

C. CHEVALET et D. BOICHARD\*

INRA Laboratoire de Génétique cellulaire BP 27  
31326 Castanet-Tolosan Cedex

\*INRA Station de Génétique quantitative et  
appliquée 78352 Jouy-en-Josas Cedex

Apports actuels et futurs des  
marqueurs génétiques dans  
l'amélioration des populations  
animales

## Sélection assistée par marqueurs

**Résumé.** L'utilisation de marqueurs génétiques polymorphes en amélioration génétique des animaux est discutée dans deux situations : dans une population issue du croisement entre deux races bien différenciées, ou dans le cadre d'une sélection intra-familiale. En général, il y aurait peu de gain à espérer par rapport à la sélection sur performances pour des caractères d'héritabilité assez élevée et faciles à mesurer. En revanche cette forme de sélection semble indiquée dans les cas suivants : sélection précoce des candidats à la reproduction, optimisation de l'introggression d'un gène marqueur, contrôle de caractères faiblement héritable ou difficiles à mesurer. Dans le cas d'une sélection intra-familiale il semble que la combinaison d'une sélection sur performances avec une sélection sur marqueurs ne soit efficace que si un très petit nombre de marqueurs (un à trois) sont simultanément pris en compte. Dans tous les cas, les effectifs d'animaux à mesurer ou à typer demeurent du même ordre de grandeur que ceux nécessités par les contrôles sur performances habituels.

La recherche de gènes majeurs et leur utilisation dans des programmes d'introggression ou de croisements est une démarche ancienne et toujours d'actualité en génétique appliquée, notamment chez la volaille. Par "sélection assistée par marqueurs" on envisage une approche plus globale, mais aussi indirecte pour exploiter un polymorphisme génétique révélé par plusieurs locus. Approche globale, parce qu'on admet d'emblée que des marqueurs nombreux sont disponibles et qu'on ne se limite pas à une région du génome. Approche indirecte parce que le polymorphisme concerne des marqueurs a priori neutres et non des gènes impliqués dans l'expression d'un caractère.

Pour utiliser comme critères de sélection les allèles présents chez les individus candidats à la sélection, on doit distinguer schématiquement deux situations : ou bien une association a été établie entre certains allèles marqueurs et des allèles (non observables) de gènes quantitatifs, ou bien seule une information concernant la position de gènes quantitatifs est disponible. Le premier cas correspond à l'existence de déséquilibres de liaison obtenus par le croisement entre deux populations fortement différenciées ; le second cas est celui réalisé à la suite d'une détection obtenue par des analyses intra-familiales.

## 1 / Sélection assistée dans une population issue d'un croisement

### 1.1 / Principe

Les premiers travaux supposant l'existence d'un déséquilibre systématique entre QTL et marqueurs ont été présentés par Soller et Beckman (1982, 1983). Récemment Lande et Thompson (1990) ont abordé cette situation plus précisément, en considérant une population obtenue par le croisement entre deux lignées, races ou populations suffisamment différentes. Ils envisagent simultanément la localisation de marqueurs liés à des QTL, et la sélection des individus selon les allèles qu'ils portent aux locus marqueurs dans les régions responsables de la variabilité des caractères quantitatifs. Dans un premier temps une analyse statistique par régression (progressive) cherche à expliquer la valeur des caractères par les allèles marqueurs. Les locus qui contribuent significativement indiquent la position des QTL. Si la carte est assez dense, un tri des locus marqueurs doit être fait pour éviter des redondances entre marqueurs proches. Aux allèles des marqueurs retenus sont alors attachées des valeurs correspondant à leurs

contributions (apparentes) au caractère. Pour chaque individu on peut alors calculer une valeur génétique en sommant les contributions des allèles qu'ils portent : cette valeur constitue l'index selon lequel la sélection est pratiquée.

