

## Utilisation des marqueurs pour l'amélioration génétique des animaux

J.M. ELSEN (1), J.J. COLLEAU (2), C. CHEVALET (3), K. MOAZAMI GOUDARZI (4), M.Y. BOSHER (5)

(1) INRA, Station d'Amélioration Génétique des Animaux, BP27, 31326 Auzeville

(2) INRA, Station de Génétique Quantitative et Appliquée, 78352, Jouy en Josas, Cedex

(3) INRA, Laboratoire de Génétique Cellulaire, BP27, 31326 Auzeville

(4) INRA, Laboratoire de Génétique Factorielle, 78352, Jouy en Josas, Cedex

(5) LABOGENA, 78352, Jouy en Josas, Cedex

**RÉSUMÉ** – Cette synthèse fait le point sur l'utilisation, pour la gestion des populations, de l'information qu'apporteront les marqueurs génétiques, et notamment les marqueurs de l'ADN. Dans un premier temps, les propriétés de ces marqueurs, mises à profit pour l'amélioration génétique, sont décrites : polymorphisme, liaison génétique et déséquilibre de liaison. L'information résultante peut être valorisée pour un contrôle de l'ensemble du génome ou pour le suivi de zones particulières à l'intérieur desquelles seraient localisés des gènes affectant les caractères d'intérêt économiques. Les applications correspondantes sont successivement détaillées et le gain potentiel apporté par le marquage génétique par rapport aux méthodes actuelles y est évalué. Dans le premier groupe figurent l'identification des génotypes, dont le contrôle de filiation, la gestion de la diversité génétique intra et entre populations, les croisements d'absorption. Dans le second groupe sont décrits l'introgression de gènes majeurs et la gestion d'un ensemble de QTLs marqués et non marqués. Il ressort de cette revue que l'utilisation des marqueurs doit être conçue de façon complémentaire et non concurrente aux méthodes classiques, et que son efficacité sera maximum dans les situations où les seules informations de performances renseignent peu ou à fort coût sur la valeur génétique des animaux.

## Using genetic markers in animal breeding

J.M. ELSEN (1), J.J. COLLEAU, C. CHEVALET, K. MOAZAMI GOUDARZI, M.Y. BOSHER

(1) INRA, Station d'Amélioration Génétique des Animaux, BP27, 31326 Auzeville

**SUMMARY** – This review describes the use, for population management, of information given by genetic markers, in particular by DNA markers. In a first part, markers properties useful for animal breeding, are defined : polymorphism, genetic linkage, linkage disequilibrium. The resulting information may allow a better control of the whole genome or a better management of specific genome areas where quantitative trait loci were detected. Corresponding applications are successively detailed and the potential gain given by molecular markers, as compared to classical methods, is evaluated. In the first group are listed genotype identification, including paternity checking, genetic resource management within and between populations and absorption crosses. In the second group are described introgression processes and the multiQTL selection. The main conclusions are that markers will be useful as a complementary tool for animal breeding, and not in substitution of traditional techniques, and that their efficiency will be maximum in these circumstances where the only phenotypic information are poor or too costly for a precise knowledge of genetic values.

## INTRODUCTION

La gestion génétique des populations comporte la gestion des rameaux (races, lignées) et celle des individus dans les rameaux. Le premier volet inclut la conservation de la diversité génétique entre populations et les programmes de croisement parmi lesquels la production de produits terminaux croisés, la création de lignées synthétiques et l'introgession de gènes à effet majeur. Le deuxième volet comporte l'identification des individus dans le cadre des contrôles de filiation et l'ensemble des programmes de sélection, cette dernière partie étant probablement la plus développée sur le plan de la réflexion technique et scientifique et sur le plan des réalisations pratiques.

Jusqu'à une date récente, l'essentiel de cette gestion génétique était basée sur des outils statistiques, pour lesquels la transmission des caractères était décrite par des modèles mathématiques très simples considérant que les caractères sont gouvernés par un nombre infini de gènes indépendants à petits effets. Quelques domaines échappaient cependant à cette règle et utilisaient des observations beaucoup plus proches des gènes : le contrôle des paternités, la caractérisation des distances génétiques entre races, la sélection sur un génotype majeur à l'aide de gènes marqueurs liés.

Avec la découverte de polymorphismes au niveau de l'ADN, et notamment de l'existence de nombreux sites microsatellites, nous disposons désormais d'un grand nombre de gènes marqueurs qui sont localisés sur des cartes génétiques propres aux différentes espèces d'animaux domestiques (Rohrer et al., 1994; Barendse et al, 1994; Crawford et al, 1994; Bishop et al, 1994) et qui pourront être utilisés pour les différents volets de la gestion génétique. Après avoir rappelé la définition et les caractéristiques fonctionnelles d'un marqueur génétique, nous montrons ici comment et dans quelles limites ils peuvent être utilisés pour la gestion génétique des populations.

### 1. PROPRIÉTÉS DES MARQUEURS GÉNÉTIQUES

Rappelons que l'information génétique est stockée dans chacune des cellules sous la forme de longues molécules

d'ADN, le génome, constituant majeur des chromosomes. L'ADN a pour éléments de bases quatre nucléotides. Les gènes sont formés par un enchaînement précis de ces nucléotides et contiennent la partie codante (traduite sous forme de protéines) du génome. Chez les espèces diploïdes (les animaux domestiques dans leur majorité sont diploïdes) l'information génétique est en double (2n chromosomes), une moitié (n chromosomes) provenant du père, l'autre de la mère. Les gamètes d'un reproducteur ne contiennent qu'un seul jeu de chromosomes (n), chaque chromosome étant pris aléatoirement au moment de la méiose parmi les deux chromosomes homologues hérités des parents.

