

# Nouvelles approches de phénotypage pour la sélection animale

S. LAGARRIGUE<sup>1,2</sup>, M. TIXIER-BOICHARD<sup>3,4</sup>

<sup>1</sup> Agrocampus Ouest, Génétique Animale, F-35000 Rennes, France

<sup>2</sup> INRA, UMR598 Génétique Animale, F-35000 Rennes, France

<sup>3</sup> INRA, UMR1313 Génétique Animale et Biologie Intégrative, F-78352 Jouy-en-Josas, France

<sup>4</sup> AgroParisTech, Génétique Animale et Biologie Intégrative, 16 rue Claude Bernard, F-75231 Paris, France

Courriel : Sandrine.lagarrigue@agrocampus-ouest.fr

L'évolution des méthodes de sélection n'est pas sans conséquence sur les besoins en matière de phénotypage ; de même les évolutions technologiques donnent accès à des phénotypes d'un genre nouveau. Dans ce contexte, nous aborderons trois volets mettant en lien phénotypage et sélection animale : l'impact de la sélection génomique sur le phénotypage, puis l'impact sur la sélection du phénotypage automatisé et/ou à haut débit, et enfin les méthodes permettant de décomposer un phénotype complexe d'intérêt socioéconomique en phénotypes élémentaires pour mieux en comprendre les processus biologiques sous-jacents.

En génétique, le concept de «phénotype» est la résultante des effets combinés du génotype et de l'environnement. En référence à la formule  $P=G+E$  bien connue en génétique, un phénotype (P) observé chez un individu donné est le fruit de la composante génétique (G) de cet individu et des effets d'environnement auquel ce dernier est soumis.

Le «phénotype» correspond aux valeurs mesurables prises par des caractères qui peuvent être très différents, de par leur niveau d'observation et les finalités de ces observations (figure 1). Nous distinguerons ainsi deux types de phénotypes (même si ces deux catégories ne sont pas totalement disjointes) :

- les phénotypes définis par leur importance socio-économique pour au moins un acteur d'une filière (éleveur, abatteur, transformateur, distributeur, consommateur, citoyen...). Les caractères déterminant ces phénotypes peuvent devenir des critères de sélection, sous réserve 1) d'être héréditaires (*i.e.* les variations des mesures (P) sont en partie dues à des variations des génotypes (G) transmises à la descendance), 2) d'être mesurables sur un grand nombre d'animaux avec un coût de mesure raisonnable. Ces caractères sont mesurés au niveau de l'animal vivant ou de ses produits plus ou moins transformés. Quelques exemples : l'efficacité alimentaire d'un animal, la fertilité d'une vache, le taux de viande vs de gras d'une carcasse de porc (rendement en viande), la solidité de la coquille d'œuf,

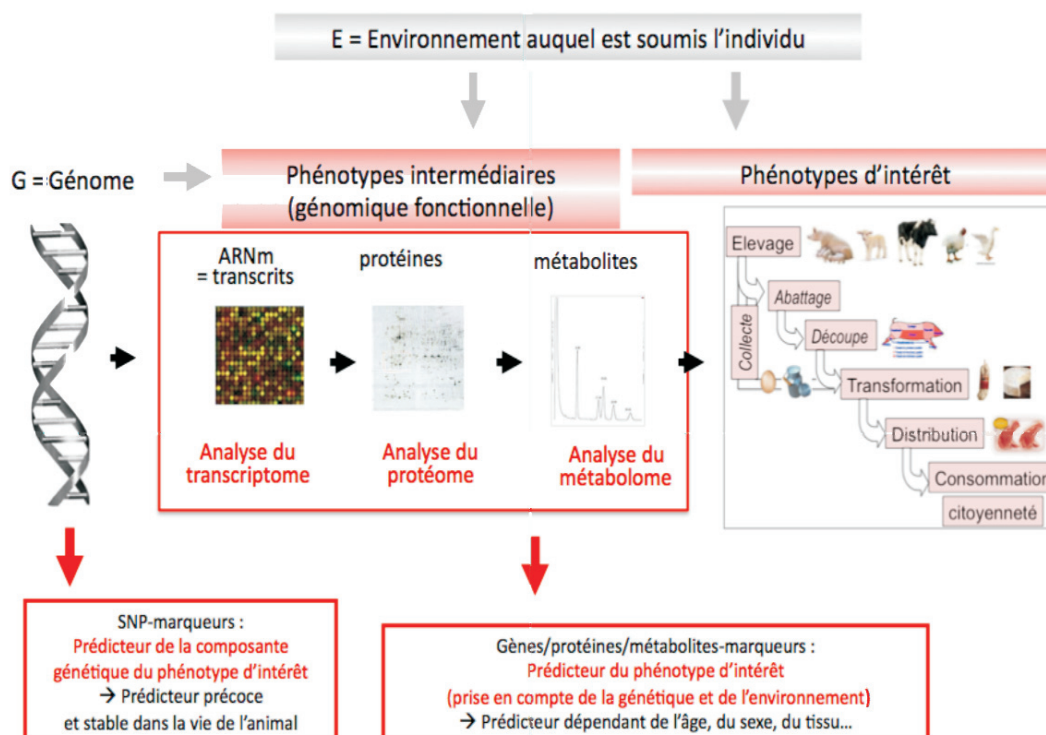
la tendreté de la viande, l'aptitude au tranchage d'un jambon, etc. Ces phénotypes complexes sont le fruit de l'action de nombreux processus biologiques ;

- les phénotypes plus élémentaires généralement observés avec la finalité de comprendre les mécanismes biologiques sous-jacents aux phénotypes plus complexes cités précédemment. Ces phénotypes correspondent à l'expression de gènes mesurée à travers la quantité de leurs transcrits, de leurs protéines présentes dans un tissu donné. Les technologies haut débit en «ome» permettent maintenant d'étudier l'expression de l'ensemble des gènes d'un génome : on parle de transcriptome ou de protéome selon que des transcrits ou des protéines d'un tissu sont quantifiés. Du fait de cette exhaustivité d'observation, ces phénotypes sont utilisés à des fins cognitives pour disséquer, comprendre sans a priori la mise en place ou la variabilité de caractères plus «intégrés» observés au niveau d'un tissu ou d'un animal. Des phénotypes anatomiques ou métaboliques (paramètres sanguins, hormones, métabolites) peuvent être également exploités à ces fins. La combinaison de ces phénotypes élémentaires permet d'avoir une vision exhaustive de l'animal, quasi clinique par analogie avec la santé humaine. Elle offre de nouvelles perspectives en biologie animale avec, entre autres, l'objectif de mieux comprendre les phénotypes d'intérêt socio-économique et ainsi de mieux les maîtriser par des leviers génétiques (sélection) ou autres

(alimentation notamment). Ces phénotypes élémentaires peuvent alors devenir un critère de sélection.

Les besoins en matière de phénotypage sont fortement impactés par l'évolution des méthodes de sélection et des systèmes d'élevage, avec la définition de nouveaux objectifs de sélection. Nous présenterons d'abord l'impact de la sélection génomique. La sélection génomique a été possible grâce à la disponibilité d'un grand nombre de marqueurs moléculaires et à la diminution importante des coûts de génotypage. Dans ce contexte nouveau, le phénotypage peut devenir un facteur limitant face aux potentialités offertes par la sélection génomique, en particulier en termes de nouveaux objectifs de sélection. En effet, les systèmes d'élevage évoluent pour répondre à de nouveaux enjeux : respect de l'environnement, santé et bien-être des animaux, robustesse ou adaptabilité de ces derniers à des conditions changeantes. La mesure des performances et donc le phénotypage doivent s'adapter à ces nouvelles conditions. Ces phénotypes peuvent également être utilisés à des fins d'analyse de la biodiversité, par exemple pour caractériser l'adaptation de races locales à un milieu particulier.

Nous verrons ensuite comment des développements technologiques contribuent à faire évoluer les méthodes du phénotypage vers une approche à plus haut débit, standardisée et automatisable.

**Figure 1.** Phénotypes : une variété de caractères qui diffèrent par leur niveau d'observation.

Nous analyserons l'apport du concept des biomarqueurs, issu de la médecine humaine, avant d'aborder les méthodes permettant de décomposer un phénotype complexe en phénotypes élémentaires, afin de relier le génotype et les caractères d'intérêt socioéconomique. De nombreuses recherches en génétique animale ont pour objet de localiser les régions chromosomiques contrôlant la variabilité de caractères d'intérêt (*Quantitative Trait Loci*, QTL), et la difficulté consiste à identifier le ou les gènes responsables de l'effet QTL d'une région. Ainsi, dans les années 2000, ont été développées de nouvelles stratégies dites de «génétique-génomique», combinant des phénotypes élémentaires et des méthodes de cartographie (cf. Le Mignon *et al* 2010). Ces stratégies visent à faciliter l'identification des mutations causales sous-jacentes à des QTL et surtout à mieux comprendre les processus biologiques impactés par ces mutations et responsables de ces caractères, pouvant ainsi ouvrir de nouvelles voies de maîtrise de ces caractères autres que la génétique.

