



HAL
open science

Analyse génétique d'un caractère quantitatif

M. Sourdioux, Sandrine Lagarrigue, M. Douaire

► **To cite this version:**

M. Sourdioux, Sandrine Lagarrigue, M. Douaire. Analyse génétique d'un caractère quantitatif. *Productions Animales*, 1997, 10 (3), pp.251-258. hal-02693262

HAL Id: hal-02693262

<https://hal.inrae.fr/hal-02693262>

Submitted on 1 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse génétique d'un caractère quantitatif

L'évaluation génétique des reproducteurs est encore essentiellement basée sur l'application de la théorie de la génétique quantitative. Cependant, depuis quelques années, la recherche des gènes gouvernant les caractères d'intérêt agronomique a été largement développée. Deux approches sont utilisées, l'une recherchant des zones du génome liées au niveau d'expression d'un caractère, l'autre s'intéressant directement à des gènes bien identifiés. Comme le montrent les résultats obtenus jusqu'à présent, ces deux approches ne donnent pas accès à la même information et doivent être utilisées de façon complémentaire.

L'amélioration génétique des espèces animales, basée sur l'application de la théorie de la génétique quantitative, repose elle-même sur un concept de déterminisme polygénique des caractères (*c'est-à-dire* un grand nombre de gènes ayant tous un effet additif et de faible importance). Des modèles biométriques fournissent un support à la prédiction de la valeur génétique des individus, sans s'intéresser aux gènes ou aux mécanismes biologiques sous-jacents déterminant les différences entre individus. Cependant, dès 1923, Sax (cité par Tanksley 1993) proposait de mettre en relation la variation d'un caractère quantitatif - taille des graines de haricot - avec un polymorphisme (phénotypique) -couleur-, jetant ainsi les bases des recherches sur le détermi-

nisme génétique des caractères polygéniques et de la sélection assistée par marqueur. Le nombre d'études de ce type est longtemps resté limité. En effet, les possibilités d'observer un polymorphisme étaient restreintes à des observations de phénotypes, aux techniques d'hémagglutination ou d'électrophorèse de protéines. Aujourd'hui, diverses techniques de biologie moléculaire permettent de mettre en évidence l'infinie variabilité génétique au niveau de l'ADN. Il devient donc envisageable de sélectionner les individus directement sur la connaissance d'une partie de leur génotype et de combiner cette information aux méthodes actuelles d'amélioration génétique (Weller et Fernando 1991, Ruane et Colleau 1995).

Résumé

Depuis quelques années, de nombreux travaux ont été entrepris pour identifier les gènes contrôlant directement une part de la variabilité de caractères quantitatifs afin d'enrichir les méthodes d'amélioration génétique. Ils ont été classés en deux types d'approche. Une première stratégie exploite la connaissance de cartes géniques et recherche des QTL (Quantitative Trait Loci), régions du génome contrôlant un caractère quantitatif. Elle permet une application à la sélection par l'utilisation de structures géniques (appelées marqueurs) associées au caractère ; c'est aussi une étape dans la découverte du gène directement responsable d'une part de la variabilité. L'autre approche s'appuie sur la connaissance biologique du caractère et s'intéresse directement à quelques gènes, supposés importants pour le niveau d'expression du caractère, qui deviennent ainsi des candidats à l'exploitation de la variabilité observée. L'analyse de ces méthodes et des résultats obtenus permet aujourd'hui de mieux percevoir leurs potentialités et leurs limites et surtout de mettre en évidence leur complémentarité dans l'analyse du génome contrôlant des caractères d'intérêt économique, voire leur similitude partielle dans l'élaboration et l'utilisation des dispositifs expérimentaux.

L'introduction d'une information moléculaire en sélection est déjà en place sur des caractères monogéniques ; citons par exemple le typage de bovins laitiers Holstein pour la BLAD (Bovine Leukocyte Adhesion Deficiency ; Shuster *et al* 1992). Le polymorphisme de quelques gènes à effet majeur sur des caractères de production est également pris en compte dans certains schémas, comme celui du gène de sensibilité à l'halothane chez le porc (Hanset et Grobet 1992, Dalens et Runavot 1993), du gène des caséines α_{S1} chez la chèvre (Manfredi *et al* 1995). La généralisation de cette pratique à des gènes de moindre effet ne sera effective que lorsque le déterminisme des principaux caractères d'importance économique sera - au moins en partie - élucidé.

Le principal enjeu des recherches de génétique moléculaire sur les espèces d'élevage est

Marqueur

Un marqueur est une séquence d'ADN polymorphe facilement repérable. Si ce marqueur est un gène, sa détection peut se faire au niveau du phénotype de l'animal ou de l'expression du gène (polymorphisme protéique). Elle peut se faire directement au niveau de l'ADN (par exemple par la technique de RFLP = polymorphisme de longueur de fragments de restriction), que ce soit pour un gène ou une séquence anonyme. Le développement de cette technique de RFLP a été à la base du regain d'intérêt pour la recherche du déterminisme des caractères quantitatifs. Depuis, les microsattellites, qui sont à la fois bien répartis dans le génome et très polymorphes, sont devenus les marqueurs de choix en cartographie génétique des espèces animales. Les marqueurs sont d'autant plus utiles dans la recherche de QTL qu'ils présentent un degré de polymorphisme élevé (multiallélisme) et que ces allèles sont facilement différenciables les uns des autres (codominance).

donc de réaliser cette analyse génétique des caractères quantitatifs. Dans un premier temps, la découverte de marqueur peut être suffisante. Ainsi, l'allèle sensible du gène de sensibilité à l'halothane en race Landrace français a pu être éradiqué par l'utilisation des marqueurs protéiques PHI (phosphohexose isomérase) et 6PGD (6-phosphogluconate déshydrogénase) (Courreau *et al* 1985). Cependant, l'efficacité de l'utilisation de marqueurs dépend du degré de liaison génétique entre ces marqueurs et le gène responsable du caractère étudié. Il est donc généralement nécessaire de déterminer les mutations causales, responsables (d'une partie) de la variabilité observée sur le caractère. Pour atteindre ce résultat, plusieurs étapes sont nécessaires, en particulier une étude de la transmission génétique conjointe (ou co-ségrégation) du ou des marqueurs et du caractère quantitatif, une recherche des mutations potentiellement en cause et une confirmation métabolique de leur rôle sur le caractère.

