

MOYENS POSSIBLES DE DÉCELER DES GÈNES INFLUANT SUR LA VARIANCE PHÉNOTYPIQUE

P. MÉRAT

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

Quoique cela ne soit pas souvent pris en considération, certains gènes ou combinaisons génétiques particuliers peuvent influencer sur la variance phénotypique d'un caractère quantitatif.

En particulier, l'hétérozygotie apportée par le croisement diminue souvent la variance attribuable à des causes non génétiques (LERNER, 1954).

Chez la Poule, divers auteurs ont observé que les coquelets hétérozygotes Kk ont un poids significativement plus variable que les homozygotes kk vers l'âge de 8-10 semaines, sans que ceci s'accompagne d'une différence appréciable des valeurs moyennes (MÉRAT, 1966).

Nous avons discuté d'autre part (MÉRAT, 1967 *a*) les conséquences que peut avoir la présence dans une population de ce type de variabilité génétique. Le comportement de tels gènes vis-à-vis de la sélection est en effet particulier et pourrait avoir des incidences pratiques quant au choix de l'intensité de sélection, ou fournir une voie nouvelle d'interprétation du « plafonnement » des performances après une sélection de longue durée.

Si l'on connaît peu d'exemples caractérisés de gènes influant sur les fluctuations d'origine extra-génétique, c'est qu'on les a peu recherchés, mais aussi qu'ils ne sont pas faciles à mettre en évidence.

Lorsqu'on dispose d'un locus « marqueur » tel K/k chez la poule, la comparaison de la variance phénotypique associée aux divers génotypes pour ce locus s'impose d'elle-même.

Si, par contre, indépendamment de tout gène marqueur, l'on cherche à déceler des gènes « de variance » non fixés dans une population, les possibilités sont restreintes. On peut tenter une expérience de sélection à l'exemple de FALCONER et ROBERTSON (1956) ou de PROUT (1963) ; mais, outre les difficultés possibles d'interprétation (1), la durée en sera longue pour la plupart des espèces domestiques.

(1) Celle, notamment, de savoir si, par sélection des phénotypes extrêmes en plus et en moins, on aura trié des allèles de plus ou moins grande variabilité, ou simplement choisi un mélange de génotypes extrêmes bien distincts.

Il paraît alors naturel de se tourner vers une analyse statistique intragénération, compte tenu de la subdivision en familles ou groupes génétiques distincts.

Nous nous placerons dans le cas où les facteurs non génétiques de variation sont assez nombreux et répartis au hasard entre ces groupes, et dans l'hypothèse de panmixie.

La constatation d'une hétérogénéité des variances intra-familles peut faire penser au type de gènes auquel nous nous intéressons, mais elle n'est pas suffisante pour prouver leur existence. En effet, il est clair que la ségrégation dans certaines familles et non dans d'autres de gènes à effet important sur la valeur moyenne du caractère étudié aboutira au même résultat. Nous montrons (MÉRAT, données non publiées) que c'est aussi le cas lorsqu'un nombre modérément grand de gènes à effet individuel relativement petit est responsable de la majeure partie de la variabilité génétique.

C'est par l'examen de la *forme* des distributions de fréquence suivant les familles, que nous suggérons un test complétant le précédent.

Nous allons en effet montrer que la ségrégation de gènes influant sur la variance phénotypique doit produire, dans les familles ou groupes où elle a lieu, des distributions de fréquences « trop pointues » ou leptokurtiques, à coefficient de « kurtosis » (KENDALL, 1959) positif.

a) *Gènes modifiant l'écart-type et non la moyenne*

Les familles présentant une disjonction pour de tels gènes constituent un mélange de populations de même moyenne (que nous prendrons égale à zéro), d'écart-types respectifs $\sigma_1, \sigma_2, \dots$, dans la proportion p_1, p_2, \dots

Nous supposons que chacune de ces populations peut être considérée comme normale.