## 1 / Impact de la sélection génomique sur le phénotypage

Classiquement, la sélection animale consiste à sélectionner des candidats sur la base de leur valeur génétique estimée à partir des mesures de leur phénotype et des phénotypes de leurs apparentés.

L'accès à des phénotypes individuels et aux généalogies est donc nécessaire sur l'ensemble de la population en sélection. Les paramètres déterminant le progrès génétique annuel sont reliés par une formule mathématique générale :  $\Delta G = I \cdot R \cdot \sigma_g / T$ , où  $I$  est l'intensité de sélection (taux de retenus parmi les candidats),  $R$  la précision de la sélection (coefficient de corrélation entre le prédicteur phénotypique et la valeur génétique),  $\sigma_g$  l'écart-type génétique du caractère, et  $T$  l'intervalle de génération. Cette formule se traduit par  $\Delta G = I \cdot h^2 \cdot \sigma_p / T$  en cas de sélection massive où seul le phénotype du candidat est utilisé pour prédire sa valeur génétique ( $h^2$  est l'héritabilité du caractère).

Avec la sélection génomique, la valeur génétique des candidats est estimée par génotypage de milliers de marqueurs couvrant le génome. La correspondance entre les génotypes aux marqueurs et le phénotype est établie au préalable par la constitution d'une population de référence de quelques milliers d'individus, qui sont à la fois phénotypés pour le phénotype d'intérêt et génotypés pour des milliers de marqueurs.

Par rapport à la sélection classique, la sélection génomique présente plusieurs avantages : précocité de l'évaluation génétique et accélération du progrès génétique, précision de l'évaluation génétique quasi identique quel que soit le sexe ou le caractère, augmentation de l'intensité de sélection qui ne dépend

que du coût de génotypage à mettre en regard avec la taille de la population disponible et le coût du reproducteur, déconnexion entre la mesure du phénotype et la sélection des candidats.

L'impact de la sélection génomique sur ces différents facteurs offre de nouvelles perspectives en matière de phénotypage, avec des situations variables selon les espèces et les structures de production.

### 1.1 / Nouvelles perspectives pour les phénotypes femelles : de la sélection vers le conseil génétique

La précocité de l'évaluation génétique permet de réduire l'intervalle de génération. Actuellement cet intervalle est élevé dans les espèces à cycle de reproduction long (cas des ruminants) ou encore pour des caractères mesurés tardivement dans la vie de l'animal (caractères de reproduction) et/ou mesurés uniquement chez les femelles. Dans ce dernier cas, il faut procéder à une évaluation des mâles sur leur descendance et/ou leurs sœurs (bovins laitiers, poules pondeuses). Dans les espèces peu prolifiques, soulignons que l'évaluation génétique des femelles est habituellement très peu précise (du fait d'un faible nombre d'apparentés disponibles) ; cette faible précision est d'autant plus marquée que le caractère est peu héritable (cas de la fertilité). De plus, la précision accrue de l'évaluation génétique sur les femelles ouvre de nou-

velles marges de progrès génétique sur les caractères peu héréditaires.

Ces nouvelles marges de progrès génétique conduisent à repenser les objectifs de sélection actuels. Par exemple, une quinzaine de caractères sont évalués chez les bovins laitiers que l'on peut décomposer en caractères de production (quantité de lait, taux protéique et butyreux...) et en caractères fonctionnels (conformation de la mamelle, longévité, fertilité, résistance aux mammites...). La production laitière connaît un progrès génétique annuel d'environ 100 kg/an/vache (Institut de l'Élevage 2011). L'augmenter davantage n'est pas forcément pertinent, étant donné la dégradation de certains caractères fonctionnels, parfois peu héréditaires, comme la fertilité, et en corrélation génétique négative avec les caractères de production. Un poids plus important peut alors être donné aux caractères fonctionnels. Pour ces caractères, une évaluation précise des femelles devient possible. Elle permettra une meilleure sélection des mères à génisses de renouvellement, qui représentaient un maillon faible dans les programmes de sélection des ruminants. Ces marges de progrès génétiques annuels, dues en particulier au raccourcissement du délai entre la définition des objectifs de sélection et leur amélioration effective sur le terrain, offrent des perspectives inédites pour introduire de nouveaux phénotypes d'intérêt et peut favoriser la prise de risques.

Au-delà de la sélection, si la baisse des coûts de génotypage le permet, une évaluation génétique quasi systématique des femelles en production ouvre des perspectives dans le domaine du conseil en élevage. Par exemple, savoir que certaines femelles ont un terrain génétique favorable aux mammites peut être utile pour la gestion du troupeau. Disposer d'un outil de gestion du troupeau *via* le diagnostic individuel invite à renouveler la réflexion des phénotypes à prendre en compte.

## 1.2 / Intérêts du phénotypage pour la sélection génomique : le concept de population de référence

L'utilisation d'une population de référence pour établir les relations «marqueurs génétiques-phénotypes» ouvre de nouvelles perspectives pour les phénotypes difficilement mesurables en routine et pour la prise en compte des interactions génotype x milieu. La seule condition est que la population de référence soit représentative de la population en sélection.

La sélection classique sur des caractères devant être mesurés sur des apparen-

tés tels que la qualité de viande du porc ou du poulet est freinée par des précisions modestes et des coûts élevés, ce qui diminue fortement l'intensité de sélection : peu de candidats à la sélection peuvent être évalués. La sélection génomique présente donc un intérêt évident pour ces caractères.

La taille de la population de référence est un facteur essentiel de la précision des évaluations génétiques des individus en sélection : le phénotypage de quelques milliers d'individus est en général nécessaire et suffisant. Cela permet d'envisager la mesure de caractères relativement coûteux sous réserve qu'ils apparaissent importants pour la filière ; les relations «marqueurs-phénotype» établies permettront ensuite, lors de l'étape de sélection, d'évaluer des dizaines de milliers d'individus sur leur seul génotype. Cette évaluation à moindre coût permet donc au final d'avoir une intensité de sélection forte et donc des gains de progrès génétique accrus. Pour des caractères d'importance majeure, on peut envisager de sacrifier les individus de la population de référence : c'est le cas des caractères nécessitant l'abattage des animaux, comme ceux relatifs à la qualité de la viande, et des caractères nouveaux, comme la résistance à une maladie, qui nécessite la prise de mesures par inoculation expérimentale dans des bâtiments protégés. Le coût d'investissement peut alors être amorti lors de l'étape de sélection avec la diffusion de mâles améliorateurs mais également avec l'utilisation de tests génétiques optimisant la sélection des femelles de renouvellement (nombreuses chez les bovins) ou la conduite d'élevage. Ces possibilités de retour sur investissement sont une nouveauté majeure de la sélection génomique, qui peut motiver des structures financièrement solides et prêtes à investir sur ce marché de la génétique animale.

Ainsi, la réflexion que doivent mener les structures de sélection sur les phénotypes à cibler est un enjeu majeur aujourd'hui pour leur maintien et leur développement dans ce domaine ultra compétitif de la sélection et, au-delà, du diagnostic moléculaire en élevage qui va probablement se développer. Nous constatons déjà l'entrée de nouveaux acteurs hors du monde agricole comme la compagnie pharmaceutique Pfizer qui propose aujourd'hui des «*Molecular Value Predictions* (MVPs®)» pour différents caractères : efficacité alimentaire, persillé et tendreté de la viande. On peut s'attendre à ce que cette compagnie, qui dispose d'une cellule «santé animale» dans laquelle se trouve maintenant une cellule «génétique animale», augmente sa gamme de produits.

Dans le cas des schémas de sélection pyramidaux, l'animal en production est issu d'un croisement, alors que le parent sélectionné est issu d'une lignée ou d'une race. L'idée serait d'utiliser comme population de référence la population en production, qui présente l'avantage d'être élevée dans les conditions de milieu réelles de production, qui sont variables, et non dans les conditions standardisées des élevages de sélection. Le principe est alors de phénotyper ces individus croisés ou les produits animaux associés et de les génotyper (l'ADN étant présent dans tout échantillon biologique, ou presque) pour établir les relations «marqueurs génétiques-phénotype». Si la densité de marqueurs est suffisante et que la population de référence est importante, ces relations peuvent alors en théorie être utilisées au niveau de l'étage de sélection sans que l'on ait recours à l'information généalogique. Une telle sélection devrait en partie supprimer les interactions génotype x environnement. Par ailleurs, pour les caractères classiquement enregistrés dans la chaîne de production, comme la qualité des carcasses de porc, le facteur limitant ne serait plus le phénotypage, mais plutôt le génotypage.

