit) vers l'homogénéité génétique. C'est donc que l'hétérogénéité ne peut être maintenue indéfiniment que s'il existe des mutations récurrentes ; mais, si leurs taux u et v sont petits par rapport à $\frac{1}{N}$, $4Nk = 4N(u + v)$ est petit, f est voisin de 1, l'hétérogénéité est « sporadique » (c'est-à-dire est maintenue seulement par la présence de temps en temps de quelques gènes rares).

(1) La valeur $1/(1 + 4Nk)$ a été donnée par S. WRIGHT (1931) pour le « modèle de l'île », comportant des immigrants depuis une population de composition constante (cf. MALÉCOT, 1950, p. 44).

II. — *Calcul du coefficient de parenté
entre groupes panmictiques communiquant par migration*

Soient deux loci tirés au hasard respectivement dans les réservoirs gamétiques pour les emplacements x et y (dans la génération F_n). Leur coefficient de parenté (probabilité d'identité) est désigné par $\varphi_{xy}(n)$. Mais les gamètes ayant apporté ces loci peuvent provenir respectivement des emplacements z et w avec les probabilités

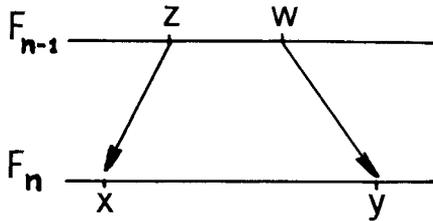


FIG. 2. — Passage de la $(n - 1)^{\text{e}}$ génération à la n^{e} génération avec tirages au hasard dans des réservoirs gamétiques d'emplacements quelconques

l_{z-x} et l_{w-y} (fig. 2) ; quand il en est ainsi, leur probabilité d'identité devient $\varphi_{zw}(n - 1)$ à condition qu'ils soient fournis par deux parents différents (tirés au sort, l'un en z et l'autre en w) ; en effet, s'ils sont fournis par le même parent P (ce qui ne peut avoir lieu que si l'emplacement w coïncide avec z) leur probabilité d'identité devient alors : $\frac{1 + \varphi_{zz}(n - 1)}{2}$ puisqu'ils ont probabilité $\frac{1}{2}$ de porter le même locus du parent P et probabilité 1 de porter respectivement les deux loci homologues de P qui sont tirés au sort indépendamment (en raison de la panmixie ⁽¹⁾) dans le réservoir gamétique local pour F_{n-1} en z .

Or, la probabilité d'être fournis par le même parent occupant l'emplacement z est, si l'on suppose que les N_z individus occupent cet emplacement sont des parents équiprobables ⁽²⁾.

$$l_{z-x} l_{z-y} / N_z$$

La probabilité d'être fourni par des parents différents occupant le même emplacement z est donc :

$$l_{z-x} l_{z-y} \left(1 - \frac{1}{N_z} \right)$$

⁽¹⁾ Dans le cas de non panmixie, il faudrait remplacer $\varphi_{zz}(n - 1)$ par une fonction $\theta_{zz}(n - 1)$ probabilité d'identité des deux loci qui s'unissent dans l'emplacement z (ou « coefficient de consanguinité en z »).

Nous avons vu au paragraphe I que cette modification est déjà nécessaire lorsque l'autofécondation est exclue, mais que la correction est insignifiante.

Nous avons montré (MALÉCOT, 1966) qu'il en est de même plus généralement, lorsqu'on exclut les croisements entre les parents les plus proches (à condition de supposer que la loi de migration parentale admette une variance).

⁽²⁾ Cela suppose que chaque individu produit des gamètes de deux sexes (cas des plantes) et que le fait d'être déjà parent (c'est-à-dire d'avoir déjà engendré un gamète utile au moins) ne modifie pas la probabilité d'être encore parent (c'est-à-dire d'engendrer encore d'autres gamètes utiles). Le nombre de gamètes utiles produites par un même individu doit donc être distribué suivant une loi de Poisson ; cela est réalisé lorsque chaque parent apporte au réservoir gamétique un grand nombre de gamètes, parmi lesquels une faible proportion, choisie au hasard, constitue les « gamètes utiles ».

En désignant toujours par $1 - k$ la probabilité de persistance de l'identité (« non bêtamutation ») sur chaque chaînon, on obtient la récurrence :

$$\varphi_{xy}(n) = (1 - k)^2 \sum_z l_{z-x} l_{z-y} \frac{1 + \varphi_{zz}(n - 1)}{2N_z} \tag{I. 1}$$

$$+ (1 - k)^2 \sum_z \left(1 - \frac{1}{N_z}\right) l_{z-y} l_{z-y} \varphi_{zz}(n - 1)$$

$$+ (1 - k)^2 \sum_z \sum_{w \neq z} l_{z-x} l_{w-y} \varphi_{zw}(n - 1).$$

$$\varphi_{xy}(n) = (1 - k)^2 \sum_z \sum_w l_{z-x} l_{w-y} \varphi_{zw}(n - 1) \tag{I. 1'}$$

$$+ (1 - k)^2 \sum_z l_{z-x} l_{z-y} \frac{1 - \varphi_{zz}(n - 1)}{2N_z}.$$

Cette récurrence linéaire non homogène admet si k est constant ⁽¹⁾ une solution indépendante de n , que nous noterons φ_{xy} , et que nous calculerons plus bas ; la différence $\varphi_{xy}(n) - \varphi_{xy}$ que nous appellerons $\psi_{xy}(n)$, est évidemment une solution de l'équation homogène associée à (1'), c'est-à-dire de l'équation déduite de (1) en remplaçant le terme $\frac{1 + \varphi_{zz}(n - 1)}{2N_z}$ par $\frac{\psi_{zz}(n - 1)}{2N_z}$; or il est facile de montrer que toute solution d'une telle récurrence tend vers zéro : elle est en effet susceptible, en posant $B_{n-1} = \sup_{(z, w)} [\psi_{zw}(n - 1)]$, de la majoration :

$$\begin{aligned} B_n &= \sup_{(x, y)} |\psi_{xy}(n)| \leq (1 - k)^2 \sum_z l_{z-x} l_{z-y} \left(\frac{1}{2N_z} + 1 - \frac{1}{N_z}\right) B_{n-1} \\ &\quad + (1 - k)^2 \sum_z \sum_{w \neq z} l_{z-x} l_{w-y} B_{n-1} \\ &\leq B_{n-1} (1 - k)^2 \left[\sum_z \sum_w l_{z-x} l_{w-y} - \frac{1}{2N_z} \sum_z l_{z-x} l_{z-y} \right]. \end{aligned}$$

Donc $B_n < (1 - k)^2 B_{n-1} < (1 - k)^{2n} B_0$, qui tend vers zéro.

