

# Début de la sélection assistée par marqueurs chez les bovins laitiers : étude par simulation

J.-J. COLLEAU

INRA, Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas Cedex

**RESUME** – On s'est intéressé au démarrage de la sélection assistée par marqueurs dans une unité de sélection laitière et l'on a cherché à chiffrer par simulation le gain génétique obtenu sur la première série de jeunes taureaux de testage qui seraient affectés par cette sélection. On a fait varier le contexte génétique du caractère sélectionné (héritabilité, part de variabilité génétique contrôlée par les gènes détectés, nombre de ces gènes) et le volume de typage réalisé dans la population. Si l'on donne la valeur 20 à l'écart-type génétique du caractère, les valeurs trouvées pour le supplément de progrès génétique sont de l'ordre de 1 à 2, ce qui peut paraître modeste mais est compréhensible en réalité (les raisons en sont données). Il est rappelé que les opérations de sélection et leurs éventuels perfectionnements, surtout chez les bovins laitiers, ne manifestent leurs effets que sur le moyen terme.

## Starting marker-assisted selection in dairy cattle: a simulation study

J.-J. COLLEAU

INRA, Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas Cedex

**SUMMARY** – The starting period of marker-assisted selection was investigated in a dairy cattle breeding organisation. The objective was to try and quantify the extra genetic gain occurring in the first batch of young bulls concerned by this new selection method. The genetic determinism of the trait selected was allowed to vary (heritability, fraction of the genetic variance controlled by the genes detected and number of these genes). The amount of typing within the whole population varied as well. If 20 stands for the value of one genetic standard deviation, then the extra gains found were often in the range 1-2. These modest values are not surprizing at all for technical reasons briefly commented. Finally, one should keep in mind that selecting dairy cattle is a long painstaking task and benefits of technical improvements occur very often in the mid or long term.

## INTRODUCTION

La découverte dans le génôme des animaux domestiques de sites à fort polymorphisme (site des marqueurs moléculaires) et le développement des méthodes de typage correspondantes ouvrent de nouvelles perspectives pour la sélection animale (Boichard et al, 1995 ; Elsen et al, 1995) de même que pour l'amélioration des plantes.

En effet, il devient possible de détecter l'existence de gènes commandant, plus significativement que d'autres, l'expression des caractères d'intérêt économique (appelés « QTL », de l'anglais « Quantitative Trait Locus »). Si l'on dispose de marqueurs en grand nombre, certains seront proches des éventuels QTLs, ce qui pourra être mis en évidence, grâce à des méthodes statistiques appropriées, en analysant les liaisons entre performances et génotypes aux marqueurs. Il est toutefois nécessaire de garder à l'esprit que chez les animaux domestiques, la détection puis l'utilisation des QTLs en sélection sont compliquées par l'histoire de ces populations (phénomène d'« équilibre de liaison » entre marqueurs et QTLs). En effet, des recombinaisons nombreuses ont pu se produire entre QTLs et marqueurs voisins. Un même allèle au marqueur (A par exemple) pourra donc être associé tantôt à un allèle favorable au QTL (Q+) et tantôt à un allèle défavorable (Q-), suivant les individus.

Les sélectionneurs fondent de grands espoirs sur l'efficacité immédiate de ce qu'on appelle habituellement « sélection assistée par marqueurs » (SAM). Les nombreux travaux de simulation théorique, concernant notamment les bovins laitiers, montrent cependant que la supériorité de ce type de schéma ne se manifeste que progressivement, sur le moyen terme (Meuwissen et Goddard, 1996 ; Ruane et Colleau, 1995, 1996 ; Spelman et Van Arendonk, 1997 ; Spelman et Garrick, 1998). Cet exposé a donc pour objet, non pas de redémontrer cette supériorité mais de montrer comment elle va s'exprimer à très court terme. On s'intéressera au niveau génétique attendu de la première série de testage de taureaux laitiers qui pourrait intégrer une sélection assistée par marqueurs, en comparaison à une série homologue qui ne l'intégrerait pas. Les ordres de grandeur et les facteurs de variation de la différence attendue seront évalués par simulation d'un schéma de sélection type.

