

Isoestérases chez le nématode *Heterodera avenae*.

II. — Modélisation appliquée à la structure génétique de deux populations

Michel GILLOIS *, Claude CHEVALET *, Dominique THOMAS **
et Antoine DALMASSO **

* I.N.R.A., Laboratoire de Génétique Cellulaire, chemin de Borde Rouge, Auzeville, 31320 Castanet-Tolosan, France.
** I.N.R.A., Station de Recherches sur les Nématodes, 123 boulevard Francis Meilland, 06602 Antibes, France.

RÉSUMÉ

Les auteurs formulent un modèle en trois parties permettant d'apprécier par électrophorèse la distribution des états géniques d'une population examinée pour son polymorphisme génétique. L'application de ce modèle à l'analyse des fréquences du locus « estérase b » de deux populations, Fr₁ et Fr₄ (Villasavary, Aude et Nuisement-sur-Cooles, Marne), d'*Heterodera avenae* élevées en microparcelles, montre que l'une des populations (Fr₁) suit la loi de Hardy-Weinberg, c'est-à-dire qu'il y a un régime de panmixie, alors que la seconde (Fr₄) présente un excédent en homozygotes. L'étude d'autres loci doit permettre de savoir si cela résulte d'un allèle nul ayant échappé aux observations et ayant faussé le nombre réel d'homozygotes, ou d'une consanguinité due à des effectifs limités. Dans le dernier cas le modèle rendrait très bien compte d'accouplements entre frère et sœur dans une proportion de 80% de la population. Cela n'est pas tellement surprenant car les œufs des *Heterodera* restent rassemblés dans le kyste, ce qui favorise aussi l'« inbreeding ».

SUMMARY

Isoesterases in the nematode Heterodera avenae. II — A genetic model applied to two populations

A general model to appreciate allelic distributions in populations analyzed for protein polymorphism was applied to two *Heterodera avenae* strains; one (Fr₁) from Villasavary (Aude), the other (Fr₄) from Nuisement-sur-Cooles (Marne). These strains have been multiplied in microplots at Rennes for five years. Data obtained on "est. b" locus (a non specific esterase) showed differences between Fr₁ and Fr₄. Fr₁ was demonstrated to be in Hardy-Weinberg equilibrium while the second presented a strong excess of homozygotes. Studies on other loci would permit to know if the "est. b" homozygote frequency have been exaggerated in Fr₄ for a null hidden allele at this locus, or if it is a true excess resulting of some inbreeding behavior in relation with low density. In this last case the model would account for brother × sister inbreeding of 80% of the Fr₄ population. In fact this is not so surprising for eggs of *Heterodera* are not widespread at egg laying, they remain aggregated in the cyst and that, added to the low density, might have favoured inbreeding.

Méthode d'analyse de la structure génétique d'une population

A cause des difficultés techniques imposées par la taille des nématodes phytophages, peu de travaux traitent de leur diversité génétique. Cette notion renferme cependant des informations capitales sur les potentialités adaptatives des espèces : distribution géographique, agressivité vis-à-vis des plantes, etc... On sait depuis peu que cette diversité n'est pas du tout du

même ordre entre un *Meloidogyne* parthénogénétique obligatoire et un *Heterodera* amphimictique, comme l'est *H. avenae*. Le mode de reproduction y est pour beaucoup. Mais la diversité est aussi affectée par d'autres facteurs liés à l'environnement ou à l'échelle géographique (faculté de dispersion estimée en amplitude et fréquence). Évaluer la diversité interne de populations d'*H. avenae* est un premier point important (Bergé *et al.* 1981); analyser le pourquoi d'une structure génétique observée peut en plus