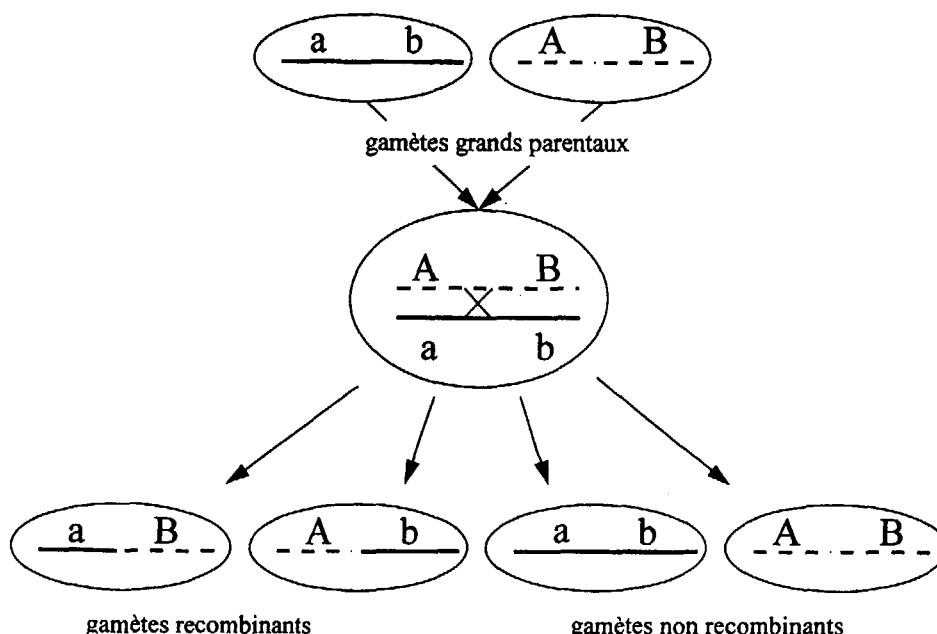
Diverses techniques de laboratoires (Levéziel et Vaiman, 1995) permettent de visualiser de petites zones du génome (désignées ci-après segments génomiques). Deux phénomènes fondamentaux permettent d'utiliser cette visualisation :

- le **polymorphisme** de ces segments, c'est à dire le fait que, pour un même emplacement sur le génome (le même *locus*), il est possible d'observer diverses formes du segment génomique correspondant (on parle alors des formes alléliques ou simplement des *allèles* de ce locus ou de ce gène), entre individus mais aussi intra individu entre chromosomes homologues. On dit dans ce cas que l'individu est *hétérozygote* (dans le cas contraire il est dit *homozygote*).

- la **liaison génétique**, c'est à dire le fait que deux locus proches l'un de l'autre sur l'ADN se transmettent ensemble lors de la méiose. Cette liaison génétique se relâche quand les locus s'éloignent, et, pendant la méiose, des phénomènes de *recombinaisons* apparaissent dans lesquels une coupure de la molécule d'ADN a lieu entre les deux locus avec une reconstitution de la molécule entière à partir de morceaux de chromosomes issus des deux chromosomes homologues (figure 1). A l'extrême, deux locus situés sur des chromosomes différents se transmettent indépendamment.

Ces deux phénomènes peuvent être mis à profit pour la gestion génétique des populations. Le **polymorphisme** est utile pour établir l'origine génétique du locus concerné : au

Figure 1 : Appariations de recombinaisons entre chromosomes homologues au cours de la méiose.



niveau individuel, puisque pour un individu hétérozygote en un locus il est possible (dans certains cas, voir plus loin) de connaître l'origine grand parentale du segment génomique qu'il transmet à ses descendants ; au niveau des populations, si l'hypothèse est admise que les formes possibles pour un même segment génomique diffèrent d'autant plus que les populations sont phylogénétiquement plus éloignées. La **liaison génétique** est utile pour généraliser les conclusions des observations sur un locus particulier à l'ensemble de l'ADN proche de ce locus : comme il y a co-ségrégation lors de la méiose du segment génomique visualisé et des zones avoisinantes, ce segment génomique devient un **marqueur** génétique de ces zones et notamment des gènes qu'elles contiennent. Dans certaines applications, les marqueurs sont utilisés pour suivre l'ensemble du génome, dans d'autres applications, ils nous renseignent sur la transmission de gènes qui leur sont liés. Nous allons décrire ces deux groupes d'applications.

## 2. MARQUAGE DE L'ENSEMBLE DU GÉNOME

L'idée de base est que l'ensemble du génome peut être approché par l'information apportée par un ensemble de marqueurs bien répartis. Ceci peut être mis à profit pour l'identification d'individus ou de groupes, la gestion de la diversité génétique entre ou intra populations ou l'organisation d'un plan d'introgression.

### 2.1. IDENTIFICATION DE GÉNOTYPES

Le **contrôle de filiation** est une des applications les plus anciennes du marquage génétique. Son principe est la recherche d'incohérences entre les génotypes en différents locus marqueurs d'un individu et de ses parents. Par exemple, un descendant de génotype M1M1 au locus *M* et de mère certaine M1M1 ne saurait avoir un individu de génotype M2M2 pour père. Le contrôle de filiation procède donc par exclusion, celle-ci étant d'autant plus probable que le nombre de locus analysés est important. Selon Hanset (1975), 15 marqueurs bialléliques (fréquence 0,5) codominants sont nécessaires pour obtenir une probabilité *a priori* d'exclusion justifiée de paternité supérieure à 0,95. En France, le contrôle de filiation est une obligation légale (taureaux d'insémination, embryons) à la charge d'une structure unique (LABOGENA) qui, en 1994, a vérifié la filiation de 17 000 bovins, 2 000 ovins, 500 caprins et 25 000 équins. Les outils utilisés sont pour l'essentiel des antigènes des globules rouges. Parallèlement à ces techniques immunologiques, des techniques biochimiques permettent de détecter le polymorphisme de protéines sanguines par électrophorèse. Ainsi, pour les bovins, 11 marqueurs sont étudiés qui contiennent de 2 jusqu'à 800 allèles chacun conférant une efficacité théorique (ou probabilité *a priori* d'exclusion) variant selon les races, de l'ordre de 0,99 pour la race Charolaise à 0,96 pour la race Holstein (Boscher et Guérin-Fauble, 1994). Des efforts portent aujourd'hui sur l'utilisation de marqueurs de l'ADN (microsatellites) pour réaliser ces contrôles. Ces marqueurs ont l'avantage de leur efficacité due à leur fort polymorphisme, de leur abondance et de leur simplicité d'emploi, puisqu'il n'est plus nécessaire de produire de réactifs immunologiques.