Toute solution de (1') tend donc vers la solution indépendante de n et unique ⁽²⁾ définie par le système :

$$\varphi_{xy} = (1 - k)^2 \sum_z \sum_w l_{z-x} l_{w-y} \varphi_{zw} + (1 - k)^2 \sum_z l_{z-x} l_{z-y} \frac{1 - \varphi_{zz}}{2N_z}. \tag{I. 2}$$

Cette récurrence sur la variable d'espace — récurrence qui (comme dans le cas d'un groupe isolé) subsiste si les sexes sont séparés, à condition de poser $1/4 N_1 + 1/4 N_2 = 1/N$ — se résout aisément dans les divers cas.

⁽¹⁾ Donc si le taux fictif de bêtamutation peut être regardé comme indépendant des fréquences : ce qui est le cas en particulier s'il n'y a aucune mutation ($k = 0$).

⁽²⁾ En effet ce qui précède montre que le système homogène associé à (2) ne peut avoir qu'une solution identiquement nulle.

I. *Cas d'une suite unidimensionnelle illimitée de groupes de même effectif* $N(N_z = N \forall z)$.

Dans ce cas la formule (2) montre que $\varphi_{x+k, y+k} = \varphi_{xy} \forall k$: φ_{xy} ne dépend que de $y - x = d$ et pourra être notée $\varphi(d)$ ⁽¹⁾.

Nous poserons
$$\begin{aligned} z &= x + r, \quad l_{z-x} = l_r \\ w &= y + s, \quad l_{w-y} = l_s \end{aligned}$$

et nous transformerons la récurrence (2) en introduisant la « fonction génératrice » définie par la série de Laurent.

$$K(\alpha) = \sum_y \alpha^{y-x} \varphi_{xy} = \sum_d \alpha^d \varphi(d)$$

qui, d'après (2), vérifie, indépendamment de l'ordre des sommations, lorsqu'il y a convergence absolue l'équation :

$$\begin{aligned} K(\alpha) &= (1 - k)^2 \sum_y \sum_z \sum_w \alpha^{y-w} \alpha^{w-z} \alpha^{z-x} l_{z-x} l_{w-y} \varphi_{zw} \\ &\quad + \frac{(1 - k)^2}{2N} \sum_y \sum_z \alpha^{y-z} \alpha^{z-x} l_{z-x} l_{z-y} (1 - \varphi_{zz}). \end{aligned}$$

En effectuant d'abord la sommation par rapport à y , on met en facteur la série

$$\sum_y \alpha^{y-w} l_{w-y} (= \sum_y \alpha^{y-z} l_{z-y})$$

qui est absolument convergente pour $|\alpha| = 1$, indépendante de w ou de z et que l'on notera $\sum l_s/\alpha^s = L(1/\alpha)$ ⁽²⁾.

On a alors :

$$\begin{aligned} K(\alpha) &= (1 - k)^2 \sum_z \alpha^{z-x} l_{z-x} \sum_w \alpha^{w-x} \varphi_{zw} L(1/\alpha) \\ &\quad + \frac{(1 - k)^2}{2N} \sum_z (1 - \varphi_{zz}) \alpha^{z-x} L(1/\alpha). \end{aligned}$$

Dans le 1^{er} terme du 2^e membre, la sommation sur w donne le facteur $K(\alpha) L(1/\alpha)$; en posant en bref $\varphi_{zz} = \varphi(0)$, et compte tenu de ce que :

$$\sum_z \alpha^{z-x} l_{z-x} = \sum_r l_r \alpha^r = L(\alpha)$$

$$K(\alpha) = (1 - k)^2 L(\alpha) K(\alpha) L(1/\alpha) + (1 - k)^2 \frac{1 - \varphi(0)}{2N} L(\alpha) L(1/\alpha)$$

$$K(\alpha) = \frac{1 - \varphi(0)}{2N} \frac{(1 - k)^2 L(\alpha) L(1/\alpha)}{1 - (1 - k)^2 L(\alpha) L(1/\alpha)}.$$

⁽¹⁾ La détermination effective, ci-dessous, d'une fonction φ_{xy} ne dépendant que de $y - x$, jointe à l'unicité établie pour la solution de (2), assure que $\varphi(d)$ est toujours la limite des $\varphi_{xy}(n)$.

⁽²⁾ C'est la fonction déjà introduite dans la 1^{re} partie, chapitre IV (fonction génératrice de la loi de migration parentale).

On peut alors, comme au chapitre IV de la partie I, raisonner, lorsque $L(\alpha)$ est un polynôme, sur les pôles du dénominateur, racines de l'équation

$$H(\alpha) = L(\alpha) L\left(\frac{1}{\alpha}\right) = \frac{1}{(1-k)^2} = 1 + k_1,$$

et ne considérer que les pôles α_1 et $\alpha_2 = \frac{1}{\alpha_1}$ très voisins de 1 (α_1 étant compris entre 0 et 1). On écrit alors :

$$K(\alpha) = \sum_j \frac{A_j}{\alpha - \alpha_j} \quad \text{et} \quad \varphi(d) = \frac{1}{2\pi i} \sum_j \int_c \frac{\alpha_j^{d-1} A_j}{\alpha - \alpha_j} d\alpha \sim A_1 \alpha_1^{d-1}$$

ce qui donne, puisque $A_1 = \lim_{\alpha \rightarrow \alpha_1} (\alpha - \alpha_1) K(\alpha) = \frac{1 - \varphi(0)}{2N} \frac{1}{-(1-k)^2} \frac{dH}{d\alpha_1}$

$A_1 \sim \frac{1 - \varphi(0)}{4N \sigma \sqrt{k_1}}$ et :

$$\varphi(d) \sim \frac{1 - \varphi(0)}{4N \sigma \sqrt{2k}} \left(1 - \frac{\sqrt{2k}}{\sigma}\right)^d$$

Mais ce calcul doit se ramener à celui du chapitre IV, partie I, puisque les covariances stationnaires s'expriment (chap. II, 3) en fonction des coefficients de parenté correspondants par la formule :

$$v(d) = E[(q_x - \bar{q})(q_y - \bar{q})] = \varphi_{xy} C(1 - C)$$

et en particulier (si $\frac{1}{N}$ est négligeable par rapport à $\varphi(0)$) :

$$v(0) = \varphi(0)(C - C^2)$$

donc la formule que nous venons d'écrire pour $\varphi(d)$ donne pour $v(d)$ la formule :

$$v(d) \sim \frac{C - C^2 - v(0)}{4N \sigma \sqrt{2k}} \left(1 - \frac{\sqrt{2k}}{\sigma}\right)^d \quad (d \geq 0)$$

qui est bien la formule déjà obtenue au chapitre IV, (partie I) en particulier lorsque $L(\alpha) = 1 - 2m + m(\alpha + 1/\alpha)$.

