

# | De la génétique à la génomique

J.-P. BIDANEL<sup>1</sup>, D. BOICHARD<sup>1</sup>, C. CHEVALET<sup>2</sup>

<sup>1</sup> INRA, UR337 Génétique Quantitative et Appliquée, F-78350 Jouy-en-Josas, France

<sup>2</sup> INRA, UMR444 Génétique Cellulaire, F-31326 Castanet-Tolosan, France

Courriel : jean-pierre.bidanel@jouy.inra.fr

L'essor de la génomique bouleverse depuis une vingtaine d'années la recherche en génétique et en biologie. L'INRA a été un acteur majeur du développement de cette discipline qui révolutionne nos connaissances sur le génome des animaux d'élevage et renouvelle les méthodes d'amélioration génétique et de gestion des populations animales.

La génétique est depuis une vingtaine d'années une discipline très médiatique, en raison notamment des espoirs de traitement des maladies chez l'Homme, mais aussi des craintes qu'elle suscite. Il est vrai que nos connaissances sur les gènes et leur fonctionnement ont explosé au cours des deux dernières décennies, avec notamment le séquençage complet du génome de l'Homme et de plusieurs autres organismes animaux, végétaux et microbiens. Pourquoi une telle révolution dans une discipline dont l'origine remonte à plus de 150 ans et à la découverte par Mendel des règles de transmission d'informations héréditaires entre générations ? Sans nul doute à la fois parce que la génétique est une discipline centrale de la biologie, mais aussi et surtout parce que les progrès conjoints de la biologie, de la physique et de l'informatique ont conduit à l'émergence d'une nouvelle discipline, la génomique, dont l'objectif est l'étude exhaustive des génomes, en particulier de l'ensemble des gènes, de leur disposition sur les chromosomes, de leur séquence, de leur fonction et de leur rôle. L'essor de la génomique a en particulier été rendu possible par un certain nombre de découvertes clés et par une nouvelle approche, systématique et technologique, de l'étude du génome, qui a nécessité le développement d'infrastructures spécialisées d'usage commun (centres de génotypage et de séquençage, centres de ressources biologiques, bases de données communes...) qui ont bouleversé la façon de mener les recherches en biologie.

Cet essor a largement été impulsé par le secteur de la génétique humaine, pour lequel les perspectives formidables, notamment sur le plan thérapeutique, qu'offrirait la génomique, ont permis de mobiliser des moyens consi-

dérables afin de décrypter le génome humain, puis d'en étudier les variations entre individus et le fonctionnement. Ces avancées ont largement contribué à celles obtenues chez les animaux d'élevage, grâce aux progrès technologiques qu'elles ont permis (mise au point de méthodes, de techniques et d'outils) et à l'apport des connaissances obtenues chez l'Homme pour l'étude des génomes animaux, en particulier ceux, relativement proches, des mammifères.

Malgré tout, les progrès obtenus chez l'Homme ne sont que rarement directement transposables chez les animaux, à la fois parce que le polymorphisme de l'ADN diffère entre espèces et que les objets d'étude sont souvent différents : maladies génétiques chez l'Homme, caractères d'intérêt zootechnique chez les animaux, même s'ils se rejoignent parfois, comme pour l'étude de la fonction de reproduction. Des programmes spécifiques ont donc été développés en vue de mettre au point des méthodes et des outils de plus en plus performants d'étude de la structure et du fonctionnement du génome, puis de les utiliser pour mieux connaître et gérer la variabilité génétique des caractères d'intérêt dans les populations d'animaux de rente. L'INRA a participé dans une proportion importante à ces programmes et largement contribué aux principales avancées scientifiques de la génétique des animaux d'élevage aux cours des deux dernières décennies. Cet article se propose de dresser un historique des principaux faits marquants en matière d'étude de la structure et du fonctionnement du génome, puis d'application de ces nouveaux outils pour la gestion des populations d'animaux d'élevage, en y décrivant la contribution de l'INRA.

## 1 / Les premiers travaux de cartographie du génome des espèces d'élevage

La localisation des gènes sur leur support physique, la molécule d'ADN, constitue un des objectifs majeurs des travaux de cartographie des génomes. Jusque dans les années 80, l'étude directe du polymorphisme de l'ADN était extrêmement difficile. Les travaux de génétique moléculaire reposaient d'une part sur l'analyse phénotypique de quelques systèmes particulièrement polymorphes comme les groupes sanguins, le complexe majeur d'histocompatibilité (CMH) ou des systèmes protéiques du sang et du lait, d'autre part sur les outils de cytogénétique (Grosclaude *et al* 1996).

### 1.1 / Les premières cartes génétiques

La situation a changé après la proposition de Botstein *et al* (1980) de constituer des cartes génétique à partir du polymorphisme lié à des différences de longueur des fragments d'ADN générés par les enzymes de restriction découvertes dans les années 70. Les premières cartes de marqueurs RFLP ont vu le jour dans les années 80 chez l'Homme et les espèces de laboratoire et ont permis les premières localisations de gènes.

Le véritable essor de la cartographie génétique des animaux d'élevage n'a toutefois eu lieu qu'à la suite de deux avancées majeures : 1/ la mise au point de la réaction en chaîne par polymérase (ou PCR, abréviation anglophone de *Polymerase Chain Reaction*), qui permet de copier en grand nombre une

séquence d'ADN donnée ; 2/ l'utilisation des marqueurs microsatellites. Ces nouveaux marqueurs, dont le polymorphisme est dû à des variations du nombre de courtes séquences répétées (2 à 4 paires de bases), présentent l'avantage par rapport aux marqueurs RFLP d'être abondants (tous les 50 à 100 kilobases chez les mammifères) et très polymorphes (6 à 7 allèles en moyenne). De plus, leur polymorphisme peut aisément être déterminé de façon automatisée, favorisant ainsi le développement d'analyses à grande échelle (Pitel et Riquet 2000). Ils sont de ce fait rapidement devenus les marqueurs de prédilection chez l'Homme, la souris et les animaux d'élevage.

Les microsatellites ont en effet été les marqueurs de référence des premières cartes génétiques des animaux d'élevage développées en Europe et aux USA à partir du début des années 90. Les cartes européennes ont été développées dans le cadre de projets associant l'INRA et les principaux autres laboratoires européens (Roslin Institute (RU), Université de Wageningen (PB) et d'Uppsala (SU)...): PiGMap (1991-1994) chez le porc, BOVMAP (1993-1995) chez les bovins, puis CHICKMAP (1995-1998) chez le poulet et SALMAP (1997-1999) chez les salmonidés (Levéziel et Crihiu 2000, Milan *et al* 2000a, Vignal 2000, Gharbi *et al* 2006). Les premières cartes génétiques étaient disponibles chez les bovins et le porc dès le milieu des années 90 (Archibald *et al* 1995). A cette même période, l'INRA développait également une carte du génome de la chèvre (Vaiman *et al* 1996a), participait au consortium international sur la carte ovine et jouait un rôle moteur dans la construction de la première carte équine (Guérin *et al* 1999). Enfin, dans les années 2000, l'INRA prend en charge le développement de cartes génétiques d'espèces moins étudiées comme le lapin, le canard et la caille (*e.g.* Chantry-Darmon *et al* 2006).