Pour des situations extrêmes dans lesquelles les accouplements ne peuvent pas être contrôlés, c'est l'identification

plutôt que le contrôle des paternités qui est recherchée, cette identification se faisant *a posteriori*, à l'aide de marqueurs génétiques. Ce cas est fréquent chez les arbres, les poissons ou les animaux sauvages pour lesquels l'utilisation des marqueurs très polymorphes apparaît comme une solution efficace. Son extension aux mammifères d'élevage pourrait parfois être envisagée.

L'**identification des lignées** en vue d'une protection des droits de leur créateur est une préoccupation importante des sélectionneurs de plantes. Ici encore, les marqueurs de l'ADN sont des candidats de choix pour établir un profil génétique spécifique de chaque lignée (Soller et Beckmann, 1983). Si l'on se réfère aux réflexions relatives à la brevetabilité du vivant, ce type d'application pourrait se développer dans le futur pour les animaux domestiques.

### 2.2. GESTION DE LA DIVERSITÉ GÉNÉTIQUE ENTRE POPULATIONS

Les généticiens essaient depuis longtemps de situer génétiquement les différentes sous populations d'une même espèce (races, lignées, rameaux) les unes par rapport aux autres. Au delà de son intérêt historique, la connaissance des filiations entre ces sous populations est nécessaire à l'organisation des programmes de **conservation de la diversité génétique**. Ces programmes ont pour ambition d'éviter la perte d'allèles influençant des caractères d'intérêt zootechnique et qui pourraient se révéler utiles dans l'avenir si les conditions économiques changent ou après un épuisement de la variabilité génétique immédiatement disponible. Dans le choix des sous populations conservées, les redondances doivent être évitées afin d'optimiser l'utilisation de ressources financières limitées.

La distance génétique entre deux sous populations peut être mesurée de diverses façons, et notamment être basée sur les informations apportées par les fréquences alléliques en différents locus. Sans entrer dans le détail algébrique des mesures de distances (Nei, 1971; Rogers, 1972; Wahlund, 1928) qui ont été proposées, indiquons que deux populations seront classées comme d'autant plus proches que les fréquences d'allèles en différents locus le seront, traduisant ainsi l'hypothèse que, par dérive, sélection ou mutation, des populations éloignées possèdent des distributions d'allèles très différentes. L'hypothèse est aussi faite que les marqueurs observés donnent une bonne image de l'ensemble du génome, ce qui suppose qu'ils soient en nombre suffisant et bien répartis. Ces principes ont été appliqués par Grosclaude et al (1990) aux races bovines françaises.

L'utilisation de telles distances posent toutefois quelques problèmes théoriques encore en discussion. La notion de distance ne prend réellement son sens que dans le cas d'une structure arborescente dans laquelle existent de réelles filiations entre les races. Or, pour les animaux d'élevage, il est vraisemblable que les croisements opérés depuis la domestication se traduisent par une structure des relations entre races en réseau plus qu'en arbre. D'un autre point de vue, dans la perspective de recherche de redondances pour les aptitudes d'élevage telle qu'évoquée plus haut, il faut s'interroger sur l'intérêt de l'utilisation de marqueurs *neutres* (génétiquement indépendants de gènes contrôlant les caractères sélectionnés) tels que les microsatellites, malgré l'intérêt de leur grand nombre, de leur fort polymorphisme et de leur codominance.

Par ailleurs, à l'image des essais réalisés dans le domaine végétal, les distances génétiques pourraient être un élément de décision dans le choix de lignées devant être croisées pour exploiter les effets d'hétérosis (supériorité de l'individu croisé par rapport à la moyenne de ses parents). En effet, l'observation de l'hétérosis est souvent expliquée par l'hypothèse d'une *dominance orientée* dans laquelle une majorité de gènes contrôlant un même caractère sont soumis à un effet de *dominance* (pour un gène G, on parle de dominance si les individus G1G2 ont une valeur supérieure à la moyenne des homozygotes G1G1 et G2G2). L'hétérosis sera d'autant plus fort que les lignées soumises au croisement seront génétiquement différentes (et homozygotes), donc que les produits croisés seront hétérozygotes en un plus grand nombre de gènes. Goddard et Ahmed (1982), dans une analyse rétrospective d'un ensemble de croisements *Bos taurus* x *Bos indicus*, ont ainsi trouvé une corrélation de 0,30 entre le poids à un an de génisses croisées et la distance génétique entre leur races d'origines.

### 2.3. GESTION DE LA DIVERSITÉ GÉNÉTIQUE INTRA POPULATION

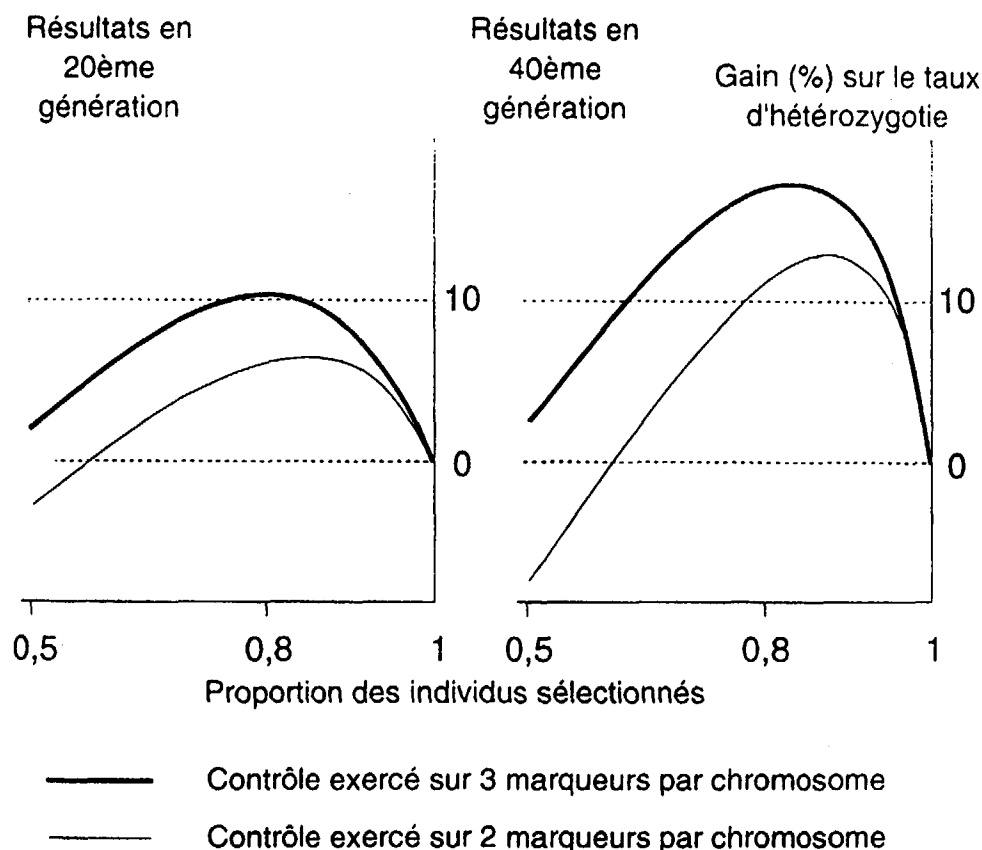
L'efficacité théorique est prouvée de systèmes de reproduction par rotations entre familles pour le maintien d'une variabilité génétique suffisante dans les petites populations en conservation. De tels systèmes ont été mis en place pour plusieurs races de faible effectif mais restent très contraignants. Chevalet (1992) a montré comment les marqueurs génétiques pouvaient aussi aider à maintenir cette variabilité. Le principe est de sélectionner les reproducteurs sur leur taux d'hétérozygotie estimé par la proportion de locus hété-

