

F. MINVIELLE

INRA Laboratoire de Génétique Factorielle
78352 Jouy-en-Josas Cedex

Les bases de la génétique quantitative

Les modèles à un et deux locus

Résumé. Intuitivement, il est raisonnable de penser que les "propriétés génétiques" d'une population sont directement fonction des gènes et des génotypes présents dans cette population et de leur transmission d'une génération à l'autre. Mais la fonction qui les lie n'est sans doute pas simple sauf peut-être si l'on a affaire à un caractère dont les variations sont dues à peu de locus (c'est-à-dire à un ou deux).

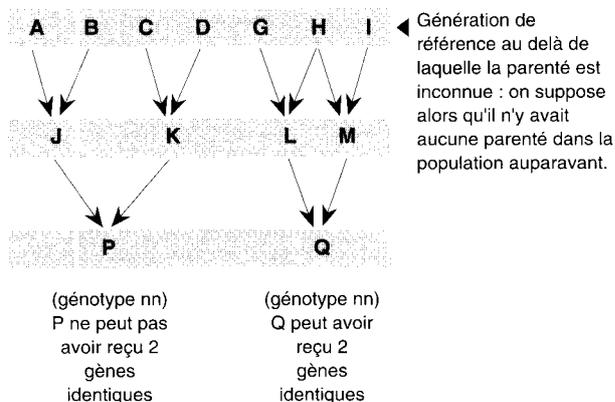
En effet, en se plaçant dans ce cas de figure, on décrit assez facilement la génétique d'un caractère au sein d'une population et surtout on développe les concepts de base qui, étendus aux caractères polygéniques (déterminés par de multiples locus), sous-tendent l'amélioration génétique. Ceci est précisément l'objectif de cet article qui développe successivement la notion de parenté et les mesures qui en découlent, puis l'étude d'un caractère quantitatif déterminé par un seul locus (monogénique) et enfin celle d'un caractère quantitatif digénique (deux locus).

1 / Mesure de la parenté

En amélioration génétique animale, l'information disponible pour évaluer chaque "candidat" à la reproduction provient des individus qui ont un certain lien de parenté avec le candidat, en incluant en premier lieu bien sûr le candidat lui-même, et qui doivent donc aussi avoir des gènes en commun avec lui. Mesurer la parenté permettra donc ultérieurement d'accorder à chaque information l'importance relative qu'elle mérite pour l'évaluation du candidat.

1.1 / La notion d'identité entre gènes

Considérons les généalogies ci-dessous où les animaux sont représentés par des lettres.



Supposons que les deux animaux P et Q à un locus donné soient homozygotes avec les deux mêmes gènes n et n. Les deux gènes ont des origines différentes

chez P, on dira alors que ces gènes sont semblables. Par contre chez Q, ils pourraient être les deux copies d'un même gène ancestral de H, une copie transmise par l'intermédiaire de L et l'autre par celui de M, les deux parents de Q. On dira alors de ces deux gènes qu'ils peuvent être identiques.

Bien entendu comme on ne peut pas être vraiment certain de l'identité de deux gènes, c'est une propriété à laquelle on attache une certaine probabilité.

1.2 / Mesures de l'apparentement

Le fait que J et K ne puissent pas, dans l'exemple, posséder de gène identique découle du fait qu'ils n'ont pas de parent commun. Leur parenté est nulle. Par contre, les animaux L et M ont entre eux une certaine parenté (ce sont un demi-frère et une demi-soeur).

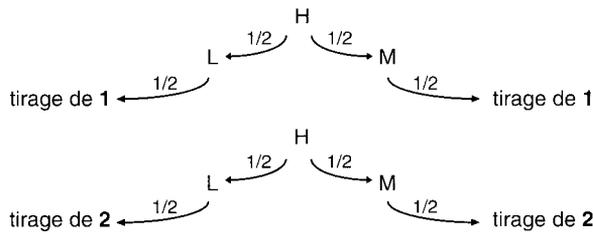
a / Coefficient de parenté f

Le coefficient de parenté f_{JK} entre les deux individus J et K est la probabilité (pour un locus quelconque) qu'un gène tiré au hasard parmi les deux portés par J et qu'un gène tiré de la même façon de K soient identiques.

Bien évidemment, pour J et K, $f_{JK} = 0$. Les animaux J et K ne sont pas apparentés.