améliorer la connaissance des mœurs reproductrices des nématodes dans la nature. Il est proposé aux nématologistes une démarche permettant une approche génétique de l'origine des phénomènes de spéciation, en partant d'un exemple puisé dans une publication précédente (Bergé *et al.* 1981). Ainsi, si l'on dispose d'échantillons recueillis en des localisations distinctes, on peut évaluer des distances génétiques entre populations (Nei, 1971, 1972, 1974) ; dans une même population on peut mesurer la variabilité génétique par un indice d'hétérozygotie et caractériser la structure génotypique par rapport aux fréquences alléliques observées. Ces études complètent, par des données génétiques, l'analyse des différences morphologiques étudiées entre populations, ou entre individus à l'intérieur d'une même population. L'analyse de la structure génétique peut apporter des informations aussi sur l'existence ou non de pression de sélection au niveau de certains loci. Considérons un locus quelconque, non lié au sexe, où l'on distingue plusieurs allèles A_1, A_2, \dots, A_r , de fréquences f_1, f_2, \dots, f_r , dans un groupe d'individus reproducteurs, à un instant donné. Si ces reproducteurs s'accouplent au hasard, sans tenir compte de leurs génotypes respectifs (panmixie), alors les génotypes des enfants nés de ces accouplements ont les probabilités suivantes :

$$\begin{aligned} (I) \quad Pr \{ A_i A_i \} &= f_i^2 & (1 \leq i \leq r) \\ Pr \{ A_i A_j \} &= 2 f_i f_j & (1 \leq i < j \leq r) \end{aligned}$$

C'est la distribution de Hardy-Weinberg qui concerne les génotypes. Si les phénotypes sont moins nombreux, à cause du phénomène de dominance, on doit exprimer les probabilités des phénotypes en faisant la somme des fréquences des génotypes correspondant à un même phénotype : ce sont ces nouvelles expressions théoriques qui doivent être utilisées dans les analyses statistiques (cf. l'étude des groupes sanguins ABO chez l'homme, Elandt-Johnson, 1971, chap. 14).

Si la population est très nombreuse et constitue une seule unité de reproduction, si à chaque génération les accouplements se font au hasard, si la population ne reçoit pas d'immigrant (ce qui est souvent le cas pour les nématodes phytophages) et si aucune pression de sélection ne s'exerce sur le locus considéré, alors ces fréquences f_i , et les fréquences génotypiques

demeurent constantes au cours des générations (loi de Hardy-Weinberg). Ici l'on dispose d'observations dans une seule génération, on a pour seule information la distribution des génotypes à un instant donné de l'histoire de la population, et il n'est pas possible d'en inférer des conclusions concernant les forces qui ont guidé l'évolution de la population. On peut seulement confronter la distribution observée à des distributions théoriques obtenues sous certaines hypothèses et proposer des interprétations aux observations.

La confrontation de la distribution observée à une distribution de Hardy-Weinberg est la première étape. Une concordance entre elles est en faveur de l'hypothèse selon laquelle les accouplements se sont faits au hasard. En revanche, cela n'indique rien en ce qui concerne la taille de la population, l'existence des migrations avec d'autres populations avoisinantes. Si les individus observés sont jeunes, cela peut suggérer l'absence de sélection gamétique intense, mais n'indique rien sur les forces de sélection que peuvent subir les zygotes entre la naissance et l'âge reproducteur.

Si un écart significatif est observé par rapport à la distribution de Hardy-Weinberg, ses causes possibles doivent être recherchées parmi les phénomènes qui ont un effet immédiat sur les fréquences génotypiques. On peut envisager une sélection gamétique intense traduite par des pourcentages anormaux de la ségrégation des gènes, ou bien des phénomènes d'incompatibilité partielle ou totale entre certains allèles, induisant une mortalité précoce de certains zygotes. De telles causes ne peuvent pas en général être distinguées par une analyse statistique ; en revanche ces phénomènes entrent dans le champ des méthodes de la génétique formelle, et doivent être élucidés avant d'être pris en compte dans un modèle de génétique des populations. Par ailleurs, on peut signaler qu'une grande partie des polymorphismes actuellement révélés par électrophorèse sont relatifs à des enzymes dont les fonctions très générales laissent penser que les gènes correspondants sont sans doute peu sujets à une sélection très intense. Si l'on suppose donc que le locus considéré obéit bien aux lois de Mendel, les causes d'écart à une distribution de Hardy-Weinberg

doivent être recherchées dans les règles d'accouplement.