rozygotes parmi un nombre donné de locus marqueurs observés. Une petite étude théorique visant à préciser le taux de sélection (qui doit résulter d'un optimum entre un choix plus intense sur le critère d'hétérozygotie et une perte minimale de variabilité pour cause de sélection), ainsi que le nombre de marqueurs nécessaires, a montré (figure 2) que ce taux de sélection doit être faible (80 à 90 pour-cent) et qu'il est inutile de typer plus de 3 marqueurs par chromosome. L'efficacité de cette technique de contrôle de la variabilité, qui est d'autant meilleure que la population est de faible effectif, est cependant limitée, avec un gain maximum de 15 pour-cent par rapport à renouvellement au hasard.

### 2.4. AIDE À L'INTROGRESSION

Parmi les gènes contrôlant un caractère possédant un intérêt zootechnique, certains ont un effet majeur sur le caractère et peuvent être démontrés par la seule observation des phénotypes. C'est le cas du gène culard en bovins (Charlier et Menissier, 1995), des gènes Callypige (Berghmans et al, 1995) et Booroola en ovins, du gène du débit de traite chez la chèvre (Le Roy et al, 1995). Dans la plupart des cas, l'allèle favorable de ces gènes ne sont présents que dans certaines races (que nous qualifierons de donneuses, D), alors que d'autres races (receveuses, R), par ailleurs très améliorées ou adaptées à leur milieu, en sont dépourvues. Pour bénéficier de l'effet favorable du gène majeur sans perdre les propriétés générales de la race receveuse, il faut donc introduire (introgresser) l'allèle favorable par un croisement simple  $F1 = R \times D$ , un premier croisement en retour ( $BC_1 = F1 \times R$ ) vers la race receveuse suivi d'une succession

Figure 2 : Gains relatifs sur le taux d'hétérozygotie en 110 locus, dus à la sélection d'une population de 50 individus sur un échantillon de 20 ou 30 marqueurs (d'après Chevalet, 1992)



**Tableau 1**  
Proportion du génome receveur R récupéré pendant une introgression (d'après Hospital et al., 1992).

Génération	Chromosome porteur du gène introgressé		Chromosome non porteur du gène	
	avec sélection	sans sélection	avec sélection	sans sélection
1	0,636	0,607	0,817	0,750
2	0,747	0,684	0,943	0,875
3	0,855	0,741	0,983	0,938
4	0,891	0,784	0,993	0,968
5	0,903	0,817	0,996	0,984
6	0,913	0,842	0,998	0,992

de croisements en retour ( $BC_i = BC_{i-1} \times R$ ), étant entendu qu'à chaque génération  $BC_i$ , seul les reproducteurs porteurs de l'allèle favorable sont retenus. A l'issue de cette série de croisements, des animaux R (ou tout au moins proche de R), porteurs de l'allèle favorable, sont donc disponibles. A titre d'exemple, le gène Booroola a été introgressé en Mérinos d'Arles et Romanov (Boomarov, 1991).

Hillel et al (1990), puis Hospital et al (1992), ont montré l'intérêt, pour accélérer le processus d'introgression, de choisir les reproducteurs  $BC_i$  des différentes générations de croisement en retour sur leurs génotypes en des locus marqueurs, en retenant les porteurs d'un nombre maximum d'allèles marqueurs spécifiques de la race R. Comme pour la gestion de la variabilité intra population, les locus marqueurs, s'ils sont nombreux et bien répartis, donnent une image de l'ensemble du génome. L'hypothèse de travail est que le polymorphisme est suffisant en ces locus pour être en mesure de repérer leur origine raciale, et que la sélection sur ces marqueurs, du fait des liaisons génétiques, entraîne une sélection générale du génome receveur. Dans le cas le plus favorable, les allèles aux différents locus marqueurs seraient systématiquement différents entre les races D et R, et il suffirait, après chaque croisement en retour, de trier les reproducteurs sur la proportion d'allèles issus de la race receveuse. Des simulations d'introgression assistée par marqueurs réalisées par Hospital et al (1992) ont montré l'efficacité de cette technique qui fait gagner 2 générations de croisements en retour par rapport à la méthode traditionnelle pour obtenir une même proportion de génome R (tableau 1). Ces auteurs ont par ailleurs montré que 2 ou 3 marqueurs par chromosome, bien répartis, sont suffisants pour obtenir ce résultat. Leur étude détaillait l'évolution de la proportion de génome R sur le chromosome porteur du gène à introgresser et sur les autres

chromosomes et proposait des pistes pour une organisation dynamique de cette introgression assistée par marqueurs.