Pour L et M par contre c'est différent. Nous avons localisé l'ancêtre commun H. Donnons les numéros 1 et 2 aux deux gènes qu'il porte au locus choisi et symbolisons les différentes étapes de la transmission des gènes jusqu'au tirage d'un gène de L et de M. Plusieurs cas sont possibles, mais deux seulement

aboutiront au possible tirage ultime de deux gènes identiques :



En conséquence, $f_{LM} = 2(1/2) \cdot 1/4 = 1/8$

Les coefficients de parenté les plus courants sont ceux pour :

un individu et lui-même	1/2
un individu et son père	1/4
des frères	1/4
des demi-frères	1/8
un individu et son grand-père	1/8

b / Coefficient de consanguinité F

Considérons P et Q et la probabilité respective que les animaux P et Q soient homozygotes pour deux gènes identiques. Ces probabilités sont appelées des coefficients de consanguinité. De par sa définition même, on voit facilement que le coefficient de consanguinité F d'un individu représente exactement la même chose que le coefficient de parenté f entre ses deux parents : c'est dans les deux cas la probabilité d'un même tirage de deux gènes identiques. Donc dans notre exemple,

$$F_P = f_{JK} = 0 \text{ et } F_Q = f_{LM} = 1/8$$

On retrouve ici la notion intuitive selon laquelle l'utilisation de reproducteurs apparentés augmente la consanguinité dans la descendance. F varie de 0 (par exemple, si il y a panmixie dans une grande population ou suite au croisement de deux races) à 1 (consanguinité complète après une vingtaine de croisements successifs frère x soeur).

c / Relation de dominance u

Que deux individus soient chacun porteur d'un gène identique indique qu'ils sont apparentés. C'est donc encore "plus" le cas si ils ont chacun une paire de gènes identiques deux à deux ! Dans ce cas, leur génotypes sont identiques au locus concerné. La probabilité de réalisation d'un tel événement est appelée la relation de dominance u. Elle est égale à 0 pour la plupart des parentés courantes, sauf pour la relation de dominance entre frères (1/4) ou entre vrais jumeaux (1). On peut montrer qu'en général, entre deux animaux P et Q, de parents respectifs J, K et L, M elle s'écrit en fonction des coefficients de parenté entre les parents des individus concernés :

$$u_{PQ} = (f_{JL} \times f_{KM}) + (f_{JM} \times f_{KL})$$

1.3 / Conclusion

Coefficient de parenté et relation de dominance mesurent les deux types d'apparentement possibles dans une population qui se perpétue de façon panmixique, sans connaître de changement de composition génétique et d'où, en particulier, la consanguinité est absente. Dans l'hypothèse, plus réaliste, de consanguinité de niveau F chez l'ancêtre commun, source de

la parenté, on s'attend à ce que soit augmenté le niveau de parenté. Ainsi entre L et M, le coefficient de parenté devient alors $(1 + F_H) 1/8$ plutôt que $1/8$. Tout à fait logiquement, il serait alors de 1, c'est-à-dire $(1 + 1)1/2$, entre un individu entièrement consanguin et "lui-même".

Enfin, par souci de simplicité, une seule source de parenté (un seul ancêtre commun) a été considérée ici. Bien sûr, en réalité, plus il existe de voies de parenté entre deux individus plus le coefficient de parenté est grand : les contributions de chaque voie sont indépendantes, elles se calculent chacune comme on l'a vu plus haut et les valeurs s'ajoutent pour donner le coefficient f recherché.

2 / Etude d'un caractère mesurable déterminé par 1 locus

2.1 / Mise en situation

Dans une population idéale (celle qui se perpétue sans connaître de changement de composition génétique) un caractère mesurable est déterminé par un locus avec deux allèles. A chacun des trois génotypes présents sont ainsi associées une valeur du caractère mesurable et une certaine fréquence.

Génotype	T_1T_1	T_1T_2	T_2T_2	
Fréquence	p^2	$2pq$	q^2	$p + q = 1$
(exemple)	0,16	0,48	0,36	$0,4 + 0,6 = 1$
Mesure	m_1	m_2	m_3	moyenne
(exemple)	8	6	2	4,88

Pour faciliter l'écriture, on a coutume de changer l'origine des mesures et de la placer à mi-chemin (le point 5 ici) entre celles, m_1 et m_3 , des deux homozygotes. On obtient ainsi les quantités correspondantes à m_1 , m_2 et m_3 , symbolisées respectivement par a, d et -a et qui prennent dans l'exemple les valeurs 3, 1 et -3.