Deux phénomènes principaux peuvent induire des corrélations entre les génotypes de deux individus qui s'accouplent : la consanguinité et l'homogamie. La consanguinité réunit dans un appariement deux individus apparentés. La cause d'une consanguinité systématique peut tenir simplement à des considérations géographiques, si par exemple des individus nés au même endroit, et donc ayant des chances d'être apparentés, se déplacent peu entre leur naissance et l'âge de la reproduction (cela est le cas de nombreux nématodes du sol). Plus généralement l'existence de colonies de petites tailles, réparties sur un même territoire et échangeant par migration quelques reproducteurs, est une forme de consanguinité systématique. La conséquence de l'appariement des parents est la naissance d'un excès d'homozygotes, puisque les parents ont des chances accrues de porter ensemble un même allèle. L'homogamie traduit le fait que les conjoints se réunissent préférentiellement en fonction de leurs génotypes ou de leur phénotypes ; elle peut être positive si le pourcentage des couples présentant un même phénotype est supérieur à celui attendu dans le cas d'accouplements au hasard, ou négative dans le cas contraire. Si l'homogamie est positive, elle se traduit par un accroissement de la proportion d'homozygotes chez les descendants, comme la consanguinité, mais cet effet peut être différentiel selon les allèles portés, alors qu'il est toujours uniforme dans le cas de consanguinité. Si l'homogamie est négative, elle conduit au contraire à un excès de génotypes hétérozygotes. Une formulation unifiée de la consanguinité et de l'homogamie a été proposée par Gillois (1964, 1966, 1967), Gillois, Bouffette et Bouffette (1969 a, b) et Croizé-Pourcelet (1970) ; elle se traduit par les expressions suivantes des probabilités des génotypes :

$$Pr \{A_i A_i\} = F1 p_{1i} + F3 p_{3i}^2 \quad (1 \leq i \leq r)$$

$$(II) \quad Pr \{A_i A_j\} = F_2 P_{2,ij} + F3 2p_{3i}p_{3j} \quad (1 \leq i < j \leq r)$$

avec :

$$\begin{aligned} \sum_{i=1}^r p_{1i} &= \sum_{i=1}^r p_{3i} = \sum_{1 \leq i < j \leq r} P_{2,ij} \\ &= F1 + F2 + F3 = 1 \end{aligned}$$

Ces expressions traduisent les conséquences d'une structure théorique d'accouplements ; elles permettent de rendre compte des tendances à l'homozygotie (loi caractérisée par une probabilité F_1 , et des fréquences théoriques p_{1i}), à l'hétérozygotie (loi de probabilité F_2 , avec des fréquences théoriques $P_{2,ij}$) et au croisement au hasard (loi de probabilité F_3 , avec des fréquences alléliques théoriques (p_{3i})). Ces formules n'ont de sens que dans l'étude de l'évolution d'une population soumise à des règles particulières d'accouplements ; elles ne peuvent pas, en général, être à la base d'une procédure d'estimation des paramètres introduits car ceux-ci sont redondants, mais elles donnent des interprétations possibles aux écarts à la distribution de Hardy-Weinberg. Les interprétations possibles, compatibles avec les données recueillies, doivent ensuite être soumises à un test expérimental complémentaire.