## 1. SCHEMA DE SELECTION SIMULE

### 1.1. ORGANISATION GÉNÉRALE

On s'intéresse à une série de testage de 100 taureaux et à leurs apparentés sur deux générations (pour diminuer le volume des calculs impliqués dans les simulations). Les pères de cette série de testage sont les 5 meilleurs choisis dans une série antérieure de 100 taureaux, elle-même issue de 5 grands-pères paternels (GPP) et de 50 grand-mères paternelles (GMP). Les mères de la série sont au nombre de 100 et sont issues de 5 grands-pères maternels (GPM) et de 50 grand-mères maternelles (GMM). Les mères sont superovulées deux fois vers 15 mois (1 taureau différent utilisé à chaque fois). Les 100 mâles à mettre au testage sont donc à choisir parmi 300 individus (3\*100).

### 1.2. INFORMATIONS UTILISÉES

#### 1.2.1. Sélection classique

Tous les taureaux sont connus d'après les performances de 90 filles de testage, puis de 1000 filles supplémentaires à la mise en service. Les grand-mères et les mères sont connues sur 2 lactations et une lactation respectivement.

#### 1.2.2. Sélection assistée par marqueurs

On suppose qu'on utilise des marqueurs très proches des QTLs contrôlant le caractère simulé. On suppose aussi qu'on est capable de discriminer dans le génotype marqueur d'un individu typé ce qui provient du père et ce qui provient de la mère. Le typage est « restreint » quand il ne concerne que l'élite de la population (GPP et leurs fils testés, pères, GPM, mères, candidats à la mise au testage). Il peut être étendu à des individus

sans ces qualifications, dans le seul but d'apporter globalement de l'information (c'est à dire améliorer la précision des index de sélection). Le « typage étendu 1 » correspond au typage supplémentaire, *a posteriori* des filles de testage des pères et des GPM, après le choix de ceux-ci. Ce typage permet d'améliorer sans délai la connaissance de la valeur des QTLs dont hérite chaque candidat à la mise en testage et donc d'affiner le choix définitif. Grâce au typage des filles de testage du père, on évalue mieux la valeur des gènes hérités du père. Grâce au typage des filles du GPM, on évalue mieux la valeur des gènes hérités de la mère, quand ceux-ci proviennent en fait du GPM. Le « typage étendu 2 » ajoute encore le typage de 100 filles de chaque GPM à sa mise en service. Du fait que ces filles sont contemporaines des mères des candidats au testage, cette information supplémentaire est immédiatement valorisable pour le choix de ceux-ci.

### 1.3. INDEX DE SÉLECTION

#### 1.3.1. Sélection classique

Elle est basée sur le BLUP actuel qui ignore l'existence d'éventuels QTLs. Avec cette méthodologie, on rappelle que les individus pleins-frères ou sœurs ont strictement le même index et ne peuvent être discriminés.

#### 1.3.2. Sélection assistée par marqueurs

Chez les animaux typés, il est possible de tracer le pédigrée exact pour un marqueur donné. On est par exemple capable de diviser les demi-frères paternels testés en deux groupes suivant l'un ou l'autre des allèles reçus du père. On peut également voir chez les jeunes taureaux mis au testage ce qui a été hérité du GPP ou de la GMP d'une part, du GPM ou de la GMM d'autre part.

A partir des pédigrées aux différents marqueurs, on compare donc les performances des animaux ayant hérité de différents marqueurs, notamment intra-famille. On aboutit à une série d'indexations, effectuées simultanément, de valeurs génétiques élémentaires. En effet, on indexe les individus pour chaque composante du génôme c'est-à-dire pour chaque QTL (en distinguant l'allèle d'origine paternelle et celui d'origine maternelle) et la globalité des autres gènes (valeur génétique polygénique). La méthode formelle est décrite par Fernando et Grossman (1989). Pour fixer les idées, s'il y a 3 QTLs, il faut alors calculer  $(2*3) + 1 = 7$  valeurs génétiques élémentaires pour un animal donné. L'index de sélection est alors la somme de toutes ces valeurs élémentaires.