Dans la population, la valeur moyenne correspondante est bien :

$$ap^2 + 2pqd - aq^2 = a(p - q)(p + q) + 2pqd = a(p - q) + 2pqd = -0,12$$

Bien sûr m_1 , m_2 et m_3 tout comme a, d et -a sont des indicateurs de la valeur moyenne de chaque génotype en ce qui concerne le caractère mesuré.

a / Valeur génotypique G

En exprimant a, d et -a par différence à la valeur moyenne pour la population, on définit alors les valeurs génotypiques G. Ainsi, pour T_1T_1 on aura :

$$G_1 = a - a(p - q) - 2pqd = 3 - (-0,12) = 3,12$$

De même pour T_1T_2 , $G_2 = 1 - (-0,12) = 1,12$ et pour T_2T_2 , $G_3 = -3 - (-0,12) = -2,88$.

D'où la description "quantitative" de la population.

	T ₁ T ₁	T ₁ T ₂	T ₂ T ₂	Moyenne \bar{G}
Fréquence	0,16	0,48	0,36	
Valeur génotypique G	3,12	1,12	-2,88	0

b / Variance génotypique V(G)

Dans quelle mesure la valeur génotypique est-elle variable dans la population ? On l'évalue en calculant la variance génotypique V(G), moyenne des carrés des écarts à la moyenne des valeurs génotypiques :

$$V(G) = 0,16 (3,12 - 0)^2 + 0,48 (1,12 - 0)^2 + 0,36 (-2,88 - 0)^2 = 5,15$$

et plus généralement

$$V(G) = p^2G_1^2 + 2pqG_2^2 + q^2G_3^2.$$

2.2 / Transmission du caractère

Sous ce modèle, G est clairement défini pour chaque génotype et V(G) pour la population. Mais un parent ne transmet pas l'intégralité de son génotype, seulement un des deux gènes qu'il possède à ce locus.

a / Effet moyen

Quelle est donc, pour chacun des deux allèles possibles, cette valeur transmissible ? On l'appelle l'effet moyen de l'allèle T₁, ou de l'allèle T₂ selon le cas.

Maintenant si nous imposons pour la reproduction de cette population un reproducteur mâle T₁T₁, qui transmettra forcément T₁ à sa descendance, alors que les femelles proviendront sans restriction de la population tout entière, la descendance aura alors une valeur génotypique moyenne différente de celle de la population de référence (0 par construction) par une quantité qui sera le reflet de la valeur de l'allèle T₁ et qu'on appelle l'effet moyen α_1 (tableau 1) :

Tableau 1. Effet moyen de l'allèle T₁

Gène reçu du mâle	T ₁		
Génotype de la descendance	T ₁ T ₁	T ₁ T ₂	T ₂ T ₂
Valeur génétique	G ₁	G ₂	G ₃
Fréquence	p 0,4	q 0,6	0 0
Valeur génotypique moyenne de la descendance du mâle et des femelles de la population	pG ₁ + qG ₂ (0,4)(3,12) + (0,6)(1,12) = 1,92		

Ainsi $\alpha_1 = pG_1 + qG_2 - 0$
 $= p[a - a(p - q) - 2pqd] + q[d - a(p - q) - 2pqd] - 0$, et, après réarrangements et simplifications
 $= q[a + d(q - p)] = 0,6[3 + 1(0,6 - 0,4)] = 1,92$

De même, pour l'allèle T₂ on aura $\alpha_2 = -p[a + d(q - p)] = -1,28$

Enfin, on remarquera que $\alpha = \alpha_1 - \alpha_2 = a + d(q - p) = 3,2$. Cette quantité α mesure l'effet moyen de la substitution de l'allèle T₁ à l'allèle T₂, dans la population de référence, sur la valeur génotypique. Plus α sera

grand, plus le changement de T₂ à T₁ changera le caractère à la génération suivante.

b / Valeur additive A (ou valeur génétique)

Suite au développement du concept d'effet moyen, il devient naturel d'utiliser α_1 , α_2 et α pour mieux décrire ce que recouvre la notion de valeur génotypique. Chaque individu "possède" les effets moyens des deux gènes de son génotype qui sont, par leur définition même, transmissibles à sa descendance.