L'intérêt de la méthode du coefficient de parenté (qui toutefois, rappelons-le, a l'inconvénient d'exclure la sélection⁽¹⁾) est de permettre :

- a) L'extension à la migration continue ;
- b) L'analyse des parts qui reviennent aux parentés des différents ordres.

Ces deux points feront l'objet l'un du paragraphe III ci-après, l'autre de la partie III (à paraître plus tard).

(1) A moins de linéariser celle-ci et de remarquer qu'elle est alors mathématiquement équivalente à des mutations définies par les mêmes constantes k et C . (1^{re} partie, chapitre II),

2. *Extension au cas d'une suite bidimensionnelle illimitée de groupes de même effectif N répartis aux sommets d'un réseau rectangulaire.*

Les emplacements des groupes sont maintenant définis par des vecteurs x, y , etc. dont chaque composante x_1, x_2, y_1, y_2 , etc. est une valeur entière. Dans ce cas, pour passer de la formule (2) ci-dessus à une fonction caractéristique, il faut introduire les deux variables α_1 et α_2 [$|\alpha_1| = 1, |\alpha_2| = 1$] et poser

$$K(\alpha_1, \alpha_2) = \sum_{y_1, y_2} \alpha_1^{y_1-x_1} \alpha_2^{y_2-x_2} \varphi_{xy} = \sum_{d_1, d_2} \alpha_1^{d_1} \alpha_2^{d_2} \varphi(d)$$

d désignant le vecteur de composantes $d_1 = y_1 - x_1$ et $d_2 = y_2 - x_2$. En posant aussi $L(\alpha_1, \alpha_2) = \sum \alpha_1^{y_1-x_1} \alpha_2^{y_2-x_2} l_{y-x}$, on obtient encore :

$$K(\alpha_1, \alpha_2) = (1 - k)^2 L(\alpha_1, \alpha_2) L\left(\frac{1}{\alpha_1}, \frac{1}{\alpha_2}\right) K(\alpha_1, \alpha_2) + (1 - k)^2 \frac{1 - \varphi(0)}{2N} L(\alpha_1, \alpha_2) L\left(\frac{1}{\alpha_1}, \frac{1}{\alpha_2}\right) \tag{2.1}$$

D'où l'on tire la fonction génératrice K , puis :

$$\varphi(d) = \left(\frac{1}{2\pi i}\right)^2 \int_{C_1} \int_{C_2} \alpha_1^{d_1-1} \alpha_2^{d_2-1} K(\alpha_1, \alpha_2) d\alpha_1 d\alpha_2 \tag{2.2}$$

les intégrations étant effectuées sur les cercles $C_1(|\alpha_1| = 1)$ et $C_2(|\alpha_2| = 1)$.

Dans le cas particulier de la migration entre colonies adjacentes

$$L(\alpha_1, \alpha_2) = 1 + m \left(\alpha_1 - 2 + \frac{1}{\alpha_1}\right) + m' \left(\alpha_2 - 2 + \frac{1}{\alpha_2}\right),$$

cas qui est un bon modèle du cas « discret non isotrope », l'inversion (2.2) peut (MALÉCOT, 1950, 1971) s'effectuer en calculant d'abord (comme dans le cas unidimensionnel) un résidu par rapport à α_1 , puis ramenant l'intégration par rapport à α_2 à une quadrature elliptique. On trouve, lorsque m, m' et k sont petits, et en posant $G(\alpha_2) = m'(\alpha_2 + 1/\alpha_2 - 2)/b$ et $1/(1 - k) = 1 + k_1$:

$$\varphi(0, d_2) \sim \frac{1 - \varphi(0, 0)}{8Ni\pi} \int_{C_2} \frac{\alpha_2^{d_2-1} d\alpha_2}{\sqrt{[k_1 + 2m - G(\alpha_2)]^2 - 4m^2}}$$

Et l'on est ramené en posant $c = -m'$,

$$b = k_1 + 2m + 2m' \quad \text{et} \quad P_i(\alpha) = \alpha^2 + 1 - (b/m')\alpha - 2(-1)^i(m/m')\alpha$$

à l'intégrale elliptique

$$I_p = \int_{C_2} \frac{\alpha_2^{p-1} d\alpha_2}{\sqrt{[b + c(\alpha_2 + 1/\alpha_2)]^2 - 4m^2}} = \frac{1}{m'} \int_{C_2} \frac{\alpha_2^p d\alpha_2}{\sqrt{P_1(\alpha_2)P_2(\alpha_2)}} \tag{2.3}$$

qui peut être calculée :

1° En la ramenant, par le changement de variable $\alpha_2 = \frac{\beta + 1}{\beta - 1}$, à une « forme réduite de Legendre », ce qui donne en particulier :

$$I_0 = \frac{4^i}{\sqrt{(k_1 + 4m)(k_1 + 4m')}} K(\alpha)$$

$K(\alpha)$ étant l'intégrale complète de 1^e espèce calculée pour

$$\sin^2 \alpha = 1 - \frac{k_1(k_1 + 4m + 4m')}{(k_1 + 4m)(k_1 + 4m')} \text{ (tables d'Abramowicz, 1964) ;}$$

on en déduit $\varphi(0, 0) = 4K(\alpha)/[4K(\alpha) + 8N\pi\sqrt{(k_1 + 4m)(k_1 + 4m')}]$.

2° En utilisant la récurrence connue :

$$c^2(I + I/p)I_{p+2} + 2bc(I + I/2p)I_{p+1} + (b^2 + 2c^2 - 4m^2)I_p + 2bc(I - I/2p)I_{p-1} + c^2(I - I/p)I_{p-2} = 0$$

qui permet de calculer les I_p ($p < 2$) en fonction de I_0, I_1, I_2 , et qui donne en parti-

culier la formule asymptotique: $I_p \sim \text{cte} \times \alpha^p / \sqrt{p}$ avec $\alpha_1 = 1 + \frac{k_1}{2m'} - \sqrt{\frac{k_1}{m'} + \frac{k_1^2}{4m'^2}}$

(dans le cas unidimensionnel, on trouverait $I_p \sim I_0 \alpha_1^p$). Mais cette formule, difficile à tester pour p grand en raison de l'extrême petitesse du coefficient de parenté correspondant, peut être très incorrecte pour p petit, comme MORTON (1970) l'a vérifié; il arrive qu'une simple exponentielle (comme dans le cas unidimensionnel) soit meilleure: c'est ce que confirment les approximations suivantes.

