

E. MANFREDI

INRA, Station d'Amélioration Génétique des Animaux, BP 27,
31326 Castanet-Tolosan cedex

e-mail : Eduardo.Manfredi@toulouse.inra.fr

7 - Utilisation des marqueurs génétiques

Intérêt et limites de la sélection intra-race assistée par marqueurs

Résumé. La prédiction de la réponse à la sélection intra-race assistée par marqueurs (SAM) varie en fonction des contextes étudiés et des approches de simulation utilisées. Les bénéfices théoriques de la SAM annoncés dans le passé ont été revus à la baisse, mais l'utilisation des typages offre des nouvelles possibilités de sélection qui doivent être raisonnées. Quand on dispose de marqueurs associés étroitement à des QTL importants, la modalité d'utilisation des typages dépend de l'objectif génétique recherché. Quand le déséquilibre de liaison entre les marqueurs et les QTL n'est pas garanti, la SAM peut devenir coûteuse et son application nécessite une évaluation économique préalable.

Plusieurs types de marqueurs ont apporté des contributions importantes à l'amélioration génétique des animaux. L'évolution dans les techniques de marquage du gène Hal et son utilisation en génétique porcine en est un bon exemple (Sellier 1998). Les typages moléculaires constituent une source d'information qui peut être utilisée pour l'amélioration génétique des animaux dans des programmes de sélection, croisement, introgression et création des races et des lignées (Elsen 2000, cet ouvrage). Ce texte présente l'utilisation de la SAM dans des programmes de sélection intra-race pour des caractères quantitatifs. Pour raisonner les modalités d'utilisation des typages génétiques dans ce contexte, on peut réaliser des expériences de sélection ou faire appel à des simulations de la réponse à la sélection. Les résultats publiés des simulations peuvent être déroutants, favorables ou non à l'utilisation des marqueurs, d'où l'intérêt de présenter, d'abord, les contextes et les approches utilisées pour prédire la réponse à la SAM. Des résultats des simulations concernant un marqueur associé à un QTL ayant un effet confirmé sur un caractère sont ensuite présentés et des extensions à des cas plus complexes sont discutées.

1 / Les contextes et les approches des simulations

Le contexte est défini par la population (taille, démographie, mode de reproduction), la modalité d'expression phénotypique des caractères (âge, sexe), l'objectif de la sélection (unicaractère ou multicaractère) et son horizon, l'état initial des polymorphismes (nombre de QTL, connus ou marqués, taux de recombinaison, hétérozygotie des

marqueurs et des QTL, importance des effets, fréquences alléliques, existence de marqueurs flanquants, etc) et, éventuellement, l'état initial d'autres effets génétiques non marqués ou polygéniques (écart type polygénique, héritabilité et corrélations génétiques). Les contributions de la SAM au progrès génétique seront plus favorables dans un contexte où la sélection classique est peu efficace et l'information moléculaire est précise (caractères à faible héritabilité et/ou exprimés dans un seul sexe, récessivité ou rareté des allèles favorables, cartes génétiques denses, etc ; voir Elsen 2000, cet ouvrage).

En général, chaque auteur propose une approche originale pour raisonner l'utilisation de la SAM en sélection, mais il est possible de schématiser les choix des modélisateurs à cinq niveaux :

(1) Le modèle génétique peut être plus ou moins détaillé, allant du modèle d'hérédité mixte simple, un QTL et des polygènes, à des modèles oligogéniques où l'on dispose des marqueurs d'un nombre fini de QTL.

(2) L'information moléculaire peut être conçue pour remplacer la sélection classique basée sur les performances (sélection génotypique) ou peut être combinée avec les performances pour une utilisation simultanée (sélection combinée). On peut concevoir également l'utilisation alternée des sélections génotypique et combinée.

(3) La transmission de gènes entre parents et descendants peut être modélisée pour des groupes d'animaux (modèle déterministe) ou pour chaque individu, en faisant intervenir le hasard dans la transmission des gènes de chaque parent vers son descendant (modèle stochastique).

(4) Si tous les auteurs comparent la réponse de la SAM à celle de la sélection classique basée sur les performances, cette comparaison peut être faite

pour une génération ou pour plusieurs générations de sélection. Dans ce dernier cas, l'horizon défini et le critère de comparaison retenu (gain génétique à un temps donné, gain cumulé ou gain cumulé et actualisé) influencent les conclusions des études.