L'étude de co-ségrégation, étape nécessaire dans cette recherche des gènes sous-jacents à la variabilité d'un caractère, est réalisée par deux approches qui ont parfois été opposées mais qui présentent une complémentarité pour bien des aspects. Une première approche consiste à rechercher des zones du génome (QTL) associées à des marqueurs (polymorphes) sans *a priori* sur le rôle de ces zones. Il s'agit d'une mise en relation statistique entre les valeurs moyennes d'un caractère et les formes alléliques d'un marqueur. Cette recherche conduit à un balayage complet du génome grâce à l'établissement de cartes géniques suffisamment denses. La deuxième approche consiste à s'intéresser aux polymorphismes de gènes candidats, c'est-à-dire de gènes dont on suspecte *a priori* le rôle sur un caractère déterminé. Elle est donc

basée sur une hypothèse préliminaire élaborée le plus souvent d'après l'observation de modifications structurales ou quantitatives des protéines codées ou des ARN transcrits, en plus des connaissances physiologiques globales du caractère.

Les premiers résultats obtenus par ces deux approches, séparément ou en combinaison, permettent aujourd'hui d'effectuer un bilan de leurs potentialités et efficacités, de préciser le type de résultats réellement obtenus et de proposer des choix de stratégies, principalement en fonction des objectifs recherchés, des populations et des dispositifs disponibles.

1 / Deux approches dont les résultats sont de natures différentes

1.1 / La recherche de QTL par une approche statistique

L'approche statistique nécessite l'établissement d'une carte génétique, réseau dense de marqueurs répartis dans l'ensemble du génome et positionnés les uns par rapport aux autres. Elle consiste alors, grâce à un dispositif expérimental adapté, à identifier les marqueurs qui co-ségrègent avec le caractère quantitatif d'intérêt. La position sur la carte génétique des quelques marqueurs ainsi identifiés délimite les QTL du caractère. Leur localisation constitue l'étape initiale d'un clonage positionnel (Collins 1992) qui consiste à isoler, de proche en proche à partir des marqueurs, de nouvelles séquences d'ADN jusqu'à la découverte du gène muté en cause. Cette approche présente l'avantage majeur de mettre en évidence des zones chromosomiques dont le rôle pouvait être totalement inconnu auparavant. Les résultats obtenus chez les végétaux, précurseurs dans ce domaine avec les travaux réalisés sur la tomate (Paterson *et al* 1988), suggèrent que les caractères quantitatifs étudiés jusqu'à présent seraient gouvernés par quelques gènes à effet fort (dont l'effet est supérieur à 10 ou 15 % de la variance phénotypique) et un plus grand nombre à faible effet (Tanksley 1993 ; figure 1). A des fins d'application, ce sont souvent ces quelques gènes à effet fort qui représentent les objectifs prioritaires. Leur mise en évidence, chez les animaux d'élevage, semble nécessiter le typage d'environ 500 individus, lorsqu'ils sont issus d'un croisement F2 de deux lignées divergentes (Bidanel *et al* 1996). De premiers QTL ont été décrits sur porcins (Andersson *et al* 1994) et bovins (Georges *et al* 1995).

Le résultat de cette approche est donc l'attribution d'un rôle dans la variabilité d'un caractère à des zones chromosomiques d'environ 20 cM. Une fois le premier repérage de ces zones chromosomiques effectué, il reste difficile de préciser la localisation exacte du ou des gènes découverts (Hyne *et al* 1995). L'in-

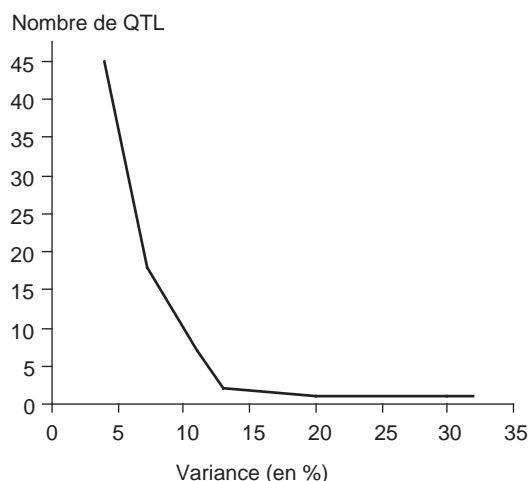
Une approche consiste à identifier des marqueurs liés au caractère intéressant et à les localiser sur la carte génétique afin de mettre en évidence les segments d'ADN qui gouvernent la variabilité du caractère.

tervalle de confiance lié au positionnement du QTL voit sa taille décroître de façon inversement proportionnelle au nombre d'individus utilisés dans le plan d'expérience, jusqu'à une limite d'environ 10 cM (Collins 1992, Darvasi *et al* 1993) soit environ 10 millions de paires de bases. Par ailleurs, pour les QTL ayant un effet relativement faible (par exemple une différence entre les deux homozygotes inférieure à 0,5 écart-type), le calcul classique de l'intervalle de confiance pourrait être fortement biaisé (Mangin *et al* 1994, Hyne *et al* 1995). Dans ces conditions, si d'autres marqueurs polymorphes sont disponibles dans les zones potentiellement intéressantes, une localisation plus fine peut être espérée grâce à la densification de la carte génétique. Mais, sans autres informations que cette position des marqueurs, la localisation et l'identification du gène responsable d'un effet sur un caractère peuvent s'avérer lentes et coûteuses (voir, par exemple, l'isolement du gène *Ob*, contrôlant l'appétit et une forme d'obésité chez la souris ; Zhang *et al* 1994).