Le coefficient de « kurtosis » γ_2 de la population résultante ⁽¹⁾ est du signe de

$$E(x^4) - 3 [E(x^2)]^2$$

x étant une valeur individuelle quelconque.

Détaillons d'après les distributions composantes et posons $x'_i = \frac{x_i}{\sigma_i}$, i étant l'indice attribué à l'une de ces distributions.

$$\text{On obtient : } E(x^4) = \sum_i p_i \sigma_i^4 E(x'^4_i) = 3 \sum_i p_i \sigma_i^4$$

Compte tenu de ce que $E(x'^2_i) = 3$ dans l'hypothèse de la normalité des distributions élémentaires.

$$\text{D'autre part : } E(x^2) = \sum_i p_i \sigma_i^2 E(x'^2_i) = \sum_i p_i \sigma_i^2$$

Compte tenu de ce que $E(x'^2_i) = 1$

$$\text{Donc } E(x^4) - 3 [E(x^2)]^2 = 3 \left[\sum_i p_i \sigma_i^4 - \left(\sum_i p_i \sigma_i^2 \right)^2 \right]$$

(1) Par définition (KENDALL, 1959) : $\gamma_2 = \frac{E(x^4)}{[E(x^2)]^2} - 3$.

La quantité entre crochets, de la forme $\sum p_i a_i^2 - (\sum p_i a_i)^2$ est toujours positive.

Le mélange de distributions envisagé donne donc obligatoirement une distribution résultante à coefficient γ_2 positif, donc « trop pointue ».

b) Cas de gènes influant à la fois sur la moyenne et l'écart-type

Soient des populations de moyennes m_1, m_2, \dots et d'écart-types respectifs $\sigma_1, \sigma_2, \dots$, mélangées dans les proportions p_1, p_2, \dots

En prenant pour origine la moyenne de la population résultant du mélange, le coefficient γ_2 de cette population est encore du signe de

$$E(x^4) - 3 [E(x^2)]^2$$

En détaillant d'après les distributions composantes et posant

$$x'_i = \frac{x_i - m_i}{\sigma_i}$$

on obtient, i étant l'indice de l'une quelconque de ces distributions :

$$E(x_i^4) = E(x'_i \sigma_i + m_i)^4$$

ou, en tenant compte de ce que

$$E(x'_i) = E(x'^3_i) = 0, E(x'^2_i) = 1, E(x'^4_i) = 3 :$$

$$E(x_i^4) = 3 \sum_i p_i \sigma_i^4 + 6 \sum_i p_i \sigma_i^2 m_i^2 + \sum_i p_i m_i^4$$

De même,
$$[E(x_i^2)]^2 = \left[\sum_i p_i \sigma_i^2 + \sum_i p_i m_i^2 \right]^2$$

La quantité étudiée peut donc s'écrire :

$$\begin{aligned} & \left[\sum_i p_i m_i^4 - 3 \left(\sum_i p_i m_i^2 \right)^2 \right] + 3 \left[\sum_i p_i \sigma_i^4 - \left(\sum_i p_i \sigma_i^2 \right)^2 \right] \\ & + 6 \left[\sum_i p_i \sigma_i^2 m_i^2 - \left(\sum_i p_i \sigma_i^2 \right) \left(\sum_i p_i m_i^2 \right) \right] \end{aligned}$$

Le premier terme est celui rencontré dans un mélange de distributions de moyennes différentes et de même variance. Il doit en général être négatif (MÉRAT, 1967 b et données non publiées).

Le deuxième terme, comme indiqué ci-dessus, est toujours positif.

Quant au troisième, son signe n'est pas évident *a priori*.

Pour conclure, un mélange de distributions de variance inégale peut produire des distributions à coefficient γ_2 positif, au moins si les différences concernant les moyennes ne sont pas très importantes (1).