## 1.2 / Les cartes chromosomiques

L'information apportée par les cartes génétiques ne permettant pas à elle seule de positionner les marqueurs ou les gènes sur les chromosomes, des travaux ont été conduits en parallèle pour obtenir cette localisation physique. Ils reposaient sur deux principaux outils : les hybrides somatiques et la carte cytogénétique (Yerle 2000). Les hybrides somatiques interspécifiques sont obtenus par fusion, dans certaines conditions, de cellules somatiques de l'espèce étudiée (lignée primaire), avec celles

d'une lignée cellulaire établie d'une autre espèce (hamster ou rat). Les lignées de cellules hybrides ainsi établies perdent aléatoirement des chromosomes ou des fragments chromosomiques de l'espèce étudiée. Une collection d'hybrides, si elle est bien faite, permet de localiser des gènes ou des marqueurs par hybridation de sondes, tests d'expression enzymatique, etc. Des collections (panels) d'hybrides ont été développées à l'INRA chez les bovins (Guérin *et al* 1994) et le porc (Yerle *et al* 1996). Distribué dans une trentaine de laboratoires européens et américains, le panel d'hybrides porcins a ainsi permis de localiser régionalement plus de 400 marqueurs sur le génome du porc. Des cartes cytogénétiques ont été établies par hybridation *in situ* d'une sonde contenant le fragment d'ADN que l'on voulait localiser sur des chromosomes entiers en métaphase. La méthode, baptisée FISH (*Fluorescence In Situ Hybridization*) lorsque la sonde est marquée par un fluorochrome et les signaux d'hybridation détectés par un microscope à fluorescence, permet une résolution de 3 à 10 cM. La résolution peut être augmentée en travaillant sur des chromosomes interphasiques ou des chromosomes étirés (peignage moléculaire, Yerle 2000).

## 1.3 / Les premières cartes comparées : émergence de la génomique comparative

Très tôt, la cartographie de gènes dans les espèces animales a révélé des «conservations de synténies», c'est-à-dire l'existence de groupes de gènes liés (portés par un même chromosome de l'espèce) dont les homologues humains étaient portés également par un seul chromosome humain. Les équipes INRA ont particulièrement participé à l'acquisition de ces données, en développant la cartographie chromosomique par les approches de cytogénétique (hybridation *in situ* de sondes de gènes sur les chromosomes).

Ces travaux de cartographie comparée ont permis de déterminer les régions chromosomiques homologues conservées dans différentes espèces, au niveau des génomes entiers, puis à des échelles de plus en plus fines. Ainsi, l'hybridation sur les chromosomes d'une espèce de sondes moléculaires complexes représentant les chromosomes entiers d'une autre espèce a permis de révéler les segments de chromosomes portant des gènes homologues dans les différentes espèces (Homme,

souris, bovin, porc ; Hayes 1995, Yerle *et al* 1996, Hayes *et al* 2003, Schibler *et al* 2006). Ces travaux se sont ensuite poursuivis en fonction de l'évolution des approches génomiques (cartographie sur hybrides d'irradiation, puis séquençage des génomes), pour déboucher sur la «génomique comparative».

Outre son intérêt en matière d'évolution des génomes, l'identification de régions homologues permet, moyennant certaines hypothèses, de transposer les connaissances acquises dans des espèces bien étudiées, comme l'Homme ou la souris, aux espèces d'élevage, moins bien connues ou, à un autre niveau, d'utiliser les résultats obtenus chez les bovins pour l'étude des génomes ovins ou caprins, relativement proches. Cette approche est aujourd'hui utilisée de façon systématique pour la recherche de mutations causales de variations génétiques, et intégrée dans les méthodes d'annotation des génomes nouvellement séquencés.

## 2 / L'essor de la génomique

Les avancées technologiques de l'analyse de l'ADN ont progressivement conduit à l'idée de ne plus travailler sur les gènes pris un par un, mais d'aborder globalement l'étude de la structure de l'ensemble du génome et de son expression : c'est le passage de la génétique à la génomique. Cette nouvelle discipline inclut l'étude de la structure des génomes (la détermination des séquences d'ADN), de son expression (établissement des répertoires des gènes exprimés – le transcriptome, et des protéines – le protéome), de sa variabilité (étude des polymorphismes), de son évolution (génomique comparative, génomique évolutive).

Les attentes suscitées par cette nouvelle approche ont conduit en France, dans le domaine animal, au lancement du programme AGENAE, initié par l'INRA et mené en partenariat avec les filières de l'élevage avec les soutiens de l'Etat (Ministère chargé de la Recherche, ANR) et de l'Europe. Il a pour objectif de développer des recherches génériques (cartographie des génomes, répertoire de transcrits, diversité génétique) et des actions de recherche finalisées chez les bovins, les porcins, la poule et la truite ([www.inra.fr/agenae](http://www.inra.fr/agenae)). En appui à ces projets de recherche, l'INRA a mis en place des structures nationales spécifiques aux animaux (centre de ressources GADIE, groupe bioinformatique

SIGENAE) complémentaires des infrastructures nationales de génomique (centres nationaux de génotypage et de séquençage à Evry, réseau des génotypes régionaux).

## 2.1 / Les cartes à haute densité : l'organisation du génome

Assez rapidement, il est apparu que la résolution des cartes génétiques, limitée à 1-3 cM du fait notamment du nombre d'événements de recombinaison dans les familles de référence utilisées, était insuffisante pour ordonner des marqueurs très proches. De nouveaux outils permettant d'obtenir des cartes à haute densité ont donc été développés à partir de la fin des années 90 : les plus couramment utilisés sont les cartes d'hybrides d'irradiation, les banques de grands fragments et, lorsqu'elles deviennent disponibles, les informations sur les séquences d'ADN elles-mêmes. Les niveaux de résolution des différents types de cartes sont présentés sur la figure 1.