La procédure d'analyse peut être la suivante. On teste d'abord l'hypothèse ($F_3 = 1$) (donc : $F1 = F2 = 0$), c'est-à-dire la distribution de Hardy-Weinberg, par un test de χ^2 . Supposant qu'on a mis en évidence r allèles, dont les relations de dominance permettent de définir k phénotypes distincts ($k \leq r(r+1)/2$), le test s'écrit :

$$\chi^2 = \sum_{\ell=1}^{l=k} \frac{(n_{\ell} - N_{\ell})^2}{N_{\ell}}$$

où : n_{ℓ} est le nombre d'individus observés de phénotype ℓ , et N_{ℓ} est le nombre théorique espéré d'individus de phénotype ℓ . Si tous les génotypes sont observables, alors $k = r(r+1)/2$, et les nombres N_{ℓ} s'écrivent :

$$N_{ii} = N f_i^2 \quad N_{ij} = 2N f_i f_j$$

où f_i est la fréquence observée de l'allèle A_i , et $N = \sum_{i=1}^k n_{\ell}$ le nombre des observations.

Si $k < r(r+1)/2$, il faut estimer les fréquences alléliques p_{3i} par la méthode du maximum de vraisemblance (Elandt-Johnson, 1971) et exprimer les nombres théoriques N_p d'après ces estimations. Le nombre de degrés de liberté est toujours égal à $k - r$. Si l'hypothèse ($F3 = 1$) est admissible, on s'en tient à l'hypothèse des accouplements au hasard.

Si cette hypothèse de panmixie est rejetée, il faut considérer la fréquence globale des homo-

zygotes. On peut la comparer à la somme des carrés des fréquences alléliques, $q = \sum_i f_i^2$, si les fréquences sont directement observables (sinon, une procédure d'estimation est nécessaire, qui exige la formulation d'une hypothèse). Il s'agit de faire un test selon une loi binomiale qui pourra être en général assimilée à une loi normale, de moyenne q et de variance $q(1-q)/N$, si N est assez grand, si le pourcentage d'homozygotes est trop élevé, l'hypothèse la plus simple à envisager est celle d'une consanguinité systématique pure. Dans ce cas les formules II se simplifient, en posant : $F1 = f$, $F2 = 0$, $F3 = 1 - f$, : $p_{1i} = p_{3i} = p_i$, et elles deviennent :

$$(III) \quad \begin{aligned} Pr \{ A_i A_j \} &= f p_i + (1 - f) p_i^2 & 1 \leq i \leq r \\ Pr \{ A_i A_j \} &= (1 - f) 2 p_i p_j & 1 \leq i < j \leq r \end{aligned}$$

On doit alors estimer, par le maximum de vraisemblance les fréquences théoriques p_i , et le coefficient de consanguinité moyen f de l'échantillon, puis reporter ces estimations dans l'expression du test χ^2 . Le nombre de degrés de liberté est diminué de 1 puisqu'on introduit un paramètre explicatif supplémentaire, il devient : $k - r - 1$. Si les observations sont conformes à une distribution du type (III), c'est un argument en faveur d'une consanguinité systématique, mais on peut aussi envisager une homogamie positive non différentielle.

Si le pourcentage d'hétérozygotes, au contraire, est trop élevé, il faut chercher à estimer $F2$, en faisant $F1 = 0$. Pour que le problème ne soit pas indéterminé, il faut préciser l'hypothèse sous-jacente, en postulant par exemple une homogamie négative non différenciée en fonction des génotypes qui s'unissent ; on écrit alors, avec $p_i = p_{3i}$:

$$(IV) \quad \begin{aligned} P_{2,ij} &= p_i \frac{p_j}{1 - p_i} + p_j \frac{p_i}{1 - p_j} = \\ &2 p_i p_j \frac{1 - (p_i + p_j)/2}{(1 - p_i)(1 - p_j)} \end{aligned}$$

Il faut à nouveau tester la conformité des observations à la distribution (II - IV), en estimant, à nouveau par le maximum de vraisemblance, les paramètres p_i et $F2$ ($F3 = 1 - F2$). Il faut noter ici que les calculs numériques

peuvent être rendus assez complexes par la forme des relations (IV). Ici encore le nombre de degrés de liberté du χ^2 est de : $k - r - 1$. La conformité des observations à cette distribution peut traduire une homogamie négative, mais aussi une sélection précoce favorable à l'ensemble des hétérozygotes.

