

## Le phénotypage des animaux : le nouveau défi ?

MONGET, P. (1), LE BAIL, P-Y (2)

(1) (2) AGENAE, laboratoire de physiologie de la reproduction et des comportements

(1) INRA UMR85, physiologie de la reproduction et des comportements, F-37380 Nouzilly, CNRS, université de Tours, haras nationaux, F-37380 Nouzilly, France

(2) INRA, UR1037 SCRIBE, campus de Beaulieu, F-35000 Rennes, France

**RESUME-** Depuis environ deux ans, la génomique a vécu un saut technologique sans précédent en matière de séquençage et de génotypage. Dans l'espèce bovine, une puce permet maintenant de génotyper n'importe quel animal avec plus de 50 000 marqueurs génétiques répartis dans le génome. Rapidement, ce type d'outil sera étendu à plusieurs centaines de milliers de marqueurs et sera généralisé aux autres espèces d'élevage. Parallèlement, les outils d'étude de l'expression des gènes (puce, protéomique, métabolomique) se développent tout aussi rapidement.

C'est la raison pour laquelle le phénotypage le plus complet et le plus fin possible des animaux devient maintenant le facteur limitant dans les études de liaison du phénotype avec le génotype ou l'expression des gènes. Les chercheurs, avec l'aide des unités expérimentales, doivent donc classer leurs animaux en fonction de toutes caractéristiques phénotypiques intéressantes, anatomiques, fonctionnelles, endocriniennes, métaboliques, comportementales etc. Il peut s'agir de caractéristiques "visibles" comme la taille au garrot, les réactions de stress, le comportement social ou sexuel, la durée de l'intervalle entre chaleurs et ovulation, mais aussi de caractéristiques moins visibles sans investigation (biopsie, dissection ...), comme la qualité de la viande, la concentration sérique en hormones ou métabolites, l'aptitude d'un sperme à la congélation etc. Au total, c'est une vision zootechnique et biologique exhaustive "quasi clinique" que le chercheur, avec l'aide des animaliers, devra désormais avoir des animaux présents dans les unités expérimentales. De plus, ces mesures doivent être standardisées pour être comparables partout dans le monde. Une action est lancée en 2009 par le département Phase sur "l'ontologie du phénotypage" à l'INRA, en étroite collaboration avec James Reecy de l'université de l'Iowa (responsable : Pierre-Yves Le Bail).

## Animal phenotyping : the new challenge ?

MONGET, P. (1), LE BAIL, P-Y (2)

(1) (2) AGENAE, laboratoire de physiologie de la reproduction et des comportements

For the last two years, genomics has undergone a technological breakthrough in sequencing and genotyping. In cattle, a chip now allows the genotyping of any animal with more than 50,000 genetic markers all over the genome. This tool will soon be extended to hundreds of thousands of markers, and will be generalised to other species. In parallel, tools dedicated to gene expression studies (chips, proteomics, metabolomics) are being developed very quickly as well.

This is the reason why phenotyping as exhaustive and precise as possible is now limiting for the study of the genotype/phenotype relationship. Researchers, with the help of experimental units, must classify animals according to every interesting phenotypic characteristic, such as anatomic, functional, endocrine, metabolic, behavioral etc... This concerns "apparent" characteristics such as size, reactions to stress, social and sexual behavior, duration of interval between heat and ovulation, as well as characteristics less apparent without investigation (biopsy, dissection...), such as meat quality, hormone and metabolite serum concentrations, sperm congelation ability etc... Overall, researchers along with animal keepers should now have a "clinical" vision of animals in experimental units. Ultimately, these measures should be standardised to be comparable all over the world. An investigation began in 2009 in the Phase Division on the ontology of phenotyping in the INRA institut, in strong collaboration with James Reecy from Iowa University (Head of the project: Pierre-Yves Le Bail).