### a) Les cartes d'hybrides d'irradiation

La construction d'hybrides d'irradiation (ou RH pour *Radiation Hybrids*)

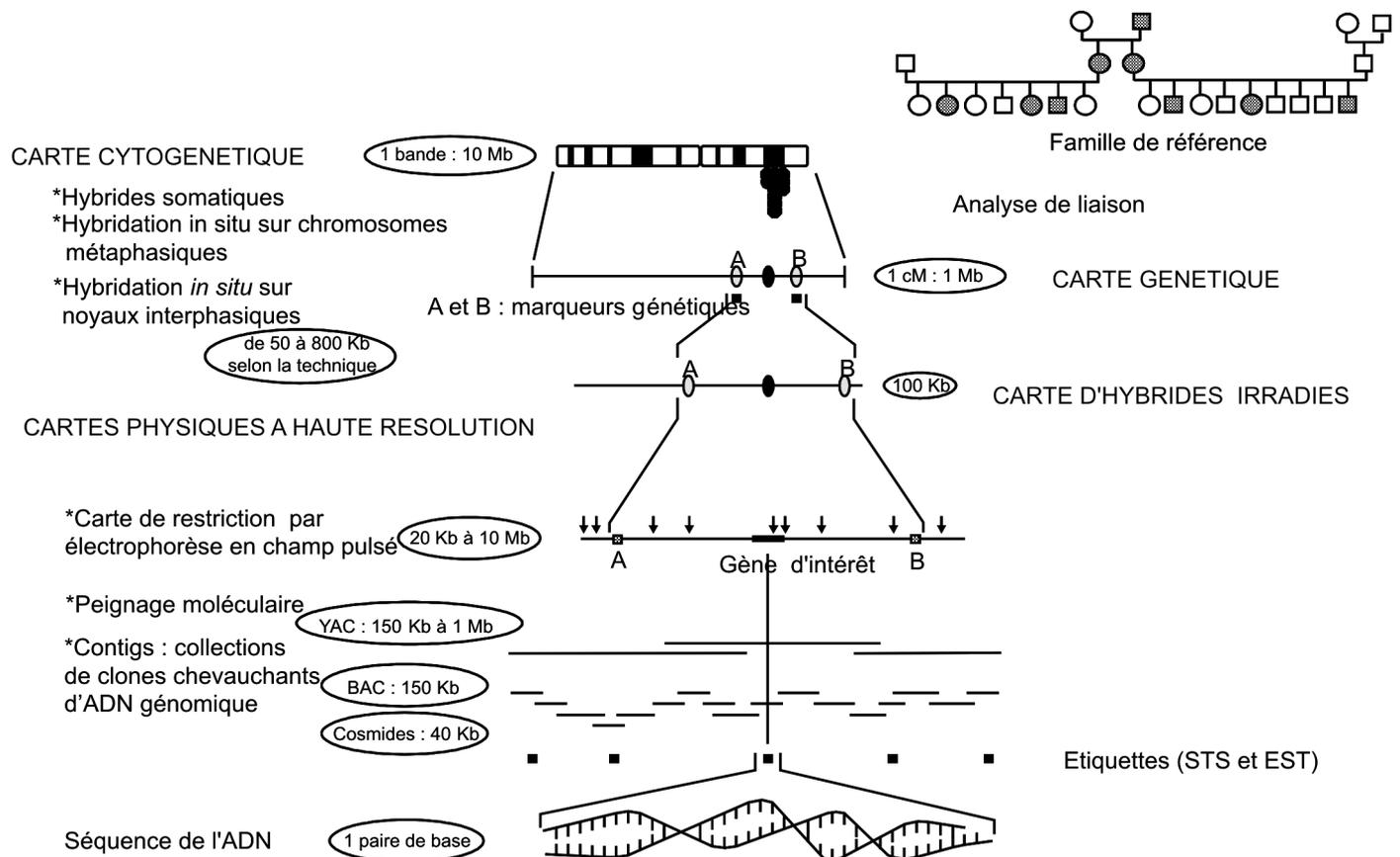
repose sur le même principe que celui des hybrides somatiques évoqué ci-dessus, mais les cellules de la lignée primaire (celles de l'espèce étudiée) sont soumises à une irradiation qui fragmente les chromosomes avant fusion avec les cellules de hamster ou de rat. La fréquence de la cassure, et donc la longueur des fragments de chromosomes et la résolution de la carte, sont fonction de l'intensité de l'irradiation. Les cartes sont établies par analyse statistique de la fréquence des cassures entre locus sur l'ensemble de la collection d'hybrides (panel RH) : plus deux locus sont proches, plus les cassures sont rares entre eux et plus ils ont de chance d'être retrouvés dans les mêmes clones. En pratique, la résolution de ces cartes est de 10 à 20 fois supérieure à celle des cartes génétiques (Yerle 2000). De plus, les locus n'ont pas besoin d'être polymorphes pour pouvoir être ordonnés. L'INRA a développé les deux panels RH disponibles chez le porc, l'un à 7000 rads en 1996 (Hawken *et al* 1999), l'autre à 12 000 rads (Yerle *et al* 2002), le panel RH du poulet (à 6000 rads, Jennen *et al* 2004, Morisson *et al* 2007), un des deux panels RH ovin (à 12 000 rads, Laurent *et al* 2007) et a participé à la construction de l'un des 4 panels RH bovins

actuellement disponibles. Ces panels permettent d'ordonner sans erreur des milliers de séquences sur un génome, ou d'établir des milliers de points d'ancrage entre génomes, permettant alors de définir précisément leur structure comparée.

### b) Les banques de grands fragments

Pour pouvoir être étudié, l'ADN d'un chromosome, de l'ordre de 150 millions de bases (Mb) chez les mammifères, doit au préalable être coupé en fragments de taille manipulable, qui sont ensuite clonés, c'est-à-dire insérés dans un vecteur, bactérie ou levure, qui permet de les conserver et de les multiplier. Jusqu'à la fin des années 80, les vecteurs utilisés, les plasmides ou les cosmides, permettaient d'insérer des fragments de quelques centaines à quelques dizaines de milliers de paires de bases et de travailler à l'échelle du gène. Au cours des années 90, la mise au point de nouvelles techniques permettant de cloner des fragments de plusieurs centaines de milliers de paires de bases dans des levures (les YAC pour *Yeast Artificial Chromosome*) ou des bactéries modifiées (BAC pour *Bacteria Artificial Chromosome*) a permis d'étudier des

Figure 1. Les différents types de cartes cytogénétiques et physiques et leurs résolutions (Yerle 2000).



petites régions chromosomiques pouvant contenir plusieurs gènes et de créer des collections ou banques de clones chevauchants contenant l'ensemble du génome des espèces étudiées. Compte tenu des chevauchements entre clones, la quantité d'ADN contenue dans une banque doit être supérieure à la taille du génome pour avoir une probabilité suffisante de retrouver un fragment d'ADN unique quelconque dans au moins un clone. Cette probabilité est par exemple de 60, 80 et 99% pour des banques équivalant à un, deux et dix génomes (soit 1, 2 et 10X). Une banque de grands fragments représente donc plusieurs dizaines de milliers de clones (30 000 clones pour une banque 10X de YAC de 1000 kb, 200 000 clones pour une banque de BAC de 150 kb) et nécessite une très bonne organisation pour assurer la gestion des collections de clones, la détection (ou criblage) et la mise à disposition des clones intéressants. Les équipes INRA ont développé dans la seconde moitié des années 1990 des banques de grands fragments de plusieurs espèces d'élevage (bovin, porc, mouton, chèvre, lapin, cheval) et un atelier de grands fragments a été créé en 1999 pour construire et gérer ces banques et celles de partenaires, les cribler et fournir des clones à la demande (Centre de Ressources Génomiques des Animaux d'Intérêt pour l'Élevage «GADIE», Jouy-en-Josas, Rogel-Gaillard 2000).