a) Dans le cas où k_1 est grand par rapport à m et m' et en prenant $m = m'$, la somme des 2 racines (inverses) de chaque trinôme $P_i(\alpha)$ de (2.3) est grande par rapport à 1, on peut donc, lorsque α décrit le cercle C de rayon 1, écrire :

$$P_1(\alpha) = (\alpha - \alpha_1)(\alpha - \alpha_2) \sim -\alpha_2(\alpha - \alpha_1)$$

et
$$P_2(\alpha) = (\alpha - \alpha_3)(\alpha - \alpha_4) \sim -\alpha_4(\alpha - \alpha_3)$$

α_2 et α_4 étant chacun une constante équivalente à la somme des racines, ce qui permet d'effectuer sur (2.3) l'approximation

$$I_p \sim \frac{1}{m} \int \frac{\alpha^p d\alpha}{\sqrt{\left(\frac{k_1}{m} + 2\right)\left(\frac{k_1}{m} + 6\right)} \sqrt{\left(\alpha - \frac{1}{k_1/m + 2}\right)\left(\alpha - \frac{1}{k_1/m + 6}\right)}}$$

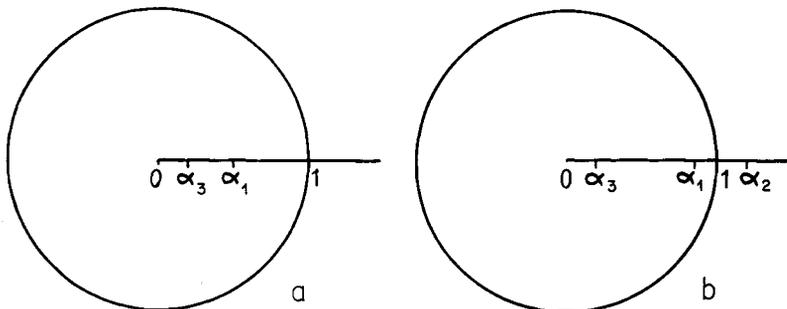


FIG. 3. — Position des racines du trinôme de l'intégrale elliptique de la formule 2.3

- a) k_1 est grand par rapport à m et m'
- b) k_1 est petit par rapport à m et m'

Puisque les racines $\alpha_1 = \frac{1}{\frac{k_1}{m} + 2}$ et $\alpha_3 = \frac{1}{\frac{k_1}{m} + 6}$ sont des réels positifs < 1 ,

(fig. 3 a), l'intégrale le long du contour C (avec la détermination positive du radical pour $\alpha = 1$) se ramène au double d'une intégrale (de droite à gauche) le long d'un segment réel, d'où :

$$I_p \sim \frac{2i}{\sqrt{(k_1 + 2m)(k_1 + 6m)}} \int_{\alpha_3}^{\alpha_1} \frac{\alpha^p d\alpha}{\sqrt{(\alpha_1 - \alpha)(\alpha - \alpha_3)}}$$

et, en posant $\alpha = \frac{\alpha_1 + \alpha_3}{2} + \frac{\alpha_1 - \alpha_3}{2} \cos \theta$

$$I_p \sim \frac{2i}{\sqrt{(k_1 + 2m)(k_1 + 6m)}} \int_0^\pi \alpha^p d\theta.$$

Or $\alpha_1 - \alpha_3$ [de l'ordre de $(m/k_1)^2$] est infiniment petit par rapport à $\alpha_1 + \alpha_3$ [de l'ordre de m/k_1]. On a donc :

$$I_p \sim \frac{2i\pi}{k_1} \left(\frac{\alpha_1 + \alpha_3}{2}\right)^p \sim \frac{2i\pi}{k_1} \left(\frac{m}{k_1}\right)^p$$

ce qui explique les résultats trouvés par MORTON cité plus haut $\frac{\varphi(p)}{\varphi(0)} = \frac{I_p}{I_0} \sim \left(\frac{m}{k_1}\right)^p$ manifeste, exactement comme dans le cas unidimensionnel, une décroissance exponentielle, et avec le même amortissement puisque l'on a :

$$1 + \frac{k}{2m} - \sqrt{\frac{k}{m} + \frac{k^2}{4m^2}} = 1 + \frac{k}{2m} \left(1 - \sqrt{1 + \frac{4m}{k}}\right) \sim \frac{m}{k}$$

En outre $I_0 \sim \frac{2i\pi}{k_1}$ d'où $\varphi(0) \sim \frac{1 - \varphi(0)}{4Nk_1} \Rightarrow \varphi(0) = \frac{1}{1 + 4Nk_1}$ ce qui coïncide avec le « modèle d'isolement complet.

b) Il n'en est plus de même dans le cas où $m = m'$ est grand par rapport à k_1 : la décroissance n'est plus exponentielle : la formule (2.3) montre que α_1 et α_2 (de produit 1, de somme $2 + \frac{k_1}{m}$) sont infiniment voisins d'un point (d'affixe 1) sur le cercle C , alors que ce n'est pas le cas de α_3 et α_4 , ce qui assure que

$$\frac{1}{\sqrt{(\alpha - \alpha_3)(\alpha - \alpha_4)}}$$

reste borné sur le cercle C alors que $\frac{1}{\sqrt{(\alpha - \alpha_1)(\alpha - \alpha_2)}}$ devient très grand au voisinage de $\alpha = 1$ (fig. 3 b). On a donc :

$$I_p \sim \frac{1}{m' \sqrt{(1 - \alpha_3)(1 - \alpha_4)}} \int_C \frac{\alpha^p - d\alpha}{\sqrt{(\alpha - \alpha_1)(\alpha - \alpha_2)}}$$

$$I_p \sim \frac{I}{m\sqrt{-4}} \int_C \frac{\alpha^{p-1} d\alpha}{\sqrt{\alpha + I/\alpha - 2 - k_1/m}}$$

d'où, en posant $\alpha = e^{i\theta}$ et $h = \frac{k_1}{4m}$ (petit) :

$$I_p \sim \frac{i}{4m} \int_{-\pi}^{+\pi} \frac{e^{ip\theta} d\theta}{\sqrt{h + \sin^2(\theta/2)}}$$

$$I_p \sim \frac{i}{2m} \int_0^\pi \frac{\cos p\theta d\theta}{\sqrt{h + \sin^2(\theta/2)}}$$

et, en posant $\theta = 2\left(\frac{\pi}{2} - \varphi\right)$:

$$(-I)^p I_p \sim \frac{i}{m} \int_0^{\pi/2} \frac{\cos 2p\varphi d\varphi}{\sqrt{h + I - \sin^2 \varphi}}$$

Ces intégrales vérifient pour $p \geq 1$ la récurrence

$$(2p + 1)I_{p+1} - 4p(2h + 1)I_p + (2p - 1)I_{p-1} = 0$$

et de plus les deux premières sont en posant $\sin^2 \alpha = (I + h)^{-1}$:

$$I_0 \sim \frac{i}{m\sqrt{I + h}} K(\alpha)$$

$$I_1 \sim \frac{i}{m} \frac{I + 2h}{\sqrt{I + h}} K(\alpha) - \frac{2i}{m} \sqrt{h + I} E(\alpha)$$

$E(\alpha)$ étant l'intégrale complète de deuxième espèce.

