

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

# Utilisation des marqueurs pour localiser les gènes responsables de la variabilité des caractères quantitatifs (QTL)

**Résumé.** Cet article rappelle tout d'abord les conditions expérimentales qui ont permis, chez les plantes, la localisation de gènes ayant des effets sur des caractères quantitatifs, en soulignant l'importance de la notion de déséquilibre de liaison. Les situations qui permettent de s'en rapprocher dans les populations animales sont ensuite évoquées : croisement entre races éloignées, localisation d'un gène majeur attribué à une mutation récente, déséquilibre aléatoire induit par la consanguinité ou déséquilibre systématique dû à la sélection, études de ségrégation intra-familiales. L'ordre de grandeur des effectifs d'animaux qu'il sera nécessaire de typer et d'observer est indiqué, en fonction du nombre et de l'importance des gènes quantitatifs à détecter.

Où sont situés les gènes responsables de la variabilité génétique d'un caractère quantitatif ? La question a été depuis longtemps envisagée en Génétique Humaine pour rechercher une liaison entre un marqueur et le gène responsable d'une maladie (Elston 1981). En Génétique Appliquée cela correspond à un gène majeur. Dans une situation stricte de Génétique Quantitative on ne connaît pas a priori de gène, la variabilité génétique ne se manifeste pas par la présence d'allèles en ségrégation mais par l'existence d'une variance génétique, mesurable par une méthode statistique. La question devient : le marqueur est-il lié à un gène responsable d'une certaine fraction de la variabilité génétique d'un caractère quantitatif ? Tant qu'on n'a disposé que de marqueurs isolés, les réponses sont demeurées souvent ambiguës, car il est impossible de trancher entre un marqueur proche d'un gène à faible effet, et un marqueur éloigné d'un gène à fort effet (Soller et Genizi 1978). On peut maintenant imaginer que l'on dispose d'une panoplie de marqueurs, bien répartis sur tous les chromosomes, assez proches les uns des autres, bien polymorphes, de telle sorte que ces ambiguïtés puissent être levées.

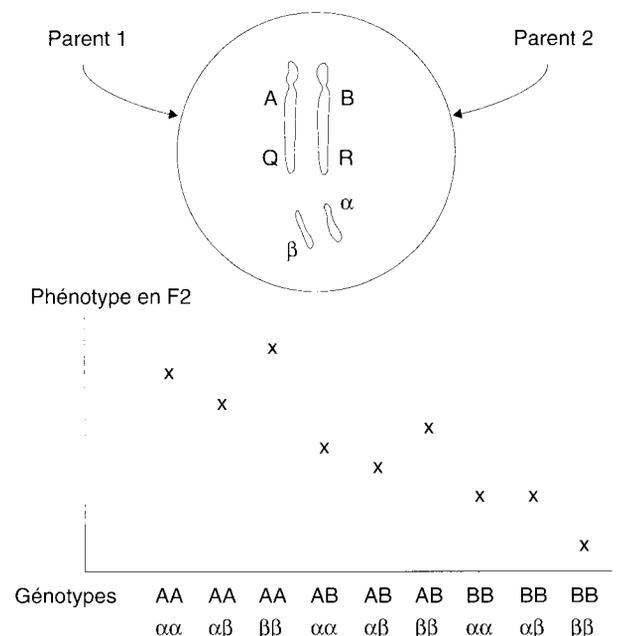
## 1 / Raison d'un succès chez les plantes

La première localisation systématique de gènes impliqués dans des caractères quantitatifs a été réussie chez la Tomate. La méthode s'appuie sur un réseau moyennement dense de marqueurs (20 cM en moyenne entre deux marqueurs), et sur une situation expérimentale qui permet d'exploiter le déséquilibre de liaison existant entre les locus marqueurs et les

locus à effets quantitatifs (Lander et Botstein 1989, Paterson *et al* 1988).