Une fois les banques constituées, l'étape suivante a consisté à aligner les grands fragments d'ADN contenus dans les BAC ou les YAC pour constituer des ensembles ordonnés de clones ou «contigs» et, *in fine*, construire des cartes physiques à haute densité de l'ensemble du génome. Différentes méthodes (analyse de fragments de restriction (*fingerprinting*), hybridation croisée, recherche de marqueurs ou fragments de séquence communs déjà positionnés, séquençage d'extrémités) ont été utilisées pour repérer les chevauchements et constituer des contigs de taille croissante afin d'obtenir progressivement l'alignement de l'ensemble des clones d'une banque. Ces investissements ont permis de donner une place significative aux équipes françaises dans les consortiums internationaux visant le séquençage complet des génomes des bovins et du porc, en participant à ces efforts internationaux (Meyers *et al* 2005, Humphray *et al* 2007 chez le porc ; Schibler *et al* 2004, Snelling *et al* 2007 chez le bovin).

### c) L'intégration des cartes

En pratique, les différents outils évoqués ci-dessus ne sont pas utilisés isolément. Le positionnement de marqueurs et de gènes sur les cartes d'hybrides d'irradiation est par exemple très utile pour l'assemblage des clones de grands fragments. De même, les cartes génétiques et d'irradiation apportent des informations complémentaires sur l'ordre entre les gènes et sur leurs distances. Des données sûres disponibles dans d'autres espèces de référence (complètement séquencées) peuvent également être utilisées pour affiner la cartographie dans une nouvelle espèce.

L'ensemble de ces opérations nécessite évidemment des moyens informatiques importants pour gérer l'ensemble des données réunies, calculer les cartes les plus probables et mettre les résultats à disposition de la communauté scientifique. Les équipes INRA ont ainsi développé des logiciels originaux de cartographie (CARTHAGENE, de Givry *et al* 2005), et créé des sites internet ouverts à la communauté internationale (carte bovine BOVMAP, cartes RH chez le porc et le poulet).

## 2.2 / Le séquençage des génomes

Le séquençage de l'ADN consiste à déterminer l'ordre d'enchaînement des nucléotides d'un fragment d'ADN donné. Jusqu'à aujourd'hui, pour les grands génomes, il est la plupart du temps réalisé par la méthode dite de Sanger, qui utilise la réaction de polymérisation de l'ADN avec une enzyme et des di-désoxyribo-nucléotides marqués par des fluorochromes différents, qui bloquent l'élongation de la molécule d'ADN et génèrent des fragments de taille variable mis en évidence par électrophorèse.

Les premiers grands projets de séquençage, chez l'Homme puis chez les autres espèces, ont d'abord concerné la partie exprimée du génome, en constituant des banques de copies (ADNc) des ARN messagers exprimés dans différents tissus, puis en séquençant des fragments de ces ADN (des «EST», acronyme pour *Expressed Sequence Tags*).

Les projets de séquençage de génomes complets ont d'abord concerné les bactéries (quelques millions de bases) avant d'aborder le séquençage des génomes eucaryotes, qui n'a été envisageable qu'en automatisant à l'extrême les opérations de préparation des

échantillons et de lecture des électrophorèses à l'aide de séquenceurs automatiques.

### a) Le séquençage des gènes exprimés

Chez les mammifères, on estime que seulement 5% du génome correspondrait à de l'ADN codant, avec un nombre de gènes qui serait compris entre 20 000 et 50 000. En outre, si le contenu en ADN du noyau de toutes les cellules d'un organisme est équivalent, on estime que chaque type cellulaire n'exprime que 10% des gènes, ceux qui sont nécessaires à sa fonction. Les premiers travaux de séquençage à grande échelle se sont donc concentrés sur cette partie du génome, a priori la plus significative. Cette exploration du transcriptome (l'ensemble des gènes transcrits en ARN) a fait appel à différentes approches pour construire les banques d'ADNc à séquencer : banques représentatives de tissus particuliers (glande mammaire, foie, ovaire, embryon – pour les travaux conduits à l'INRA) ; banques représentatives des gènes exprimés de façons différentes dans différentes conditions (méthodes subtractives «SSH») ; banques multi-tissus cherchant à représenter une forte proportion des gènes exprimés par un organisme (Hatey 2000). De très grandes collections de séquences «EST» ont ainsi été produites (par millions chez l'Homme, par centaines de milliers chez le poulet, le bovin, le porc, la truite...). Cependant, on sait que les processus complexes de transcription de l'ARN, liés à l'existence de plusieurs promoteurs et de sites de polyadénylation, à des épissages alternatifs, etc. aboutissent à la fabrication de différents ARN messagers à partir d'un seul gène. Le séquençage des EST, représentant en général un fragment seulement de l'ADNc, est insuffisant pour acquérir une connaissance approfondie du transcriptome : des projets spécifiques de séquençage des ADNc, «pleine longueur» ont ainsi été conduits (notamment au Génoscope, le Centre National de Séquençage à Evry). Cette connaissance s'est révélée déterminante dans l'annotation des séquences complètes du génome.

L'exploration du transcriptome est le moyen de connaître la réalité de l'expression du génome, elle permet ensuite de caractériser le profil d'expression d'un gène dans plusieurs tissus, et inversement d'établir des profils d'expression des différents tissus pour un ensemble de gènes : c'est l'objet de la génomique fonctionnelle.

Les équipes de l'INRA se sont investies massivement pour produire des séquences EST chez le bovin, le poulet, le porc et surtout la truite arc-en-ciel dans le cadre du projet AGENAE, au début des années 2000. Cela a permis de construire des répertoires importants de gènes chez ces espèces pour réaliser des «puces à ADN» et entamer des études de génomique fonctionnelle (cf. <http://www.inra.fr/agenae/doc.html>). Dans le cas de la truite, l'apport INRA représente la moitié des données réunies au niveau international, à égalité avec l'USDA aux Etats-Unis. De plus, l'ensemble des données EST des animaux d'élevage est mis à disposition de la communauté scientifique au moyen d'un système d'information (SIGENAE, [www.sigena.org](http://www.sigena.org)) développé dans le cadre AGENAE avec l'aide des partenaires de l'élevage (filières bovines et piscicole) et de l'Europe.

#### b) Le séquençage génomique

Le premier génome eucaryote séquencé complètement – celui de la levure – a été réalisé dans le cadre d'un grand projet européen dans le courant des années 90, en combinant les données acquises par un consortium de laboratoires. Des génomes eucaryotes de «petite taille» ont ensuite été séquencés en raison de leur valeur de modèle (le nématode *Caenorhabditis elegans*, la plante *Arabidopsis thaliana*).