Enfin, si les formulations (III) et (IV) ne sont pas satisfaisantes, ou si la proportion des homozygotes paraît compatible avec sa fréquence théorique $q = \sum f_i^2$ sans que (I) ne soit globalement satisfaisant, aucune hypothèse simple ne peut être avancée. Une vérification du déterminisme génétique de la variation enzymatique considérée peut être envisagée. Un ajustement des données à un modèle du type (II), avec $F1$, $F2$ et $F3$ non nuls, et avec des relations particulières du type de (IV), peut être envisagé mais, il ne pourra pas en général conduire à une interprétation claire des observations.

Exemple d'application : Analyse des populations Fr_1 et Fr_4 au locus estérase b chez *Heterodera avenae*

DONNÉES EXPÉRIMENTALES

Bergé *et al.* (1981) ont caractérisé les différents allèles par leur distance de migration par rapport à l'un d'eux arbitrairement désigné 1,0 pris comme référence et dénommés b 0,70 - b 0,90 - b 1,10 et b 1,20. Ci-dessous, ils seront désignés respectivement par 1-2-3-4 et 5.

Le tableau 1 donne les effectifs observés des différents génotypes et le tableau 2 les fréquences observées f_i ($i = 1, \dots, 5$) des différents allèles.

VÉRIFICATION DE L'HYPOTHÈSE DE PANMIXIE

On commence par tester si les distributions des génotypes sont compatibles avec la panmixie (Hardy-Weinberg), $F3 = 1$. Si l'hypothèse est correcte, ces distributions doivent être proches de celle donnée par f_1^2 , $2 f_1 f_2$, etc...

Ces distributions théoriques sont, en nombre d'individus dans chaque classe, arrondis à 10^{-2} près ; les nombres théoriques d'individus qui devraient être observés pour 100 individus pris au hasard en fonction des fréquences alléliques sont donnés au tableau 3.

Tableau 1

Effectifs observés des différents génotypes dans les populations Fr_1 et Fr_4

Observed frequencies of the genotypes in populations Fr_1 and Fr_4

	Génotypes		Effectifs observés	
	allèle/allèle	Nomenclature Bergé et al.	Observed frequencies Fr_1	Fr_4
Homozygotes :	1/1	0,70/0,70	8	0
	2/2	0,90/0,90	1	11
	3/3	1,00/1,00	38	1
	4/4	1,10/1,10	0	11
	5/5	1,20/1,20	0	21
Hétérozygotes :	1/2	0,70/0,90	6	0
	1/3	0,70/1,00	17	0
	1/4	0,70/1,10	1	0
	1/5	0,70/1,20	1	0
	2/3	0,90/1,00	14	6
	2/4	0,90/1,10	1	15
	2/5	0,90/1,20	0	13
	3/4	1,00/1,10	8	3
	3/5	1,00/1,20	5	2
	4/5	1,10/1,20	0	17
Totaux : homozygotes			47	44
hétérozygotes			53	56

On fait alors deux tests de compatibilité : d'une part un test de χ^2 d'autre part un test portant sur la fréquence des homozygotes.

Test de χ^2

— Fr_1 : $k = 15$; $L = 5$; $v = 10$

Le calcul donne $\chi^2 \simeq 9,46$: il n'y a pas d'écart significatif. L'hypothèse d'identité des valeurs observées et des valeurs théoriques est acceptée.

— Fr_4 : $k = 10$, $r = 4$ (seulement quatre fréquences non nulles), $v = 6$; on trouve : $\chi^2 = 14,04$, l'écart est significatif (à 5%). L'hypothèse d'identité entre valeurs observées et théoriques est cette fois rejetée.