(5) Les règles proposées pour l'utilisation de l'information moléculaire peuvent être fixes (sélection statique) ou variables (sélection dynamique) au cours des générations de sélection.

Les choix relatifs aux niveaux (1) et (2) sont très liés au contexte et une approche réaliste et exhaustive doit être préférée pour tirer des conclusions pratiques. Par exemple, pour (1), l'hypothèse d'un nombre élevé d'allèles au QTL expliquant une proportion importante de la variance génétique donnera des résultats plus favorables à la SAM que l'hypothèse d'un QTL bi-allélique avec un effet faible sur le caractère. Pour (2), limiter la SAM à la seule utilisation des index combinant performances et typages peut sous-estimer la contribution des informations moléculaires quand d'autres possibilités d'utilisation existent.

Les choix (3), (4) et (5) interviennent quel que soit le contexte. Les modèles déterministes et stochastiques sont complémentaires et donnent des résultats cohérents. Le modèle déterministe sert généralement à étudier des grandes tendances et il se prête bien à l'optimisation. Le modèle stochastique permet d'affiner ces premiers résultats pour tirer des conclusions pratiques. Pour évaluer la contribution de la SAM, on doit préférer l'approche multi-génération et dynamique car le bénéfice tiré de l'utilisation des typages évolue en même temps que les fréquences alléliques aux QTL et, en conséquence, les règles d'utilisation des polymorphismes doivent changer au cours du temps.

2 / La SAM pour un cas simple

La situation discutée correspond à la sélection d'un caractère sous l'influence de polygènes et d'un QTL associé étroitement à un marqueur (recombinaison négligeable). Dans ce cas favorable à la SAM, des simulations uni-génération de Smith (1967) donnaient des bénéfices de 60 % en sélection massale pour un caractère d'héritabilité moyenne influencé par un QTL expliquant la moitié de la variabilité génétique totale. Gibson (1994) a montré, par des simulations multi-génération, que ces premiers résultats encourageants correspondent à des bénéfices dans le court terme, mais que, dans le long terme, la SAM donne des gains génétiques inférieurs et des taux de consanguinité supérieurs à ceux produits par la sélection sur performances. Ces résultats conduisent à des conclusions trop pessimistes. D'une part, ces simulations limitent l'utilisation des polymorphismes à la combinaison des typages et des performances dans un index combiné, alors que l'on pourrait concevoir d'autres modalités d'utilisation des typages. D'autre part, la définition de l'index combiné est figée dans le temps. Dekkers et Van Arendonk (1998) et Manfredi *et al* (1998) ont montré que l'utilisation des typages, dans une SAM dynamique où les poids relatifs donnés au QTL et aux polygènes dans l'index combiné évoluent dans le temps, donne des résultats supérieurs à la sélection classique en terme de vitesse de fixation des allèles favorables au QTL, d'évolution des

moyennes polygéniques et des gains génétiques totaux pour tout horizon donné. En conclusion, l'information nouvelle fournie par les typages est en théorie bénéfique pour la sélection si elle est utilisée de façon exhaustive et raisonnée.

En pratique, le calcul des index combinés nécessite des considérations statistiques et génétiques. En statistique, l'indexeur doit pouvoir estimer sans biais et avec une bonne précision les effets du QTL et les distinguer des effets polygéniques et du milieu. Souvent, seulement une minorité des animaux en contrôle de performances est typée, d'où la nécessité d'attribuer aux animaux sans typages des probabilités d'appartenance aux génotypes. Les méthodes proposées vont d'un simple aménagement du BLUP (Israel et Weller 1998) - on utilise l'information des parents typés pour attribuer des génotypes aux descendants sans typage - aux approches itératives faisant intervenir les performances et des généalogies complexes dans la détermination des génotypes manquants (Meuwissen et Goddard 1997). Du côté génétique, il faut s'assurer que les poids relatifs de ces effets dans l'index combiné produisent les gains souhaités pour l'horizon de sélection défini, sachant que, quand un QTL est en ségrégation, la sélection sur performances n'assure pas nécessairement le meilleur compromis entre gains à court et à long terme.