Le génome de l'Homme a été le premier grand génome complexe séquencé dans le cadre du Projet Génome Humain (PGH), qui a été lancé en 1989 et s'est achevé en 2004, 3 ans après la première publication d'une séquence brute en 2001. Initialement réalisé par un consortium international de centres de séquençage, dont le Centre National de Séquençage d'Evry en France chargé du séquençage du chromosome 14, il était basé sur le séquençage systématique de grands fragments (BAC ou YAC) assemblés en parallèle. A partir de 1998, le PGH a subi la concurrence d'une société privée, Celera Genomics, utilisant une technique de séquençage de fragments aléatoires dite de «shotgun», basée sur une robotisation extrême et un assemblage purement informatique des séquences. En pratique, les 2 projets ont achevé la production de la première séquence brute au même moment (*International Human Genome Sequencing Consortium* (IHGSC) 2001, Venter *et al* 2001).

Une fois la séquence du génome humain achevée, les centres de séquen-

çage ont entrepris le séquençage d'autres génomes complexes en tirant parti de leur expertise et des connaissances acquises chez l'Homme. Après ceux de la souris, du rat, les génomes de la poule, du chien, du bovin et du cheval sont considérés comme séquencés et le génome du porc, du mouton et du lapin sont en cours. Les projets utilisent de façon combinée le séquençage classique et par Shotgun en tirant parti de toutes les informations disponibles sur les génomes déjà séquencés.

Avant d'aborder le séquençage complet des génomes, des régions d'importance majeure avaient fait l'objet d'un séquençage génomique : par exemple, la région du complexe majeur d'histocompatibilité du porc a été séquencée en France au Génoscope dans une collaboration avec une équipe mixte INRA/CEA (Renard *et al* 2006). Ensuite, la France n'a pas participé directement au séquençage complet du génome des animaux d'élevage, mais les équipes INRA ont contribué de façon importante aux phases préliminaires. La réalisation de cartes à haute densité, des cartes RH, la construction et la cartographie physique des banques de grands fragments (BAC), le séquençage partiel de ces BAC (réalisé au Génoscope) leur ont valu de participer aux consortiums internationaux de séquençage. Au-delà même de ces apports initiaux, cette expertise se révèle toujours importante dans les étapes de validation des assemblages des séquences, notamment chez le poulet avec un travail en cours poursuivi par une équipe INRA en collaboration avec le Centre National de Génotypage (CNG, Evry) et avec le soutien de l'ANR.

#### c) Vers le très haut débit

De nouvelles technologies de séquençage à très haut débit et à moindre coût sont apparues récemment, avec un objectif tel que «séquencer un génome humain pour 1000 dollars», alors que le coût total du projet Génome Humain a été estimé à 2,7 milliards de dollars (valeur 1991), que le seul séquençage du chromosome 14 par le Génoscope a coûté environ 10 millions d'euros, et que le coût actuel du séquençage d'un génome entier de mammifère est de 20 à 50 millions de dollars. Ces nouvelles technologies ouvrent des possibilités encore inexplorées, notamment pour l'analyse de la variabilité génétique au sein des populations et pour l'analyse du transcriptome, et laissent penser que «toutes» les espèces animales d'élevage pourront bénéficier d'un séquençage complet, une difficulté résidant tout de

même dans l'assemblage de ces séquences.

## 2.3 / Génotypage : l'analyse du polymorphisme

La recherche des polymorphismes génétiques, la cartographie génétique, sont des activités «classiques» en génétique qui n'ont pas attendu l'avènement de la génomique. Cependant la possibilité d'aborder la question globalement au niveau des génomes complets a progressivement renouvelé les façons d'appréhender le polymorphisme des populations. Les possibilités de détecter systématiquement des variants de l'ADN, partout dans le génome et avec une densité élevée (avec la découverte des marqueurs microsatellites notamment), et les développements de la technologie (PCR, séquençage de l'ADN) ont permis de changer complètement d'échelle : la carte génétique humaine produite par le laboratoire Généthron (Weissenbach *et al* 1992) doublait le nombre de marqueurs disponibles en apportant un millier de nouveaux marqueurs très informatifs, puis en comptait rapidement plus de 5000 (Dib *et al* 1996) permettant de localiser facilement les gènes impliqués dans les maladies héréditaires.

La recherche animale à l'INRA a pris très vite la même voie, en développant des marqueurs dans le cadre des programmes européens, en entreprenant les grands projets de détection de QTL dans le courant des années 90 (cf. § 3), et en mettant en place des ateliers de génotypage à haut débit (LABOGÉNA, plate-forme Génopole de Toulouse).

#### Un nouveau monde : les «SNP»

Cependant, l'appréhension du polymorphisme connaît actuellement une importante mutation. Chez l'Homme, la disponibilité de la séquence complète du génome, l'existence de collections très importantes de séquences fragmentaires de gènes (les EST), la redondance des séquences produites pour réaliser l'assemblage de la séquence complète, ont permis de révéler un nombre très élevé de sites variables, au niveau de la population mondiale. Quand la variation concerne une seule base nucléotidique, on parle d'un «SNP» (acronyme pour *Single Nucleotide Polymorphism*). Des millions de tels SNP ont été révélés le long du génome humain : c'est la nouvelle classe de marqueurs génétiques désormais utilisée en génétique humaine. Le projet HAPMAP a permis de décrire

l'essentiel des variations actuelles du génome humain et de caractériser chaque zone du génome par les fréquences des haplotypes observés (*International HapMap Consortium*, Frazer *et al* 2007). La comparaison de ces fréquences avec certains traits phénotypiques, comme des résistances aux maladies, permet de suggérer l'existence de régions chromosomiques où des allèles favorables ont été sélectionnés (Sabeti *et al* 2007). Parallèlement à ces découvertes, les technologies d'analyse du polymorphisme connaissent un nouveau changement d'échelle : chez l'Homme, différentes technologies offrent la possibilité de faire un génotypage sur 500 000 à 1 000 000 de marqueurs en une seule expérience. Ces technologies s'accompagnent d'une diminution drastique du coût du génotype, d'au moins d'un facteur 100, ouvrant ainsi des perspectives importantes dans l'analyse de la variabilité des populations.

Dans les espèces animales, cette approche a débuté avec la disponibilité progressive de séquences du génome (nombreuses EST, début du séquençage complet), par un projet pilote concernant le chromosome 3 bovin typé sur 1536 marqueurs chez plus de 2000 animaux (collaboration INRA-CNG). D'autres programmes ont suivi avec une analyse « pangénomique » en cours sur 60 000 marqueurs pour préparer une nouvelle version de la Sélection Assistée par Marqueurs (SAM) chez les bovins laitiers (cf. § 4), et divers projets chez la poule, le porc et le cheval.