Ainsi, dans le cadre des programmes de sélection pyramidaux, la sélection génomique ouvre des perspectives nouvelles concernant les caractères observables sur les chaînes de production et de transformation, ce qui peut apporter un avantage aux sociétés qui ont intégré à la fois l'aval et l'amont d'une filière. Ces perspectives sont encore purement théoriques. Un projet pilote a été financé en 2010 par l'Agence Nationale de la Recherche (ANR) pour évaluer l'intérêt d'établir les relations «marqueurs-phénotype» sur des individus croisés *versus* des lignées pures chez le porc et la poule pondeuse.

## 2 / Nouvelles technologies destinées à l'automatisation du phénotypage

Que ce soit en sélection classique ou en sélection génomique, certains caractères sont encore difficiles à phénotyper de façon standardisée, répétable et rapide ; citons l'efficacité alimentaire, les caractères de résistance aux maladies et de comportement, des caractères liés à la qualité nutritionnelle des produits ou encore à la qualité organoleptique de la viande qui nécessite un jury de dégustation pour les évaluer. Pour ces caractères aujourd'hui difficiles d'accès, voire pour de nouveaux caractères qui présenteraient les mêmes freins, innover dans les technologies de mesure ou identifier



des prédicteurs plus faciles à mesurer sont des enjeux importants. L'exploration sans a priori des phénotypes élémentaires constitue une voie pour identifier de tels prédicteurs.

Le phénotype quel qu'il soit nécessite un protocole de mesure précis, permettant la transparence et la reproductibilité de la collecte de l'information. Le programme INRA «ATOL» sur l'ontologie des phénotypes animaux vise à standardiser la nomenclature des caractères et la description des protocoles de mesure (Hurtaud *et al* 2011). En effet, de nombreux protocoles plus ou moins lourds sont réalisés dans les unités expérimentales ou en ferme et produisent des données de qualité mais difficiles à regrouper si les conditions de mesure n'ont pas été harmonisées. Cette hétérogénéité limite considérablement la valorisation de ces données et empêche le passage au haut débit. La définition de protocoles de référence est donc une étape vers le haut débit. Il faut ensuite pouvoir automatiser l'acquisition de données. Deux grandes voies de standardisation peuvent être envisagées : d'une part l'utilisation de l'électronique et de la robotisation pour les phénotypes complexes sur l'animal vivant, d'autre part les technologies en «omes» également automatisées pour les phénotypes élémentaires.

## 2.1 / Électronique et robotisation

L'automatisation suppose d'investir pour développer des prototypes et d'utiliser l'identification électronique des animaux. L'objectif est de réaliser des mesures individuelles en grand nombre sur des animaux élevés en groupes (phénotypage dit «horizontal»). Un exemple typique est fourni par l'évolution de l'élevage des poules pondeuses vers des systèmes au sol ou encore par la nécessité de mesurer la consommation individuelle d'un poulet de chair au sol. La corrélation entre des performances mesurées en cage ou au sol peut être faible, en particulier pour l'efficacité alimentaire du poulet (N'Dri *et al* 2007). Des distributeurs automatiques d'aliments existent déjà pour certaines espèces et sont utilisés dans des stations expérimentales ou de testage et des automates sont en cours de développement pour les oiseaux. En revanche, les nids électroniques, permettant d'enregistrer le numéro de la poule, l'heure de la ponte et pourquoi pas le poids de l'œuf pondue, sont en cours d'expérimentation en France et ont déjà fait l'objet d'essais en Allemagne depuis 2006. Certains sélectionneurs de poulets de chair annoncent qu'ils utilisent déjà des balances réparties dans la cellule d'élevage, sur lesquelles les animaux vont se percher et se peser ainsi tout

seuls. Le Syndicat des Sélectionneurs Avicoles et Aquacoles Français (SYSAAF) réalise des essais d'identification électronique sur plusieurs espèces de volailles, avec des résultats encore hétérogènes en fonction des espèces et des milieux d'élevage. Le taux de perte et la fiabilité de lecture sont des paramètres déterminants.

À côté de ces automates, on dispose aussi d'appareils permettant de mesurer à distance la température corporelle d'un animal et d'enregistrer les déplacements d'un animal dans un groupe, ce qui permet d'envisager un phénotypage du comportement. L'intérêt de ces mesures est aussi de supprimer la manipulation des animaux, source de stress pour l'animal et d'accidents pour les animaliers. L'animal idéal serait alors celui qui se phénotype tout seul. Le développement du phénotypage électronique suppose des investissements importants qu'il faut amortir sur un grand nombre de mesures, par exemple dans le cas de la sélection génomique. C'est un choix stratégique auquel il convient de réfléchir dans le cadre des stations expérimentales ou des élevages dédiés, sous contrat ou privés.

## 2.2 / Technologies en «omes»

Grâce à l'émergence des technologies de génomique fonctionnelle, des phénotypes d'un genre nouveau deviennent accessibles (figure 1). Comme ces technologies permettent d'accéder aux mécanismes sous-jacents, on parle aussi de phénotypage vertical. Parce que ces technologies sont automatisées, les coûts de mesure deviennent raisonnables.

Dans le domaine du transcriptome, les méthodologies d'estimation de la quantité de transcrits dans un tissu donné sont en pleine évolution. Les méthodes de séquençage à haut débit vont dans un avenir très proche remplacer les méthodes originelles basées sur le principe de l'hybridation entre les ARN d'un tissu et un support miniaturisé (appelé «puce à ADN») contenant la collection des gènes de l'espèce.

Contrairement aux méthodes actuelles basées sur un processus multi-étapes non automatisable, les méthodes de séquençage offrent de nouvelles perspectives, en fournissant des données moins bruitées et potentiellement en plus grand nombre, du fait de l'automatisation presque complète du processus.

Parce qu'il est possible d'amplifier les transcrits pour obtenir en amont des quantités suffisantes en vue de leur étude, la microdissection laser est un raffinement technologique qui permet

de travailler sur un type cellulaire au sein d'un tissu composite. Bien que difficile à mettre en place et encore peu développée, cette approche permet d'analyser de façon beaucoup plus ciblée et donc précise l'effet d'une mutation ou d'un facteur de milieu sur l'expression des gènes dans un tissu (Bevilacqua *et al* 2010).

Dans le domaine de la protéomique, les méthodes d'analyse de l'empreinte spectrale (*Mass Spectral Fingerprinting*) gagnent en précision et en résolution et sont utilisées en médecine humaine sur les fluides biologiques comme les expectorations, afin de diagnostiquer un cancer du poumon (Lewis *et al* 2010).

L'analyse spectrale permet aussi de caractériser les phénotypes fins que sont les métabolites. Une illustration en est donnée par le programme PhénoFinLait, lancé en 2008 par les différents acteurs des filières laitières bovines, caprines et ovines ([www.phenofinlait.fr](http://www.phenofinlait.fr)). Ce programme vise à étudier la composition fine du lait et ses facteurs de variation (génétique, alimentation...). Techniquement, l'estimation en routine de la composition en acides gras et de certaines protéines majeures du lait est aujourd'hui possible grâce à la mise au point de méthodes fondées sur les spectres Moyen Infra-Rouge (MIR) disponibles dans les laboratoires du Contrôle Laitier. La finalité de ce programme est de mettre en relation la composition fine du lait, nouvellement accessible sur de grands effectifs, avec d'une part les conditions de milieu (alimentation des animaux en particulier) et d'autre part le génotype des animaux, dans le but d'établir les relations «marqueurs-phénotype» relatives aux différents acides gras et protéines majeures du lait. Ce programme nécessite donc une collecte massive de données et d'échantillons dans les élevages et de données spectrales dans les laboratoires, sur un dispositif de quelques milliers d'individus pour chaque espèce. Les résultats obtenus devraient permettre une maîtrise de ces nouveaux caractères par les leviers que sont la sélection et l'alimentation, et ce dans les trois espèces considérées.

La standardisation des méthodes d'analyse des ARN (transcriptome), protéines (protéomique) ou de métabolites (métabolomique) ouvre ainsi des perspectives importantes pour l'étude des fluides biologiques, qui a le grand avantage d'être peu invasive et de permettre un suivi régulier des performances au cours de la vie de l'animal. Les progrès dans ce domaine seront sans doute apportés par le développement des micro-méthodes.

