

NOTE

Méthode de prise en compte de la consanguinité dans un modèle simple de simulation de performances

J.-L. FOULLEY et C. CHEVALET *

*I.N.R.A., Station de Génétique quantitative et appliquée,
Centre de Recherches zootechniques
F 78350 Jouy-en-Josas*

**I.N.R.A., Laboratoire de Génétique cellulaire, Centre de Recherches de Toulouse,
B.P. 12, F 31320 Castanet-Tolosan*

Résumé

Cette note présente une méthode de prise en compte de la consanguinité dans la simulation de valeurs génétiques à partir de valeurs parentales. Par rapport aux formules négligeant la consanguinité adoptées par WILLHAM & THOMSON (1970) dans le cow-game et sous l'hypothèse d'un déterminisme génétique additif, cette méthode revient à changer l'écart-

type résiduel $\sqrt{\frac{1}{2}} \cdot \sigma_G$ en $\sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2}\right)} \cdot \sigma_G$; f_P , f_M étant les coefficients de consanguinité du père et de la mère respectivement et σ_G l'écart-type génétique. La même modification s'applique à la simulation d'un caractère corrélé.

Dans le cow game (WILLHAM & THOMSON, 1970), sont simulées à chaque génération, les valeurs génétiques additives et phénotypiques des individus issus des accouplements planifiés par l'expérimentateur.

Pour un caractère, le modèle de génération de l'effet génétique direct s'écrit :

$$G_I = \frac{1}{2} (G_P + G_M) + \sqrt{\frac{1}{2}} \cdot \alpha \cdot \sigma_G \quad (1)$$

$$P_I = \mu + G_I + E_I \quad (2)$$

où :

- G_P , G_M , G_I sont les valeurs génétiques additives des père, mère et individu respectivement ;
- μ , σ_G la moyenne et l'écart-type génétique du caractère ;
- E_I l'ensemble des effets de milieu affectant l'individu I ;
- α le tirage d'une loi normale réduite $N(0, 1)$.

Ce modèle ne prend pas en compte la réalisation d'accouplements consanguins. Pour y remédier, MALLARD et DOUAIRE ont introduit une dépression de consanguinité sur la valeur phénotypique individuelle P selon la formule :

$$P_{I(F)} = P_{I(O)} - K F_I \quad (3)$$

où F_I est le coefficient de consanguinité de l'individu I ; K la dépression de consanguinité.

Dans ce modèle, aucune modification explicite n'est apportée aux valeurs génétiques telles qu'elles sont définies en (1).

L'objet de cette note est de formuler une proposition de calcul de celles-ci dans le cas simple où les caractères ont un déterminisme génétique strictement additif et où les variables aléatoires introduites présentent une loi conjointe normale.

La dépression de consanguinité étant liée aux effets de dominance et d'épistasie, on est logiquement conduit, dans le cadre du modèle choisi, à ne plus faire intervenir celle-ci, mais à prendre en compte une modification de la variance génétique.

Sous nos hypothèses, et, compte tenu des formules développées notamment par WRIGHT (1922) et MALECOT (1948), la matrice des variances et covariances entre valeurs génétiques additives s'écrit :

	G_I	G_P	G_M
G_I	$(1 + f_I) \sigma_G^2$	$2\varphi_{IP} \sigma_G^2$	$2\varphi_{IM} \sigma_G^2$
G_P		$(1 + f_P) \sigma_G^2$	$2\varphi_{PM} \sigma_G^2$
G_M			$(1 + f_M) \sigma_G^2$

où f et φ sont les coefficients de consanguinité et de parenté de MALECOT.

Il en découle immédiatement l'expression du meilleur prédicteur de G_I à partir des valeurs génétiques parentales :

$$\hat{G}_I = b_P G_P + b_M G_M \quad (4)$$

avec :

$$\begin{pmatrix} b_P \\ b_M \end{pmatrix} = \begin{pmatrix} 1 + f_P & 2\varphi_{PM} \\ 2\varphi_{PM} & 1 + f_M \end{pmatrix}^{-1} \begin{pmatrix} 2\varphi_{IP} \\ 2\varphi_{IM} \end{pmatrix} \quad (5)$$

Sachant, par ailleurs que :

$$\begin{aligned} \varphi_{PM} &= f_I \\ 2\varphi_{IP} &= f_I + \frac{f_P}{2} + \frac{1}{2} \\ 2\varphi_{IM} &= f_I + \frac{f_M}{2} + \frac{1}{2} \end{aligned} \quad (6)$$

Il vient :

$$b_P = b_M = \frac{1}{2}$$

On retrouve donc la relation très générale :

$$\widehat{G}_I = E(G_I/G_P, G_M) = \frac{1}{2} G_P + \frac{1}{2} G_M \quad (7)$$

Pour générer G_I , il suffit alors de calculer \widehat{G}_I selon la formule (7) et de lui ajouter une résiduelle S_I de variance $\text{var } S_I = \text{var } G_I - \text{var } \widehat{G}_I$.