Le calcul par récurrence, fait pour différentes valeurs de α , donc de $h = \frac{k_1}{4m}$, a fourni, pour la variation en fonction de p de $\text{Log } I_p$, les courbes ci-jointes (fig. 4), en trait plein. Il est représenté en même temps, en tirets, et à une translation verticale près pour faciliter la comparaison, l'une des approximations suivantes :

a) le logarithme de l'expression asymptotique établie au 2^o, soit

$$\log \frac{\alpha^p}{\sqrt{p}} \sim \text{Log}(I - 2\sqrt{h})^p - \frac{1}{2} \text{Log } p$$

L'approximation s'avère satisfaisante pour les plus petites valeurs de $\frac{k_1}{4m}$, celles $\leq 0,003$ qui correspondent à $\alpha = 89^\circ$ et $\alpha = 87^\circ$ (fig. (4a) et (4b)).

b) Le logarithme de $(I - 2\sqrt{h})^p$, qui correspond à l'expression asymptotique simplement exponentielle obtenue ci-dessus dans le cas unidimensionnel ($\sigma = \sqrt{2m}$ représentant toujours l'écart-type de la composante de la migration parallèlement à l'un des axes de coordonnées). Pour les valeurs $\alpha = 85^\circ$ et $\alpha = 80^\circ$ cette approximation est satisfaisante (fig. 4 c et 4 d). Comme celles des données numériques de MORTON (1970) qui correspondent à $\frac{k_1}{m}$ petit rentrent dans cette marge ($\frac{k_1}{4m} \geq 0,02$),

il n'est pas surprenant que MORTON ait conclu alors à la supériorité de la simple exponentielle. Mais ces courbes montrent que cette supériorité cesse quand k_1 est encore plus petit ($\frac{k_1}{4m} \leq 0,003$) que les valeurs considérées par cet auteur. Ces différences de comportement expliquent la longue controverse qu'il signale (p. 581).

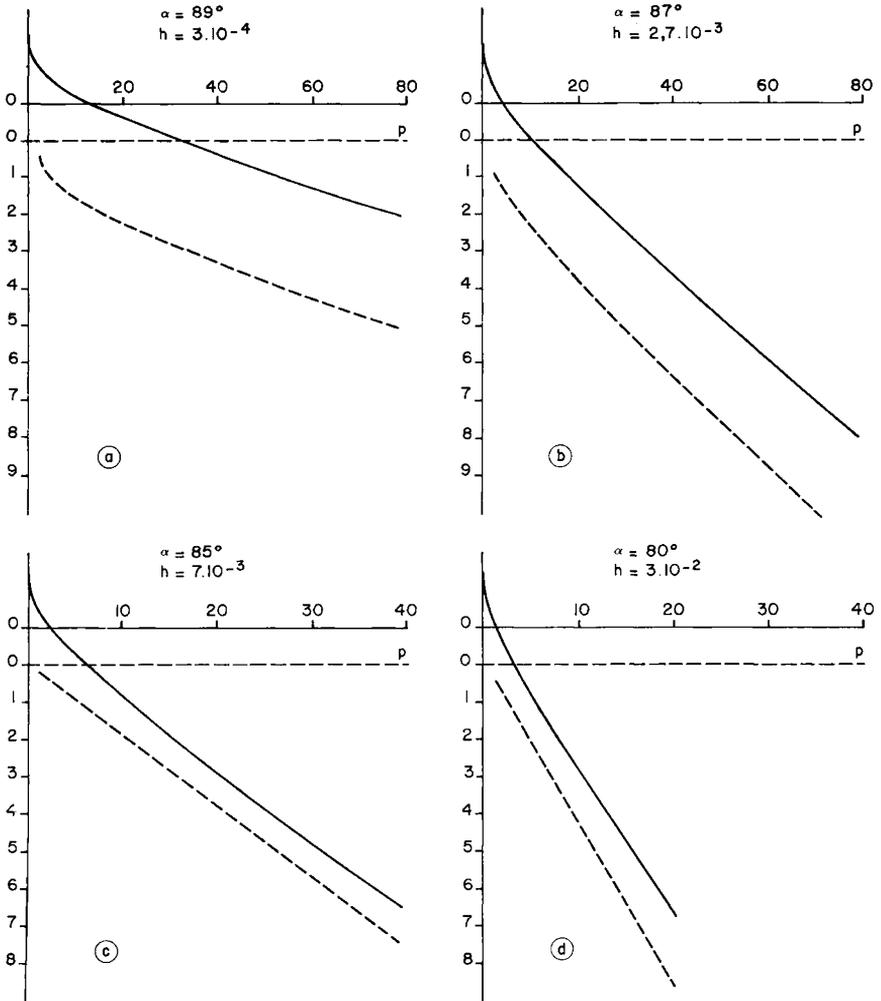


Fig. 4. — *Décroissance comparée du coefficient de parenté et d'une fonction exponentielle ou quasi exponentielle*

(la comparaison est facilitée par la représentation des logarithmes de ces fonctions, à une translation verticale près pour faciliter la comparaison)

- a et b ————— $\log I_p$
- - - - - $\log ((1-2 \sqrt{h})^p) - \frac{1}{2} \log p$
- c et d ————— $\log I_p$
- - - - - $\log ((1-2 \sqrt{h})^p)$

Remarque :

La formule $\varphi(0, \rho) = \frac{1 - \varphi(0,0)}{8Ni^\pi} I_\rho$ donne en particulier

$$\varphi(0,0) = \frac{1}{1 + 8N^\pi m^{1/K(\alpha)}}$$

dont le tableau ci-dessous fournit quelques valeurs :

h	$3 \cdot 10^{-4}$	$2,7 \cdot 10^{-3}$	$7 \cdot 10^{-3}$	$3 \cdot 10^{-2}$
α	89°	87°	85°	80°
$K(\alpha)$	5,43	4,33	3,83	3,15
$8/K(\alpha)$	1,47	1,84	2,08	2,53

$\varphi(0,0)$ dépend donc peu de k_1 lorsque k_1 est petit par rapport à m (il en allait tout autrement dans le cas a)).