### 3 / Application du concept de biomarqueurs à la prédiction du phénotype

Depuis le développement des approches de génomique fonctionnelle, de nouveaux programmes de recherche ont été mis en œuvre. Ils visent à identifier quelques transcrits, protéines et ou métabolites dont la quantité est un bon prédicteur de la mesure d'un phénotype d'intérêt difficilement mesurable.

L'expression consacrée est celle de «biomarqueurs» d'un phénotype donné. En quelque sorte, le biomarqueur (ou phénotype moléculaire) prédit le phénotype observable sur l'animal tout comme les SNP-marqueurs prédisent la valeur génétique d'un caractère d'un animal (figure 1). Soulignons que le pouvoir prédictif d'un biomarqueur, au sens statistique du terme, est considéré comme fiable si le biomarqueur a été validé dans une population indépendante de celle dans laquelle il a été trouvé. L'identification de tels phénotypes est en plein essor en génétique humaine. Un biomarqueur peut permettre de diagnostiquer une maladie complexe («marqueurs-diagnostic»), de prédire l'évolution d'une maladie chez un patient («marqueurs-pronostic») ou encore de prédire la réponse de ce patient à une thérapie donnée («marqueurs-thérapie prédictive») (Rizzi *et al* 2008, Weichselbaum *et al* 2008). Des programmes similaires peuvent être conduits en génétique animale pour obtenir des phénotypes moléculaires utilisés comme prédictifs d'un caractère d'intérêt jusqu'alors difficilement mesurable. Ainsi, des protéines sont proposées comme des biomarqueurs de la tendreté de la viande chez les bovins (Guillemin *et al* 2011).

Comme il est indiqué dans la figure 1, un biomarqueur, qu'il soit un transcrit, une protéine ou un métabolite est la résultante de l'action du génotype de l'individu et de l'environnement auquel il est soumis. Il permet donc de prédire un phénotype complexe d'un individu, alors que les marqueurs génétiques (SNP-marqueurs) permettent de prédire uniquement la composante génétique d'un phénotype. Différentes retombées de ces biomarqueurs peuvent alors être envisagées en sélection et au-delà, en conseil ou encore pour la caractérisation de l'adaptation des races locales.

#### 3.1 / Sélection classique des reproducteurs

Une sélection classique de reproducteurs améliorateurs peut être envisagée

à partir de la mesure d'un biomarqueur sur chacun des reproducteurs et leurs apparentés. Comme pour tout critère de sélection, cette mesure devra être validée par l'estimation de son héritabilité et de sa corrélation génétique avec l'objectif de sélection, avant d'être choisie. Chez la poule pondeuse par exemple, une mutation est connue pour produire une odeur désagréable dans les œufs en fonction de la teneur en choline et sinapine de l'aliment, très liée à la présence de colza. Pour continuer d'utiliser du colza dans la ration et éviter toute odeur dans les œufs, il est nécessaire de sélectionner dans les lignées commerciales les animaux contre cette mutation. À défaut de pouvoir utiliser un test génétique pour cela, il est possible de détecter cette odeur grâce à un dosage de la substance biochimique qui en est responsable. Le dosage se fait dans les œufs pour les femelles et dans les fèces pour les mâles, après les avoir soumis quelques jours à une alimentation riche en choline.

L'utilisation de tels gènes ou métabolites rencontre toutefois des limites que l'on retrouve également pour les caractères classiques, et qui sont dues à l'expression temporelle et spatiale d'un phénotype quel qu'il soit (figure 1). Ainsi, cette expression peut être tardive dans la vie de l'animal, retardant d'autant l'évaluation génétique, ou encore être sexo-dépendante, diminuant la précision de l'évaluation génétique. Cette expression peut également être spécifique d'un tissu, ce dernier pouvant parfois s'avérer difficilement accessible, compromettant alors la possibilité de mesure.

#### 3.2 / Sélection génomique des reproducteurs

Les biomarqueurs peuvent également être utiles dans les programmes de sélection génomique. Ils peuvent être utilisés de manière transitoire comme prédictifs de la mesure du phénotype à un coût raisonnable permettant un phénotypage sur un très grand nombre d'individus. Imaginons que nous ayons identifié des biomarqueurs pour la résistance à une maladie précise. De tels gènes permettraient donc de phénotyper une population de référence de plus grande taille que ne le permettrait une mesure classique par inoculation expérimentale.

Par ailleurs, l'identification de tels gènes peut conduire à préciser un phénotype complexe et contribuer alors à accentuer la différence phénotypique entre génotypes ; cela permet une détection plus précise des marqueurs génétiques impliqués. Citons l'exemple du gène majeur nommé *RN*, ayant un effet défavorable sur le rendement à

la cuisson de la viande de porc et pour lequel une différence de trois écarts-types est mise en évidence avec la mesure du phénotype RTN (Rendement Technologique Napole). Cette différence est de sept écarts-types si l'on utilise un phénotype plus élémentaire qui est le potentiel glycolytique, lequel a permis une cartographie beaucoup plus précise du gène majeur *RN* (Le Roy *et al* 2000). Différents programmes de recherche de biomarqueurs et de QTL pour des caractères coûteux à mesurer sont en cours de réalisation dans les espèces d'élevage. Citons par exemple les programmes Qualvigene ou QualViVol sur la qualité des viandes (Allais *et al* 2010, Le Bihan-Duval *et al* 2011).

#### 3.3 / Et au-delà d'une application pour la sélection

L'identification de biomarqueurs d'un caractère peut avoir des retombées dans les domaines de la sélection mais aussi du diagnostic au sens large, ce qui permet d'optimiser des prises de décision à différents maillons d'une filière. Comme nous l'avons déjà évoqué, cette double retombée peut avoir des conséquences pour la définition des phénotypes à cibler. Par exemple, avoir un prédictif de l'aptitude au tranchage d'un jambon cuit permettrait d'orienter les pièces avant transformation vers le circuit de transformation le plus approprié afin de limiter les pertes au niveau des produits finis. Au niveau de l'élevage, avoir un prédictif «métabolique» presque infaillible des chaleurs et simple d'utilisation permettrait d'optimiser la réussite des inséminations (Hayhurst *et al* 2009, Lovendahl *et al* 2009).

L'adaptation des animaux à des conditions changeantes devient une priorité de recherche. Les races locales sont généralement considérées comme adaptées à des milieux très divers, mais les étudier dans un environnement standard ne permet pas d'identifier en quoi elles sont adaptées à leur milieu d'origine. Il faut soit reproduire ces milieux en élevage expérimental, soit essayer de caractériser le phénotype *in situ*. Il serait pour cela intéressant de disposer d'un panel de biomarqueurs validés expérimentalement dans des conditions définies avant d'être étudiés sur des prélèvements non invasifs de fluides biologiques faciles à obtenir sur les animaux *in situ*. L'analyse de ces biomarqueurs serait alors intéressante à combiner avec la recherche de signatures de sélection, approche entièrement génétique qui exploite l'histoire des populations mais n'utilise pas les phénotypes individuels (Gautier *et al* 2009).

## 4 / Décomposition d'un phénotype complexe en phénotypes élémentaires : caractérisation fine de QTL

### 4.1 / Contexte

Depuis la fin des années 90, de nombreux QTL pour différents caractères ont été identifiés dans les espèces animales. En 2010, on peut recenser 2344 QTL relatifs à 185 caractères chez la vache, 4928 QTL associés à 499 caractères chez le porc, 1290 QTL pour 164 caractères chez la poule, 84 QTL associés à 30 caractères chez le mouton (<http://www.animalgenome.org/cgi-bin/QTLdb/>). À titre de comparaison, 4287 QTL ont été localisés chez la souris (<http://www.informatics.jax.org/phenotypes.shtml>). Bien que de nombreux QTL aient été décrits, l'identification des mutations causales à ces QTL est encore difficile, quelques dizaines seulement ayant été caractérisées, correspondant à moins de 1% des QTL cartographiés. Pourtant, identifier ces mutations causales présente plusieurs intérêts : *i*) augmenter la précision de l'évaluation génomique ; *ii*) analyser très rapidement si une mutation causale est en ségrégation dans des populations, voire des espèces différentes de celle dans laquelle elle a été découverte ; *iii*) faciliter son introgression dans une population où elle ne serait pas présente ; *iv*) produire des connaissances fondamentales sur les mécanismes moléculaires impliqués dans la variabilité d'un caractère d'intérêt.