La figure 1 illustre la situation de déséquilibre obtenue par le croisement de deux lignées pures. Aux marqueurs pris en compte, comme aux locus quantitatifs recherchés, les lignées pures sont caractérisées par des allèles différents. Chez la plante hybride F1, tous les allèles d'un parent sont portés par un même lot chromosomique, que ce soient les allèles des marqueurs (A,  $\alpha$  pour le parent 1 ; B,  $\beta$  pour le parent 2)

**Figure 1.** Situation de déséquilibre créé par le croisement de 2 lignées pures.



ou les allèles du locus quantitatif (Q issu du parent 1 ; R issu du parent 2). Si les locus (A,B) et (Q,R) sont proches, les gamètes produits par une plante F1 sont en majorité des types parentaux [AQ] ou [BR], tandis qu'une faible proportion (égale au taux de recombinaison) sont des types recombinés [AR] et [BQ]. Par conséquent, si dans les descendance F2 la comparaison entre les plantes de génotypes [AA], [AB], ou [BB] indique une différence significative, on peut conclure que au voisinage du locus marqueur (A,B), il existe un locus dont les allèles ont des effets différents sur le caractère. De plus les différences entre les moyennes des plantes des trois génotypes permettent d'évaluer l'effet quantitatif de la substitution d'un allèle par un autre, et de tester si l'un des allèles quantitatifs est dominant ou récessif par rapport à l'autre. En réalité, en travaillant avec une carte de marqueurs assez dense, on ne considère pas simplement des génotypes en un locus, mais des haplotypes correspondant à des locus adjacents : cela permet d'avoir des résultats avec des effectifs moindres, et de localiser les locus quantitatifs entre deux marqueurs.

Au moins pour certaines espèces diploïdes et autogames, cette méthodologie est très puissante. Elle permet des localisations très fines des gènes quantitatifs, elle fournit une mesure quantitative des effets des allèles, et les déséquilibres entre marqueurs et allèles quantitatifs peuvent être ensuite directement utilisés pendant quelques générations pour introduire un allèle quantitatif d'une variété dans une autre.

L'existence du déséquilibre initial est une condition nécessaire à la mise en œuvre de cette méthodologie. Si l'on considère un marqueur ( $\alpha,\beta$ ) sur un autre chromosome (figure 1), les ségrégations indépendantes répartissent les allèles quantitatifs Q et R entre les génotypes ( $\alpha,\alpha$ ), ( $\alpha,\beta$ ) et ( $\beta,\beta$ ). Il en serait de même si les individus F1 de la figure 1 étaient né d'accouplements au hasard entre individus d'une même grande population "en équilibre", où il n'y a pas de déséquilibre de liaison. Si par exemple on suppose qu'il y a équilibre de liaison entre un locus marqueur (A,B) et le locus quantitatif (Q,R), les fréquences des génotypes se maintiennent, au cours des générations, égales aux produits des fréquences alléliques, quel que soit le taux de recombinaison (r) entre les locus. Dans ce cas, aucune association entre le marqueur et le locus quantitatif ne peut être détectée en comparant les moyennes des individus de différents génotypes au locus marqueur.

La recherche des conditions d'apparition d'un déséquilibre de liaison est donc une étape capitale pour entreprendre une recherche des gènes quantitatifs.

## 2 / Sources d'un déséquilibre de liaison dans une population animale

Comme on ne peut créer de lignées pures chez les animaux, sauf chez la souris, il faut soit tirer parti de situations particulières soit sélectionner des situations expérimentales pour isoler des sous-populations au sein desquelles existe un déséquilibre de liaison.

### 2.1 / Croisements entre races (exemple : porcs LW x MS)

Un croisement révélant des différences génétiques importantes pour plusieurs caractères quantitatifs constitue une situation expérimentale analogue au croisement de lignées. Une différence notable est cependant que dans chacune des deux races, il existe une variabilité génétique en général importante, qui peut se manifester aussi bien aux locus marqueurs qu'aux locus quantitatifs : en chaque locus on s'attend à trouver plusieurs allèles dans chaque population. En outre la différence génétique entre les deux populations peut correspondre seulement à des fréquences alléliques différentes, et non à la présence d'allèles différents et spécifiques des populations. Il semble dans cette situation qu'une première exigence expérimentale serait de choisir des marqueurs tels que chaque population soit caractérisée par des ensembles différents d'allèles, de sorte que dans la descendance on puisse déterminer sans ambiguïté l'origine des gènes marqueurs. En ce qui concerne les locus quantitatifs, on ne peut rien affirmer : notamment, si des allèles quantitatifs sont communs aux deux populations, les déséquilibres de liaison ne seront pas maximaux. La conséquence statistique est, à effectifs égaux, une perte de puissance : on augmente le risque de manquer une association marqueur-caractère quantitatif.