L'on ne voit guère, *a priori*, de causes susceptibles d'aboutir à ce résultat, autres que la ségrégation dans certaines familles de gènes à effet appréciable sur la variance, au moins si l'on suppose l'absence de questions d'échelle résultant de limites imposées

(1) Pour fixer les idées, si l'on a un mélange en proportion égale de deux distributions de même moyenne, dont l'une a une variance double de l'autre, γ_2 pour la population résultante est égal à + 0,33.

à la variation, ou de corrélations entre moyenne et variance, hypothèse dont la vraisemblance sera confirmée si les distributions sont sensiblement normales dans certaines familles ou groupes génétiques.

La ségrégation de gènes à effet important sur la valeur moyenne d'un caractère tendra à rendre γ_2 négatif (MÉRAT, 1967b), le plus souvent du moins.

Quant à des gènes dont l'effet propre serait de produire des distributions trop « pointues », ils ne paraissent pas *a priori* devoir être une possibilité très vraisemblable ou du moins très courante.

Le test de l'écart à zéro de certains coefficients γ_2 repose sur l'hypothèse « nulle » de la distribution normale du coefficient observé g_2 dans chaque groupe avec une moyenne nulle et une variance :

$$s^2 g_2 = \frac{24 n (n - 1)^2}{(n - 3) (n - 2) (n + 3) (n + 5)}$$

n étant l'effectif de l'échantillon (SNEDECOR, 1959).

Ceci n'est possible que pour des effectifs relativement importants dans chaque famille, la normalité n'étant obtenue que dans ces conditions ⁽¹⁾. Cependant, le test de GEARY (1935) pourrait, à défaut, être employé (d'après KENDALL, 1959).

Malgré cette limitation pratique, il nous a paru intéressant de suggérer cette possibilité en l'absence d'autres méthodes connues.

Reçu pour publication en octobre 1967.

SUMMARY

POSSIBLE MEANS TOWARDS THE DETECTION OF GENES WITH AN INDIVIDUAL EFFECT ON THE PHENOTYPIC VARIANCE

In a random-mating population, the observation of heterogenous intra-family variances is not a suitable test to detect genes with an individual effect on the phenotypic variance. A method consisting in the detection of positive coefficients of « kurtosis » in certain families is therefore suggested. It requires rather large numbers in each family.

RÉFÉRENCES BIBLIOGRAPHIQUES

- FALCONER D. S., ROBERTSON A., 1956. Selection for environmental variability of body size in mice. *Zeit. Indukt. Abst. Vererb. Lehre*, **87**, 385-391.
- GEARY R. C., 1935. The ratio of the mean deviation to the standard deviation as a test of normality. *Biometric*, **27**, 310.
- KENDALL M. G., 1959. *The advanced theory of statistics*. II. Griffin and Co, London.
- LENER I. M., 1954. *Genetic homeostasis*. Oliver and Boyd, London.
- MÉRAT P., 1966. Contribution à l'étude de la « valeur sélective » associée à quelques gènes chez la poule domestique. I. *Ann. Biol. anim. Bioch. Biophys.*, **7**, 79-104.
- MÉRAT P., 1967 a. Les gènes influant sur la variance d'un caractère quantitatif et leurs répercussions possibles sur la sélection. *Ann. Génét.* (sous presse).
- MÉRAT P., 1967 b. L'étude des distributions de fréquences en vue de déceler la présence de gènes à effet important. *Ann. Biol. anim. Bioch. Biophys.* (sous presse).
- PROUT T., 1963. The effects of a history of stabilizing selection on sensitivity to foreign environments. *C. R. II^e (abstr.) Congrès Génétique*, **1**, 165.
- SNEDECOR G. W., 1959. *Statistical methods*. Iowa State College Press.

⁽¹⁾ On pourrait aussi, sans recours à l'hypothèse de normalité, dénombrer les groupes dont le coefficient g_2 est significativement positif au seuil 5 ou 1 p. cent, grâce aux tables de PEARSON et HARTLEY (1966, *Biometrika Tables for Statisticians*, p. 207-208). Ceci permettrait des effectifs plus restreints.