## 2.4 / La génomique fonctionnelle

La génomique fonctionnelle cherche à appréhender l'expression de l'ensemble des gènes dans un tissu, à un stade particulier ou dans des conditions physiologiques particulières. Dans une cellule, l'information portée par l'ADN est transcrite sous forme d'ARN puis traduite en protéines. En parlant de génomique fonctionnelle, on généralise les approches usuelles de la génétique moléculaire en se proposant d'examiner simultanément l'expression de tous les gènes. Des programmes d'étude globale de l'expression des gènes à l'échelle des ARN (programmes transcriptome) ainsi que des protéines (programmes protéome) sont actuellement mis en place. Ils constituent en France, pour les animaux d'élevage, un des volets majeurs du programme AGENAE.

Les approches « transcriptome » permettent 1/ de cibler l'étude des génomes à la seule portion codante, 2/ de cibler les tissus responsables de l'expression du caractère étudié, et 3/ de prendre en compte l'existence d'une variabilité dans les produits de l'expression des gènes non identifiables au niveau de l'ADN. Les différentes technologies d'étude du transcriptome développées (Hatey 2000) nécessitent toutes : 1/ la sélection d'un tissu voire d'un type cellulaire, 2/ l'extraction de la population des ARN transcrits dans ces cellules, et 3/ la synthèse sous forme d'ADN d'un brin complémentaire synthétique de ces ARN (ADNc). Différentes techniques permettent ensuite de comparer deux à deux des populations d'ARN correspondant à des situations physiologiques ou pathologiques très différentes. D'autres techniques d'étude systématique ont pour objectif d'identifier « tous » les gènes exprimés dans un type cellulaire, et de pouvoir ainsi visualiser dans leur ensemble des cascades de signalisation et les interactions entre gènes.

La technique actuellement la plus utilisée est l'analyse systématique de l'expression des gènes à l'aide de bibliothèques ordonnées : son principe est de fixer sur un support (membranes, lames de verres...) des sondes ADN représentatives de chaque gène d'un génome (ou d'une sélection de gènes correspondant à une fonction physiologique particulière) ; ce sont les « puces à ADN ». La population d'ADNc issue de l'échantillon étudié est alors hybridée sur ce support : les signaux d'hybridation obtenus permettent d'identifier les transcrits présents dans le tissu, et l'intensité de ces signaux permet de déterminer le niveau d'expression du gène (Hatey 2000). Dans les années à venir, l'exploitation des données du transcriptome et l'utilisation de ces technologies, combinée aux approches de cartographie génétique, devraient permettre de progresser dans la cartographie et l'identification de certains QTL (cf. § 3.2).

Les programmes AGENAE de séquençage d'EST (cf. § 2.2 a) ont permis de construire des puces au Centre de Ressources GADIE mises à disposition de la communauté (<http://www-crb.jouy.inra.fr/>), et d'offrir les outils d'analyse informatique (groupe SIGENAE, <http://www.sigena.org/>).

Comme en matière de génotypage et de séquençage, l'évolution technologique est très rapide. Chez l'Homme, il

existe des puces capables de repérer tous les gènes potentiellement exprimés, distribués par de grands groupes industriels susceptibles de proposer bientôt les mêmes services pour les espèces d'élevage. L'évolution en cours des techniques de séquençage à très haut débit pourrait aussi ouvrir de nouvelles possibilités d'analyse du transcriptome.

## 2.5 / Bioinformatique

La génomique génère des masses considérables de données, qu'il faut gérer, maintenir, et – surtout – interpréter pour leur donner du sens. Les données concernant les séquences de gènes et de génomes sont gérées par de grands centres internationaux (bases de données EMBL, GenBank et DDBJ, gérées, respectivement, par l'*European Bioinformatics Institute* de Cambridge pour l'Europe, le *National Center for Biology Information* pour les USA et le *National Institute of Genetics* pour le Japon) qui offrent également un certain nombre d'outils de visualisation et d'analyse multibases comme « ENSEMBL » (<http://www.ebi.ac.uk/ensembl/>) ou « ENTREZ » (<http://www.ncbi.nlm.nih.gov/>).

Un des objectifs majeurs de la bioinformatique est d'aider à l'annotation des génomes, c'est-à-dire à la recherche des parties significatives dans les séquences complètes (gènes et leurs régions régulatrices, gènes codant pour des protéines ou des ARN), et à la prédiction de la fonction des gènes repérés. Les méthodes font appel à des approches informatiques originales, et s'appuient de plus en plus sur les concepts de la phylogénie des espèces qui montrent l'intérêt de disposer des séquences d'un nombre suffisant de génomes d'espèces ayant des positions particulières dans l'arbre phylogénétique (Corpet et Chevalet 2000). L'annotation reste un problème difficile, parce qu'il existe peu d'espèces où l'on dispose de données expérimentales sûres sur l'expression de nombreux gènes (seule la levure a fait l'objet d'un véritable programme systématique de ce type), et que la prédiction de la fonction d'un gène se fait donc en général par analogie avec ce qu'on connaît de la fonction d'un homologue de ce gène dans une autre espèce.

En France, l'INRA a mis en place les serveurs SIGENAE pour gérer l'ensemble des données produites par les projets AGENAE, accueillir également les données de projets européens, et offrir un certain nombre d'outils d'ana-

lyse : visualisation des séquences de gènes et d'EST, analyse d'expression obtenues sur les puces à ADN, recherche de polymorphisme, visualisation des similarités entre gènes homologues d'espèces différentes, etc. Parmi les productions originales des équipes de bioinformatique de l'INRA (départements MIA, GA et SPE), on peut citer les logiciels CARTHAGENE (cartographie, De Givry *et al* 2005), EUGENE (prédiction des gènes, Foissac *et al* 2003), Iccare (comparaison des gènes exprimés entre espèces, Muller *et al* 2004), Narcisse (comparaison multi-échelle de génomes complètement séquencés, Courcelle *et al* 2008).

### 3 / Cartographie et identification de QTL et de gènes

La localisation, puis l'identification de locus responsables de variations mesurables des caractères d'intérêt constitue un des objectifs majeurs des travaux sur la structure du génome. La localisation primaire d'un gène peut être réalisée par l'étude d'un gène candidat individuel – on reste alors dans une approche de génétique classique – ou par une démarche génomique de recherche systématique de gènes ou de QTL en «balayant» le génome à l'aide d'un réseau de marqueurs couvrant l'ensemble des chromosomes (Bidanel *et al* 1996, Le Roy et Elsen 2000). Une fois les gènes localisés, leur identification requiert en général une étape préalable de cartographie fine afin de réduire les intervalles de localisation à 1-2 cM, puis la recherche des gènes et des polymorphismes candidats dans ces intervalles de localisation (Eggen 2000). Ces étapes n'ont jusqu'à présent pas eu le caractère systématique sur l'ensemble du génome de l'étape de primo-localisation, mais n'en font pas moins largement appel aux outils de génomique.