Comme :

$$\begin{aligned} \text{var } G_I &= (1 + f_I) \sigma_G^2 \\ \text{var } \widehat{G}_I &= \frac{1}{2} \left(1 + 2f_I + \frac{f_P + f_M}{2} \right) \sigma_G^2 \end{aligned} \quad (8)$$

on a :

$$\text{var } S_I = \frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right) \sigma_G^2$$

En présence de consanguinité et sous l'hypothèse d'un déterminisme génétique additif, la formule (1) est donc modifiée comme suit :

$$G_I = \frac{1}{2} (G_P + G_M) + \sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right)} \cdot a \cdot \sigma_G \quad (9)$$

Cette formule montre bien que la prédiction de \widehat{G}_I par $G_I = \frac{1}{2} (G_P + G_M)$ est d'autant plus précise que les parents sont plus consanguins ; à la limite si $f_P = f_M = 1$ (pas de variabilité génétique chez chacun des parents, G_I est exactement égal à $\frac{1}{2} (G_P + G_M)$). Par ailleurs, l'expression du carré de la corrélation entre G_I et \widehat{G}_I

$$\text{soit } R_{G_I, \widehat{G}_I}^2 = 1 - \frac{1}{2} \cdot \frac{1 - \frac{f_P + f_M}{2}}{1 + f_I}$$

fait apparaître clairement l'influence simultanée de f_I , f_P et f_M sur la relation existant entre ces 2 variables.

A titre d'illustration sont figurées au tableau 1 les valeurs de F , $\text{var } G$, $\text{var } \widehat{G}$, $\text{var } S$ et R^2 dans le cas d'accouplements répétés entre frère et sœur.

TABLEAU 1

Evolution de la variance génétique (var G) d'un caractère à déterminisme additif, de celles de son prédicteur parental (var \widehat{G}) et de l'erreur (var S) ainsi que de la précision de la prédiction (R^2) dans le cas d'accouplements répétés entre frère et sœur
Evolution of genetic variance (var G), of the variance of its predictors in fonction of parental values (var \widehat{G}) and of the residual variance (var S) of an additive genetic character in case of repeated brother-sister matings

Génération	F (a)	var G	var \widehat{G}	var S	R^2
1	0	1	1/2	1/2	1/2
2	1/4	5/4	3/4	1/2	3/5
3	3/8	11/8	1	3/8	8/11
4	1/2	3/2	19/16	5/16	19/24
etc.					

(a) $f_i = \frac{1}{4} (1 + 2 f_{i-1} + f_{i-2})$ pour $i \geq 2$.

Les germains de la première génération sont issus de parents non apparentés.

Globalement, la covariance génétique additive augmente dans la population avec le degré de consanguinité mais la variance résiduelle (à valeurs génétiques fixées des parents) diminue.

Ce modèle se généralise aisément à la simulation de caractères corrélés en particulier à celle de l'effet génétique maternel. Dans le cow game, un caractère corrélé (noté par un astérisque *) est généré comme suit :

$$G_I^* = \frac{1}{2} (G_P^* + G_M^*) + \sqrt{\frac{1}{2}} (\alpha CGG_1^* + \beta CGG_2^*) \tag{10}$$

où :

- $CGG_1^* = \frac{Cov(G, G^*)}{\sigma_G}$
- $CGG_2^* = \sqrt{Var G^* - CGG_1^{*2}}$

• α est la valeur du tirage $N(0, 1)$ effectué en (1) et β un autre tirage $N(0, 1)$ indépendant.

Si on pose :

$$G_I = \widehat{G}_I + S_I$$

$$G_I^* = \widehat{G}_I^* + S_I^*$$

avec :

$$S_I = \alpha \sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right)} \cdot \sigma_G$$

$$S_I^* = C_1 \alpha + C_2 \beta \tag{12}$$

Le problème revient à déterminer C_1 et C_2 tels que :

$$\left. \begin{aligned} C_1 &= \frac{\text{Cov}(S_I, S_I^*)}{\sigma_{S_I}} \\ C_2 &= \sqrt{\text{Var } S_I^* - C_1^2} \end{aligned} \right| \quad (13)$$

En présence de consanguinité, et, toujours dans l'hypothèse d'un caractère à déterminisme génétique additif, la covariance génétique entre 2 caractères mesurés sur un même individu s'écrit :

$$\text{Cov}(G_I, G_I^*) = (1 + f_I) \text{Cov}_o(G, G^*) \quad (14)$$

où $\text{Cov}_o(G, G^*)$ est la covariance entre caractères dans la population non consanguine.