L'utilisation des typages en dehors des index combinés offre des possibilités additionnelles pour la sélection. Alors que des phases de sélection génotypique peuvent s'insérer dans un schéma de sélection massale, le cas souvent étudié correspond à des schémas plus complexes, par exemple insérer une sélection génotypique de jeunes mâles destinés à l'insémination artificielle des ruminants laitiers. Dans ce contexte, on peut concevoir trois modalités d'utilisation des typages : un index combiné pour la sélection des adultes, des accouplements entre parents en fonction des génotypes au QTL et une présélection génotypique de jeunes mâles avant leur mise en testage sur descendance. Cette présélection a un effet bénéfique direct sur la vitesse de fixation des allèles favorables et un effet bénéfique indirect sur l'évolution des moyennes polygéniques car l'on peut alors attribuer un moindre poids au QTL dans l'index combiné pour la sélection des parents (Manfredi *et al* 1998). Dans le cas discuté ici d'un marqueur associé avec un QTL, la présélection peut se faire intra-famille ou dans l'ensemble des mâles. L'efficacité de la sélection intra-famille dépendra de la prolificité de l'espèce et des techniques de reproduction utilisées. La sélection dans l'ensemble des mâles nés peut conduire à une fixation plus rapide des allèles favorables au QTL, mais il faut considérer le risque d'une augmentation de la consanguinité et d'une éventuelle diminution de l'intensité de sélection chez les mères.

3 / La SAM quand le déséquilibre de liaison n'est pas garanti

Dans une population animale ouverte, le déséquilibre de liaison détecté ne peut être garanti qu'au sein des familles et il se dégrade au cours du temps. Toutefois, les simulations donnent les mêmes tendances que dans le cas simple d'association entre QTL et marqueurs, les résultats étant fortement

influencés par les contextes et les approches des simulations. Les marqueurs sont moins utiles quand la sélection classique est très efficace et leur utilisation doit être raisonnée si l'on veut trouver un équilibre entre gains à court et à long terme. Avec une approche déterministe, Ollivier (1998) a montré que l'utilisation des marqueurs dans des index combinés n'améliore que légèrement la précision de la sélection. Avec une approche stochastique, Ruane et Colleau (1995) ont trouvé que, pour des caractères d'héritabilité moyenne, les bénéfices de la SAM étaient modestes et l'évolution des moyennes polygéniques était défavorable par rapport à l'indexation BLUP basée sur performances. Quand ils ont ajouté une sélection génotypique de jeunes mâles intra-famille dans un schéma MOET (Multiple Ovulation and Embryo Transfer) pour l'amélioration de caractères exprimés dans un sexe, la SAM s'est révélée plus efficace que la sélection classique en termes de gains polygéniques et d'augmentation de la fréquence des allèles favorables au QTL (Ruane et Colleau 1996).

Depuis la proposition de Fernando et Grossman (1989) pour construire des index estimant les effets des allèles au QTL et les effets polygéniques, des recherches se poursuivent pour aboutir à des méthodes d'indexation combinée en routine. L'incertitude sur la co-transmission des marqueurs et des QTL, du fait d'une possible recombinaison et de la non informativité des familles, est aggravée par les typages manquants dans une indexation en routine des grandes populations en contrôle de performances (Bink *et al* 1998). De plus, les méthodes d'indexation estiment les effets génétiques en supposant connues les variances des effets polygéniques et des effets alléliques et, en pratique, on doit disposer des estimées fiables de ces paramètres. Ces problèmes limitent à présent la mise en place des applications en routine.

Des phases de sélection génotypique précoce intra-famille ont été proposées pour augmenter l'intensité de sélection au QTL et, éventuellement, diminuer les intervalles entre générations. La sélection génotypique intra-famille s'insère avant l'acqui-

sition des performances des descendants et permet de choisir ceux ayant le plus de chance d'avoir reçu des allèles favorables au QTL. L'efficacité de cette sélection dépend de la taille des familles et des modalités d'utilisation des marqueurs pour prédire les génotypes au QTL chez les descendants. Chez les bovins laitiers, les bénéfices d'une telle sélection ont été estimés à 30 % (Kashi *et al* 1990) et 10 % (Mackinnon et Georges 1998), mais sans tenir compte des pertes dues à des réductions d'intensité de sélection chez les mères. Une évaluation économique est nécessaire avant toute application de la SAM car les opérations pratiques permettant d'associer, pour chaque reproducteur, un allèle marqueur à un allèle du QTL peuvent devenir coûteuses dans des programmes de sélection en routine (Soller et Medjugorac 1999).