## INTRODUCTION

Depuis environ deux ans, la génomique en médecine humaine a vécu un saut technologique sans précédent en matière de séquençage et de génotypage. Les échelles ont changé. On ne met plus plusieurs années pour séquencer un génome entier de vertébré pour plusieurs millions d'euros, mais quelques jours pour quelques milliers d'euros. On ne génotype plus les patients avec quelques milliers de marqueurs génétiques, mais quelques millions, sur une lame de quelques centimètres carré en s'appuyant sur des moyens informatiques de traitement de plus en plus performants. La frontière est d'ailleurs floue entre génotypage et re-séquençage, cette dernière technique permettant de facto un génotypage exhaustif

En médecine humaine, ces technologies et les données qu'elles ont apportées ont été rapidement mises en regard des tableaux cliniques aussi complets que possible des patients, et des relations génotype-phénotype ont pu être établies, sans pour autant lever toutes les ambiguïtés. Les

résultats obtenus marquent l'ère de la "Genome Wide Association" entre un syndrome pathologique et quelques marqueurs génétiques dispersés dans le génome, identifiés parmi des centaines de milliers. Ces approches ouvrent la voie à des traitements préventifs ou curatifs individualisés, tenant compte du fond génétique unique de chaque patient. Parallèlement, les outils d'étude de l'expression des gènes (puce, protéomique, métabolomique) se développent tout aussi rapidement.

Nos espèces d'élevage peuvent bénéficier maintenant de cette révolution. Dans l'espèce bovine par exemple, une puce permet maintenant de génotyper n'importe quel animal avec plus de 50 000 marqueurs génétiques répartis dans le génome (pour revue, Goddard et Hayes, 2009). Cet outil est déjà largement utilisé dans les programmes de cartographie fine des QTLs chez les bovins laitiers. Rapidement, il sera étendu à plusieurs centaines de milliers de marqueurs, et sera généralisé aux autres espèces d'élevage, et sera orienté vers différents types de

production (lait, viande, œufs, prolificité...) ou pour mieux adapter les races (lignées) aux conditions d'élevages (bien-être, utilisation de nouveaux nutriments, diminution des rejets organiques).

## 1. LES MESURES PHENOTYPIQUES EFFECTUEES

Le phénotypage le plus complet et le plus fin possible des animaux deviendra donc très rapidement le facteur limitant dans les études de liaison du phénotype avec le génotype ou l'expression des gènes. Les chercheurs, avec l'aide précieuse des unités expérimentales, doivent donc classer leurs animaux en fonction de toutes caractéristiques intéressantes (anatomiques, fonctionnelles, endocriniennes, métaboliques, comportementales etc.) en sachant que les caractères phénotypiques (traits) observés d'un individu, qui constituent son phénotype, ne dépendent pas seulement de son génotype (des allèles de ses gènes) mais aussi des conditions environnementales dans lesquels s'expriment ses gènes.

Il peut s'agir de caractéristiques "visibles" comme la taille au garrot, les réactions de stress, le comportement social ou sexuel, la durée de l'intervalle entre chaleurs et moment d'ovulation. Il peut s'agir aussi de caractéristiques moins "visibles", nécessitant une intervention plus ou moins invasive (biopsie, dissection...), comme la structure de la viande, la concentration sérique en AMH (témoin de la population de follicules ovariens en croissance), l'aptitude d'un sperme à la congélation etc. Il sera important de prendre en compte certains paramètres susceptibles d'introduire des biais dans l'interprétation des mesures, comme "l'histoire" de l'animal (métabolisme, comportement), ainsi que son pedigree, afin de vérifier l'aspect familial du caractère mesuré et d'en évaluer l'héritabilité éventuelle. Au total, c'est une vision quasi clinique que le chercheur devra désormais avoir des animaux présents dans les unités expérimentales.

## 2. L'ONTOLOGIE DU PHENOTYPAGE

L'ensemble de ces mesures sera réalisé partout dans le monde. De la même façon qu'en médecine humaine, l'évaluation des paramètres sanguins, ou les dosages hormonaux ou métaboliques, sera standardisée pour être comparables d'un laboratoire ou d'un pays à un autre, toutes les mesures qui seront réalisées sur les animaux devront être standardisées selon un cahier des charges très précis, commun à tous les laboratoires. Cette "ontologie du phénotypage" actuellement en cours à l'INRA, en collaboration avec James Reecy de l'université de l'Iowa, mobilise plus d'une cinquantaine de chercheurs et d'ingénieurs.

### 2.1. CAHIER DES CHARGES (BESOINS STANDARDISES DES DEFINITIONS)

Ce cahier des charges est constitué de plusieurs volets :

la nécessité d'établir des relations génotype / phénotype précises et d'associer un phénotype à chaque allèle,

l'ambition de pouvoir repérer les constantes fonctionnelles des groupes de gènes au cours de l'évolution afin d'en tirer des règles générales au sein des vertébrés,

l'ambition de développer des outils de modélisation et de prédiction des phénotypes en fonction des conduites d'élevage à partir de bases de données standardisées,

la construction de modèle *in silico* mimant la physiologie cellulaire et animale,

le besoin stratégique de faire reconnaître les spécificités des produits locaux issus de l'élevage au sein des échanges mondialisés.