(14) est un cas particulier de la formule :

$$\text{Cov}(G_I, G_J^*) = 2\varphi_{IJ} \text{Cov}_o(G, G^*)$$

A partir de ces formules et sachant que $\text{Cov}(S_I, S_I^*)$ est égal à :

$$\text{Cov}(G_I, G_I^*) - \text{Cov}(\widehat{G}_I, \widehat{G}_I^*)$$

il vient :

$$\text{Cov}(S_I, S_I^*) = \frac{1}{2} \text{Cov}_o(G, G^*) \left(1 - \frac{f_P + f_M}{2} \right)$$

En utilisant l'expression de la variance résiduelle donnée en (8), on obtient :

$$\begin{aligned} C_1 &= \frac{\text{Cov}_o(G, G^*)}{\sigma_o G} \sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right)} \\ C_2 &= \sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right) \left(\sigma_o^2 G^* - \frac{\text{Cov}_o(G, G^*)^2}{\sigma_o G} \right)} \end{aligned}$$

(l'indice o indiquant la valeur du paramètre dans la population non consanguine).

Soit :

$$G_I^* = \frac{1}{2} (G_P^* + G_M^*) + \sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right)} (\alpha \text{CGG}_{1(o)}^* + \beta \text{CGG}_{2(o)}^*) \quad (15)$$

Sous ces hypothèses, la prise en compte de la consanguinité se traduit simplement par le changement dans les formules (1) et (10) de :

$$\sqrt{\frac{1}{2}} \text{ par } \sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right)}$$

Cette proposition peut apparaître en définitive extrêmement simpliste puisqu'elle ne prend en compte que l'incidence de la consanguinité sur la distribution des valeurs génétiques additives et ne fait pas intervenir d'effet dépressif.

Cela impliquerait de partir d'un modèle plus complexe incluant les situations d'identité des gènes et les autres composantes de la variance génotypique (GILLOIS, 1964). En effet, ces paramètres sont nécessaires pour exprimer les covariances entre apparentés et donc pour calculer le prédicteur de la valeur génotypique d'un descendant. Ce procédé devient très compliqué, notamment d'un point de vue numérique, à cause du calcul des coefficients d'identité. Une autre possibilité consisterait à simuler, non plus les valeurs génétiques, mais des génotypes à partir de tirages mendéliens dans les génotypes des parents. La valeur génotypique serait alors calculée à partir des valeurs fixées au départ des génotypes réalisés aux différents loci.

Par ailleurs, l'application d'une procédure simple combinant une dépression ($-KF_I$) sur la moyenne et la génération de valeurs génétiques additives selon la méthode développée précédemment ne peut être davantage retenue. On sait, en effet, que pour un caractère soumis en partie à des effets de dominance, la variance génétique additive diffère de l'expression adoptée ici (ROBERTSON, 1952 ; FALCONER, 1960 ; CHEVALET & GILLOIS, 1978). Notre démarche a donc le mérite d'être cohérente avec l'hypothèse d'un déterminisme génétique additif sous-jacent. Elle n'est pas non plus totalement irréaliste pour des caractères peu affectés par la consanguinité tels que ceux de valeur bouchère considérés dans le cow game. Par contre, pour la productivité numérique, on pourrait procéder comme MALLARD et DOUAIRE en diminuant la probabilité de survie des individus consanguins.

Reçu pour publication en mai 1981.

Summary

A method for taking into account inbreeding in a simple model of simulated performances

This paper presents a method for taking into account inbreeding when genetic values of individuals are simulated on the basis of the parental values. Compared to formulae used by WILLHAM and THOMSON (1970) in the cow game and, under the assumption of an additive

genetic determinism, this method leads to change the residual standard deviation $\sqrt{\frac{1}{2}} \cdot \sigma_G$

into $\sqrt{\frac{1}{2} \left(1 - \frac{f_P + f_M}{2} \right)} \cdot \sigma_G$; f_P , f_M being the inbreeding coefficients of the sire and

dam respectively and σ_G the genetic standard deviation. The same modification has to be applied when simulating a correlated trait.

Références bibliographiques

- CHEVALET C., GILLOIS M., 1978. Inbreeding depression and heterosis : Expected means and variances among inbred lines and their crosses. *Ann. Génét. Sél. anim.*, **10**, 73-98.
- FALCONER D.S., 1960. *Introduction to quantitative genetics*. Oliver and Boyd, London.
- GILLOIS M., 1964. *La relation d'identité en génétique*. Thèse, Fac. Sci., Paris, 294 p.
- MALÉCOT G., 1948. *Les mathématiques de l'hérédité*. Masson et C^{ie}, Paris, 63 p.
- MALLARD J., DOUAIRE M. (sans date). *Jeu de VO : simulateur de sélection bovine*. E.N.S.A. Rennes (ronéoté).
- ROBERTSON A., 1952. The effect of inbreeding on the variation due to recessive genes. *Genetics*, **37**, 189-207.
- WILLHAM R.L., THOMSON G., 1970. *Instructions for use of beef cattle simulation program*, Iowa State University (Mimeo).
- WRIGHT S., 1922. Coefficients of inbreeding and relationship. *Am. Nat.*, **56**, 330-338.