Conclusions

Comme prévu par Chevalet et Boichard (1992), l'utilisation des typages en sélection intra-race est loin de l'approche indirecte et exhaustive où des milliers des marqueurs remplaceraient les performances pour la sélection. Pour la sélection combinée, la recherche se concentre sur des méthodes statistiques expérimentales permettant d'appréhender la complexité des données disponibles en routine avant de passer à des applications. L'utilisation des marqueurs dans des phases indépendantes de sélection est déjà appliquée. Actuellement, la multiplicité et l'évolution rapide des objectifs des filières animales posent des problèmes à la sélection classique car la prolifération de caractères à sélectionner entraîne des augmentations des coûts du contrôle de performances et des réductions potentielles de gains génétiques pour chaque caractère. Dans ce contexte, la contribution des typages génétiques est prometteuse, mais elle doit être raisonnée en termes de gains génétiques et économiques pour atteindre l'objectif fixé.

Références

- Bink M., Van Arendonk J., Quaas R., 1998. Breeding value estimation with incomplete marker data. *Genetics Selection Evolution*, 60, 45-58.
- Chevalet C., Boichard D., 1992. Sélection assistée par marqueurs. INRA Productions Animales, n° hors série « Éléments de génétique quantitative et application aux populations animales », 291-294.
- Dekkers J.C., Van Arendonk J.A., 1998. Optimizing selection for quantitative traits with information on an identified locus in outbred populations. *Genetical Research*, 71, 257-275.
- Elsen J.M., 2000. Sélection et introgression assistées par marqueurs. INRA Productions Animales, numéro hors série « Génétique moléculaire : principes et application aux populations animales », 233-237.
- Fernando R., Grossman M., 1989. Marker assisted selection using best linear unbiased prediction. *Genetics Selection Evolution*, 21, 467-477.
- Gibson J.P., 1994. Short term gain at the expense of long term response with selection of identified loci. In : *Proceedings of the 5th World Congress of Genetics Applied to Livestock Production*, 21, 201-204.
- Israel C., Weller J.I., 1998. Estimation of candidate gene effects in dairy cattle populations. *Journal of Dairy Science*, 81, 1653-1662.
- Kashi Y., Hallerman E., Soller M., 1990. Marker-assisted selection of candidate bulls for progeny testing programmes. *Animal Production*, 51, 63-74.
- Mackinnon M.J., Georges M., 1998. A bottom up approach to marker assisted selection. *Livestock Production Science*, 54, 229-250.
- Manfredi E., Barbieri M., Fournet F., Elsen J.M., 1998. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genetics Selection Evolution*, 30, 127-148.
- Meuwissen T.H., Goddard M.E., 1997. Estimation of effects of

Quantitative Trait Loci in large complex pedigrees. *Genetics*, 146, 409-416.

Ollivier L., 1998. The accuracy of Marker-Assisted Selection for Quantitative Traits Within Populations in Linkage Equilibrium. *Genetics*, 148, 1367-1372.

Ruane J., Colleau J.J., 1995. Marker assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genetical Research*, 66, 78-83.

Ruane J., Colleau J.J., 1996. Marker assisted selection for a sex-limited character in a nucleus breeding population. *Journal of Dairy Science*, 79, 1666-1678.

Sellier P., 1998. Genetics of meat and carcass traits. In : M. F. Rothschild, A. Ruvinsky (eds), *Genetics of the Pig*, Chapter 14. CAB International.

Smith C., 1967. Improvement of metric traits through specific genetic loci. *Animal Production*, 9, 349-358.

Soller M., Medjugorac I., 1999. A successful marriage : making the transition from Quantitative Trait Locus Mapping to Marker-Assisted Selection. In : J. C. M. Dekkers, S. J. Lamont, M. F. Rothschild (eds), *From Jay L. Lush to genomics: visions for animal breeding and genetics*, 85-96. Iowa University Press, Ames, Iowa, USA.