12 / Une recherche ciblée sur des gènes identifiés

L'approche « gène candidat » présente l'avantage de s'intéresser directement au polymorphisme de gènes ou de leurs séquences promotrices. Ainsi, le problème du passage d'une région chromosomique de grande taille (d'un point de vue moléculaire) au gène responsable est éliminé. C'est aussi l'inconvénient majeur d'une telle approche puisqu'elle nécessite une connaissance au moins partielle d'un gène dont on connaît la fonction dans le métabolisme étudié. Cela représente aujourd'hui environ 2 000 cibles sur une estimation de 50 000 gènes (Soller 1994). Une piste intéressante pour la recherche du déterminisme des caractères est l'étude des gènes dont on connaît un allèle ayant un effet très important sur le caractère, comme par

Figure 1. Nombre de QTL découverts chez la tomate en fonction de la part de variance prise en compte par chacun d'entre eux. Synthèse des résultats obtenus sur onze caractères quantitatifs (d'après Tanksley 1993).



exemple un allèle nul, c'est-à-dire un allèle ne permettant pas la synthèse de la protéine (Robertson 1985). On suppose alors l'existence d'allèles intermédiaires entre l'allèle normal et l'allèle muté connu, et que les variations dites qualitatives et quantitatives sont contrôlées par les mêmes loci. Quelques exemples corroborent cette hypothèse comme l'action du gène de synthèse de la gibbéréline sur la hauteur du maïs (Beavis *et al* 1991, Touzet *et al* 1995), la découverte d'allèles à différents niveaux d'expression, jusqu'à l'allèle nul, pour le gène de la caséine α_{s1} de chèvre (Martin et Grosclaude 1993), ou encore la concordance entre un QTL (*Mob2*) et le gène *Ob* dans la variation quantitative du dépôt adipeux chez la souris (Warden *et al* 1995).

L'approche « gène candidat » peut être plus particulièrement efficace lorsque le polymorphisme testé n'est pas neutre mais modifie la séquence de la protéine ou se situe dans une région promotrice ou régulatrice. La relation de causalité n'est cependant pas facile à établir. A titre d'exemple, dans la recherche du déterminisme génétique de l'hypertension artérielle, le polymorphisme observé sur le gène de la 11 β -hydroxylase (cinq mutations conduisant à cinq substitutions d'acides aminés) serait lui-même responsable de l'effet (Cicila *et al* 1993) alors que le polymorphisme du gène de la rénine, étudié depuis 1989, ne serait qu'un marqueur. En effet, les résultats d'associations entre les allèles *s* et *r* du gène de la rénine et la pression artérielle divergent selon les croisements raciaux effectués (Rapp *et al* 1994). Hormis l'existence d'interactions entre les allèles étudiés et une fraction du génome des races croisées, ces différences de résultats ne s'expliquent que par la différence de phase de liaison entre les allèles *s* et *r* et les loci des mutations causales. Ainsi, sans confirmation fonctionnelle, seule une association stable, quel que soit l'échantillon d'animaux utilisé, entre un polymorphisme de ce type et le caractère, permet de considérer ce polymorphisme comme directement responsable d'une modification du caractère.

13 / La complémentarité des deux approches

Dans le but d'identifier avec précision un gène responsable du niveau d'un caractère, les deux approches peuvent être utilisées de façon très complémentaire. L'approche QTL est susceptible de donner accès aux 95 % des gènes non encore clonés, mais on est également loin d'avoir testé l'effet du polymorphisme potentiel des 2 000 gènes candidats possibles sur leur métabolisme respectif. Les deux approches peuvent donc toutes les deux servir de point de départ, l'approche QTL fournissant un point d'ancrage par une cartographie physique, l'approche « gène candidat » par des arguments fonctionnels ou métaboliques. De nombreux caractères zootechniques, sur les espèces les plus importantes, pourraient souvent bénéficier de ces deux voies d'accès : c'est le cas, par exemple, de l'étude de l'engraissement chez le poulet

L'autre approche, plus biologique, vise à évaluer le rôle de gène(s) identifié(s) dans la variabilité du caractère considéré.

(Douaire et Sourdioux 1994). Lorsque les deux approches sont utilisées pour le même caractère, des résultats convergents tendent à confirmer la pertinence des résultats de chacune. Ainsi, la première utilisation d'une carte bovine pour la recherche de QTL sur les performances laitières a conduit à isoler un fragment comprenant les gènes des caséines (Georges *et al* 1995) par ailleurs déjà étudiés par une approche « gène candidat » (NG-Kwai-Hang *et al* 1983, Cowan *et al* 1990, Bovenhuis *et al* 1992, Cowan *et al* 1992, Bovenhuis et Weller 1994). Les deux approches peuvent également se heurter à la même difficulté, comme le montre l'étude d'un fragment du chromosome 2 du rat où une région QTL contient cinq gènes candidats (Deng *et al* 1994) dans le déterminisme de la variabilité de la pression artérielle. Dans ce cas, aucune des deux approches ne permet encore de distinguer avec certitude le ou lesquels de ces 5 gènes est impliqué.