### 2.2 / Mutations récentes (exemples : Hal, Booroola)

Si l'on suppose qu'un gène majeur correspond à une mutation récente et unique dans une population, on peut espérer trouver une association entre ce locus et des marqueurs voisins. En effet si la mutation est unique, l'allèle mutant est associé à l'origine avec un haplotype particulier pour les gènes voisins. Aussi longtemps que les recombinaisons ne l'auront pas complètement détruit, il caractérisera les individus mutants. Inversement des individus présentant cet haplotype ne sont pas nécessairement porteurs de la mutation, mais il constitue une indication très forte s'il est défini par plusieurs marqueurs fortement liés. Du fait des recombinaisons, l'allèle mutant au locus majeur se trouve progressivement associé à des haplotypes différents : la découverte de ce type d'association n'est donc possible que dans un délai relativement court après l'apparition de la mutation. On peut se ramener artificiellement à cette situation en considérant la descendance d'un reproducteur porteur de la mutation.

### 2.3 / Dérive aléatoire et consanguinité

Sans sélection et dans une grande population, on attend des déséquilibres nuls entre deux locus quelconques. Cela reste vrai en moyenne si la population est petite. Dans une petite population de N individus, il n'y a place que pour un nombre fini d'allèles à chaque locus : partant du maximum de 2N à la génération initiale, leur nombre décroît rapidement pour aboutir à la fixation au bout d'un temps moyen de l'ordre de  $4N$  générations. Sur un segment correspondant à un taux de recombinaison r, le nombre de segments différents (allèles apparents pour les haplotypes sur ce segment) est également limité. Si, par exemple, avec 5 locus polymorphes ayant 4 allèles par locus (microsatellites) le long d'un segment de 5 centiMorgans, on ne trouve qu'une dizaine d'haplo-

types distincts, parmi les 1024 possibles, il existe nécessairement des déséquilibres de liaison, qui peuvent être exploités pour rechercher des associations. Cet effet ne semble utilisable que pour des liaisons très fortes.

## 2.4 / Sélection

Une source de liaison systématique est la sélection. Le phénomène est susceptible d'amplifier des déséquilibres apparus par hasard. Si un gène ou une petite région chromosomique est responsable de la variabilité d'un caractère quantitatif qui a été sélectionné depuis plusieurs générations, la sélection des allèles favorables entraîne la sélection corrélative de certains allèles en des locus voisins neutres. Ce mécanisme est susceptible d'amplifier, mais non de créer, des déséquilibres de liaison entre allèles quantitatifs et allèles neutres en des locus voisins.

## 2.5 / Etudes intra-familiales (analyse de ségrégation)

Si l'on suppose qu'il y a équilibre de liaison entre les marqueurs et les QTL, seule une analyse intra-famille est pertinente. Un schéma simple, utilisant des familles de demi-frères est illustré en figure 2. Dans la descendance d'un reproducteur hétérozygote pour un marqueur (A,B), on compare les phénotypes des descendants qui ont reçu l'allèle (A) ou l'allèle (B). Une différence significative indique qu'un locus quantitatif variable a ségrégré conjointement avec le marqueur. Dans cette approche on n'attribue pas d'effet quantitatif à un allèle marqueur : ces effets ne peuvent être définis qu'intra famille, car un même

allèle marqueur (A) peut être associé à un allèle favorable dans une famille et défavorable dans une autre. Dans cette approche un certain nombre de familles sont non informatives, pour chaque marqueur : celles où le père est homozygote au locus marqueur (on ne peut pas classer les descendants) ou au locus quantitatif (les descendants ont tous reçu le même allèle quantitatif de leur père). La méthode est d'autant plus efficace que l'on dispose de marqueurs très polymorphes et codominants, et que le caractère analysé a une héritabilité plus élevée. Une application particulière de cette méthode (Weller) en race Holstein (Cowan *et al*) a permis de détecter un QTL de la quantité de lait à l'aide du polymorphisme du gène de la prolactine : chaque famille de demi-frères est constituée par les quelques centaines de taureaux qui sont fils d'un taureau-élite, et chaque fils est testé sur descendance sur 50 à 100 filles. Le caractère étudié (index sur descendance) a une héritabilité élevée, et le nombre d'animaux à typer est restreint aux taureaux.

## Conclusion

Le choix d'une méthode et du matériel animal dépendent de l'usage attendu de ces localisations de QTL, car les informations obtenues ne sont pas équivalentes.

La méthode utilisant le déséquilibre induit par le croisement identifie des régions et associe des effets aux allèles marqueurs, valables pour les croisements réalisés seulement : ces régions peuvent ne révéler aucune variabilité quantitative dans une autre population. Les effectifs nécessaires (Lande et Thompson 1990) en admettant une couverture "assez dense" sont relativement abordables, ils dépendent de l'héritabilité et du nombre (ou de l'importance) des QTL responsables de la variation du caractère (tableau 1).