Plusieurs facteurs peuvent expliquer cette difficulté à détecter les mutations causales sous-jacentes à chaque QTL : *i*) le nombre de gènes pouvant varier de quelques unités à quelques centaines d'unités dans les régions QTL ; *ii*) la méconnaissance de la fonction de nombreux gènes, exemple du gène majeur RN découvert cinq ans après sa localisation chromosomique (Milan *et al* 2000) ; *iii*) la nature de la mutation qui peut affecter les parties régulatrices ou codantes des gènes, exemple de la mutation PIS découverte six ans après sa localisation chromosomique avec la délétion d'une région de 11,7 kb située entre deux gènes PISRT1 et FOXL2 dont elle affecte l'expression (Pailhoux *et al* 2001, Pannetier *et al* 2005) ; *iv*) la nature du «gène» impliqué qui peut être un gène dit non codant, comme les miRNA qui affectent l'expression de gènes ; *v*) le type de causalité : on recherche en général une mutation causale par région QTL alors que plusieurs mutations peuvent s'y trouver et présenter de surcroît des interactions, ce qui rend plus complexe encore l'analyse.

Les cartes de SNP à très haute densité disponibles couplées à des dispositifs animaux de grands effectifs permettent aujourd'hui une cartographie beaucoup plus fine des régions QTL. Néanmoins, même de taille réduite, une région QTL peut contenir encore plusieurs gènes ainsi que de très nombreux polymorphismes (en moyenne une mutation toutes les 500 bases). Aussi, parallèlement à ces approches de cartographie fine de QTL, la mise en place d'approches de génomique fonctionnelle (comme le transcriptome, le protéome ou le métabolome d'un tissu) offre-t-elle de nouvelles perspectives pour observer différents niveaux de régulation de la cellule (ARNm, métabolites, protéines) et apporter des informations complémentaires. Une stratégie possible est d'identifier des ARNm, protéines ou métabolites dont les quantités diffèrent significativement entre des lignées divergentes ou entre des individus extrêmes au sein de familles ou populations utilisées pour la détection de ces QTL. Cette stratégie présente néanmoins deux limites. Les lignées ou races utilisées divergent bien souvent pour différents caractères. De nombreux processus biologiques peuvent donc être identifiés sans être précisément affectés à un caractère donné (sans compter les effets de la dérive génétique particulièrement forts pour les petites populations expérimentales). De plus, pour un caractère donné, différentes régions QTL sont en général détectées. Il est donc impossible d'affecter à chacun de ces QTL les processus biologiques qu'ils impactent spécifiquement. Le paragraphe 4.2 décrit une des stratégies combinant génomique fonctionnelle et cartographie de QTL qui ne présente pas les limites évoquées plus haut. Cette stratégie (la plus utilisée) est basée sur l'identification de gènes d'un tissu associé au caractère d'intérêt, et dont l'expression est régulée par la région QTL d'intérêt. Une telle région est alors appelée «région eQTL» pour région contrôlant un caractère quantitatif de type expressionnel. Une région eQTL est détectée par analyse génétique comme un QTL, mais au lieu d'utiliser des données de performances d'élevage, l'analyse repose sur des données d'expression de gènes. D'autres stratégies utilisant les données d'expression dans le cadre de la recherche de QTL sont possibles et sont détaillées dans les articles de Le Mignon *et al* (2009 et 2010).

### 4.2 / Identification de gènes ayant un QTL d'expression (eQTL) colocalisant avec un QTL d'intérêt

#### a) Principe

La ou les mutations causales sous-jacentes à un QTL d'intérêt peuvent

dans certains cas avoir un effet sur la variation de l'expression des transcrits ou des protéines issus d'un ou de plusieurs gènes, variation qui à son tour peut conduire à la variation du caractère d'intérêt (figure 2A). Identifier de tels gènes pour chacune des régions QTL d'intérêt peut donc apporter une information fonctionnelle nouvelle sur la mutation recherchée, en particulier sur la fonction du gène porteur de la mutation causale ou affecté par cette mutation, permettant ainsi de cibler le meilleur candidat parmi tous les gènes de la région. Celui-ci contrôle non seulement le caractère d'intérêt mais aussi des caractères plus élémentaires (l'expression de gènes) qui sont alors des phénotypes intermédiaires entre mutation recherchée et caractère d'intérêt. De tels gènes doivent donc remplir deux conditions (figure 2A) : 1) permettre l'observation d'une région eQTL colocalisée avec la région QTL d'intérêt ; 2) avoir une fonction en lien avec le caractère d'intérêt. Ce second critère est important. En effet, les régions QTL et eQTL sont localisées de manière plus ou moins grossière et peuvent donc contenir plusieurs gènes et plusieurs mutations, certaines affectant l'expression de gènes qui n'ont aucun lien avec le caractère d'intérêt (figure 2A, encadrés gris). Une des limites ici est la connaissance très partielle des fonctions des gènes d'un génome (cf. § 3.3).

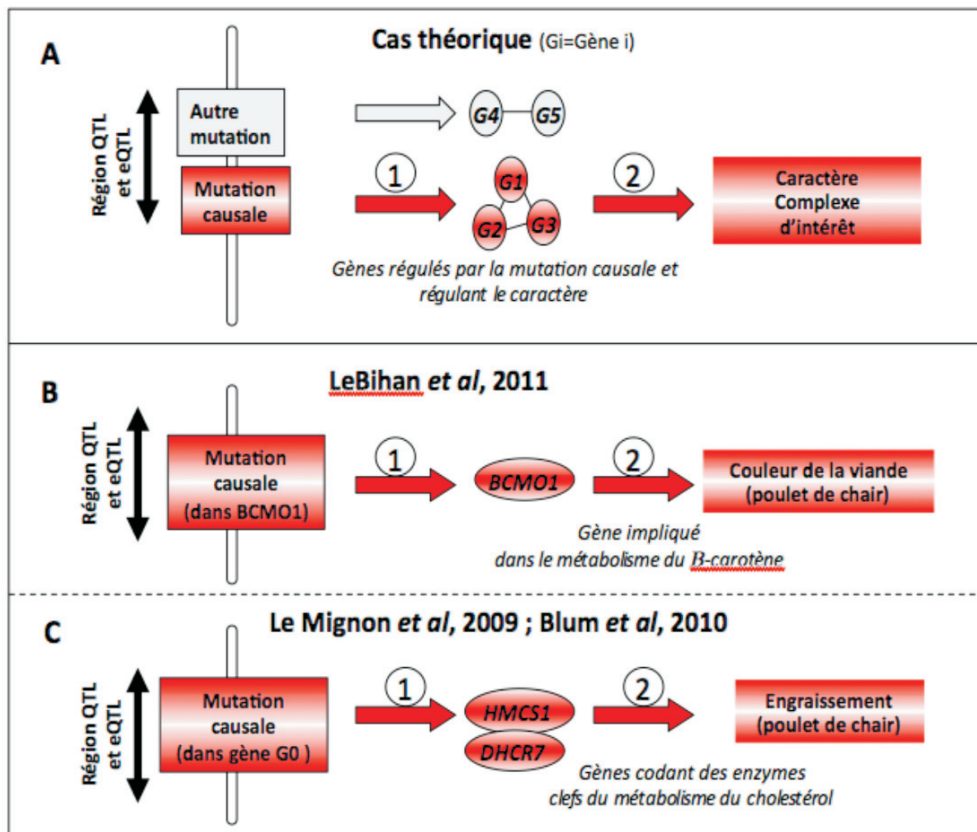
Pour identifier les gènes répondant à ces critères, deux situations peuvent être distinguées. La première consiste à se concentrer sur un ou quelques gènes localisés dans la région QTL. Cette démarche est pertinente si un des gènes de la région a une fonction directement liée au caractère. La seconde, alternative à la première, ne fait aucun a priori et prend en compte l'ensemble des gènes du génome. Illustrons ces deux situations/approches par deux exemples tirés des espèces d'élevage (figure 2B et 2C), même si ces exemples sont encore peu nombreux.