La somme de ces effets moyens est la valeur génétique ou valeur additive de l'individu en question.

Le tableau 2 donne les valeurs additives pour a = 3, d = 1 et pour quatre fréquences différentes de l'allèle T₁. Il illustre ainsi des propriétés importantes de la valeur additive :

Tableau 2. Valeur additive de l'individu. Exemple pour 4 fréquences différentes de l'allèle T₁

Génotype	valeur additive	p=0	p=0,1	p=0,4	p=0,7
T ₁ T ₁	A ₁ = 2 α = 2q α	8	6,84	3,84	1,56
T ₁ T ₂	A ₂ = $\alpha_1 + \alpha_2 = (q - p)\alpha$	4	3,04	0,64	-1,04
T ₂ T ₂	A ₃ = 2 α_2 = -2p α	0	-0,76	-2,56	-3,64
	$\alpha = a + d(q - p)$	4	3,80	3,20	2,60

La valeur additive A (tout comme les effets moyens) n'est pas une constante. Elle dépend de la fréquence des allèles dans la population : ainsi le meilleur génotype T₁T₁ voit sa valeur additive diminuer comme T₁ devient plus fréquent dans la population. Ceci est logique; quand la population est presque entièrement de génotype T₁T₁, l'impact de la substitution d'un allèle T₂ par un allèle T₁ supplémentaire est bien moindre que lorsque les allèles T₁ sont rares.

La valeur additive est donc fonction du génotype de l'individu et de la composition génétique de la population à laquelle il appartient.

On montrera que la valeur additive se transmet à moitié d'un individu à sa descendance en prenant comme exemple le cas du génotype T₁T₂ qui produira en proportions égales des gamètes avec l'allèle T₁ et des gamètes avec l'allèle T₂ :

Gamète de T ₁ T ₂	T ₁	T ₂	
Gamète de la population			
T ₁	1/2 p	1/2 p	
T ₂	1/2 q	1/2 q	
Génotype de la descendance	T ₁ T ₁	T ₁ T ₂	T ₂ T ₂
	p/2	p/2 + q/2	q/2
		= 1/2	

La valeur additive moyenne de la descendance d'un individu de génotype T_1T_2 (de valeur additive A_2) sera donc

$$(2q\alpha)p/2 + (\alpha_1 + \alpha_2)/2 + (-2p\alpha)q/2 = (\alpha_1 + \alpha_2)/2 = A_2/2$$

(D'où, par extension, la notion que deux fois la performance moyenne des descendants d'un géniteur évalue la valeur additive de ce géniteur). On vérifierait simplement que, par construction, la valeur additive moyenne \bar{A} de la population est égale à 0.

Par contre, la valeur additive d'un descendant particulier du père T_1T_2 n'est pas forcément $(\alpha_1 + \alpha_2)/2$. En effet, la valeur additive d'une progéniture d'un père T_1T_2 sera fonction du "tirage" (tableau 3).

Tableau 3. Valeur additive d'une progéniture d'un père T_1T_2 .

Génotype	Fréquence	Valeur additive	=	1/2	+	Aléa	
				Valeur du père		de méiose	
T_1T_1	$p/2$	$2\alpha_1$	=	$(\alpha_1 + \alpha_2)/2$	+	$(3\alpha_1 - \alpha_2)/2$	
T_1T_2	$1/2$	$\alpha_1 + \alpha_2$	=	$(\alpha_1 + \alpha_2)/2$	+	$(\alpha_1 + \alpha_2)/2$	
T_2T_2	$q/2$	$2\alpha_2$	=	$(\alpha_1 + \alpha_2)/2$	+	$(-\alpha_1 + 3\alpha_2)/2$	
Et en moyenne :			$(\alpha_1 + \alpha_2)/2$	=	$(\alpha_1 + \alpha_2)/2$	+	0

Il y aura généralement une différence entre la valeur additive d'un individu et celle prédite à partir de celles de ses parents (pris au sens large); on l'appelle l'aléa de méiose. Si en moyenne l'aléa de méiose disparaît, il contribue cependant à la variation des valeurs additives entre individus de même "parent" !

c / Variance additive (ou variance génétique)

A ces valeurs additives différentes d'un génotype à l'autre correspond une variation, la variance génétique ou variance additive $V(A)$, qui est la moyenne des carrés des écarts à la moyenne des valeurs additives :

$$V(A) = 0,16(3,84 - 0)^2 + 0,48(0,64 - 0)^2 + 0,36(-2,56 - 0)^2 = 4,92$$

Plus généralement $V(A) = 2pq\alpha^2$.