Tests des fréquences des homozygotes

Tableau 2

Fréquences des différents allèles dans les populations Fr_1 et Fr_4

Relative frequencies of the alleles in populations Fr_1 and Fr_4

Allèles	Fréquence Relative frequency	Fr_1	Fr_4
1	f_1	= 0,205	0
2	f_2	= 0,115	0,280
3	f_3	= 0,600	0,065
4	f_4	= 0,050	0,285
5	f_5	= 0,030	0,370

Si la loi de Hardy-Weinberg est satisfaisante, la fréquence des homozygotes doit être de : Σf_i^2 , soit 0,4187 pour Fr_1 (somme des fréquences théoriques d'homozygotes), et 0,3008 pour Fr_4 .

Les fréquences observées, soit 0,47 pour Fr_1 et 0,44 pour Fr_4 devraient être compatibles avec le résultat de 100 tirages binomiaux avec les probabilités théoriques précédentes. Comme 100 est assez grand, on peut admettre l'approximation par la loi de Gauss. Les écarts-types respectifs de la moyenne de 100 tirages sont alors de 0,049 pour Fr_1 et de 0,046 pour Fr_4 .

Tableau 3

Répartition théorique de 100 individus dans les différents génotypes
Theoretical percentages of individuals in the different genotypes

Allèles	Fr ₁					Allèles	Fr ₄				
	1	2	3	4	5		1	2	3	4	5
× 1	42,0	4,72	24,6	2,05	1,23	× 1 (absent)	0	0	0	0	0
2		1,32	13,8	1,15	0,69	2		7,84	3,64	15,96	20,72
3			36	6	3,6	3			0,42	3,71	4,81
4				0,25	0,3	4				8,12	21,09
5					0,09	5					13,69

La valeur observée pour Fr₁ est compatible, celle observée pour Fr₄ est incompatible (au seuil de 1%). En effet l'écart réduit :

$$\varepsilon = \frac{\text{pourcentage observé} - q}{\sqrt{\frac{q(1-q)}{100}}}$$

vaut $\frac{0,47 - 0,4187}{0,049} = 1,04$ pour Fr₁

La table de l'écart réduit (loi normale) indique la compatibilité entre les valeurs théoriques et observées pour Fr₁. Pour Fr₄ $\varepsilon = 3,03$, il y a un écart significatif et l'hypothèse de l'équilibre de Hardy-Weinberg n'est pas vérifiée dans ce cas.

Conclusion

L'hypothèse de panmixie est satisfaisante pour Fr₁, mais pas pour Fr₄ : pour cette seconde population l'excès significatif de génotypes homozygotes fait penser à une consanguinité systématique, ce qui est fort possible compte tenu des conditions d'élevage, ou à l'existence d'un allèle nul non repéré par l'observateur. On va donc essayer d'estimer le taux de consanguinité (F1 > 0, F2 = 0), puis voir si la distribution théorique tirée de cette valeur estimée est compatible avec les observations, dans la mesure où on écarte la deuxième éventualité. Cette dernière sera l'objet du test sur l'hypothèse « allèle nul ».

ESTIMATION DES PARAMÈTRES POUR Fr₄,
 DANS L'HYPOTHÈSE DE CONSANGUINITÉ

Le plus simple est de supposer que les génotypes admettent des probabilités de la forme (où $f = F1$) :

$$(A) \quad P_{ii} = f p_i + (1 - f) p_i^2$$

$$P_{ij} = 2(1 - f) p_i p_j \quad (i \neq j)$$

et d'estimer les cinq paramètres :

$$f \quad (0 < f < 1), p_2, p_3, p_4, p_5 \quad (\sum p_i = 1)$$

($p_1 = 0$, en effet l'allèle 0,70 n'est pas présent dans la population Fr₄). On peut aussi imaginer de représenter la situation en disant qu'une proportion λ d'individus est issue d'accouplements frère \times sœur (conduisant à une consanguinité individuelle φ de 25% pour ces individus), et une proportion $(1 - \lambda)$ d'individus issus d'accouplements entre partenaires non apparentés. En admettant que les gènes sont tous tirés dans une même réserve gamétique caractérisée par les probabilités p_i , cela donnerait comme formules théoriques :