III. — Cas de la migration continue

Le modèle « à migration continue » est une approximation du modèle « discret non panmictique » que nous allons définir en faisant $N = 1$ dans le modèle de II 1° et 2° : supposer ainsi chaque groupe panmictique réduit à un individu, c'est supposer que le réservoir gamétique pour x est alimenté en proportions différentes par les gamètes provenant des différents individus : aussi deux gamètes utiles provenant du même emplacement z auront donc le même géniteur, ce que traduit le remplacement de N par 1 dans le 2° membre de (I.I) : mais deux gamètes utiles provenant de deux géniteurs distincts proviendront par là même de deux emplacements z et z' distincts avec des probabilités l_{z-x} et $l_{z'-x}$ en général différentes.

Il sera commode de ne plus noter les emplacements d'origine par des abscisses entières mais de les supposer répartis aux sommets d'un réseau rectangulaire ⁽¹⁾ avec les équidistances $\frac{1}{\delta_1}$ et $\frac{1}{\delta_2}$, δ_1 et δ_2 désignant les densités linéaires parallèlement aux deux axes et $\delta_1 \delta_2$ la densité superficielle que nous noterons δ .

Chaque emplacement z occupé par un individu (et un seulement, rappelons-le) sera défini par deux coordonnées z_1 et z_2 appartenant respectivement à des progressions arithmétiques de raisons $\frac{1}{\delta_1}$ et $\frac{1}{\delta_2}$. Nous changerons de notations en désignant dorénavant par $l_{z-x} \Delta z$ la probabilité pour un gamète utilisé sur l'emplacement x de parvenir de l'emplacement z auquel nous associons l'aire $\Delta z = \frac{1}{\delta}$ ⁽²⁾, pour que celle-ci soit occupée par un seul individu.

⁽¹⁾ A titre d'exemple ; les notations et le passage à la limite se transposent aisément pour le cas unidimensionnel ou le cas tridimensionnel.

⁽²⁾ On a maintenant $\sum_z l_{z-x} \Delta z = 1$, la sommation étant étendue à tous les emplacements.

Les formules se récrivent alors comme suit :

$$\begin{aligned} \varphi_{xy} &= (\mathbf{I} - k)^2 \sum_{zw} l_{z-x} l_{w-y} \varphi_{z w} \Delta z \Delta w \\ &+ (\mathbf{I} - k)^2 \sum_z \frac{\mathbf{I} - \varphi_{zz}}{2\delta} l_{z-x} l_{z-y} \Delta z \quad (1) \end{aligned} \quad (3.1)$$

et l'on peut, lorsque les fonctions l_{z-x} et φ_{zw} varient peu sur une étendue égale à l'aire $\Delta z = \frac{\mathbf{I}}{\delta}$ et sont regardées comme des fonctions bornées et intégrables de variables continues z_1 et z_2 , faire l'approximation ⁽²⁾ :

$$\begin{aligned} \varphi_{xy} &= (\mathbf{I} - k)^2 \iint \varphi_{zw} l_{z-x} l_{w-y} dz dw \\ &+ (\mathbf{I} - k)^2 \frac{\mathbf{I} - \varphi(0)}{2\delta} \int l_{z-x} l_{z-y} dz. \end{aligned} \quad (3.2)$$

Par analogie avec le cas continu, nous pouvons introduire (à la place des fonctions génératrices) les transformées de Fourier :

$$\begin{aligned} \iint e^{iv_1(z_1-x_1) + iv_2(z_2-x_2)} l_{z-x} dz &= G(v_1, v_2) = G(v) \\ \iint e^{iv_1(z_1-w_1) + iv_2(z_2-w_2)} \varphi_{zw} dz &= H(v_1, v_1) = H(v) \end{aligned}$$

v désignant le vecteur de composantes v_1 et v_2 dans l'espace dual.

En multipliant les deux membres de (3.2) par $e^{iv_1(y_1-x_1) + iv_2(y_2-x_2)}$ et intégrant par rapport à y_1 et y_2 , on obtient :

$$H(v) = (\mathbf{I} - k)^2 G(v) K(v) G(-v) + \frac{\mathbf{I} - \varphi(0)}{2\delta} (\mathbf{I} - k)_2 G(v) G(-v) \quad (3.3)$$

formule qui n'est autre que la formule (2.1) où l'on poserait :

$$K(e^{iv_1}, e^{iv_2}) = H(v) \text{ et } L(e^{iv_1}, e^{iv_2}) = G(v)$$

On peut alléger les écritures en posant :

$$(\mathbf{I} - k)^2 G(v) G(-v) = M(v)$$

et l'on obtient :

$$H(v_1, v_2) = H(v) = \frac{\mathbf{I} - \varphi(0)}{2\delta} \frac{M(v)}{\mathbf{I} - M(v)} \quad (3.4)$$

(1) Nous remplaçons l'écriture d'un des facteurs Δz par l'écriture $\frac{\mathbf{I}}{\delta}$ parce que la dernière somme ne pourra faire apparaître qu'une intégrale *simple*.

(2) Il est important de noter qu'il ne s'agit jamais que d'une approximation : on ne peut faire tendre Δz vers zéro puisque la densité δ reste finie.

On peut alors obtenir φ_{zw} — qui est fonction du vecteur $w - z = d$ de composantes d_1 et d_2 — par la formule d'inversion :

$$\varphi_{zw} = \varphi(d) = \varphi(-d) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{iv_1 d_1 + iv_2 d_2} H(v_1, v_2) dv_1 dv_2 \quad (3.5)$$

Les calculs seront poussés plus loin dans le cas isotrope (§ 5).