L'approche « gène candidat » semble aussi être une étape logique entre la délimitation d'une zone génomique de taille importante (QTL) et la recherche des gènes et mutations réellement impliqués. Par exemple, l'isolement d'une zone chromosomique associée à l'hypertension artérielle et proche du gène de l'ACE (Angiotensine Converting Enzyme) (Jacob *et al* 1991) conduit maintenant à s'intéresser directement à ce gène. A l'avenir, l'approche statistique devrait donc être utilisée comme une étape préliminaire pour les espèces dont le génome a été l'objet d'un important travail de cartographie. L'isolement du QTL proche d'un gène candidat sera un argument supplémentaire à l'étude isolée de ce gène. Dans cette optique, il apparaît donc essentiel de

poursuivre les efforts de cartographie des gènes fonctionnels, la cartographie comparée permettant d'aider à localiser des QTL et des gènes candidats pour les espèces dans lesquelles la carte génique est moins avancée.

2 / Les dispositifs expérimentaux et leur exploitation

Un des paramètres conditionnant la mise en évidence des associations gène (ou marqueur)-caractère réside dans la réalisation d'un dispositif expérimental adapté. Les deux approches, QTL ou gènes candidats, utilisent le plus souvent le même type de plan d'expérience. L'approche « gène candidat » permet cependant de s'affranchir sous certaines conditions de quelques contraintes qui sont inévitables dans une approche QTL.

2.1 / Plans de croisement et analyses intra-familles

Les dispositifs expérimentaux et méthodes statistiques destinés à la recherche de QTL font l'objet de très nombreuses études depuis une quinzaine d'années. Plusieurs synthèses bibliographiques font une revue plus ou moins complète des dispositifs élaborés depuis lors (Tanksley 1993, Goffinet *et al* 1994, Kinghorn *et al* 1994, Visscher et Haley 1995). Le point essentiel de l'ensemble des dispositifs réside dans la création ou l'observation des déséquilibres de liaison entre le(s) marqueur(s) et les gènes à effet quantitatif recherchés. Le croisement de lignées pures (homozygotes) très éloignées pour un caractère (intercroisement ou rétro-croisement) est le moyen le plus efficace pour créer artificiellement un déséquilibre de liaison. La recherche végétale a pu la première exploiter ce type de dispositif. La situation est différente chez les animaux pour lesquels l'absence de lignées homozygotes est une contrainte supplémentaire. L'utilisation de lignées homogènes et présentant des performances très éloignées pour un caractère donné est cependant possible. Le programme français de recherche de QTL chez le porc -croisement Large White x Meishan (Gellin et Chevalet 1994, Bidanel *et al* 1996) illustre parfaitement l'utilisation de ces dispositifs chez l'animal. Cependant, la sélection de telles lignées à cette seule fin expérimentale s'avérerait trop longue et trop coûteuse pour la plupart des espèces.

Quoiqu'il en soit, il sera souvent difficile de se passer d'une analyse intra-population. En effet, dans l'approche réalisée par croisement, la variabilité mise en évidence est principalement de type inter-lignée (inter-race). Les allèles ayant un effet positif sur le caractère ont de grandes chances d'être déjà fixés dans les races sélectionnées actuellement exploitées (Smith et Smith 1993, Van der Beek *et al* 1995), bien que l'opportunité d'identifier des allèles favorables dans des populations non sélectionnées apparaisse pour certains caractères (reproduction, résistance aux mala-

Déséquilibre de liaison et phase de liaison

Il y a déséquilibre de liaison si la fréquence des gamètes (P_{AB}) porteurs de deux gènes A et B non allèles est différente du produit des fréquences de ces deux derniers (P_A, P_B). Ainsi, la valeur D du déséquilibre est définie par :

$$D = P_{AB} - P_A \times P_B$$

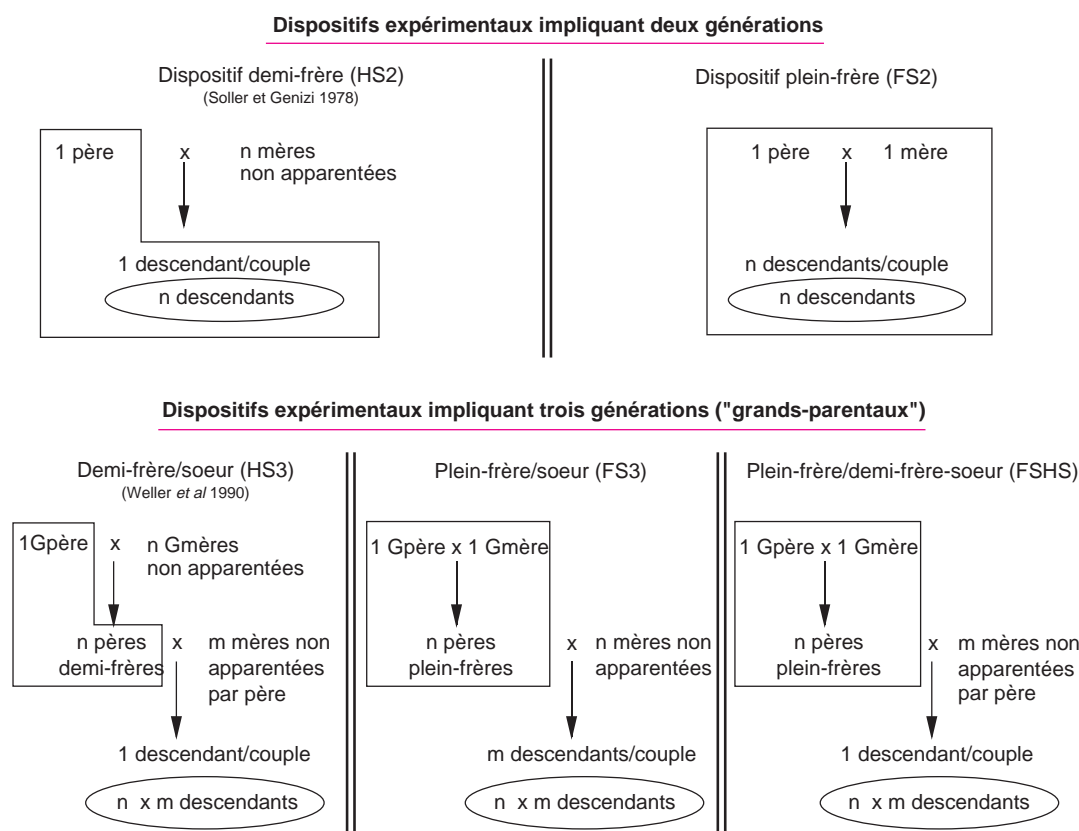
$$\text{ou également par : } D = P_{AB} \times P_{ab} - P_{Ab} \times P_{aB}$$