## 1.2 / Résultats théoriques

L'efficacité de cette méthode est évaluée par le rapport de la réponse attendue d'une sélection selon un index combinant les marqueurs et les données quantitatives, à la réponse attendue selon un index calculé d'après les seules performances. Cette efficacité relative est d'autant plus faible que l'héritabilité du caractère est plus grande. Ainsi chez les Bovins la sélection pour la production de lait selon un index sur descendance ne verrait pas son efficacité améliorée par l'usage de marqueurs (Chevalet *et al* 1984a). A la génération suivante, en revanche, la sélection des jeunes taureaux à mettre en testage, faite actuellement sur ascendance, pourrait être complétée par la prise en compte des allèles marqueurs reçus des parents. L'apport de marqueurs est plus important pour des caractères à faible héritabilité : dans ce cas en effet, dès qu'un allèle est reconnu favorable, son observation chez un individu apporte une information bien plus précise que la performance individuelle. Cette tendance reste vérifiée dans le cas d'une sélection combinée sur valeurs individuelle et familiale (tableau 1). Le gain de précision a cependant un coût d'autant plus élevé que l'héritabilité est plus faible, car un plus grand nombre d'individus doivent être élevés et testés. La partie inférieure du tableau 1 où l'on suppose fixé le nombre d'individus disponibles illustre cette restriction, et montre l'incertitude de ces prédictions qui dépendent du nombre de QTL.

## 1.3 / Discussion

Les prédictions de Lande et Thompson reposent sur l'exploitation d'un déséquilibre initial entre locus marqueurs et QTL, qui s'estompe au fil des générations. Par ailleurs la sélection induit des modifications des effets quantitatifs qu'on attache aux allèles. Un autre risque à terme d'une sélection assistée par marqueurs est de sélectionner, avec une grande efficacité, certains allèles favorables au moment de l'estimation, au détriment d'autres allèles dont l'effet n'était pas significatif parce que trop rares, ou dont l'effet, soumis à des interactions épistatiques, peut devenir favorable quand la constitution génétique a changé en d'autres locus. Pour remédier à ces fluctuations, les auteurs suggèrent de procéder à une ré-estimation systématique des effets des allèles marqueurs, à chaque génération.

## 2 / Sélection assistée intra-population

Si, dans une population, on doit admettre que les locus marqueurs et les locus quantitatifs sont en équilibre de liaison, seule une analyse intra-familiale est possible pour localiser les QTL. On ne peut alors attacher un effet quantitatif à un allèle marqueur, sauf à l'intérieur de chaque famille. Pour tirer parti des marqueurs en sélection, on peut aussi modifier l'écriture des index de sélection en tenant compte de l'information qu'ils apportent sur les structures de parenté.

**Tableau 1.** Efficacité relative de la sélection combinée (marqueurs et performances) comparée à la sélection sur performances. On admet que 50% de la variance génétique est expliquée par les marqueurs (d'après Lande et Thompson, 1990).

$h^2$	0,20			0,10			0,025		
HS	1,35			1,5			1,6		
I-Max	1,7			2,2			4,5		
nQTL	5	11	23	5	11	23	5	11	23
I-1000	1,6	1,5	1,2	2	1,7	na	2,7	na	na
I-10000	1,7	1,7	1,6	2,2	2,2	2,1	4	4	3

$h^2$  : héritabilité du caractère.

HS : cas d'une sélection combinant la valeur individuelle et la moyenne de famille des demi-frères, valeur maximale théorique.

I-Max : sélection sur valeur individuelle, valeur maximale théorique.

I-1000 et I-10000 : sélection sur valeurs individuelles, pour des populations d'effectifs totaux de 1000 et 10000 individus.

nQTL : nombre de QTL supposés.

na : méthode non applicable, effectif trop faible pour détecter une liaison.

## 2.1 / Estimations intra-famille des effets des allèles marqueurs

On se place dans la même structure expérimentale que pour la détection des relations marqueurs-QTL, avec plusieurs familles de demi-frères. Dans la descendance d'un père hétérozygote en un marqueur, on peut alors estimer l'effet correspondant à la substitution d'un allèle par l'autre, en classant les demi-frères selon l'allèle reçu du père. Par rapport à l'expérience de détection, la seule économie concerne le nombre de marqueurs à prendre en compte, puisqu'on peut se limiter à ceux des régions pertinentes. Cette forme de sélection assistée nécessite les mêmes moyens expérimentaux que la sélection sur performances, avec des effectifs à optimiser pour maximiser le gain total attendu de la sélection sur performances et sur marqueurs.