L'ensemble de ces critères a conduit les responsables du département PHASE de l'Inra à s'investir dans la mise en place d'un référentiel des caractères phénotypiques des animaux de rente qui puisse être reconnu au niveau international. L'ambition de cette base ontologique est de s'adresser tant aux généticiens pour leurs programmes de sélection, qu'aux physiologistes à des fins de modélisation et de biologie comparée. Elle doit être pertinente dans le choix des caractères phénotypiques retenus, précise dans la définition de ces caractères, explicite sur les conditions et les techniques de mesure.

### 2.2. BASES EXISTANTES

Les premières bases ontologiques à avoir été créées, et faisant référence au sein de l'ensemble de la communauté scientifique, concernaient la taxonomie où le nom latin de l'espèce est universellement reconnu, où la description de l'espèce tient lieu de définition, et où la classification phylogénétique hiérarchise les relations entre taxons. Aujourd'hui, il existe de nombreuses bases ontologiques accessibles en ligne pouvant apporter des informations intéressantes plus ou moins directement les caractères phénotypiques, certaines traitant de la typologie des caractères (NCBO Bioportal, *PATO-Phenotypic Quality Ontology...*), de biochimie (ChEBI...), d'anatomie (*Digital Morphology, Zebrafish Anatomical Ontology...*), de santé (*Disease Ontology, Infectious Disease Ontology, Mouse Pathology...*), de fonctions biologiques (*Gene Ontology, Mouse Genome Informatics...*). Celles traitant plus directement de phénotypage (*Human trait ontology, Mammalian Trait Ontology...*) sont peu nombreuses et s'adressent plus généralement à l'homme, ou aux modèles animaux utilisés pour étudier les maladies humaines. L'INRA a donc décidé de s'associer au projet ATO (*Animal Trait Ontology*), développé par le programme « *Animal Genome Research Program* » de l'USDA et destiné à l'élaboration d'une ontologie de traits normalisés pour les animaux de rente (Hughes et al., 2008) en accord avec ses propres objectifs. Des démarches sont également en cours pour donner une dimension européenne à cette stratégie (Reecy et al., 2009).

### 2.3. QUESTIONS RELATIVES À LA STRUCTURATION D'UNE BASE ONTOLOGIQUE (ATO)

La plus grande difficulté rencontrée lors de l'élaboration d'une base ontologique est de savoir cerner le champ de ses utilisations et de ses objectifs. Une base trop fournie en caractères insuffisamment pertinents, mal orientée et mal structurée conduit à un outil trop complexe qui rebute les utilisateurs potentiels. D'autre part, dans ce type d'approche la recherche de l'exhaustivité n'est pas envisageable et souvent constitue un frein à l'élaboration de l'outil : un référentiel démarré avec un nombre de caractères restreint, mais régulièrement alimenté, lié à des définitions courtes et précises, a fait le succès d'une base comme *Gene Ontology*.

Se voulant opérationnelle, la base ATO-Phase s'est résolument portée sur une hiérarchie structurée autour des

grandes productions animales et les fonctions associées : 1) bien-être et adaptation au milieu d'élevage, 2) alimentation nutrition 3) reproduction et fertilité, 4) production laitière et mamelle, 5) qualité des tissus (musculaire et adipeux). Certains caractères peuvent être associés à plusieurs productions comme les aspects liés au développement embryonnaire (reproduction et fertilité vs. qualité des tissus). Les caractères phénotypiques explicatifs de chacune des productions sont eux-mêmes composés de différents domaines relatifs par exemple, dans le cas de la « production laitière et mamelle », à l'anatomie (nombre de trayons, volume de la mamelle ...), la physiologie (durée de la lactation, production laitière journalière, prolifération cellulaire de la glande mammaire ...), et la biochimie (composition du lait en acides gras, en protéines...) Le caractère phénotypique ne correspond pas à une granulométrie particulière définie a priori. Par exemple, dans le cas de la « qualité des tissus », le caractère peut tout aussi bien porter sur la structure de la carcasse, sur la tendreté de la viande, sur la cellularité du muscle ou sur l'acidité intracellulaire. Ce qui compte, c'est le lien avéré du caractère avec la qualité du tissu visé, la pertinence du caractère étant laissée à l'appréciation des experts du domaine.