où $P_{AB}, P_{ab}, P_{Ab}, P_{aB}$ sont les fréquences des quatre combinaisons gamétiques.

Dans une population en équilibre de liaison, $D = 0$. Lorsque le déséquilibre est total (Dmax), il n'y a que deux types gamétiques dans la population : A est toujours associé à B et a à b, ou bien A est toujours associé à b et a à B. La phase de la liaison, qui caractérise l'association préférentielle d'un allèle du locus A avec un allèle du locus B, est identique pour tous les individus de la population. Cependant, aux cours des générations, les événements de recombinaison entre les deux loci A et B vont réduire le déséquilibre jusqu'à l'annuler, d'autant plus rapidement que la distance entre les deux gènes est grande.

Différents phénomènes peuvent être à l'origine d'un déséquilibre de liaison dans une population, comme la dérive génétique, qui se traduit par la fixation inéluctable d'allèles dans des populations de faible effectif, ou l'effet fondateur, dû au nombre réduit de reproducteurs à l'origine d'une population. La sélection, par l'utilisation d'un nombre restreint de reproducteurs, est également un facteur d'apparition de déséquilibre, comme la fusion entre deux populations si les fréquences géniques sont différentes pour les deux loci considérés. Dans une population en équilibre, un déséquilibre de liaison peut s'observer intra-famille. Ainsi, la phase de la liaison est identique pour tous les individus d'une même famille (aux événements de recombinaison près), alors qu'elle diffère entre les familles.

Figure 2. Dispositifs expérimentaux de recherche de QTL intra-lignée (d'après Van der Beek et al 1995). Ils sont classés selon le nombre de générations qu'ils nécessitent. Les abréviations HS et FS correspondent respectivement à demi-frère (half sib) et plein-frère (full sib). Les animaux dont le génotype est à déterminer sont encadrés, ceux dont le phénotype est à déterminer sont encerclés.



dies...) et pour certaines souches (Meishan...). Lorsque l'analyse s'effectue à l'intérieur d'une population en ségrégation, un déséquilibre de liaison peut être exploité intra-famille (Soller et Genizi 1978, Weller et al 1990, Luo 1993). Plusieurs dispositifs sont alors envisageables (figure 2). Certains sont plus spécialement adaptés à une production ou à une espèce, comme par exemple le dispositif « petite-fille » de Weller et al (1990), élaboré initialement pour l'étude de la production laitière chez les bovins. Il est nécessaire pour chaque espèce, en fonction de ses caractéristiques zootechniques, et pour le caractère étudié, en fonction du mode de mesure disponible, de raisonner un dispositif adapté.

2.2 / Puissance et fiabilité des analyses statistiques dans la recherche de QTL

L'identification d'un QTL se résume dans un premier temps à sa localisation et à l'estimation de son effet à partir des informations fournies par les marqueurs. L'hypothèse d'existence d'un QTL a tout d'abord été testée point par point, c'est-à-dire marqueur par marqueur (Soller et al 1976), puis une méthode de cartographie par intervalle (Lander et Botstein 1989) a permis d'augmenter la puissance des dispositifs tout en diminuant le risque de première espèce (risque de conclure à l'existence d'un QTL alors qu'il n'y en a pas). Il s'agit,

dans cette méthodologie, de tester l'hypothèse de présence d'un QTL entre deux marqueurs (dits flanquants). Son utilisation est largement facilitée par la découverte de nombreux marqueurs très polymorphes et bien répartis dans le génome (microsatellites). Plus récemment, l'utilisation de l'information fournie par l'ensemble des marqueurs présents sur un groupe de liaison a été proposée pour améliorer encore la puissance des tests effectués (Haley et al 1994, Kearsley et Hyne 1994).

Quel que soit le dispositif mis en place, les méthodes statistiques développées pour tester l'existence d'un QTL et estimer son effet sont basées soit sur l'utilisation du maximum de vraisemblance, soit sur l'utilisation des moindres carrés (régression, analyse de variance). Les différences de performances de ces deux méthodologies sont peu marquées. Les résultats obtenus sont similaires, tant sur la position du QTL que sur l'estimation de son effet (Lander et Botstein 1989, Haley et Knott 1992, Rebai et al 1995). Le léger avantage en terme de puissance accordé aux méthodes basées sur le maximum de vraisemblance semble être compensé par la plus grande simplicité des méthodes basées sur les moindres carrés et la possibilité de tester dans ce dernier cas des hypothèses complexes (présence de plusieurs QTL sur un groupe de liaison, interaction entre QTL, ...) (Goffinet et al 1994, Hyne et Kearsley 1995).

Les dispositifs expérimentaux sont basés sur la création ou la recherche de déséquilibres de liaison entre marqueur(s) et gène(s) recherché(s), c'est-à-dire d'associations préférentielles entre les allèles au locus du marqueur et les allèles du gène du caractère quantitatif.