### 1.4. COMPARAISON DES DEUX TYPES DE SÉLECTION

Cette comparaison est faite en simulant des individus avec leurs valeurs génétiques et leurs performances. En partant des mêmes GPP et des mêmes fils testés de GPP, des mêmes GMP, des mêmes GPM, GMM et mères, on a alors deux schémas possibles : le classique ou celui assisté par marqueurs. Il y a donc beaucoup d'animaux communs entre les deux schémas : c'est bien ainsi que les choses se passeront en pratique. On calcule ensuite la différence moyenne de niveau génétique vrai des deux séries « jumelles » de testage, en procédant à 100 tirages indépendants des ancêtres communs. *Le résultat est ensuite exprimé dans une échelle de mesure où un écart-type génétique vaut 20.* C'est en effet la valeur de l'écart-type génétique de l'index synthétique laitier (INEL), dont la gamme de variation est très familière aux professionnels de la sélection des bovins laitiers.

### 1.5. FACTEURS DE VARIATION DE LA COMPARAISON

Il y a d'abord (cf. plus haut) le volume de typage effectué dans la population (restreint, étendus 1 et 2) et ensuite le déterminisme génétique du caractère considéré. En effet, l'héritabilité globale peut varier (on prendra des valeurs de 10 % ou 30 %). La part de variabilité contrôlée par l'ensemble des QTLs peut varier aussi (on prendra 25 %, 50, ou 75 %). Et finalement, le nombre de QTLs impliqués dans l'expression du caractère pourra également varier (on prendra 1, 2 ou 3). Si le second paramètre vaut par exemple 50 % et le troisième 2, cela veut

dire qu'il existe deux QTLs expliquant chacun 25 % de la variabilité génétique.

## 2. RESULTATS ET DISCUSSION

Au tableau 1, les résultats obtenus dans le schéma restreint montrent, comme on pouvait s'y attendre, que pour un nombre donné de QTLs, la réponse augmente quand leurs effets sont plus importants. Quand l'ensemble des QTLs contrôlent une fraction donnée de la variabilité totale, la réponse diminue d'autant plus que ceux-ci sont plus nombreux : la raison est que l'indexation devient de moins en moins précise sur chacun d'eux. La réponse absolue de la sélection assistée par marqueurs est plus importante en cas d'héritabilité élevée mais ce n'est plus le cas quand on l'exprime en valeur relative, en proportion de la réponse obtenue avec le schéma classique.

**Tableau 1**  
Différence moyenne entre sélection assistée et classique  
(un écart-type génétique = 20)

Héritabilité (%)	% QTLs	1 QTL	2 QTLs	3 QTLs
10%	25%	0.48	0.19	0.18
	50%	1.74	1.26	0.62
	75%	3.50	2.30	1.49
30%	25%	0.80	0.46	0.30
	50%	1.94	1.49	1.01
	75%	3.60	2.28	1.80

Le tableau 2 montre l'effet de l'extension du typage dans la situation où il y a 2 QTLs. Son intérêt relatif est d'autant plus important que les conditions d'extériorisation des QTLs sont difficiles, c'est à dire quand l'héritabilité et (ou) la part de variabilité génétique expliquée par les QTLs sont plus faibles.