A gauche, mutations dans la région QTL dont en rouge, la mutation causale responsable du caractère d'intérêt, 1) effet de ces mutations sur des gènes (gènes *Gi*) indiqués au centre, 2) effets de ces gènes *Gi* sur le caractère d'intérêt, la fonction de ces gènes en lien avec le caractère étant indiquée en italique à droite.

b) Première approche : analyse de l'expression des gènes localisés dans la région QTL d'intérêt

Prenons l'exemple de la région QTL localisée sur le chromosome 11 contrôlant la couleur de la viande chez le poulet de chair (Nadaf *et al* 2007). Une



**Figure 2.** Identification de gènes ayant un QTL d'expression (eQTL) colocalisant avec un QTL d'intérêt.

(A) Cas théorique indiquant des liens possibles entre mutations d'une région QTL, expression génique et caractère d'intérêt ; (B) Un exemple dans lequel la mutation causale est dans la région régulatrice d'un gène dont elle contrôle l'expression, cette variation d'expression étant responsable (en partie) de la variation du caractère ; (C) Un exemple dans lequel la mutation causale qui se trouve dans un gène G0 encore non déterminé, contrôle l'expression d'un ou plusieurs gènes, cette variation d'expression étant responsable (en partie) de la variation du caractère.

analyse des gènes localisés dans la région QTL a révélé l'existence d'un gène, *BCMO1*, qui code une enzyme clef de la dégradation du  $\beta$ -carotène, pigment dont l'effet sur la couleur des tissus est bien connu. Dès le départ, ce gène représentait un bon candidat à la fois positionnel et fonctionnel pour le QTL d'intérêt. La deuxième condition explicitée ci-dessus étant remplie, il s'agissait ensuite de vérifier si ce gène pouvait être lui-même régulé par un eQTL colocalisé avec la région du QTL d'intérêt. Une analyse génétique l'a rapidement confirmé (figure 2B). *BCMO1* possède finalement trois propriétés intéressantes qui suggèrent sa possible implication dans la variabilité du caractère : il est physiquement localisé dans la région QTL, il est fonctionnellement en lien avec le caractère et il est régulé par un eQTL recouvrant l'intervalle du QTL. Cette dernière observation indique également que la mutation recherchée se trouve dans les parties régulatrices du gène. Le séquençage de ces régions régulatrices (promoteur) sur des animaux de génotypes variés au QTL a permis de mettre en évidence deux mutations. Des expérimentations de biologie moléculaire ont montré par la suite que celles-

ci avaient bien un effet sur l'expression du gène (Le Bihan-Duval *et al* 2011). Ces mutations causales impliquées dans la couleur de la viande ont fait l'objet d'une demande de brevet (Numéro EP2161345 disponible sur Internet à l'adresse : <https://data.epo.org/publication-server>). Ainsi, l'existence dans une région QTL d'intérêt d'un gène ayant une fonction en relation avec le caractère d'intérêt peut considérablement motiver l'adoption de cette première approche. L'hypothèse est alors que ce gène possède la mutation causale recherchée et que celle-ci impacte son expression et en conséquence les variations du caractère. On parle alors de *cis*-eQTL. Dans cet exemple, moins de deux années se sont écoulées entre la primo-localisation de la région QTL et la détection de la mutation causale associée, témoignant de l'efficacité de l'approche dans le cas d'un phénotype précisément mesuré comme la couleur de la viande. La réussite des travaux portant sur *BCMO1*, impliqué dans la couleur de la viande, montre que les approches de génomique fonctionnelle peuvent dans certains cas faciliter grandement la découverte de mutations causales. Cependant, selon le contexte, il n'est pas toujours aussi facile de trou-

ver une mutation causale. Au sein d'une région QTL, il existe de nombreux gènes dont la(les) fonction(s) est(sont) encore inconnue(s) et parmi ceux dont la fonction est connue, il n'existe malheureusement pas toujours de gènes possédant une fonction en lien évident avec le caractère d'intérêt. La seconde approche s'impose alors.

#### c) Deuxième approche : analyse de tous les gènes du génome

L'hypothèse implicite dans cette approche est que la mutation causale recherchée dans la région QTL d'intérêt ne modifie pas nécessairement l'expression du gène qui la porte et peut conduire à une variation d'expression d'autres gènes que ceux de la région QTL, ces gènes cibles affectant alors le caractère d'intérêt. Si le gène causal ne fait pas partie de la liste des gènes régulés par la région QTL, alors la mutation causale se trouve dans la région codante de ce gène et non dans ses régions régulatrices comme dans la première approche. Ce gène peut coder par exemple un facteur de transcription, qui, en raison d'un changement d'affinité, va impacter l'expression des gènes cibles qu'il régule.

Pour déterminer cette signature fonctionnelle, il est nécessaire d'effectuer une analyse de liaison pour chacun des niveaux d'expression des gènes du génome (déterminés par analyse de transcriptome), afin d'identifier ceux qui ont un eQTL colocalisant avec la région QTL. Cette approche sans a priori ne nécessite aucune connaissance préalable concernant les effets fonctionnels de la mutation recherchée. Elle peut également être appliquée aux protéines ou aux métabolites, afin d'augmenter les chances d'observer les événements fonctionnels impactés par la mutation causale au QTL d'intérêt. Cette mutation peut en effet faire varier l'expression de gènes non pas au niveau de l'ARNm mais au niveau de la protéine qu'ils codent ou plus en aval, au niveau des métabolites. On parle alors de pQTL (QTL contrôlant la variation d'une protéine) ou de mQTL (QTL contrôlant un métabolite). La colocalisation d'une région eQTL (ou pQTL ou mQTL) avec la région QTL d'intérêt peut ainsi apporter des informations fonctionnelles précieuses permettant de déterminer le meilleur gène candidat positionnel et fonctionnel dans la région QTL. Cette méthodologie de «génétique génomique» appelée en anglais «*genetical genomics*» a été mise en œuvre les toute premières fois par Brem *et al* (2002) et Schadt *et al* (2003), respectivement sur la levure et la souris. Elle

est actuellement en plein essor et concerne aussi bien les espèces végétales, animales modèles (souris) que l'Homme (Le Mignon *et al* 2010).

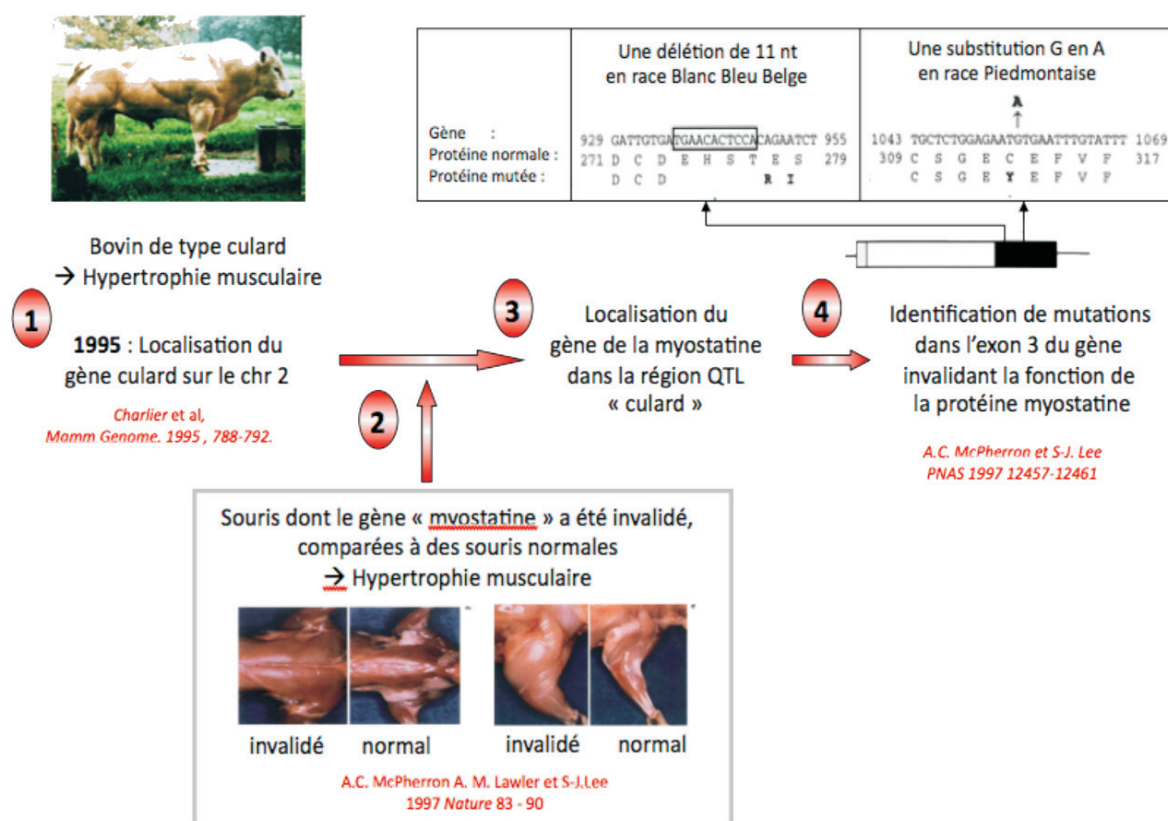
À titre d'exemple, prenons les travaux portant sur un QTL contrôlant l'engraissement chez le poulet de chair, qui est localisé sur le chromosome 5 de poule (Lagarrigue *et al* 2006). Une étude transcriptomique a donc été mise en œuvre sur le même dispositif d'animaux que l'analyse QTL, pour tenter d'identifier sans a priori des gènes contrôlés par la région QTL, gènes qui d'après leurs fonctions (cellulaires, métaboliques ou autres) devraient permettre de mieux caractériser la région QTL. Plusieurs dizaines de gènes impliqués dans des métabolismes variés ont ainsi été identifiés. Ils n'interviennent probablement pas tous dans les phénotypes intermédiaires entre la mutation recherchée et le caractère d'engraissement. Certains d'entre eux sont certainement contrôlés par la mutation causale (considérant ici le cas le plus simple d'une seule et unique mutation causale), d'autres par d'autres mutations proches de celle-ci, comme il est indiqué dans la figure 2A, encadrés gris. Parmi ces gènes, deux gènes attirent aussitôt l'attention puisqu'ils codent des enzymes clés de la synthèse du cholestérol bien connu pour intervenir dans les mécanismes de l'en-