Maximum de vraisemblance et cartographie par intervalle

La cartographie par intervalle, proposée par Lander et Botstein en 1989 est la méthodologie actuellement la plus employée pour la recherche de QTL (Quantitative Trait Loci). L'analyse statistique est basée sur une méthode du maximum de vraisemblance. L'hypothèse d'existence d'un gène à effet quantitatif entre deux marqueurs est testée par le calcul du rapport entre la probabilité d'observation de la ségrégation sous l'hypothèse de liaison entre marqueurs et QTL ($\theta < 0,5$) et la probabilité d'observation de ces mêmes données sous l'hypothèse d'absence de liaison ($\theta = 0,5$), θ étant le pourcentage de recombinaison. Pour des raisons pratiques, le \log_{10} de ce rapport de vraisemblance (« lodscore ») est en fait calculé et comparé à un seuil $\log_{10} A$ où A est voisin de $(1-\beta)/\alpha$ avec α l'erreur de type I et β l'erreur de type II. Le choix de la valeur seuil est important afin de minimiser le risque de détermination de fausses liaisons. Un seuil égal à 3 est proposé en génétique humaine, correspondant à un risque α compris entre 1 : 1 000 et 1 : 10 000. Le calcul du seuil peut être modulé en fonction de la taille du génome, de la densité des marqueurs, du dispositif expérimental et du type d'action considéré (additivité, dominance) au locus quantitatif. Cependant, la prise en compte des seules erreurs de type I n'est pas nécessairement la contrainte première à appliquer à ce type d'analyse. En effet, dans le cas où l'approche est préliminaire, ou bien lorsque l'analyse est réalisée simultanément sur plusieurs dispositifs, un risque de première espèce élevé peut être accepté pour réduire les erreurs de type II et ainsi s'assurer de la mise en évidence des principaux QTL, même si le risque d'identifier de faux positifs augmente.

2.3 / Vers des plans d'expériences communs aux approches QTL et gènes candidats

Tous les dispositifs visant à rechercher des QTL à partir d'une carte génétique, mais également certains des plans d'expérience relatifs à l'étude d'un gène candidat prennent en compte, ou créent artificiellement, un déséquilibre de liaison entre le marqueur polymorphe et le gène ou entre un site polymorphe du gène candidat et le locus de la mutation causale (de ce même gène). Seule l'hypothèse préalable que le polymorphisme étudié dans le cadre d'une approche gène candidat est directement responsable d'un effet sur le caractère, ou que la population étudiée n'est pas en équilibre de liaison, peut permettre de s'affranchir de dispositifs de croisement ou d'analyse intra-famille, relativement lourds. L'analyse statistique attribue alors un effet direct au polymorphisme testé dans des modèles simples d'analyse de variance ou de régression (cf. par exemple Nielsen *et al* 1995), ou dans des modèles prenant en compte l'origine mixte -polygénique et gène candidat - du déterminisme du caractère (modèle animal) (cf. par exemple Rothschild *et al* 1996). L'attribution d'un effet au polymorphisme est une hypothèse qui peut être considérée comme largement justifiée dans le cas de l'étude de variants protéiques, de polymorphismes d'ADN sur des zones traduites, ou lorsque des variations de taille ou de quantité d'ARNm sont connues. Dans les nombreux autres cas, l'attribution d'un effet au polymorphisme candidat testé peut être abusive. Un site polymorphe même proche (voire à l'intérieur) d'un gène connu devra être considéré comme un marqueur - marqueur de la mutation causale - et traité comme tel, dans un dispositif en croisement ou intra-famille. La simple proximité entre le site polymorphe étudié et la mutation réellement en cause n'est

pas garante de l'existence d'un déséquilibre de liaison total sur la population.

Les deux approches, dans la majorité des cas, sont donc à conduire sur des dispositifs similaires et, compte tenu de la lourdeur de ceux-ci, ils devraient être exploités par les deux approches. Ainsi, l'introduction d'un marqueur de gène candidat dans une analyse de type QTL permet d'avoir rapidement une information sur la valeur de ce candidat ; de même, l'existence d'un QTL au voisinage d'un gène candidat peut conduire à exploiter un marqueur de ce gène, peut-être plus difficilement utilisable que les marqueurs usuels de la carte, mais plus proches -donc plus efficaces - de la mutation causale.

Loin de s'opposer, les deux approches sont bien complémentaires : les résultats propres à chacune sont des aides et des contrôles de l'autre, qui doivent être utilisés pour accroître l'efficacité générale des dispositifs mis en œuvre.

Conclusion

A l'avenir, la majorité des zones chromosomiques gouvernant un caractère quantitatif devrait être découverte grâce à l'étude de la ségrégation de marqueurs équirépartis sur le génome. Des découvertes complémentaires peuvent être attendues d'une approche ciblant, sans cette étude préalable, le polymorphisme de gènes candidats, dans le cadre de fonctions physiologiques relativement bien délimitées (lipogénèse, synthèse protéique du lait ...) ou pour des espèces en retard dans l'établissement des cartes ou ne présentant pas de matériel expérimental propice à la réalisation de dispositif de croisement.

Tous les exemples aujourd'hui disponibles permettent de définir les points de stratégie essentiels pour la découverte des loci responsables d'une partie de la variabilité d'un caractère. En effet, deux étapes, qui dépassent le clivage des approches « QTL » ou « gène candidat », sont à considérer :

- démontrer la co-ségrégation d'un polymorphisme et d'un caractère, que ce soit par une approche cartographique ou fonctionnelle (mais ceci ne suffit pas à la démonstration du rôle de l'allèle étudié) ;

- démontrer l'existence d'une relation de cause à effet entre une modification de séquence d'ADN et la modification biochimique (ARN, protéine) ou physiologique responsable de l'effet. Cette étape nécessite la mise en œuvre d'analyses de type biologique pour être menée à bien. Elle est importante puisqu'elle seule permettra d'utiliser, pour l'amélioration génétique, la partie du génome directement responsable de la variabilité d'un caractère.