**Tableau 2**  
Différence moyenne entre sélection assistée et classique  
selon le volume de typage

Héritabilité (%)	% QTLs	Typage restreint	Typage étendu 1	Typage étendu 2
10%	25%	0.19	0.50	0.68
	50%	1.26	1.38	1.74
30%	25%	0.46	0.50	0.65
	50%	1.49	1.46	1.70

Les paramètres génétiques réels (rappelons qu'ici nous les avons faits varier a priori) sont malheureusement encore assez

mal connus dans la pratique car difficiles à quantifier avec précision. Les études actuelles de détection n'ont pas pour l'instant été effectuées que sur un nombre limité de familles paternelles de taureaux testés (Boichard et al, 1999). Elles semblent indiquer, en faisant la moyenne pour plusieurs caractères différents, que la part de variance génétique contrôlée par l'ensemble des QTLs détectés serait de l'ordre de 40 % et que ceux-ci seraient au nombre d'environ 4.

Les ordres de grandeur des supériorités génétiques liées à la SAM et montrées au tableau 2, ne paraîtront certainement pas spectaculaires (en gros, 1 à 2 points). Il convient d'abord de les situer en perspective par rapport au progrès génétique enregistré chaque année dans un schéma classique (4-5 points) ou par rapport au bénéfice que pourrait apporter l'utilisation d'une technique quadruplant la prolificité des mères de la série de testage (2 points). Il convient ensuite d'être conscient d'une série de phénomènes génétiques. La sélection classique permet elle aussi de sélectionner les QTLs favorables et elle est même supérieure à la SAM pour la sélection des autres gènes (nous avons retrouvé ce phénomène comme les autres auteurs). Par ailleurs, les différentes sources de gènes présents dans le groupe des candidats au testage sont indexées avec une précision très intégrale. Par exemple, la valeur des QTLs hérités est connue avec une bonne précision s'ils proviennent du GPP et une précision pratiquement analogue à celle de la sélection classique s'ils proviennent de la gmm. La meilleure efficacité de la SAM en situation de typage « étendu » confirme par ailleurs que la précision moyenne de l'indexation des QTLs dépend bien des effectifs typés.

## CONCLUSION

Cette étude de simulation montre que le tout début de la sélection assistée par marqueurs (SAM) chez les bovins laitiers sera accompagné d'un progrès modeste. Les progrès relatifs enregistrés à moyen terme devraient être plus substantiels, particulièrement pour les caractères à faible héritabilité (reproduction, résistance aux maladies) qui prennent de l'importance vu le contexte économique (Colleau et Phocas, 1994). En effet, l'accumulation des typages à toutes les générations permettra d'évaluer toutes les sources de gènes avec une précision croissante. C'est pourquoi cette étude de simulation est en cours d'extension à une période où la SAM serait instaurée depuis plusieurs générations. Les bénéfices attendus à terme ne pourront être recueillis que si les schémas de sélection respectent une organisation rigoureuse et utilisent l'ensemble des moyens aptes à promouvoir les progrès génétiques, c'est à dire la connaissance du génôme et les techniques de reproduction intensive (Colleau, 1998).

**Boichard D., Elsen JM, Le Roy P., Chevalet C., 1995.** Renc. Rech. Ruminants, 2, 139-144.

**Boichard D., 1999.** Ren. Rech. Ruminants, 6.

**Colleau J.J., 1998.** Proc. 6<sup>th</sup> WCGALP, Armidale, Australia, vol. 25, 419-422.

**Elsen J.M., Colleau J.J., Chevalet C., Moazami Goudarzi K., Boshier M.Y., 1995.** Renc. Rech. Ruminants, 2, 145-153.

**Fernando R.I., Grossman M., 1989.** Genet. Sel. Evol., 21, 467-477.

**Meuwissen T.H.E., Goddard M.E., 1996.** Genet. Sel. Evol., 28, 161-176.

**Ruane J., Colleau J.J., 1995.** Genet. Res., 66, 71.

**Ruane J., Colleau J.J., 1996.** J. Dairy Sci, 79, 1666-1678.

**Spelman R.J., Van Arendonk J.A.M., 1997.** J. Dairy Sci, 80, 3399-3410.

**Spelman R.J., Garrick D.J., 1998.** J. Dairy Sci, 81, 2942-2950.