### 3. MARQUAGE DE ZONES SPÉCIFIQUES DU GÉNOME

Les gènes ne forment qu'une partie du génome, et seule une partie de ces gènes présente un polymorphisme qui explique la variabilité génétique des caractères d'intérêt zootechnique. Par de simples observations statistiques (voir section 2.4) ou en s'aidant de marqueurs génétiques (Boichard et al, 1995), il est possible de mettre en évidence la ségrégation de certains de ces gènes à effet quantitatif. Les marqueurs, qui sont nécessaires pour repérer les QTL (Quantitative Trait Loci), ou zones du génome porteuses de ces gènes, peuvent aussi aider à fixer les allèles favorables dans les populations. Dans la première partie, nous avons indiqué que l'utilisation des marqueurs pour la gestion génétique est basée 1) sur leur polymorphisme et 2) sur leur liaison génétique avec les QTL. Un troisième phénomène, le **déséquilibre de liaison**, est également nécessaire pour s'aider des marqueurs génétiques dans la gestion des QTL. En se limitant pour simplifier au cas bi-allélique, on dit que deux segments génomiques (les locus M et Q) sont en équilibre de liaison dans un groupe d'animaux si la proportion de gamètes porteurs de l'allèle M1 est la même quel que soit l'allèle au locus Q. Dans le cas d'un déséquilibre de liaison entre M et Q, l'allèle M1 sera plus souvent associé à l'allèle Q1 (par exemple) et l'allèle M2 à Q2. Dès lors, une sélection de reproducteurs porteurs par exemple du génotype M1M1 augmentera de façon indirecte la fréquence de l'allèle Q1 dans la génération de leurs descendants. Si M est un locus marqueur et Q un QTL, on dispose ainsi d'une méthode indirecte de sélection des gènes favorables en ce dernier locus.

**Tableau 2**  
Evolution de la fréquence de l'allèle introgressé au cours des générations de croisements en retour (d'après Soller et Plotkin-Hazan, 1977).

Génération	sans sélection des porteurs	Avec sélection sur le génotype en 1 marqueur à 10cM (1)	Avec sélection sur le génotype en 2 marqueurs flanquants à 10cM
1	0,25	0,81	0,98
2	0,12	0,73	0,97
3	0,06	0,66	0,96
4	0,03	0,59	0,95
5	0,01	0,53	0,96

(1) cM : centiMorgan, mesure de la distance entre le marqueur et le gène majeur.

Le déséquilibre de liaison global au niveau d'une population ne s'observe que dans un certain nombre de circonstances, mais il est toujours présent dans la descendance directe de reproducteurs doubles hétérozygotes (en  $M$  et  $Q$ ). Il peut donc être utilisé pour la gestion d'un gène majeur ou plus généralement pour la sélection d'un ensemble de QTL affectant un ou des caractères zootechniques.

### 3.1. GESTION D'UN GÈNE À EFFET MAJEUR

La situation diffère selon que l'allèle favorable est présent ou non dans la population. Dans le premier cas, la fréquence de l'allèle favorable augmentera de façon automatique si 1) le caractère contrôlé par le gène majeur est un des objectifs de sélection de cette population 2) l'allèle favorable n'a pas d'effet adverse sur d'autres caractères 3) il n'y a pas de surdominance (c'est à dire, dans le cas biallélique, si les individus  $Q_1Q_2$  ne sont pas meilleurs que les  $Q_1Q_1$  et les  $Q_2Q_2$ ). Cette évolution automatique favorable des fréquences alléliques sera toutefois d'autant plus lente que l'allèle favorable sera initialement rare et son effet faible. Dans un certain nombre de cas, les marqueurs pourraient donc contribuer à accélérer la fixation de l'allèle favorable.

Quand l'allèle favorable est absent de la population, il faut, comme évoqué plus haut, l'introduire depuis une race exotique porteuse. A chaque génération, en l'absence de sélection, la proportion de reproducteurs porteurs de l'allèle favorable diminue de moitié (tableau 2). Il faut donc sélectionner, parmi l'ensemble des individus nés de chaque génération, les porteurs de cet allèle. Si le génotype au locus majeur peut être déterminé dans le jeune âge, dans les deux sexes et sans erreur par la simple observation du phénotype, les marqueurs n'ont pas d'utilité. Dans les autres cas, ils peuvent aider à l'introgession. Ainsi, le gène Booroola, qui a un effet considérable sur la taille de portée de brebis, n'est visible que chez les femelles, et après leur puberté. Au cours de l'introgession de ce gène dans les races Mérinos d'Arles en France et Romney en Nouvelle Zélande, il a donc été nécessaire de tester sur descendance les jeunes béliers (de génotype inconnu ++ ou F+) de chaque génération de croisement en retour en les accouplant à des brebis des races d'accueil (par hypothèse toutes ++):

les génotypes des filles, toutes ++ (père ++) ou à 50 pcent F+, 50 pcent ++ (père F+), étant déterminés par observation des niveaux d'ovulation. Avec la découverte d'un marqueur de ce gène (Montgomery et al, 1993 ; Lanneluc et al, 1994) on peut maintenant envisager d'éviter ce testage sur descendance long et coûteux.

L'utilisation d'un déséquilibre de liaison au niveau de la population ne sera possible que dans un nombre limité de cas. L'exemple du gène de sensibilité à l'halothane est de ce point de vue caractéristique, avec un très fort déséquilibre entre ce gène et son marqueur Phi en race Landrace, chez laquelle l'allèle défavorable a été éliminé rapidement par sélection sur ce marqueur, alors que dans d'autres lignées cette sélection n'a pas été possible du fait du relatif équilibre de liaison entre ces locus. Un déséquilibre est d'autant plus vraisemblable que le gène majeur et son marqueur sont proches, et que la population est génétiquement plus ouverte aux apports de sang extérieur. De ce point de vue, l'introgession d'un gène d'une race à l'autre est une situation favorable : on peut s'attendre à ce que les allèles marqueurs des fondateurs de l'introgession soient rares dans la population receveuse. Supposons, pour simplifier, que ces fondateurs sont tous de génotype marqueur  $M_D M_D$ , et les individus de la race receveuse tous  $M_R M_R$  ( $M_R$  et  $M_D$  regroupant l'ensemble des allèles rencontrés dans ces deux groupes). Les accouplements entre reproducteurs  $BC_i$  et reproducteurs de la race receveuse R, seront du type  $M_R M_D \times M_R M_R$  au locus marqueur, ce qui permettra de ne retenir dans leur descendance que les individus  $M_R M_D$ , a priori hétérozygotes  $Q_R Q_D$  au locus majeur. L'efficacité de ce procédé sera réduit 1) par l'existence d'allèles marqueurs en commun entre les deux races 2) par les recombinaisons entre les locus marqueurs et majeur. Pour éviter ce phénomène, il convient d'utiliser des marqueurs aussi proches que possible du locus majeur et/ou de trier sur les génotypes en un couple de marqueurs encadrant le gène (Soller et Plotkin Hazan, 1977). Le tableau 2 montre l'efficacité de cette technique. Quels que soient le ou les marqueurs utilisés, des vérifications par test de la descendance devront être organisées au moins pour la dernière génération d'introgession.