Cette variance est extrêmement importante pour l'amélioration génétique puisqu'elle représente la variation attribuable aux différences de nature génétique qui sont transmissibles à la descendance.

2.3 / Décomposition de la valeur génotypique $G = A + D$

La valeur génotypique d'un individu, dans le contexte du modèle à un gène, inclut logiquement sa valeur additive mais doit reconnaître également la contribution possible d'interactions entre les deux gènes présents.

a / Valeur de dominance D

On symbolise les effets de l'interaction entre les deux gènes présents par D et on les évalue par la différence $D = G - A$,

d'où les valeurs respectives pour les trois génotypes :

$$D_1 = G_1 - A_1 = -0,72$$

$$D_2 = G_2 - A_2 = 0,48$$

$$D_3 = G_3 - A_3 = -0,32$$

En général on aurait :

$$D_1 = -2q^2d$$

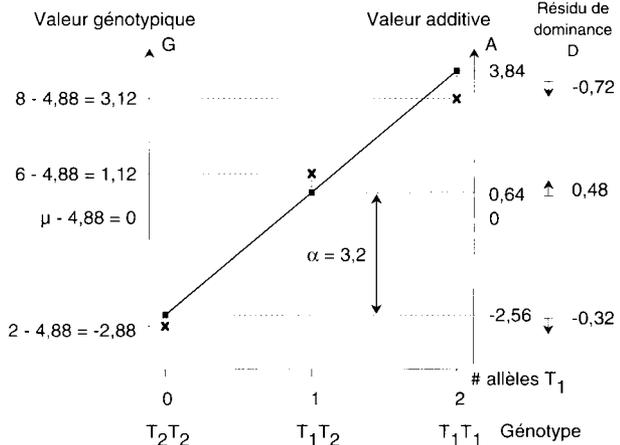
$$D_2 = 2pqd$$

$$D_3 = -2p^2d$$

En moyenne, ici aussi, $\bar{D} = 0$.

On appelle également ces valeurs des résidus de dominance et cela avec raison. En effet, l'idée derrière la décomposition de G en A et D est d'"expliquer" le plus possible les valeurs génotypiques découlant des mesures par les valeurs additives sous-jacentes et d'attribuer le reste aux interactions.

En fin de compte, cette approche est de nature essentiellement statistique, comme le décrit la figure ci-dessous :



La relation linéaire entre la valeur G et le nombre d'allèles favorables T_1 n'est pas complète : on observe une certaine différence, D , entre la valeur génotypique G (x) observée et la valeur A correspondante (•) prédite par cette relation linéaire.

b / Variance de dominance $V(D)$

De la même façon que pour $V(G)$ et $V(A)$, les valeurs de dominance des trois génotypes induisent une variation, la variance de dominance

$$V(D) = 0,16(-0,72 - 0)^2 + 0,48(0,48 - 0)^2 + 0,36(-0,32 - 0)^2 = 0,23$$

Plus généralement $V(D) = (2pqd)^2$

Enfin, et surtout, on remarquerait que $V(D) = V(G) - V(A)$, ou exprimé autrement que $V(G) = V(A) + V(D)$. A la décomposition $G = A + D$ correspond exactement celle de la variance.

Cela revient à dire que la covariance entre les variables A et D est nulle : on a décomposé la valeur génotypique G en deux variables A et D qui sont indépendantes l'une de l'autre.

2.4 / Interprétation génétique de la décomposition

a / Approche classique

Il est tentant de donner à A et à D une signification biologique. En effet, on remarque que si $d = 0$, c'est-à-dire si l'hétérozygote est intermédiaire par rapport aux deux homozygotes, les résidus D sont nuls et on a donc $G = A$ et $V(G) = V(A)$. A cette additivité génétique correspondent donc les valeurs additives A et la variance additive $V(A)$. Au contraire si d est différent de 0, il y a dominance au sens génétique du terme et celle-ci induit alors des valeurs D et une variance $V(D)$ non nulles.