L'introduction de ces modifications a été formalisée par Fernando et Grossman (1989), dans le cadre du BLUP et du Modèle Animal. L'apport quantitatif reste à évaluer, et le gain principal espéré porte sur l'amélioration du choix des animaux de la génération suivante.

## 2.2 / Prise en compte des marqueurs dans l'expression de la ressemblance entre apparentés

On peut imaginer une autre façon d'utiliser l'information apportée par les marqueurs, sans chercher à estimer d'effets associés aux allèles dans chaque famille. A proximité d'un marqueur on connaît plus précisément les probabilités d'identité des gènes que dans une région non marquée : dans une famille de demi-frères issus d'un père hétérozygote [AB] en un locus, les individus qui ont reçu A sont génétiquement plus proches entre eux qu'avec ceux de la même famille qui ont reçu B. Ceux qui ont reçu le même allèle sont apparentés comme des pleins-frères, tandis que ceux qui ont reçu des allèles différents sont non apparentés. Ceci ne vaut que pour la région mar-

quée par ce locus, mais si des QTL y sont situés et contribuent à une proportion  $p$  de la variance génétique, les corrélations génétiques entre deux demi-frères de la famille valent :  $0,25(1-p) + 0,5p$ , ou  $0,25(1-p)$ , selon qu'ils ont ou non reçu le même allèle marqueur. L'estimation de la valeur génétique d'un descendant de la famille aura alors trois ou quatre termes (valeur individuelle, valeur moyenne des individus ayant reçu le même allèle, valeur moyenne des individus ayant reçu un autre allèle, et éventuellement moyenne des individus dont l'allèle parental est indéterminé), au lieu de deux termes (valeur individuelle et moyenne de famille).

Selon cette idée on peut donc estimer plus précisément les valeurs génétiques, en calculant ces corrélations génétiques modifiées selon les marqueurs observables. Dans le cas des familles de demi-frères, la méthode est vraisemblablement identique à celle de Fernando et Grossman, elle est en principe généralisable à tout type de parenté (Chevalet *et al* 1984b).

### 2.3 / Discussion

Un facteur limitant des analyses intra-familiales est le pourcentage de familles informatives où une ségrégation des marqueurs est observable. Ce taux est d'autant plus élevé que les marqueurs disponibles sont plus polymorphes.

Les effectifs nécessaires pour mettre en oeuvre une sélection assistée intra-population impliquent autant de moyens que la sélection quantitative, par rapport à laquelle l'apport des marqueurs semble minime pour des caractères d'héritabilité assez grande. Dans le cadre du testage sur descendance, des gains significatifs peuvent être attendus pour la sélection parmi

**Tableau 2.** Prise en compte dans une famille de marqueurs pour prédire la valeur génétique d'un individu. On suppose que des QTL pour le caractère sont localisés à proximité de deux marqueurs (A,B) et (X,Y). A chaque locus l'effet de la région est apprécié par des moyennes familiales partielles correspondant à des "parentés locales".

Père	A    B		X    Y							
Descendants demi-frères, allèles reçus du père :										
A	A	A	B	B	B	?	?	?	?	
X	X	Y	X	Y	?	X	Y	X	?	
Performances :										
P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>5</sub>	P <sub>6</sub>	P <sub>7</sub>	P <sub>8</sub>	P <sub>9</sub>	P <sub>10</sub>	P <sub>11</sub>
Moyennes utiles pour la prédiction de la valeur génétique de l'individu (1) :										
locus (A,B) :										
"pleins-frères" :										
$(P_1 + P_2 + P_3 + P_4) / 4$										
"non apparentés" :										
$(P_5 + P_6 + P_7) / 3$										
"demi-frères" :										
$(P_8 + P_9 + P_{10} + P_{11}) / 4$										
locus (X,Y) :										
"pleins-frères" :										
$(P_1 + P_2 + P_4 + P_7 + P_9) / 5$										
"non apparentés" :										
$(P_3 + P_5 + P_8) / 3$										
"demi-frères" :										
$(P_6 + P_{10} + P_{11}) / 3$										

les descendants. L'index sur ascendance permet un choix entre les familles, mais non entre les descendants d'une même famille. La mesure d'effets attachés à des marqueurs, au sein de chaque famille permettrait au contraire de mettre en oeuvre une sélection intra-famille (tandis que les marqueurs ne peuvent pas être utilisés pour une sélection entre familles).