Une autre difficulté réside dans le choix entre des référentiels spécifiques d'une espèce, incluant dans la définition du caractère les conditions et les modes de mesure, (donc plus directement opérationnelles pour les généticiens), ou entre des référentiels génériques aux vertébrés permettant une approche comparative des fonctions (croissance, reproduction, comportement...) intéressant aussi les physiologistes. Si cette dernière approche est plus attractive d'un point de vue cognitif, elle se heurte souvent aux divergences phylogénétiques et convergences fonctionnelles apparues au cours de l'évolution. Par exemple, si la prolactine de poisson est d'un point de vue moléculaire bien reliée à celle des mammifères, elle ne joue aucun rôle évidemment dans la lactation comme chez les mammifères. Dans la comparaison du développement précoce entre mammifères et vertébrés inférieurs, doit-on prendre comme caractère commun un stade morphologique apparemment similaire comme le stade blastula mais fonctionnellement peu pertinent, ou/et le MZT (*Maternal Zygotic Transition*) qui intervient à des stades très variables (deux cellules chez la souris vs. mille chez le xénope) ? Il n'y a donc pas de règle générale dans la manière de choisir et de définir les caractères phénotypiques génériques, mais il faudra trouver, au cas par cas, le compromis entre la satisfaction intellectuelle et la finalité affichée de la base. Si l'on choisit de développer une base ontologique générique des caractères phénotypiques, des bases ontologiques des conditions d'élevage et de mesures séparées devront lui être associées pour permettre la constitution de bases de données exploitables à des fins de sélection ou de modélisation. En transposant, quand cela est possible, les caractères phénotypiques retenus pour les animaux de rente à des espèces de laboratoire à cycle de vie court, et pour lesquels les connaissances sont plus étendues (souris, rat, poisson zèbre ...), les connaissances des liens existant entre génotype et phénotype devraient être accélérés.

## 2.4. LOGISTIQUE

La mise en correspondance et la compatibilité de la base avec des ressources existantes, doivent être une préoccupation constante. Pour être pertinente et reconnue internationalement, la base doit s'appuyer sur un réseau dense d'experts couvrant l'ensemble de ses champs thématiques. Ce réseau doit être structuré avec un nombre de « curateurs » (personne chargée de la collecte et de la validation de la définition des caractères phénotypiques avant leur introduction dans une base d'ontologie) réduits et spécialisés afin d'éviter l'apparition de redondances et / ou de synonymies préjudiciable à la crédibilité de la base. Ce collectif doit porter à la fois des compétences très en amont tout en étant capable de les positionner dans un cadre plus appliqué. Les chercheurs et ingénieurs de l'INRA, comme ceux de l'USDA, qui ont cette double culture, sont donc particulièrement désignés pour construire un tel projet. Mais leur mobilisation doit être constante pour l'enrichir et le faire évoluer sous peine de voir la base se déprécier au cours du temps. Dans un premier temps, les institutions et les bénéficiaires de recherches appliquées susceptibles de s'appuyer sur la base devront soutenir financièrement et sous forme de projets scientifiques, la mobilisation de la communauté scientifique autour de cet outil pour permettre son appropriation, son développement et son utilisation par les chercheurs et les ingénieurs des sciences animales.

## CONCLUSION

Le phénotypage des animaux est un défi que toutes, unités expérimentales et unités de recherche, doivent relever selon un cahier des charges très précis et standardisé, afin de rattraper le train du génotypage à haut débit. Le but est d'établir le plus exhaustivement et le plus précisément possible des relations génotype / phénotype les plus solides possibles.

**Goddard, M.E., Hayes, B.J. 2009.** *Nat. Rev. Genet.*, 10, 381-391

**Hughes L. M., Bao J., Hu Z.-L., Honavar V., Reecy J. M. 2008.** *J. Anim. Sci.*, 86, 1485-1491

**Reecy J.M., Park C.A., Hu Z.-L., Hulsegge I., Van Der Steen H., Hocquette J.F., 2009.** *Book of Abstracts of the 60th Annual Meeting of the European Association for Animal Production, Barcelona 2009, p. 319. Theatre 1*