D'où bien évidemment l'origine des appellations A et D, et surtout l'interprétation intuitive inverse :

"l'importance relative de $V(D)$ et de $V(A)$ renseigne sur celle qu'a la dominance par rapport à l'additivité dans le déterminisme génétique du caractère".

b / Validité de l'approche

L'interprétation génétique qui vient d'être proposée doit être adoptée avec prudence. En effet si l'additivité induit $V(A)$ et si la dominance induit bien $V(D)$, l'inverse, c'est-à-dire : "l'existence de $V(A)$ implique l'additivité génétique et celle de $V(D)$ implique la dominance génétique", n'est pas automatique.

On s'en convaincra facilement à partir de cet exemple numérique :

Génotype	T_1T_1	T_1T_2	T_2T_2
Fréquence	0,25	0,50	0,25
Mesure	16	8	4
Mesure centrée	6	-2	-6

On a dans ce cas $V(A) = 18$ et $V(D) = 1$

Pourtant le déterminisme génétique n'est pas ici additif : les deux gènes ont des effets multiplicatifs (de valeur 4 pour T_1 et de valeur 2 pour T_2) ! (On a en effet les mesures $4 \times 4 = 16$, $4 \times 2 = 8$ et $2 \times 2 = 4$ pour les trois génotypes).

2.5 / Covariance entre apparentés

Des animaux apparentés X et Y doivent logiquement avoir une certaine ressemblance entre leurs valeurs génotypiques G_X et G_Y (et donc bien sûr aussi entre leurs productions P_X et P_Y). On peut déjà se douter que cette ressemblance dépendra de leur degré de parenté, c'est-à-dire de leur coefficient de parenté et de leur relation de dominance.

On mesure la ressemblance de leurs valeurs par la covariance entre apparentés.

a / Définition et propriétés de la covariance

La covariance entre deux variables G_X et G_Y , $Cov(G_XG_Y)$, indique dans quelle mesure G_X et G_Y varient d'une manière liée.

- $Cov(G_XG_Y) = 0$ signifie que G_X et G_Y sont indépendantes : connaître la valeur prise par G_X ne renseigne alors pas sur celle prise par G_Y .

- Si G_X et G_Y représentent en vérité la même variable (X et Y sont le même animal), on a alors $Cov(G_XG_Y) = V(G_X) = V(G_Y)$, la variance génétique.

- Enfin, la covariance entre deux sommes est la somme des covariances élémentaires : ainsi $G_X = A_X + D_X$ et $G_Y = A_Y + D_Y$, d'où

$$Cov(G_XG_Y) = Cov(A_X + D_X, A_Y + D_Y) = Cov(A_XA_Y) + Cov(D_XD_Y) + Cov(A_XD_Y) + Cov(D_XA_Y)$$

On a vu précédemment que la décomposition $G_X = A_X + D_X$ avait pour propriété (intéressante !) d'entraîner $Cov(A_XD_X) = 0$. Par conséquent, on aura $Cov(A_XD_Y) = Cov(D_XA_Y) = 0$, d'où la décomposition générale de la covariance entre génotypes individuels :

$$Cov(G_XG_Y) = Cov(A_XA_Y) + Cov(D_XD_Y)$$

b / Ecriture générale de la covariance entre génotypes individuels

Une démonstration stricte est possible à partir de l'écriture de G en fonction des effets moyens et des résidus de dominance ainsi que des mesures de la parenté entre les individus. On aboutit à :

$$Cov(G_XG_Y) = 2f_{XY} V(A) + u_{XY} V(D)$$

Par exemple, avec la généalogie : $X \rightarrow Y$,

on a, on l'a déjà vu, $A_X = A$ et $A_Y = 1/2 A$ (parent et descendant).

$$\text{Donc } Cov(A_XA_Y) = Cov(A, 1/2A) = 1/2 Cov(A, A) = 1/2 V(A) = Cov(G_XG_Y)$$

De la même manière, chez des demi-soeurs X et Y de même père Z, avec $A_Z = A$
 $A_X = 1/2 A$ et $A_Y = 1/2 A$ et
 $Cov(A_XA_Y) = Cov(1/2A, 1/2A) = 1/2 \cdot 1/2 Cov(AA) = 1/4 V(A) = Cov(G_XG_Y)$.

En effet, ici $f_{XY} = 1/8$ et $u_{XY} = 0$.

Enfin, chez des clones ou des jumeaux monozygotes
 $A_X = A_Y = A$ et $D_X = D_Y = D$.