Remerciements

Ce travail a été financé par la société Bétina sélection et l'Association Nationale de la Recherche Technique (ANRT) dans le cadre d'une convention CIFRE, l'Ecole Nationale Supérieure Agronomique de Rennes et l'Institut National de la Recherche Agronomique.

Références bibliographiques

- Andersson L., Haley C.S., Ellegren H., Knott S.A., Johansson M., Andersson K., Andersson-Eklund L., Edfors-Lilja I., Fredholm M., Hansson I., Hakansson J., Lundström K., 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science*, 263, 1771-1774.
- Beavis W.D., Grant D., Albertsen M., Fincher R., 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theor. Appl. Genet.*, 83, 141-145.
- Bidanel J.P., Bonneau M., Chardon P., Elsen J.M., Gellin J., Le Roy P., Milan D., Ollivier L., 1996. Etablissement et utilisation de la carte génétique porcine. *INRA Prod. Anim.*, 9, 299-310.
- Bovenhuis H., Weller J.I., 1994. Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers. *Genetics*, 137, 267-280.
- Bovenhuis H., Van Arendonk J.A.M., Korver S., 1992. Associations between milk protein polymorphisms and milk production traits. *J. Dairy Sci.*, 75, 2549-2559.
- Cicila G.T., Rapp J.P., Wang J.M., Lezin E.St., Chung Ng S., Kurtz T.W., 1993. Linkage of 11 β -hydroxylase mutations with altered steroid biosynthesis and blood pressure in Dahl rat. *Nature genetics*, 3, 346-353.
- Collins F.S., 1992. Positional cloning : let's not call it reverse anymore. *Nature genetics*, 1, 3-6.
- Courreau J.F., Sellier P., Boulard J., Breton T., Goullieux P., Guérin G., 1985. Marqueurs sanguins (Phi et Pgd) et sensibilité à l'halothane chez le porc Landrace français. *Journées Rech. Porcine en France*, 17, 95-104.
- Cowan C.M., Dentine M.R., Ax R.L., Schuler L.A., 1990. Structural variation around prolactin gene linked to quantitative traits in an elite Holstein sire family. *Theor. Appl. Genet.*, 79, 577-582.
- Cowan C.M., Dentine M.R., Coyle T., 1992. Chromosome substitution effects associated with κ -casein and β -lactoglobulin in Holstein cattle. *J. Dairy Sci.*, 75, 1097-1104.
- Dalens M., Runavot J.P., 1993. Test moléculaire pour le dépistage du gène de sensibilité à l'halothane chez le porc. *Techni-porc*, 16-1-93, 17-21.
- Darvasi A., Weinreb A., Minke V., Weller J.I., Soller M., 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated map. *Genetics*, 134, 943-951.
- Deng A.Y., Dene H., Rapp J.P., 1994. Mapping of a quantitative trait locus for blood pressure on rat chromosome 2. *J. Clin. Invest.*, 94, 431-436.
- Douaire M., Sourdioux M., 1994. Markers for fat deposit. *Proc. 9th European Poultry Conf. Aug. 7-12*, 2, 364-367.
- Gellin J., Chevalet C., 1994. Stratégie d'établissement des cartes géniques. Exemple du porc. *Genet. Sel. Evol.*, 26, Suppl. 1, 35-51.
- Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., Pasquino A.T., Sargeant L.S., Sorensen A., Steele M.R., Zhao X., Womack J.E., Hoeschele I., 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics*, 139, 907-920.
- Goffinet B., Beckmann J., Boichard D., Causse M., Charcosset A., Chevalet C., Christophe C., Colleau J.J., Demenais F., Durel C.E., Elsen J.M., Foulley J.L., Gallais A., Götz K., Hospital F., Kremer A., Lorieux M., Lefort-Buson M., Le Roy P., Loisel P., Mangin B., Maurice A., Perrier X., Pons O., Rebaï A., Rodolphe F., San Cristobal M., Vu Tien Khang J., 1994. Méthodes mathématiques pour l'étude des gènes contrôlant des caractères quantitatifs. *Genet. Sel. Evol.*, 26, Suppl. 1, 95-105.
- Haley C.S., Knott S.A., 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, 69, 315-324.
- Haley C.S., Knott S.A., Elsen J.M., 1994. Mapping quantitative trait loci in crosses between outbred lines using least square. *Genetics*, 136, 1195-1207.
- Hanset R., Grobet L., 1992. Sélection assistée par marqueur. *Bull. GTV*, 92-5, 71-75.
- Hyne V., Kearsey M.J., 1995. QTL analysis : further uses of "marker regression". *Theor. Appl. Genet.*, 91, 471-476.
- Hyne V., Kearsey M.J., Pike D.J., Snape J.W., 1995. QTL analysis : unreliability and bias in estimation procedure. *Molecular Breeding*, 1, 273-282.
- Jacob H.J., Lindpaintner K., Lincoln S.E., Kusumi K., Bunker R.K., Mao Y.P., Ganten D., Dzau J., Lander E.S., 1991. Genetic mapping of a gene causing hypertension in the Stroke-Prone Spontaneously Hypertensive rat. *Cell*, 67, 213-224.
- Kearsey M.J., Hyne V., 1994. QTL analysis : a simple 'marker-regression' approach. *Theor. Appl. Genet.*, 89, 698-702.
- Kinghorn B., Van Arendonk J.A.M., Hetzel J., 1994. Detection and use of major genes in animal breeding. *Agbiotech. News and Info.*, 6, 297N-302N.
- Lander E.S., Botstein D., 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121, 185-199.
- Luo Z.W., 1993. The power of two experimental designs for detecting linkage between a marker locus and a locus affecting a quantitative character in a segregating population. *Genet. Sel. Evol.*, 25, 249-261.
- Manfredi E., Ricordeau G., Barbieri M.E., Amigues Y., Bibé B., 1995. Génotype caséine α_{s1} et sélection des boucs sur descendance dans les races Alpine et Saanen. *Genet. Sel. Evol.*, 27, 451-458.
- Mangin B., Goffinet B., Rebaï A., 1994. Constructing confidence intervals for QTL location. *Genetics*, 138, 1301-1308.
- Martin P., Grosclaude F., 1993. Improvement of milk protein quality by gene technology. *Livest. Prod. Sci.*, 35, 95-115.