graissement (figure 2C). Quel(s) gène(s) compris dans la région QTL pourraient alors réguler l'expression de ces deux enzymes qui à leur tour impacteraient l'état d'engraissement ? La réponse à cette question n'est à ce stade pas encore évidente. Il peut être alors intéressant de regarder si un métabolite tel que le cholestérol est également régulé par la région QTL. Si tel est le cas, on recueille successivement des informations précieuses qui permettent petit à petit d'émettre une hypothèse de plus en plus précise sur la fonction du gène causal recherché dans la région QTL. À ce stade, ces approches sont encore limitées par la connaissance partielle des fonctions des gènes d'un génome.

#### 4.3 / Vers une meilleure connaissance de la fonction des gènes : le phénotypage, autre élément clé

Le séquençage du génome d'une espèce donnée permet la localisation génomique de l'ensemble des gènes de l'espèce. Le génome de la plupart des espèces d'élevage d'intérêt économique étant en passe d'être disponible, la connaissance du contenu en gènes d'une région QTL ne sera bientôt plus un frein. Cependant, nombre de ces gènes sont révélés par des prédictions bio-

**Figure 3.** Démarche ayant permis l'identification de mutations responsables du phénotype «culard» chez les bovins allaitants.





informatiques et n'ont pas de fonction connue. De vastes programmes de recherche visant à invalider chacun de ces gènes et à caractériser les phénotypes qui en résultent sont en cours de réalisation dans les espèces modèles : par exemple, *Arabidopsis* pour les espèces végétales, la souris pour les espèces animales (Magnol *et al* 2011). Les génomes contenant au moins 20 000 à 40 000 gènes, ces programmes sont gigantesques à mettre en place, et là aussi le phénotypage est un facteur limitant s'il ne peut être mené à haut débit et à un coût raisonnable. En revanche, l'enjeu est de taille : la caractérisation phénotypique systématique de ces collections de dizaines de milliers de mutants devrait conduire à une accélération sans précédent dans la découverte des fonctions des gènes. Ce type de recherches systématiques, et à première vue très fondamentales, aura alors des conséquences immédiates dans le cadre de la recherche des mutations responsables de caractères d'intérêt agronomique. La découverte du gène responsable du phénotype «culard» en bovin viande en est un parfait exemple (figure 3).

En 1995, le gène majeur «culard» est localisé sur le chromosome 2 bovin (Charlier *et al* 1995). En 1997, des souris dans lesquelles le gène codant la myostatine a été invalidé présentent une augmentation de la masse musculaire (McPherron *et al* 1997). Ce gène muté reproduisant le phénotype observé chez les bovins devient donc un excellent candidat fonctionnel pour le phénotype d'intérêt. Il est alors cartographié dans

le génome bovin et se trouve dans la région QTL «culard», ce qui fait de lui un bon candidat à la fois fonctionnel et positionnel. Une analyse de ses polymorphismes entre animaux culards et non culards a alors permis d'identifier la mutation causale de type faux sens dans l'exon 3 : il s'agit d'une délétion de 11 bases en race Blanc Bleu Belge et d'une substitution en race Piémontaise (Grobet *et al* 1997, McPherron et Lee 1997). Cet exemple correspond au délai le plus court qu'il y ait eu dans les espèces d'élevage entre localisation d'un QTL et identification de la mutation causale (2 ans), mais il s'agissait d'un gène à effet quantitatif majeur. Le facteur clé de cette réussite a été l'invalidation d'un gène précis chez la souris et l'observation du phénotype en résultant. Depuis, différentes mutations responsables du phénotype culard ont été identifiées dans ce gène dans différentes races bovines à viande. Elles ont chacune donné lieu au développement d'un test génétique permettant ainsi de les diagnostiquer très précocement.

## Conclusion

Le phénotypage est en pleine évolution, grâce à la combinaison de différentes techniques, qui vont de l'électronique et de la robotique aux technologies en -omique. La sélection génomique va permettre de repenser les critères actuels de sélection, voire en introduire de nouveaux, et ce avec des situations variables selon les espèces et

les structures de programmes de sélection et de production. L'essor de la sélection génomique est à l'origine d'une demande forte de caractères plus précis. En effet, plus le caractère est précis et proche du mécanisme causal, plus la détection d'un QTL est précise et plus l'identification du gène responsable devient facile dans une région du génome délimitée. La mise en place de populations de référence assez nombreuses justifie le développement de nouvelles méthodes de phénotypage à haut débit, pour une large palette de caractères, avec une nécessaire maîtrise des coûts. Les concepts d'eQTL et de biomarqueurs sont applicables aux espèces d'élevage et offrent des perspectives nouvelles. Pour être utilisés en sélection, ces biomarqueurs devront être héréditaires et corrélés à l'objectif de sélection. D'une manière générale, le passage à un phénotypage plus prédictif et à haut débit devra s'accompagner d'une évaluation précise des investissements à réaliser sur des plates-formes spécialisées permettant d'optimiser les coûts de mise en œuvre.

## Remerciements

Ce texte est en grande partie le fruit de discussions avec différents collègues. Il a été en particulier inspiré par des échanges avec Benjamin Basso, Elisabetta Giuffra, François Guillaume, Yannick Lecozler, Pascale Le Roy, Patrice Martin, Sandrine Mignon-Grasteau que nous souhaitons ici remercier.

## Références

- Allais S., Leveziel H., Payet-Duprat N., Hocquette J.F., Lepetit J., Rousset S., Denoyelle C., Bernard-Capel C., Journaux L., Bonnot A., Renand G., 2010. The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds. *J. Anim. Sci.*, 88, 446-454.
- Bevilacqua C., Makhzami S., Helbling J.C., Defrenaix P., Martin P., 2010. Maintaining RNA integrity in a homogeneous population of mammary epithelial cells isolated by Laser Capture Microdissection. *BMC Cell Biology*, 11, doi: 10.1186/1471-2121-11-95.
- Brem R.B., Yvert G., Clinton R., Kruglyak L., 2002. Genetic dissection of transcriptional regulation in budding yeast. *Science*, 296, 752-755.
- Charlier C., Coppieters W., Farnir F., Grobet L., Leroy P.L., Michaux C., Mni M., Schweser A., Vanmanshoven P., Hanset R., 1995. The mh gene causing double-muscling in cattle maps to bovine Chromosome 2. *Mamm. Genome*, 6, 788-792.
- Gautier M., Flori L., Riebler A., Jaffrézic F., Laloë D., Gut I., Moazami-Goudarzi K., Foulley J.L., 2009. A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC Genomics*, 10:550 doi:10.1186/1471-2164-10-550.
- Grobet L., Martin L.J., Poncet D., Pirotin D., Brouwers B., Riquet J., Schoeberlein A., Dunner S., Menissier F., Massabanda J., Fries R., Hanset R., Georges M., 1997. A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nat. Genet.*, 17, 71-74.
- Guillemin N., Jurie C., Cassar-Malek I., Hocquette J.F., Renand G., Picard B., 2011. Variations in the abundance of 24 protein biomarkers of beef tenderness according to muscle and animal type. *Animal*, 5, 885-894.
- Hayhurst C., Flint A.P.F., Lovendahl P., Woolliams J.A., Royal M.D., 2009. Genetic variation of metabolite and hormone concentration in UK Holstein-Friesian calves and the genetic relationship with economically important traits. *J. Dairy Sci.*, 92, 4001-4007.
- Hurtaud C., Meunier-Salaün M.C., Bugeon J., Fatet A., Reichstadt M., Valancogne A., Vernet J. Hue I., Le Bail P.Y., Dameron O., Reecy J., Park C., 2011. ATOL : a new ontology for livestock Workshop New technologies and new challenges for breeding and herd management of ICAR 22-24 juin 2011, Bourg-en-Bresse, France.
- Institut de l'Élevage, 2011. Bilan génétique des races bovines laitières 2010. [http://www.inst-elevage.asso.fr/spip.php?page=article\\_espace&id\\_espace=935&id\\_article=20029](http://www.inst-elevage.asso.fr/spip.php?page=article_espace&id_espace=935&id_article=20029).
- Lagarigue S., Pitel F., Carre W., Abasht B., Le Roy P., Neau A., Amigues Y., Sourdoux M., Simon J., Cogburn L., Aggrey S., Leclercq B., Vignal A., Douaire M., 2006. Mapping quantitative trait loci affecting fatness and breast muscle weight in meat-type chicken lines divergently selected on abdominal fatness. *Genet. Sel. Evol.*, 38, 85-97.
- Le Bihan-Duval E., Nadaf J., Berri C., Pitel F., Graulet B., Godet E., Leroux S.Y., Demeure O., Lagarrigue S., Duby C., Cogburn L.A., Beaumont C.M., Duclos M.J., 2011. Detection of a Cis eQTL Controlling BCMO1 gene expression leads to the identification of a QTG for chicken breast meat color. *PLoS One*, 6, e14825.
- Le Mignon G., Desert C., Pitel F., Leroux S., Demeure O., Guernec G., Abasht B., Douaire