Donc $Cov(G_XG_Y) = V(A) + V(D) = V(G)$, car $f_{XY} = 1/2$ et $u_{XY} = 1$ (toujours en absence de consanguinité).

Ainsi, si l'on calcule $V(A)$ et $V(D)$ pour l'exemple numérique développé au paragraphe 2.2., avec les fréquences respectives $p = 0,1, 0,4$ et $0,7$, on obtient alors les covariances présentées au tableau 4.

Comme les variances, les covariances dépendent des fréquences des allèles dans la population. Elles sont plus grandes quand les fréquences sont intermédiaires.

Tableau 4. Variances additive et de dominance et covariances. Exemple pour 3 fréquences différentes de l'allèle T_1

Fréquence de T_1	$p = 0,1$	$p = 0,4$	$p = 0,7$
Variance additive	2,60	4,92	2,84
Variance de dominance	0,03	0,23	0,18
Covariance (père, fils)	1,30	2,46	1,42
Covariance (sœurs)	1,31	2,52	1,47
Covariance (demi-sœurs)	0,65	1,23	0,71
Covariance (clones)	2,63	5,15	3,02

2.6 / Intérêt et limite du modèle

Ce modèle à un locus avec deux allèles aura permis de montrer que pour tout individu, la valeur génotypique, pouvait être décomposée en une partie "transmissible" à la descendance et en un résidu non transmis. Il motive ainsi l'approche principale de l'amélioration génétique, la sélection intra-population qui s'appuiera sur la fraction transmissible A et sa variation V(A).

Il démontre aussi que A et V(A) sont directement fonction de la composition génétique de la population et du caractère concerné.

Enfin, il illustre à travers la voie de transmission la plus directe, parent-descendant, le fait que l'information sur des apparentés peut servir à évaluer génétiquement un individu particulier.

Néanmoins, il demeure limité puisqu'il ne considère qu'un seul gène, qu'il n'intègre pas les effets du milieu et qu'il néglige par conséquent forcément les effets, sur le caractère quantitatif, des interactions entre les gènes qui constituent le génotype d'une part, entre les gènes et le milieu d'autre part.

3 / Modèle "complet" : deux locus et le milieu E

D'un façon générale, le milieu, E, agit sur l'expression d'un caractère quantitatif. Ses effets combinés à ceux du génotype de l'animal induisent une performance, une production P et l'on suppose alors

$$P = G + E$$

Maintenant, si le caractère est déterminé par deux locus (ou plus) une nouvelle source d'interaction génétique est possible, celle entre gènes différents. Parallèlement, la valeur génotypique G est alors sous l'influence d'un facteur supplémentaire que l'on appelle I, la valeur d'épistasie :

$$G = A + D + I$$

3.1 / Valeur génotypique $G = A + D + I$

Comme pour le modèle à un seul locus, la portion transmissible de la valeur génotypique est représentée par la seule valeur additive A. Mais celle-ci est maintenant, pour chaque animal, la somme des effets moyens des quatre gènes (deux par locus) de son génotype. Ici aussi, la valeur additive présente un intérêt opérationnel évident.

De la même façon, D est la somme des deux valeurs de dominance, une pour chaque locus.

Maintenant, le nouveau "résidu" est la valeur d'épistasie, I, qui s'obtiendrait par différence entre G et A + D.

Que représente I ? C'est la contribution globale de trois types d'interaction, schématisés à la figure 1 avec deux locus T et R portés par des chromosomes différents. Autant dire que l'on décomposera rarement I ailleurs que dans cet exemple.

Comme précédemment, la décomposition juste effectuée entraîne celle de la variance génotypique :

$$V(G) = V(A) + V(D) + V(I) \text{ ou plus complètement} \\ = V(A) + V(D) + V(AA) + V(AD) + V(DD).$$

Ici encore, seule la variance additive V(A) représente la variation due à des différences génétiques entre individus qui sont transmises à leur descendance.