Dans le cas où les locus sont en équilibre de liaison au niveau de la population, il est toujours possible d'utiliser

Tableau 3

Evolution de la moyenne d'une population selon trois méthodes de sélection (d'après Zhang et Smith, 1992).

Génération	Sélection assistée par marqueur (SAM)	Sélection classique (BLUP)	Sélection combinée (BLUP + SAM)
0	0,00	0,00	0,00
1	0,64	1,04	1,09
2	0,87	1,62	1,81
3	0,98	2,12	2,40
4	1,11	2,27	2,79
5	1,26	2,40	3,06
6	1,36	2,76	3,31
7	1,42	3,01	3,54
8	1,48	3,24	3,70
9	1,51	3,59	3,85

L'héritabilité du caractère est de 0,10.

La population est une cinquième génération de croisements entre deux lignées homozygotes.

**Tableau 4**  
**Progrès génétique attendu dans une sélection assistée par marqueur, en pourcentage du schéma sans marqueur**  
**(d'après Meuwissen et Van Arendonk, 1992).**

Part de la variance génétique intra famille expliquée par les marqueurs	Schéma de sélection		
	conventionnel	noyau ouvert	noyau fermé
0,00	100,0	100,0	100,0
0,05	100,8	111,6	109,4
0,10	102,1	120,8	117,8

Sélection intra famille avant la mise en testage des taureaux laitiers.

le **déséquilibre de liaison intra famille**. En effet, aux événements de recombinaison près, un reproducteur M1Q1/M2Q2 transmettra systématiquement Q1 avec M1 et Q2 avec M2. Un tri intra famille sur le génotype marqueur sera donc efficace. A titre d'exemple, supposons qu'un QTL de production laitière soit détecté dans une population bovine. A partir des données du testage sur descendance d'un taureau hétérozygote M1M2 au marqueur de ce QTL, le génotype global (la *phase*) de ce taureau peut être établi (M1Q1 sur le premier chromosome et M2Q1 sur le second ou M1Q1/M2Q2 ou M1Q2/M2Q1 ou M1Q2/M2Q2), ce qui permet de ne mettre en testage sur descendance que ses fils porteurs de l'allèle marqueur associé à l'allèle quantitatif favorable. L'efficacité du procédé est réduit par les recombinaisons entre les locus *M* et *Q*, par un polymorphisme insuffisant au marqueur ne permettant pas de connaître la transmission de l'allèle marqueur (cas des pères non hétérozygotes au marqueur, cas d'un fils M1M2 née de mère M1M2) et par l'imprécision de la détermination du génotype en *Q* avec la seule observation des phénotypes quantitatifs.

Pour des marqueurs très proches du QTL, la démarche pourrait être généralisée aux générations suivantes : ayant établi les phases pour chaque reproducteur d'une génération ancestrale, il est théoriquement envisageable, grâce au suivi continu de la transmission des marqueurs et à l'enregistrement des performances, de suivre l'évolution des phases chez les reproducteurs des différentes générations. Une extension de la méthode "BLUP modèle animal" d'évaluation des reproducteurs incluant cette information sur la transmission des allèles marqueurs a ainsi été proposée par Fernando et Grossman (1989).

### 3.2. GESTION CONJOINTE D'UN ENSEMBLE DE QTLS MARQUÉS ET NON MARQUÉS

L'utilisation des marqueurs pour cette gestion globale sera sans doute une application beaucoup plus fréquente pour les animaux domestiques que l'introgession assistée par marqueurs d'un gène unique telle que décrite précédemment. Elle a déjà été simulée par quelques auteurs. Leurs évaluations chiffrées de cette sélection assistée par marqueurs montrent une grande diversité qui s'explique essentiellement par la diversité des situations envisagées. Il est possible d'ores et déjà de cerner les meilleurs créneaux d'utilisation de cette technique.

#### 3.2.1. Utilisation d'un déséquilibre de liaison global entre marqueurs et QTL.

Ce déséquilibre existe notamment à la suite de la création d'une population par hybridation de deux lignées. On peut

alors, au moins partiellement, sélectionner les animaux de cette population hybride d'après le nombre d'allèles marqueurs "favorables" (en réalité associés à des allèles quantitatifs favorables) qu'ils portent. Les auteurs ont ainsi proposé de trier les animaux sur leur *score moléculaire* dont les coefficients sont ceux de la régression multiple des phénotypes d'une première génération sur le nombre d'allèles favorables en chaque locus marqueur.

Lande et Thompson (1990) ont ainsi montré que l'intérêt de cette forme de sélection assistée par marqueurs est maximale quand l'héritabilité du caractère est faible et quand les QTL expliquent une grande part de la variabilité génétique. Le progrès génétique pourrait être multiplié par 2 ou 3 dans les situations les plus optimistes. Ces résultats ont été réévalués par Zhang et Smith (1992 et 1993) en tenant compte de la disparition progressive, avec les générations, du déséquilibre de liaison du fait des recombinaisons. Une des situations simulée par ces auteurs est reportée dans le tableau 3. La complémentarité entre le score moléculaire et l'indice BLUP classique est claire.

#### 3.2.2. Utilisation d'un déséquilibre de liaison intra famille

La plupart de temps pour les animaux domestiques, les déséquilibres de liaison ne s'observeront qu'intra famille, le même allèle d'un marqueur lié à un QTL étant associé à différents allèles quantitatifs d'une famille à l'autre et d'un ancêtre à l'autre, ce qui complique grandement l'évaluation des reproducteurs : il faut alors généraliser à plusieurs QTL la démarche de Fernando et Grossman (1989) en intégrant l'ensemble des données quantitatives et sur les marqueurs de toutes les générations mesurées. L'idée de base est que les informations sur les génotypes marqueurs permettent d'affiner les coefficients de la matrice de corrélations entre valeurs génétiques des apparentés (Chevalet et al, 1984 ; Fernando et Grossman, 1989 ; Goddard, 1992 ; Van Arendonk et al, 1994 ; Wang et al, 1995) : à proximité du locus du marqueur *M*, deux frères se ressembleront plus qu'en moyenne s'ils ont reçu le même allèle de leur père, moins s'ils ont reçu des allèles différents.