#### 3.1 / Primo-localisation de QTL et de gènes

##### a) Approche «gène candidat»

L'approche «gène candidat» consiste à rechercher des polymorphismes dans des locus dont on suppose qu'ils affectent le caractère étudié, et à estimer les effets de ces polymorphismes. Le choix d'un gène candidat peut être basé sur des arguments fonctionnels, la connaissance du déterminisme d'un phénotype similaire dans une autre espèce ou son appartenance à une famille de gènes connus. Simple et peu coûteuse, cette

approche a donné lieu à un nombre considérable de publications au cours des deux dernières décennies. Elle n'a toutefois fourni qu'un nombre limité de résultats sur le déterminisme génétique des caractères complexes. On peut toutefois par exemple citer la mise en évidence d'un polymorphisme du gène BMP15 associé au taux d'ovulation chez les brebis de race Lacaune (Bodin *et al* 2007). En revanche, l'approche gène candidat s'est montrée extrêmement efficace pour l'identification de polymorphismes associés à des caractères à déterminisme plus simple, en particulier les gènes de coloration de la robe chez les mammifères ou du plumage chez les oiseaux. On peut en particulier citer les résultats obtenus récemment à l'INRA sur les gènes Agouti (Hiragaki *et al* 2008, Nadeau *et al* 2008), Tyrosinase (Chang *et al* 2006), mélanophiline (Vaez *et al* 2008) ou Silver (Gunnarsson *et al* 2007) chez la volaille, MC1R (Rouzaud *et al* 2000), Agouti (Girardot *et al* 2006) et Silver (Girardot *et al* 2003) chez le bovin ou encore Agouti chez le cheval (Rieder *et al* 2001).

##### b) Primo-localisation de «balayage» du génome

La détection d'un gène ou d'un QTL à l'aide d'un réseau de marqueurs est basée sur la ségrégation conjointe du gène et de marqueurs proches en déséquilibre de liaison. La position du gène et ses effets sur la distribution du caractère analysé sont ensuite estimés à partir de la position et des effets «apparents» des marqueurs sur les caractères (Le Roy et Elsen 2000). En utilisant un réseau de marqueurs régulièrement espacés sur l'ensemble du génome et un dispositif expérimental approprié, on détecte de façon quasi certaine un gène majeur en ségrégation ou les QTL en ségrégation dans la population étudiée dès lors que la taille et donc la puissance du dispositif expérimental sont suffisantes. L'approche est par contre coûteuse et nécessite des moyens de génotypage et de phénotypage importants.

Les premiers programmes ont porté sur la localisation de gènes à effets majeurs. L'INRA a ainsi localisé dès 1995 le gène RN (Milan *et al* 1995) chez le porc ou en 1996 le gène PIS chez la chèvre (Vaiman *et al* 1996b), les gènes Cou Nu et Polydactylie chez le poulet (Pitel *et al* 2000).

En parallèle, les premiers programmes de détection de QTL ont été mis en place, dont un programme bovin et un

programme porcin à l'INRA, avec l'obtention des premières localisations de QTL à partir du milieu des années 90.

Dix ans plus tard, le bilan est impressionnant. L'INRA a mis en place plus de 20 programmes de détection de QTL dans les différentes espèces animales qui relèvent de ses activités. Au total, plus de 1500 QTL porcins et près de 1000 QTL bovins sont recensés dans les bases de données internationales de l'Iowa State University (<http://www.animalgenome.org/QTLdb/>, Hu *et al* 2007). Ces nombres continueront à augmenter dans le futur, dans la mesure où de nombreux programmes restent encore à analyser. Moins de la moitié des programmes INRA sont à ce jour publiés et recensés dans les bases de données (tableau 1).

Ceci étant, le nombre de QTL identifiés permet d'en dégager quelques caractéristiques générales. Ils ont dans la plupart des cas été identifiés à partir de modèles génétiques très simples à un locus. Leur localisation est en général peu précise (10 à 40 centimorgans). Des QTL ont été identifiés pour la quasi-totalité des caractères étudiés, même ceux à très faible héritabilité comme les caractères de reproduction, ou fortement sélectionnés comme les caractères de production laitière ou de composition de la carcasse. La distribution des effets de QTL est en forme de L, avec de nombreux QTL à effets faibles et quelques QTL à effets importants. Plusieurs régions chromosomiques sont communes à des populations différentes, mais de nombreux QTL restent spécifiques à une population.

#### 3.2 / Cartographie fine de QTL et de gènes

L'intervalle de localisation des QTL est incompatible avec une démarche de recherche de gènes candidats ou de clonage positionnel. Des programmes de cartographie fine ont été développés depuis 10 ans afin de réduire l'intervalle de localisation de QTL ou de gènes. L'objectif est de trouver des marqueurs proches et en déséquilibre de liaison avec le locus recherché à l'aide des cartes à haute résolution décrites ci-dessus et de populations ayant «accumulé» suffisamment d'événements de recombinaisons dans l'intervalle de localisation du gène. Les premiers programmes INRA de cartographie fine ont été mis en place pour la localisation fine des gènes majeurs RN chez le porc et «sans corne» chez la

**Tableau 1.** Les principaux programmes de cartographie de QTL conduits à l'INRA. (Programmes ayant donné lieu à une publication avant mars 2008)