3.2 / Production ou valeur phénotypique $P = A + D + I + E$

Sous le modèle à deux locus, dans la population "idéale", le milieu vient simplement ajouter un effet supplémentaire E dont la valeur est aléatoire, varie d'un individu à l'autre et est sans relation avec le potentiel génétique. Cela revient à dire que l'effet du milieu est parfois favorable, parfois défavorable; en moyenne sa contribution est nulle. Par contre le milieu contribue à la variation des productions mesurées et on a :

$$V(P) = V(A) + V(D) + V(I) + V(E)$$

3.3 / Covariance entre apparentés (pour un même caractère)

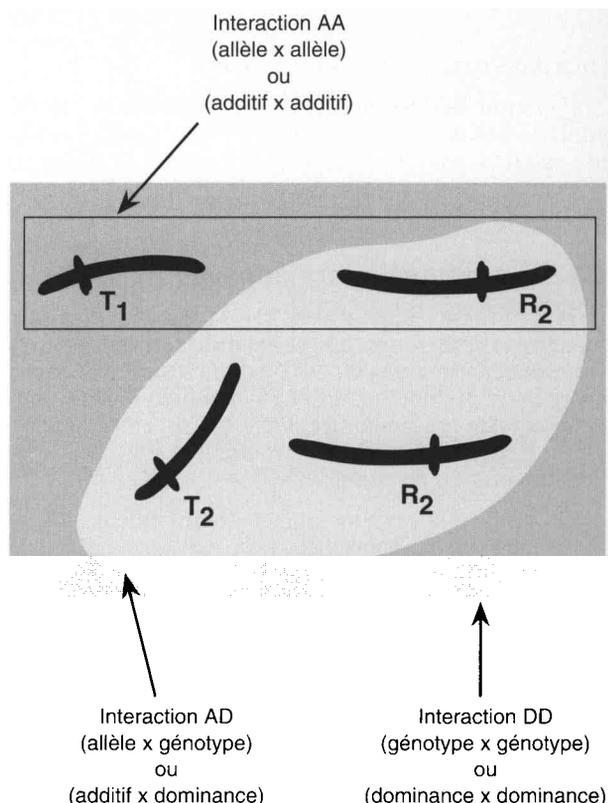
La formule présentée pour un locus se généralise "naturellement" à deux locus et la covariance s'écrit alors

$$\text{Cov}(G_X G_Y) = 2f_{XY}V(A) + u_{XY}V(D) + [2f_{XY}]^2V(AA) + 2f_{XY}u_{XY}V(AD) + u_{XY}^2V(DD),$$

toujours en fonction des seules composantes de la variance génotypique et des mesures de la parenté.

Précisons bien encore que cette covariance mesure la ressemblance qui existe "en moyenne" entre les valeurs génotypiques, pour un même caractère quan-

Figure 1. Les 3 types d'interaction avec 2 locus T et R portés par des chromosomes différents.



titatif, chez des animaux d'apparement connu, dans une population bien définie.

Conclusion

La description d'un caractère quantitatif comme étant le résultat des effets de deux locus et du milieu, aura permis de mettre en évidence l'importance opérationnelle toute particulière de la valeur additive pour l'individu et de la variance additive pour la population. Complétée par l'apport des mesures de la parenté, cette modélisation aura autorisé en plus le développement d'une façon générale d'évaluer la ressemblance entre les potentiels génétiques d'animaux différents par la covariance entre apparentés.

D'ores et déjà, l'extension intuitive de ces notions à des caractères et des populations plus réalistes laisse deviner les fondements de l'amélioration génétique actuelle.

Bien entendu, cette modélisation est, par bien des côtés, très simple. Ainsi nous rappellerons que la population de référence était "idéale". En particulier, les deux locus déterminant le caractère n'étaient pas liés, ce qui représente une certaine simplification de la réalité. Si au contraire, les deux locus sont physiquement liés ou si, ce qui revient au même, il existe dans la population une certaine association entre gènes particuliers à cause des pressions systématiques (comme la sélection) qu'elle subit, le déséquilibre de linkage ainsi induit contribue selon le cas à augmenter ou diminuer la variance additive, d'une façon qui peut ne pas être négligeable. Enfin, le modèle suppose des effets de l'environnement sans interaction avec ceux des gènes. Dans le cas contraire, on aurait alors en quelque sorte une sous-population par milieu.

Il reste à voir comment les concepts décrits ici résistent et s'adaptent à une description plus "réaliste" du caractère quantitatif et de la population.

Références bibliographiques

Falconer D.S., 1989. Introduction to quantitative genetics. Longman Londres, 3ème édition. 438 pp.

Ollivier L., 1981. Eléments de génétique quantitative. Masson Paris. 152 pp.

Minvielle F., 1990. Principes d'amélioration génétique des animaux domestiques. INRA Paris. 211 pp.