IV. — *Application : variance d'un groupe local, dans le cas de migration continue*

Si l'on envisage un groupe local constitué par s individus tous voisins de l'emplacement x (dans une génération F_n), ce n'est plus, en toute rigueur, un « groupe panmictique » (puisque il n'existe pas de parents « autochtones et équiprobables », mais seulement des parents ayant subi des migrations de probabilités différentes) ; mais supposons que le coefficient de parenté varie peu lorsqu'on se déplace dans le voisinage de deux emplacements x et y , donc qu'il reste, dans ces voisinages, sensiblement égal à φ_{xy} ; en particulier, pour deux loci pris dans le voisinage de x , leur coefficient de parenté sera sensiblement φ_{xx} ; dès lors, à un groupe local de s individus ($2s$ loci) se trouvent attachées $2s$ aléatoires X (« aléatoires gamétiques ») dont les coefficients de corrélation deux à deux sont égaux à φ_{xx} ; le calcul fait au chap. II, 3 montre que, dans le cas asymptotiquement stationnaire où la loi *a priori* de X_j lui assigne les

valeurs 1 ou 0 avec les probabilités C et $1 - C$, la variance de la fréquence $\frac{\sum_{j=1}^{j=2s} X_j}{2s}$ du gène a dans le groupe local est :

$$E \left[\left(\frac{\sum_j X_j}{2s} - C \right)^2 \right] = \frac{[\sum_j (X_j - C)^2]}{4s^2} = \frac{1}{2s} E [(X_j - C)^2] + \frac{2s(2s - 1)}{4s^2} E[(X_j - C)(X_k - C)] = \varphi_{xx}C(1 - C) + \frac{1}{2s} (1 - \varphi_{xx})C(1 - C).$$

Lorsque $2s$ est assez grand pour que $\frac{1}{2s}$ soit négligeable par rapport à φ_{xx} ⁽¹⁾, la variance de la fréquence locale est indépendante du nombre s d'individus considérés et se réduit — comme dans le cas discret — à $\varphi_{xx}C(1 - C)$; les estimations expérimentales de la variance de la fréquence locale peuvent donc être comparées à l'évaluation théorique de φ_{xx} . Naturellement, le calcul de la covariance entre les fréquences de deux groupes locaux dans les emplacements x et y fournit encore le même résultat $\varphi_{xy}C(1 - C)$, et les estimations expérimentales du coefficient de corrélation entre les fréquences locales peuvent être comparées à la corrélation théorique $\varphi_{xy}/\varphi_{xx}$.

Nous allons développer complètement le calcul dans le cas de migration bidimensionnelle isotrope.

(1) Cela n'est pas toujours réalisable : rappelons que le voisinage de l'individu d'emplacement x doit être assez restreint pour que φ y varie peu : il n'est pas toujours possible de trouver s individus dans un tel voisinage.

V. — Cas particulier de la migration isotrope

Il est possible de particulariser le modèle bidimensionnel continu en supposant que la probabilité $l_{z-x} dz$ ne dépende que de la valeur absolue du vecteur $z - x$, ou encore du module r du complexe $z_1 - x_1 + i(z_2 - x_2) = r e^{i\theta}$; nous poserons alors

$$l_{z-x} = l(r) \text{ et } dz = dz_1 dz_2 = r dr d\theta \text{ (1)}.$$

On sait qu'alors $G(v)$, au lieu d'être définie comme double transformée de Fourier de l_{z-x} , peut, de façon équivalente, être définie comme la « transformée de Hankel »

$$G(v) = 2\pi \int_0^{+\infty} J_0(r \sqrt{v_1^2 + v_2^2}) l(r) r dr$$

$G(v)$ ne dépendant que de $\sqrt{v_1^2 + v_2^2}$ — que nous noterons $|v|$ — il en est de même d'après (3.4), de $H(v)$; la formule d'inversion (3.5) montre alors que $\varphi(d)$ ne dépend de $d_1 = z_1 - w_1$ et $d_2 = z_2 - w_2$ que par la distance $\sqrt{d_1^2 + d_2^2}$ que nous appellerons a (il était d'ailleurs prévisible que $\varphi(d)$ serait comme $l(z - x) = l(r)$, invariante par rotation); en changeant de notation, (3.5) s'écrit :

$$\varphi(a) = \frac{1}{2\pi} \int_0^{+\infty} J_0(a|v|) H(v) |v| dv$$

donc, d'après (3.4), et en posant toujours $(1 - k)^2 G(v) G(-v) = M(v)$:

$$\varphi(a) = \frac{1 - \varphi(0)}{4\pi\delta} \int_0^{+\infty} J_0(a|v|) \frac{M(v)}{1 - M(v)} |v| dv \quad (3.6)$$

Application à la distribution K.

Nous avons appelé distribution K la distribution isotrope pour laquelle $l(r) dz = l(r) r dr d\theta$ est, à une puissance de r près, proportionnelle à la fonction de Bessel-Macdonald $K_{b-1}(hr)$ avec $b > 0$, de telle façon que $G(v)$ ait l'expression simple $(1 + \frac{v^2}{h^2})^{-b}$ (qui entraîne que cette famille de lois est fermée pour toutes les convolutions, chaque convolution se traduisant simplement par l'addition des paramètres b correspondants : propriété bidimensionnelle analogue à celle de la loi Γ unidimensionnelle dont la probabilité élémentaire est $\frac{1}{\Gamma(b)} (hr)^{b-1} e^{-hr} d(hr)$ et la fonction caractéristique $(1 - \frac{iv}{h})^{-b}$).

Il suffit alors de prendre pour probabilité élémentaire de la distance r :

$$2\pi l(r) r dr = \frac{1}{2^{b-1} \Gamma(b)} K_{b-1}(hr) (hr)^b d(hr)$$

pour obtenir pour transformée de Hankel :

$$G(v) = 2\pi \int_0^{+\infty} J_0(r|v|) r l(r) dr = (1 + v^2/h^2)^{-b}.$$

(1) $l(r)$ est la densité de probabilité dans le plan, donc la densité de probabilité de la distance r est $2\pi r l(r)$.

La probabilité élémentaire s'étudie quand $r \rightarrow 0$ et quand $r \rightarrow \infty$, par les formules

$$(hr)^{b-1} K_{b-1}(hr) \sim 2^{b-2} \Gamma(b-1) \quad (r \rightarrow 0) \quad (b > 1)$$

$$K_{b-1}(hr) \sim \left(\frac{\pi}{2hr}\right)^{\frac{1}{2}} e^{-hr} \quad (r \rightarrow +\infty).$$

La dernière formule montre que la densité de probabilité de la distance r est, quand $r \rightarrow \infty$, équivalente à

$$\frac{\sqrt{\pi/2} h}{2^{b-1} \Gamma(b)} (hr)^{b-1/2} e^{-hr}.$$

Si $b > 1$, l'avant-dernière formule montre que la densité de probabilité $l(r)$ dans le plan tend quand $z \rightarrow 0$ vers $h^2/4\pi(b-1)$; la distribution bidimensionnelle de la migration est donc « en cloche ». Mais la distribution unidimensionnelle de r a l'allure d'un χ^2 car $2\pi r l(r)$ tend alors vers zéro avec r .

$$\text{Si } 0 < b < 1, K_{b-1}(hr) = K_{1-b}(hr) \sim \frac{1}{2} \Gamma(1-b) \left(\frac{hr}{2}\right)^{b-1}$$

et

$$l(r) \sim \frac{h^2 \Gamma(1-b)}{4\pi \Gamma(b)} \left(\frac{hr}{2}\right)^{2b-2}.$$

La densité $l(r)$, quoique infinie pour $r = 0$, est donc intégrable dans le plan; et la densité $2\pi r l(r)$ de la distance r est nulle ou infinie pour $r = 0$ suivant que $b > \frac{1}{2}$

ou $b < \frac{1}{2}$; lorsque $b = \frac{1}{2}$, cette dernière densité est celle de la loi exponentielle.