Dans les deux approches les modifications suggérées ne semblent intéressantes que pour un ou deux marqueurs seulement. D'une part la complexité numérique des calculs devient prohibitive au delà de deux marqueurs, dans la méthode de Fernando et Grossman. D'autre part, selon le deuxième point de vue, on voit dans l'exemple d'une famille de demi-frères que, si l'on considère plusieurs marqueurs, la combinaison des différentes moyennes introduites selon les allèles transmis aux différents locus, tendra rapidement vers la moyenne générale de la famille. Les modifications suscitées par les marqueurs semblent donc s'estomper quand on augmente le nombre de marqueurs (tableau 2).

Ces quelques indications préliminaires suggèrent que dans une approche intra-population, si l'on ne suppose pas l'existence de déséquilibre de liaison, seulement un ou deux marqueurs puissent être utilement introduits pour améliorer les méthodes de sélection. Ceci peut être intéressant pour des caractères dont le déterminisme génétique est gouverné de façon majoritaire par un ou deux gènes.

### Conclusion

Dans une situation où existent d'importants déséquilibres de liaison, l'usage de marqueurs paraît susceptible d'accroître très significativement la précision de l'estimation des valeurs génétiques, notamment pour les caractères à faible héritabilité. L'efficacité d'une sélection sur marqueurs exige néanmoins des précautions : une sélection trop rapide en faveur de certains allèles peut compromettre la réponse à plus long terme ; en présence d'interactions épistatiques, la valeur génétique associée à un allèle peut changer complètement et même s'inverser après quelques générations ; d'une façon générale les effets associés aux allèles marqueurs doivent être contrôlés et réévalués régulièrement ou pour toute transposition d'une population à une autre.

Sans déséquilibre, il semble que, pour être efficace, l'usage de marqueurs doive se limiter à une seule région chromosomique, à condition que celle-ci contribue de façon importante à la variance génétique. Sur plusieurs générations on pourrait tenir compte successivement de plusieurs marqueurs. Les perspectives concerneraient donc plutôt des caractères pour lesquels un gène "quasiment majeur" est impliqué, plutôt que des caractères gouvernés par de nombreux gènes aux contributions équivalentes.

Les méthodes classiques de la génétique quantitative sont très efficaces pour améliorer un ou deux caractères suffisamment héritables. Mais faute de pouvoir être mesurés économiquement ou techniquement, d'autres caractères peu héritables ou considérés comme secondaires, mais qui interviennent dans l'économie des productions, ne sont que très partiellement pris en compte. Le marquage de gènes intervenant dans leur déterminisme permettrait d'exercer un contrôle certainement plus efficace, notamment pour des caractères sous contrôle oligogénique comme les résistances aux maladies.

## Références bibliographiques

---

- Chevalet C., de Rochambeau H., et Vu Thien Khang J., 1984a. in : *Insémination Artificielle et Amélioration Génétique : Bilan et Perspectives critiques* (J.M. Elsen et J.L. Foulley, ed.), Les Colloques de l'INRA, Vol. 29 : 229-245, INRA Publ., Versailles.
- Chevalet C., Gillois M., and Vu Thien Khang J., 1984b. *Génét. Sél. Evol.*, 16 : 431-444.
- Fernando R.L. and Grossman M., 1989. *Genet. Sel. Evol.*, 21 : 467-477.
- Lande R. and Thompson R., 1990. *Genetics*, 124 : 743-756.
- Soller M. and Beckmann J.S., 1982. *Second World Congress of Genetics Applied to Livestock Production* (Madrid), Garsi ed., Vol. 6 : 396-404.
- Soller M. and Beckmann J.S., 1983. *Theor. Appl. Genet.*, 67 : 25-33.