M., Le Roy P., Lagarrigue S., 2009. Using transcriptome profiling to characterize QTL regions on chicken chromosome 5. *BMC Genomics*, 10, 575.

Le Mignon G., Blum Y., Demeure O., Le Bihan-Duval E., Le Roy P., Lagarrigue S., 2010. Stratégies utilisant des données d'expression pour la cartographie fine de QTL. *INRA Prod. Anim.*, 23, 434-358

Le Roy P., Elsen J.M., Caritez J.C., Talmant A., Juin H., Sellier P., Monin G., 2000. Comparison between the three porcine RN genotypes for growth, carcass composition and meat quality traits. *Genet. Sel. Evol.*, 32, 165-186.

Lewis P.D., Lewis K.E., Ghosal R., Bayliss S., Lloyd A.J., Wills J., Godfrey R., Kloer P., Mur L.A.J., 2010. Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum. *BMC Cancer*, 10, DOI: 10.1186/1471-2407-10-640.

Lovendahl P., Chagunda M., O'Connell J., Friggens N., 2009. Genetics of fertility indicators based on behaviour and progesterone in milk cattle. *Practice*, 17, 7-12.

Magnol L., Monestier O., Vuillier-Devillers K., Wagner S., Cocquemot O., Chevallier M.C., Blanquet V., 2011. A sensitised mutagenesis screen in the mouse to explore the bovine genome: study of muscle characteristics. *Animal*, 5, 663-671.

McPherron A.C., Lee S.J., 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proc. Natl Acad. Sci.*, 94, 12457-12461.

McPherron A.C., Lawler A.M., Lee S.J., 1997. Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member. *Nature*, 387, 83-90.

Milan D., Jeon J.T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundstrom K., Reinsch N., Gellin J., Kalm E., Roy P.L., Chardon P., Andersson L., 2000. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science*, 288, 1248-1251.

Nadaf J., Gilbert H., Pitel F., Berri C.M., Feve K., Beaumont C., Duclos M.J., Vignal A., Porter T.E., Simon J., Aggrey S.E., Cogburn L.A., Le Bihan-Duval E., 2007. Identification of QTL controlling meat quality traits in an F2 cross between two chicken lines selected for either low or high growth rate. *BMC Genomics*, 8, 155.

N'dri A.L., Sellier N., Tixier-Boichard M., Beaumont C., Mignon-Grasteau S., 2007. Genotype by environment interactions in relation to growth traits in slow growing chickens. *Genet. Sel. Evol.*, 39, 513-528.

Pailhoux E., Vigier B., Chaffaux S., Servel N., Taourit S., Furet J.P., Fellous M.,

Grosclaude F., Cribiu E.P., Cotinot C., Vaiman D., 2001. A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nat. Genet.*, 29, 453-458.

Pannetier M., Renault L., Jolivet G., Cotinot C., Pailhoux E., 2005. Ovarian-specific expression of a new gene regulated by the goat PIS region and transcribed by a FOXL2 bidirectional promoter. *Genomics*, 85, 715-726.

Phénofinlait : [www.phenofinlait.fr/](http://www.phenofinlait.fr/)

Rizzi F., Belloni L., Crafa P., Lazzaretti M., Remondini D., Ferretti S., Cortellini P., Corti A., Bettuzzi S., 2008. A novel gene signature for molecular diagnosis of human prostate cancer by RT-qPCR. *PLoS One*, 3, e3617.

Schadt E.E., Monks S.A., Drake T.A., Lusis A.J., Che N., Colinao V., Ruff T.G., Milligan S.B., Lamb J.R., Cavet G., Linsley P.S., Mao M., Stoughton R.B., Friend S.H., 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature*, 422, 297-302.

Weichselbaum R.R., Ishwaran H., Yoon T., Nuyten D.S., Baker S.W., Khodarev N., Su A.W., Shaikh A.Y., Roach P., Kreike B., Roizman B., Bergh J., Pawitan Y., van de Vijver M.J., Minn A.J., 2008. An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer. *Proc. Natl Acad. Sci.*, 105, 18490-18495.

## Résumé

Le phénotype correspond aux valeurs mesurables prises par des caractères choisis pour leur intérêt socio-économique ou cognitif. Les besoins de phénotypage dépendent de l'évolution des méthodes de sélection et des systèmes d'élevage ainsi que des développements technologiques permettant une approche standardisée, à haut débit et automatisable. En matière de sélection animale, la mise en place de la sélection génomique introduit le concept de population de référence sur laquelle des phénotypes difficilement mesurables en routine peuvent être obtenus pour établir les relations entre marqueurs génétiques et performance. L'association marqueurs-phénotype ouvre également la voie au diagnostic individuel en appui à la gestion du troupeau. La robotisation et l'identification électronique individuelle supposent des investissements importants mais offrent des perspectives très nouvelles pour le phénotypage de caractères comme l'efficacité alimentaire ou le comportement. De façon complémentaire, les technologies de génomique fonctionnelle et l'analyse de l'empreinte spectrale des protéines permettent d'accéder aux mécanismes sous-jacents et d'affiner la définition du phénotype à l'échelle moléculaire. Le concept de biomarqueur capable de prédire le phénotype est en plein essor en médecine humaine et pourra aussi s'appliquer aux animaux d'élevage. Plus le phénotypage sera proche du mécanisme d'action, plus la détection des gènes contrôlant la variation du phénotype sera précise. En particulier, la recherche de régions du génome (eQTL) contrôlant l'expression d'un gène permet d'explorer les mécanismes responsables de la variabilité de caractères complexes. On attend donc de grands progrès dans l'identification des gènes qui sous-tendent les QTL grâce aux progrès conjoints du phénotypage et du séquençage du génome.

## Abstract

### *New approaches of phenotyping for animal breeding*

Phenotyping requires measuring traits chosen for their socio-economic importance or their cognitive value. Needs for phenotyping depend on the trend in breeding methods and production systems, and also on technological developments allowing for high throughput standardised approaches at a reasonable cost. Considering animal breeding, implementation of genomic selection introduces the concept of reference population, where specific phenotypes not routinely available can be obtained to associate genetic markers with performance. Association between markers and phenotypes also open the way towards individual diagnostic to support herd management. Robotics and individual electronic identification require important investments but offer quite new prospects for phenotyping traits such as feed efficiency or behaviour. Complementary tools are provided by technological developments in functional genomics and by spectral analysis of proteins, which can unravel underlying mechanisms and refine the definition of phenotypes at the molecular scale. The concept of biomarker has emerged in human medicine in order to predict phenotype and is expected to be applied also to farm animals. The closer the phenotyping is from the biological mode of action, the more accurate will be the detection of genes determining phenotypic variation. More particularly, the search for genome regions controlling gene expression (eQTL) makes possible to explore the mechanisms responsible for the variability of complex traits. Great progress can thus be expected for the identification of genes underlying QTL thanks to simultaneous improvements in phenotyping and genome sequencing.

LAGARRIGUE S., TIXIER-BOICHARD M., 2011. Nouvelles approches de phénotypage pour la sélection animale. In : Numéro spécial, Amélioration génétique. Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M. (Eds). *INRA Prod. Anim.*, 24, 377-386.