La distribution K est commode pour représenter non seulement la loi de migration « parentale » (déplacement des gamètes depuis le lieu de naissance des parents jusqu'au lieu de naissance des enfants) mais aussi la « loi de migration ancestrale sur p générations » qui est une convolution de telles lois (alors que la loi Γ , utilisée par CAVALLI-SFORZA (1963) pour représenter la migration dans le plan, n'est pas « fermée pour la convolution »).

Certes, la loi normale isotrope jouit de la même propriété. Mais elle ne dépend que d'un paramètre (« paramètre d'échelle » σ), alors que la loi K dépend (comme d'ailleurs la Loi Γ) non seulement d'un « paramètre d'échelle » $1/h$, mais aussi d'un paramètre de dimensionalité b ; il est donc plus facile de l'ajuster aux migrations expérimentales, sur une ou plusieurs générations.

Nous avons montré (MALÉCOT, 1966) que, dans le cas d'une migration obéissant à la loi K (comme, d'ailleurs, dans le cas d'une loi normale isotrope ou même d'une loi quelconque à variance finie) une expression asymptotique, lorsque la distance a tend vers l'infini, du coefficient de parenté $\varphi(a)$ défini par la formule (3.6), est :

$$\varphi(a) \sim \text{cte} \times K_0\left(\frac{a\sqrt{8k}}{\sigma}\right) \sim \text{cte} \times e^{-\frac{a\sqrt{8k}}{\sigma}} a^{-\frac{1}{2}}$$

σ^2 désignant le deuxième moment de la « distance maritale » (telle que l'a définie CAVALLI-SFORZA, c'est-à-dire le double de la variance de la migration parentale).

Nous constatons donc que la migration continue fournit, *pour les grandes valeurs de la distance*, la même décroissance que la migration discontinue. Mais il n'en est pas de même pour les petites distances puisque la loi continue n'a plus alors de signification concrète.

Reçu pour publication en décembre 1971.

SUMMARY

GENETICS OF NATURAL DIPLOID POPULATIONS FOR ONE LOCUS

II. — STUDY OF KINSHIP COEFFICIENT.

With same notations as in part I, the *a priori* variances and covariances of gene *a* in panmictic groups of finite number *N* related by migration are now [chap. II, 3°] calculated via kinship coefficients.

The kinship coefficient φ_{xy} between site *x* and site *y* is the *a priori* probability of identity (by descent from a common ancestor) of two loci randomly drawn in sites *x* and *y*; the notion of « betamutation » (transition to independence, with probability *k*, in each transmission of a locus to the following generation) allows to take, as the alternative to « identity without any betamutation », the stochastic independence, which then has probability $1 - \varphi_{xy}$.

The calculation of φ_{xx} — which is also the coefficient of inbreeding of a diploid in the panmictic group of site *x* — allows (chap. III) to obtain again (after multiplying by $C - C^2$) the results obtained in part I for the intergroup variance.

The decrease, when $y - x = d$ increases, of the kinship coefficient $\varphi_{xx} = \varphi(d)$, and of covariance and correlation [proportional to $\varphi(d)$], is exponential in the unidimensional case [chap. III, II, 1°] with only two parameters, *k* and σ^2 (migration variance). In case of discrete bidimensional migration [II, 2], the numerical difference with unidimensional case remains small at low distance, except if k/σ^2 is very small, then the tendency to zero at large distance is somewhat quicker.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ABRAMOWICZ et STEGUN, 1964 : handbook of mathematical fonctions
- CAVALLI-SFORZA, 1963. In J. SUTTER, *Entretiens de Monaco en sciences humaines*, Monaco, Sciences humaines, 139 à 158.
- GILLOIS M., 1964. *La relation d'identité en génétique*. Thèse Fac. Sci. Paris (ronéotée).
- GILLOIS M., 1966. La relation de dépendance en génétique. *Ann. Inst. Henri Poincaré*, **2**, 3, section B, 261 à 278.
- KIMURA M., 1968 *a*. Evolutionary rate at the molecular level. *Nature*, **217**, 624 - 626.
- KIMURA M., 1968 *b*. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet. Res.* **11**, 247 - 269.
- MALÉCOT G., 1948. *Les mathématiques de l'hérédité*. Masson, Paris, 60 p.
- MALÉCOT G., 1950. Quelques schémas probabilistes sur la variabilité des populations naturelles. *Ann. Univ. Lyon, Sciences, Section A*, **13**, 37.
- MALÉCOT G., 1951. Un traitement stochastique des problèmes linéaires (mutation, linkage, migration) en génétique de populations. *Ann. Univ. Lyon, Sciences, Section A*, **14**, 79.
- MALÉCOT G., 1954. Sur les modèles stochastiques linéaires asymptotiquement stationnaires. *Ann. Univ. Lyon, Sciences, Section A*, **17**, 19.
- MALÉCOT G., 1955. The decrease of relationship with distance. *Cold Spring Harbor Symposium on Quantitative Biology*, 52.
- MALÉCOT G., 1959. Les modèles stochastiques en génétique de population. *Publications de l'Institut de Statistique de Paris*, **8**, F 3, 173.

- MALÉCOT G., 1963. Migration et parenté génétique moyenne. In J. SUTTER, *Entretiens de Monaco en sciences humaines*, Monaco, Sciences Humaines, 205 - 212.
- MALÉCOT G., 1966 a. Identical loci and relationship. *Proc. V Berkely Symposium*, 1965, 317.
- MALÉCOT G., 1966 b. Probabilités et hérédité. *Travaux et documents de l'I.N.E.D.*, cahier n° 47, P. U. F., Paris, 1 vol.
- MALÉCOT G., 1969. *The mathematics of heredity*. Freeman, San Francisco, 86 p.
- MALÉCOT G., 1969. Consanguinité panmictique et consanguinité systématique. *Ann. Génét. Sé. anim.*, **1**, 237- 242.
- MALÉCOT G., 1971. Structure géographique et variabilité d'une grande population. *Excerpta Medica*, *III/8* (sous presse).
- MORTON N., 1970. Isolation by distance in artificial populations. *Genetics*, **66**, 569.
- WRIGHT S., 1931. Evolution in mendelian populations. *Genetics*, **16**, 97.
-