Le cas où un seul QTL est marqué a été détaillé par Ruane et Colleau (1995a, 1995b). Ces auteurs ont considéré un gène quantitatif d'effet relativement faible (expliquant 1/8 ou 1/4 de la variance génétique) et une population de petite taille, proche de celle des noyaux MOET (basés sur la superovulation et le transfert des embryons). ou des lignées de sélection porcines. L'efficacité de la sélection assistée par marqueur, comparée à la situation standard où les animaux sont choisis à l'aide du BLUP modèle animal, est maximale (tableau 4) pour les caractères d'héritabilité

Tableau 5

Efficacité relative (en pourcentage) de la sélection assistée par marqueurs utilisant un BLUP modèle animal étendu par rapport à la sélection standard (d'après Ruane et Colleau, 1995 a, b)

N° génération	Mâles et femelles mesurés			Femelles mesurées		
	$h^2=0,5$	$h^2=0,25$	$h^2=0,1$	$h^2=0,5$	$h^2=0,25$	$h^2=0,1$
1	100 *	100	101	102	100	103
2	100	102	104	104	104	109
3	100	101	104	103	105	112
6	100	101	101	102	103	106

\* Cette colonne contient des valeurs qui ne sont pas des références 100 strictes mais qui correspondent à des arrondis.

faible, et/ou les caractères qui ne s'expriment que chez une partie des individus (les mâles, par exemple). Ces résultats sont tout à fait en accord avec les observations de Lande et Thompson (1990) dans le cas d'un déséquilibre de liaison global (cf. 3.2.1). Les surcroûts de progrès génétique sont toutefois modérés et transitoires puisqu'inférieurs à 6 % au bout de 6 générations, et le plus souvent de l'ordre de 2 à 3 %.

Le cas où **plusieurs QTL** sont marqués et où certains animaux n'ont pas de performances a été envisagé par Kashi et al (1990), Meuwissen et Van Arendonk (1992), Brascamp et al (1993). La situation étudiée est celle d'une présélection sur la valeur génétique laitière des taureaux avant leur mise en testage sur descendance. Kashi et al (1990) trouvaient une augmentation de progrès génétique pouvant atteindre 20 à 30 pcent en négligeant cependant la réduction d'intensité de sélection sur les mères à taureaux et la réduction de variabilité génétique. Meuwissen et Van Arendonk (1992) montraient que la sélection assistée par marqueurs devrait être particulièrement efficace dans le cas des schémas de sélection MOET. Dans ces schémas en effet, le marquage du génome permet une sélection intra famille (deux frères ont un score moléculaire différent dès qu'ils héritent, du même parent, des allèles marqueurs différents), qui est impossible par ailleurs, la sélection des taureaux étant essentiellement une sélection sur ascendance. Le tableau 5 montre que des gains supérieurs à 20 pcent sont obtenus. Enfin, une première évaluation positive de la rentabilité de ce type de sélection assistée par marqueurs a été publiée par Brascamp et al (1993). Il faudra confirmer ces résultats et les étendre à d'autres schémas de sélection.

## CONCLUSION

Dans un certain nombre de cas, les marqueurs génétiques apparaissent dès aujourd'hui comme un outil utile pour la gestion génétique des populations. Ils sont la seule solution pour des applications telles que la vérification des paternités ou le calcul des distances entre populations. Ils enrichissent nos moyens d'action pour d'autres volets, valorisation des gènes à effet majeur ou sélection classique. L'efficacité et

plus encore la rentabilité de ces opérations doivent être confirmées. Dés maintenant, il est clair que les marqueurs génétiques, s'ils peuvent largement contribuer à une meilleure gestion des populations, ne sauraient être considérés comme une alternative aux méthodes actuelles, basées sur les informations phénotypiques et généalogiques : les deux approches sont complémentaires. L'utilité des marqueurs génétiques apparaîtra surtout dans les situations où les données classiques sont peu informatives vis à vis des valeurs génétiques, ou très coûteuses à obtenir : sélection sur des caractères exprimés dans un seul sexe, ou demandant l'abattage des animaux, caractères d'héritabilité faible... La sélection assistée par marqueurs sera d'autant plus efficace que ces marqueurs «expliqueront» une part plus importante de la variabilité génétique.

La qualité des marqueurs est très directement liée à l'intensité de leur liaison avec les zones particulières (gènes majeurs, QTL) qu'ils marquent et à leur polymorphisme. Dans beaucoup de cas, la pleine efficacité ne sera atteinte que lorsque ces gènes à effets quantitatifs seront isolés et leurs formes alléliques responsables des variations de performances identifiées. La **construction de génotypes** ou réunion, dans un même individu, d'un certain nombre d'allèles favorables en différents QTLs contrôlant plusieurs caractères peut être défendue comme un objectif technique pour le long terme. Cette technique, qui est considérée comme réaliste en amélioration des plantes, demande que les gènes intéressants soient assemblés par des introgressions successives concernant un ou plusieurs d'entre eux. Ceci n'est pas encore envisageable dans le cas des animaux pour lesquels les coûts d'élevage sont élevés et les intervalles de générations longs, et ne le sera que grâce à une artificialisation très importante des méthodes de reproduction pour disposer d'une pression de sélection considérablement accrue (Georges et Massey, 1991). Le nombre de combinaisons de génotypes aux différents locus devant être considérable, les choix des combinaisons à tester devront reposer sur des *a priori* statistiques (connaissances des effets des QTL) et biologiques (connaissance des gènes impliqués).