- N-G-Kwai-Hang K.F., Hayes J.F., Moxley J.E., Monardes H.G., 1983. Association of genetic variants of casein and milk serum proteins with milk, fat, and protein production by dairy cattle. *J. Dairy Sci.*, 67, 835-840.
- Nielsen V.H., Larsen N.J., Agergaard N., 1995. Association of DNA polymorphism in the growth-hormone gene with basal-plasma growth-hormone concentration and production traits in pigs. *J. Anim. Breed. Genet.*, 112, 205-212.
- Paterson A.H., Lander E.S., Hewitt J.D., Peterson S., Lincoln S.E., Tanksley S.D., 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature*, 335, 521-529.
- Rapp J.P., Dene H., Deng A.Y., 1994. Seven renin alleles in rats and their effects on blood pressure. *J. Hypertens.*, 12, 349-355.
- Rebai A., Goffinet B., Mangin B., 1995. Comparing power of different methods for QTL detection. *Biometrics*, 51, 87-99.
- Robertson D.S., 1985. A possible technique for isolating genic DNA for quantitative traits in plants. *J. Theor. Biol.*, 117, 1-10.
- Rothschild M., Jacobson C., Vaske D., Tuggle C., Wang L., Short T., Eckardt G., Sasaki S., Vincent A., McLaren D., Southwood O., Van der Steen H., Mileham A., Plastow G., 1996. The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc. Natl. Acad. Sci.*, 93, 201-205.
- Ruane J., Colleau J.J., 1995. Marker assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genet. Res., Camb.*, 66, 71-83.
- Shuster D.E., Kehrli M.E.Jr., Ackermann M.R., Gilbert R.O., 1992. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. *Proc. Natl. Acad. Sci.*, 89, 9225-9229.
- Smith C., Smith D.B., 1993. The need for close linkages in marker-assisted selection for economic merit in livestock. *Anim. Breed. Abstr.*, 61, 197-204.
- Soller M., 1994. Marker assisted selection - an overview. *Anim. Biotech.*, 5, 193-207.
- Soller M., Genizi A., 1978. The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating population. *Biometrics*, 34, 47-55.
- Soller M., Brody T., Genizi A., 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Appl. Genet.*, 47, 35-39.
- Tanksley S.D., 1993. Mapping polygenes. *Ann. Rev. Genet.*, 27, 205-233.
- Touzet P., Winkler R.G., Helentjaris T., 1995. Combined genetic and physiological analysis of a locus contributing to quantitative variation. *Theor. Appl. Genet.*, 91, 200-205.
- Van der Beek S., Van Arendonk J.A.M., Groen A.F., 1995. Power of two- and three-generation QTL mapping experiments in an outbred population containing full-sib or half-sib families. *Theor. Appl. Genet.*, 91, 1115-1124.
- Visscher P.M., Haley C.S., 1995. Utilizing genetic markers in pig breeding programmes. *Anim. Breed. Abstr.*, 63, 1-8.
- Warden C.H., Fisler J.S., Shoemaker S.M., Wenn P.Z., Svenson K.L., Pace M.J., Lusk A.J., 1995. Identification of four chromosomal loci determining obesity in a multifactorial mouse model. *J. Clin. Invest.*, 95, 1545-1552.
- Weller J.I., Fernando R.L., 1991. Genomic genetics and plant genetic improvement. In : L.B. Schook, H.A. Lewin and D.C. McLaren (eds), *Gene mapping techniques and applications*, 201-230. M. Dekkers, New York.
- Weller J.I., Kashi Y., Soller M., 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.*, 73, 2525-2537.
- Zhang Y., Proenca R., Maffei M., Barone M., Leopold L., Friedman J.M., 1994. Positional cloning of the mouse obese gene and its human homologue. *Nature*, 372, 425-432.

Abstract

Genetic analysis of a quantitative trait.

For several years, many studies have been undertaken in order to identify the genes which actually control the variability of quantitative traits. Mainly two strategies have been developed for this purpose. One relies on the genetic map knowledge and searches for the so-called QTL (quantitative trait loci). It gives rise to an early application to selection process by the use of markers associated to the trait ; it is also a first step towards the location of genes actually involved in part of the trait variability determinism. The

other one is supported by the biological knowledge of the trait and focus on few genes -candidate genes - supposed to be responsible for part of the trait variability. At the moment the analysis of these methods and their results allow to make out their abilities and limits. Moreover, it brings to the fore their complementary nature in the genetic analysis of quantitative traits, and their similarities in part of the experimental designs.

Sourdioux M., Lagarrigue S., Douaire M., 1997. Analyse génétique d'un caractère quantitatif. *INRA Prod. Anim.*, 10 (3), 251-258.