$$(P_{ii} = (1 - \lambda) p_i^2 + \lambda (\varphi p_i + (1 - \varphi) p_i^2)$$

(avec $\varphi = 1/4$)

$$(P_{ij} = (1 - \lambda) 2 p_i p_j + \lambda \{(1 - \varphi) 2 p_i p_j\}$$

($i \neq j$)

soit :

$$(B) \quad P_{ii} = \lambda \varphi p_i + (1 - \lambda \varphi) p_i^2$$

$$P_{ij} = (1 - \lambda \varphi) 2 p_i p_j \quad (i \neq j)$$

On voit que les deux modèles théoriques (A) et (B) ne peuvent pas être dissociés par l'expérience réalisée, leurs paramètres étant liés par : $f = \lambda \varphi$. On a donc estimé f selon le modèle A, par la méthode du maximum de vraisemblance.

On trouve : $\hat{f} = 0,20$

et les probabilités associées p_i ont pour estimations :

$(p_1 = 0)$ $\hat{p} = 0,282$ contre $0,280$ (fréquence observée)

$\hat{p}_3 = 0,068$ $0,065$
 $\hat{p}_4 = 0,287$ $0,285$
 $\hat{p}_5 = 0,363$ $0,370$

(si l'on interprète ces résultats selon le modèle avec $\varphi = 1/4$, c'est à-dire avec des accouplements frère \times sœur, cela donne une estimation $\hat{\lambda} = 0,80$, car $\hat{f} = \varphi \lambda$).

Les probabilités théoriques des génotypes selon ces estimations, données par (B) sont données au tableau 4.

Tableau 4

Probabilités théoriques des génotypes Theoretical probabilities of the genotypes					
	1	2	3	4	5
1	0	0	0	0	0
2		0,12	0,031	0,13	0,164
3			0,017	0,031	0,039
4				0,124	0,167
5					0,178

Le test de conformité du χ^2 ne donne pas d'écart significatif des observations à cette distribution ($\chi^2 = 5,91$ pour 5 degrés de liberté). Par ailleurs, la fréquence théorique des homozygotes devient ici égale à 0,439, c'est-à-dire exactement la fréquence observée.

En conclusion, ces valeurs estimées \hat{f} et \hat{p}_i (qui sont d'ailleurs très proches des fréquences observées f_i) rendent très bien compte des observations.

ESTIMATION DES PARAMÈTRES DE Fr_4 DANS L'HYPOTHÈSE D'UN CINQUIÈME ALLÈLE NUL

Si l'on suppose l'existence d'un allèle nul, (noté ici A_1 , quoique différent de b 0,70) ne se traduisant par aucune bande sur le gel, les individus de phénotypes (A_i) ($i = 2, 3, 4, 5$) peuvent être de génotypes $A_i A_i$ ou bien $A_1 A_i$. On a alors en tout 11 classes phénotypiques observables, dont les fréquences théoriques sont les suivantes dans l'hypothèse de panmixie :

$$\begin{aligned} Pr \{(A_1)\} &= Pr \{A_1 A_1\} = p_1^2 \\ Pr \{(A_i)\} &= Pr \{A_i A_i\} + Pr \{A_1 A_i\} \\ (V) \quad &= p_i^2 + 2p_1 p_i \quad (i > 1) \\ Pr \{(A_i A_j)\} &= Pr \{A_1 A_j\} = 2 p_i p_j \\ &\quad (1 < i < j \leq 5) \end{aligned}$$

L'estimation, selon le maximum de vraisemblance, des cinq fréquences théoriques p_i donne :

$\hat{p}_1 = 0,078$ probabilité de l'allèle nul
 $\hat{p}_2 = 0,259$
 $\hat{p}_3 = 0,061$
 $\hat{p}_4 = 0,265$
 $\hat{p}_5 = 0,337$

L'ajustement des données au modèle (V) est alors très bon : $\chi^2 = 6,77$ pour 6 degrés de liberté. La fréquence théorique des homozygotes apparents est alors de 0,404 ; en accord avec la fréquence observée de 44/100.