## RÉFÉRENCES

- BARENDSE W., ARMITAGE S.M., KOSSAREK L.M., SHALOM A., KIRKPATRICK B.W., RYAN A.M., CLAYTON D., LI L., NEIBERGS H.L., ZHANG N., GROSSE W.M., WEISS J., CREIGHTON P., MCCARTHY F., RON M., TEALE A.J., FRIES R., MCGRAW R.A., MOORE S.S., GEORGES M., SOLLER M., WOMACK JE., HETZEL D.J.S., 1994. *Nature Genet.*, 6, 227-235.
- BERGHMANS S. COCKETT N., GEORGES M., 1995. 2èmes Rencontres Recherches Ruminants. Paris, 13-14 dec 1995.
- BISHOP M.D., KAPPES S.M., KEELE J.W., STONE R.T., SUNDEN S.L.F., KAWKINS G.A., TOLDO S.S., FRIES R., GROSZ M.D., YOO J., BEATTIE C.W., 1994. *Genetics*, 136, 619-639.
- BOICHARD D., ELSÉN J.M., LE ROY P., CHEVALET C., 1995. 2èmes Rencontres Recherches Ruminants. Paris, 13-14 dec 1995.
- BOSCHER M.Y., GUERIN-FAUBLEE V., 1994. *Le Point Vétérinaire*, 25, 157, 841-844.
- BOOMAROV, 1991. In ELSÉN J.M., BODIN L., THIMONIER J. (Eds) 2nd Int. Workshop Major genes for reproduction in sheep. *Les colloques de l'INRA*, 57, 431-440.
- BRASCAMP E.W., VAN ARENDONK J.A.M., GROEN A.F., 1993. *J. Dairy Sci.*, 76, 1204-1213.
- CHARLIER C. MENISSIER F., 1995. 2èmes Rencontres Recherches Ruminants. Paris, 13-14 dec 1995.
- CHEVALET C., 1992. *INRA Prod. Anim.*, n° hors série Elements de génétique quantitative et application aux populations animales, 295-297.
- CHEVALET C., GILLOIS M., KHANG J.V.T., 1984. *Genet. Sel. Evol.*, 16, 431-444.
- CRAWFORD A.M., DODDS K.G., EDE A.J., PIERSON C.A., MONTGOMERY G.W., GARMONSWAY H.G., BEATTIE A.E., DAVIES K., MADDOX J.F., KAPPES S.W., STONE R.T., NGUYEN T.C., PENTY J.M., LORD E.A., BROOM J.E., BUITKAMP J., SCHWAIGER W., EPPLÉN J.T., MATTHEW P., MATTHEWS M.E., HULME D.J., BEH K.J., MCGRAW R.A., BEATTIE C.W., 1994. *Genetics*, 140, 703-724.
- FERNANDO J.L., GROSSMAN M., 1989. *Genet. Sel. Evol.*, 21, 467-477.
- GEORGES M., MASSEY J.M., 1991. *Theriogenology*, 35, 151-159.
- GODDARD M.E., 1992. *Theor. Appl. Genet.*, 83, 878-886.
- GODDARD M.E., AHMED A.M., 1982. 2nd World Congress on Genetics applied to livestock Production, Madrid, 4-8 oct 1982.
- GROSCLAUDE F., AUPÉTIT R.Y., LEFÈVRE J., MERIAUX J.C., 1990. *Genet. Sel. Evol.*, 22, 317-338.
- HANSET R., 1975. *Ann. Méd. Vét.*, 119, 71-80.
- HILLEL J., SCHAAP T., HABERFELD A., JEFFREYS A.J., PLOTSKY Y., CAHANER A., LAVI U., 1990. *Genetics*, 124, 783-789.
- HOSPITAL F., CHEVALET C. MULSANT P., 1992. *Genetics*, 132, 1199-1210.
- KASHI Y., HALLERMAN E., SOLLER M., 1990. *Anim. Prod.*, 51, 63-74.
- LANDE R., THOMPSON R., 1990. *Genetics*, 124, 743-756.
- LANNELUC I., DRINKWATER R.D., ELSÉN J.M., HETZEL D.J.S., NGUYEN T.C., PIPER L.R., THIMONIER J., HARRISON B., GELLIN J., 1994. *Mammalian Genome*, 5, 26-33.
- LE ROY P., ELSÉN J.M., RICORDEAU G., BOUILLON J., MANFREDI E. CHASTIN P., MONOD F., 1995. 2èmes Rencontres Recherches Ruminants. Paris, 13-14 dec 1995.
- LEVEZIEL H., VAIMAN D., 1995. 2èmes Rencontres Recherches Ruminants. Paris, 13-14 dec 1995.
- MEUWISSEN T.H.E., VAN ARENDONK J.A.M., 1992. *J. Dairy Sci.*, 75, 1651-1659.
- MONTGOMERY G.W., CRAWFORD A.M., PENTY J.M., DODDS K.G., EDE A.J., HENRY H.M., PIERSON C.A., LORD E.A., GALLOWAY S.M., SCHMACK A.E., SISE J.A., SWARBRICK P.A., HANRAHAN V., BUCHANAN F.C., HILL D.F., 1993. *Nature Genetics*, 4, 410-414.
- NEI M., 1971. *Am. Nat.*, 105, 385-398.
- ROGERS J.S., 1972. *Studies in Genetics VII*, Univ. of Texas Pub., 7213, 145-153.
- ROHRER G.A., ALEXANDER L.J., KEELE J.W., SMITH T.P., BEATTIE C.W., 1994. *Genetics*, 136, 231-245.
- RUANE J., COLLEAU J.J., 1995a. *Genet. Res.*, sous presse.
- RUANE J., COLLEAU J.J., 1995b. *J. Dairy Sci.*, soumis.
- SOLLER M., BECKMANN J., 1983. *Theor. Appl. Genet.*, 67, 25-33.
- SOLLER M., PLOTKIN-HAZAN J., 1977. *Theor. Appl. Genet.*, 51, 133-137.
- VAN ARENDONK J.A.M., TIER B., KINGHORN B.P., 1994. *Genetics*, 137, 319-329.
- WAHLUND S., 1928. *Hereditas*, 11, 139-168.
- WANG T., FERNANDO R.L., VAN DER BEEK S., GROSSMAN M., VAN ARENDONK J.A.M., 1995. *Genet. Sel. Evol.*, 27, 251-274.
- ZHANG W., SMITH C., 1992. *Theor. Appl. Genet.*, 83, 813-820.
- ZHANG W., SMITH C., 1993. *Theor. Appl. Genet.*, 86, 492-496.