Dans les populations animales, l'approche intra familiale semble la mieux appropriée, malgré sa puissance faible. Elle est susceptible de révéler des régions importantes dans la population étudiée, mais l'évaluation des effets doit être mise à jour continuellement. Les effectifs nécessaires sont en moyenne plus élevés. Le schéma de Weller permet de limiter le nombre d'animaux à tester à l'ordre du millier, grâce au testage sur descendance, pour détecter un QTL assez important (effet de 0,5 écart-type phénotypique). Dans une situation où seules les performances individuelles des descendants sont mesurées, il faut

**Figure 2.** Analyse intra-familiale. On considère un marqueur avec 2 allèles (A, B), proche d'un locus quantitatif à 2 allèles (Q, R). Les descendants sont classés selon l'allèle (A) ou (B) qu'ils ont reçu du père, et les phénotypes moyens de chaque catégorie sont comparés dans chaque famille ; tous les descendants ne peuvent pas en général être ainsi classés, si l'on ignore quel allèle paternel a été transmis (origine inconnue notée ?).

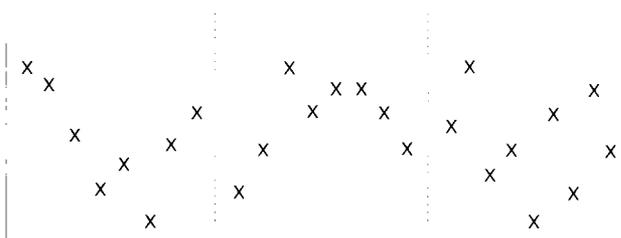
Pères hétérozygotes au marqueur :



Allèle reçu du père :

A A A B B B ? ?    A A B B B B ? ?    A A A A B B B ? ?

Performances observées chez ces descendants :



**Tableau 1.** Effectifs des individus à élever et à typer pour détecter une association entre un marqueur et un QTL, dans une population en déséquilibre issue d'un croisement (d'après Lande et Thompson 1990). On suppose que les marqueurs disponibles sont susceptibles de rendre compte de la moitié de la variabilité génétique totale. Entre parenthèses : effet moyen d'un QTL (en écart-type phénotypique), fonction de l'héritabilité et du nombre supposé de QTL.

Nombre de QTL	$h^2 = 0,3$	$h^2 = 0,1$	$h^2 = 0,025$
1	140 (0,7)	400 (0,4)	1 600 (0,2)
4	700 (0,35)	2 000 (0,2)	8 000 (0,1)
8	1 800 (0,25)	5 000 (0,14)	20 000 (0,07)

tester environ 20000 animaux répartis en 1500 familles de demi-frères pour détecter un QTL correspondant à 0,2 écart-type, avec un marqueur très proche.

Un gène impliqué dans l'expression d'un caractère quantitatif peut être localisé seulement s'il présente plusieurs formes allèles correspondant à des niveaux différents du caractère : le caractère doit présenter une variabilité génétique entre les individus observés. D'autres gènes importants dans le déterminisme du caractère peuvent ne pas être découverts par cette méthode si tous les individus observés portent le

même allèle. Mais inversement l'information de position obtenue dans une population demeure pertinente pour l'espèce car elle désigne des régions chromosomiques impliquées dans l'expression du caractère. On peut même penser à transposer ces informations entre espèces dans la mesure où les travaux de cartographie comparative montrent une conservation locale de l'organisation génétique. Ainsi, la découverte et la localisation de QTL chez la Souris - espèce où existent des lignées consanguines et dont la carte est très dense - pourraient-elles orienter la recherche de QTL homologues chez d'autres mammifères.

### Références bibliographiques

- Elstron R. C., 1981. Segregation Analysis. in : *Advances in Human Genetics*, Vol 11 (H. Harris and K. Hirschhorn, ed.), Plenum Publishing Corporation, New York, pp. 63-120.
- Lande R. and Thompson R., 1990. *Genetics*, 124 : 743-756.
- Lander E.S. and Botstein D., 1989. *Genetics*, 121 : 185-199.
- Paterson A.H., Lander E.S., Hewitt J.D., Paterson S., Lincoln S.E., and Tanksley S.D., 1988. *Nature*, 335 : 721-726.
- Soller M. and Genizi , 1978. *Biometrics*, 34 : 47-55.
- Weller J.I., Khashi Y. and Soller M., 1989. *J. Dairy Sci.*, 73 : 2525-2537.