L'hypothèse d'un cinquième allèle, nul, présent dans la population Fr_4 est donc aussi bien conforme aux observations que celle d'une consanguinité systématique.

Discussion

Le polymorphisme génétique associé au locus d'isoestérase b de deux populations Fr_1 et Fr_4 a été soumis à une analyse critique, selon un modèle qui permet de savoir si les états géniques sont en accord avec la distribution de Hardy-Weinberg ou s'il y a des excès en homozygotes ou hétérozygotes.

Le locus « est. b » de la population Fr_1 compte au moins cinq allèles qui sont en équilibre avec ladite distribution. Par contre la population Fr_4 qui ne présente que quatre allèles pour ce même locus a un excès d'homozygotes. L'hypothèse selon laquelle une partie de la population est soumise à une consanguinité individuelle de 25% (résultat d'accouplements entre frère et sœur) permet une nouvelle approximation des fréquences théoriques de ces allèles, qui est cette fois en conformité avec les fréquences observées.

On ne peut toutefois écarter la possibilité d'un allèle supplémentaire nul qui aurait contribué à gonfler artificiellement le nombre d'homozygotes. L'étude d'autres loci devrait apporter des éclaircissements sur ce point. Un nouvel excès d'homozygotes serait en faveur d'un régime de consanguinité, à l'opposé une distribution en accord avec l'équilibre de Hardy-Weinberg impliquerait un cinquième allèle nul

pour le locus « est. b ». Dans le cas d'un matériel autre que *H. avenae* on proposerait des expériences de croisements entre deux phénotypes homozygotes distincts : l'analyse de la descendance permettrait de trancher aussi facilement.

RÉFÉRENCES

- BERGÉ, J.B., DALMASSO, A., PERSON, F., THOMAS, D. & RIVOAL, R. (1981). Isoestérases chez *Heterodera avenae* : polymorphisme chez différentes races françaises. *Revue Nématol.*, 4 : 99-105
- CROIZÉ-POURCELET, J. (1970). Etude de l'évolution des petites populations homogames par la théorie de la contrainte. *Ann. Génét. Sél. anim.*, 2 : 37-52.
- ELANDT-JOHNSON, R.C. (1971). *Probability models and statistical methods in genetics*. New York, Wiley, 592 p.
- GILLOIS, M. (1964). *La relation d'identité en génétique*. Thèse doctorat d'Etat Fac. Sci. Paris, 294 p.
- GILLOIS, M. (1966). La relation de dépendance en génétique. *Ann. Inst. Henri Poincaré*, B 2 : 261-278.
- GILLOIS, M. (1967). Les lois conjointes des variables aléatoires génétiques. *Ann. Génét.*, 10 : 203-206.
- GILLOIS, M., BOUFFETTE, J. & BOUFFETTE, A.R. (1969a). Etude des populations d'effectif limité, homogames, phénotypiques et panmictiques. *Ann. Inst. Henri Poincaré*, B 5 : 69-86.
- GILLOIS, M., BOUFFETTE, J. & BOUFFETTE, A.R. (1969b). Covariance génotypique a priori dans les populations homogames. *Ann. Inst. Henri Poincaré*, B 5 : 87-89.
- NEI, M. (1971). Identity of genes and genetic distance between populations. *Genetics*, 68 : 47 p.
- NEI, M. (1972). Genetic distance between populations. *Am. Naturalist*, 106 : 283-292.
- NEI, M. & ROYCHOUDHURY, A.K. (1974). Sampling variances of heterozygosity and genetic distance. *Genetics*, 76 : 379-390.

Accepté pour publication le 3 octobre 1980.