Population / structure familiale <sup>1</sup>	Marqueurs <sup>2</sup>	Nom du programme : Caractères (références)
BO : Charolaise, Limousine, Blonde d'Aquitaine, 114 pères, 3554 veaux	26 SNP 21 MS	QUALVIGENE : Croissance, carcasse, caractéristiques du muscle, qualité de la viande (Renand <i>et al</i> 2007)
BO : Holstein, Normande, Montbéliarde 14 pères, 1554 taureaux	169 MS	24 caractères de production, de morphologie et d'élevage (Boichard <i>et al</i> 2003)
BO : F2 Holstein x Normande : 10 ♂ F0, 24 ♀ F0, 10 ♂ F1, 70 ♀ F1, 854 vaches F2	180 MS	Production, transformation en fromage, mammites, facilité de traite, morphologie, coloration de la robe (Larroque <i>et al</i> 2007)
BO : Holstein, 10 pères, 515 taureaux	150 MS	Fertilité mâle, Quantité et qualité de la semence (Basso <i>et al</i> 2005)
CA : F2 entre lignées divergentes sur l'immobilité tonique, 310 F2	248 AFLP	Comportement, caractéristiques neuro-endocriniennes (Beaumont <i>et al</i> 2005)
CA : F2 entre lignées divergentes, 10 pères, 30 mères, 434 ♀ F2	72 MS	Croissance, consommation alimentaire, comportement, production et composition des œufs (Minvielle <i>et al</i> 2005, 2006)
LA : F2 entre lignées 2066, Orylag and Laghmere, 8 pères, 187 F2	178 MS	Croissance, fourrure (Chantry-Darmon <i>et al</i> 2006)
OV : backcross Sarde x Lacaune 10 pères, 100 brebis	142 MS	Production et composition du lait (Barillet <i>et al</i> 2005)
OV, Basco-béarnaise, Lacaune et Manech, 22 pères, 783 descendants	102 MS	Production et composition du lait (Barillet <i>et al</i> 2005)
OV : Romanov, 2 familles, 152 animaux	121 MS	Sensibilité à la tremblante (Cosseddu <i>et al</i> 2002)
OV : INRA 401, 30 pères, 852 mères, 1310 descendants	1	Sensibilité à la tremblante, production (Vitezica <i>et al</i> 2007)
PO : F2 Meishan x Large White 6 ♂ F0, 6 ♀ F0, 6 ♂ F1, 23 ♀ F1, 1103 porcs F2	136 MS	PORQTL : Croissance, carcasse, qualité de la viande, comportement, précocité sexuelle, prolificité (Bidanel <i>et al</i> 2001, Désautés <i>et al</i> 2002, Milan <i>et al</i> 2002, Quintanilla <i>et al</i> 2003, Bidanel <i>et al</i> 2008)
PO : Large White, Landrace ; 8 pères, 239 ♂ / ♀	144 MS	Reproduction femelle, Croissance, Carcasse (Tribout <i>et al</i> 2008)
PO: F2 Duroc x Large White 10 ♂ F1, 32 ♀ F1, 775 animaux	91 MS	Qualité de la viande (gras intra-musculaire, composition en acides gras...), croissance, carcasse (Sanchez <i>et al</i> 2007)
PO: Large White (LW) x Pietrain) x LW 16 pères, 700 descendants	114 MS	Consommation et comportement alimentaire, croissance, carcasse, qualité de la viande (Gilbert <i>et al</i> 2008)
PO: Backcross (MeLiM x Duroc) x Duroc (D) 3 ♂, 6 ♀ (F1), 21 ♀ 5 ♂ (D), 331 animaux	153 MS	Sensibilité au mélanome dans les 3 mois suivant la naissance (Du <i>et al</i> 2007)
PO: Backcross Meishan x Large White 47 pères, 5078 descendants	4 à 8 MS selon région	Croissance, carcasse, qualité de la viande (Demeure <i>et al</i> 2005, Sanchez <i>et al</i> 2006)
PO : Diverses populations 16 + 4 pères 1448 + 382 descendants	16 MS SNP	BIOMARK Croissance, carcasse, qualité de la viande, aplombs, tétines, cortisolémie (Milan <i>et al</i> 2008)
VO : F2 entre lignées divergentes pour la croissance, 50 familles pleins frères, 698 F2	108 MS	Croissance, carcasse, qualité de la viande (Nadaf <i>et al</i> 2007)
VO : F2 entre lignées maigres et grasses, 5 ♂ F1, 38 ♀ F1, 585 F2	129 MS	Composition de la carcasse (Lagarrigue <i>et al</i> 2006)
VO : 186 F2 et 80 backcross entre lignées résistantes et sensibles à la salmonellose,	103 et 135 MS	Sensibilité à la salmonellose (Tilquin <i>et al</i> 2005)

<sup>1</sup> BO = bovin ; CA = caille ; LA = lapin ; OV = ovin ; PO = porc ; VO = volaille.

<sup>2</sup> MS = microsatellite ; SNP = Polymorphisme simple nucléotide ; AFLP = Polymorphisme de longueur de fragments amplifiés.

chèvre. Dans les deux cas, une localisation très fine (de l'ordre de 100 kb) a été possible grâce à l'existence de familles ressources de qualité, la construction de banques de BAC locales, la densification en marqueurs des régions candidates et la possibilité d'inférer avec une très bonne probabilité le génotype des animaux au QTL à partir de données phénotypiques (Milan *et al* 1995, Vaiman *et al* 1996b).

Lorsque les effets des QTL sont plus modestes, il est plus difficile de connaître le génotype des animaux au QTL et des familles ressources de plus grande taille sont nécessaires. Dans le cas de populations croisées, la meilleure solution n'est pas de produire de nouveaux animaux F2 ou «backcross», mais de réaliser plusieurs générations d'intercroisements ou de croisements en retour vers l'une des populations

grand-parentales. L'INRA a ainsi mis en place plusieurs programmes de type AIL (*Advanced Intercross Lines*), notamment chez la poule et chez la caille. Chez le porc, une variante a été utilisée ; elle consiste à sélectionner les individus porteurs de recombinaison dans l'intervalle de localisation d'un QTL et à les croiser en retour vers l'une des populations grand-parentales pour tester sur descendance la présence

ou non d'une ségrégation au QTL. Elle a permis de réduire par un facteur de 5 à 10 l'intervalle de localisation de QTL situé dans 4 régions chromosomiques (Demeure *et al* 2005, Sanchez *et al* 2006).

Une approche alternative consiste à tirer parti des événements de recombinaison ancestraux, afin de s'abstraire du temps nécessaire à la production de générations supplémentaires. Cette approche est basée sur l'hypothèse qu'au sein d'un isolat génétique (comme une race ou une lignée), des individus porteurs d'un même allèle au QTL ont tous hérité cet allèle d'un ancêtre commun. Le chromosome original (dans lequel cette mutation est apparue) est transmis au cours des générations aux événements de recombinaison près. A la  $n^{\text{ème}}$  génération, la recherche d'une portion chromosomique identique chez des animaux présentant le phénotype associé à cette mutation permet de définir la région où rechercher le gène impliqué (figure 2). Plus le nombre de générations séparant les différents animaux issus d'un même fondateur est important, plus le segment chromosomique identique par descendance (IBD) contenant le gène est petit. Ce type d'approche, ou des

approches similaires combinant l'information apportée par l'analyse de liaison et de déséquilibre de liaison, sont conduits à l'INRA notamment pour la cartographie fine de QTL chez les bovins laitiers à l'aide de marqueurs microsatellites (Gautier *et al* 2006) ou, plus récemment, d'un réseau de marqueurs SNP à haute densité (Druet *et al* 2008).

L'arrivée de puces SNP pangénomiques à très haute densité bouleversera les méthodes de détection de QTL dans les prochaines années. Il devient en effet possible de regrouper les deux étapes de primo-localisation et de cartographie fine, l'une des conditions nécessaires, la forte densité en marqueurs, étant remplie. Il convient cependant de mettre en place des dispositifs expérimentaux adaptés et de grande taille pour obtenir la résolution nécessaire.

Les puces d'expression peuvent également être utilisées pour la cartographie génétique. Une première approche consiste à utiliser des données d'expression comme phénotypes et détecter des e-QTL ou QTL d'expression. Le fait d'être proche de l'expression des gènes et de disposer de phénotypes en

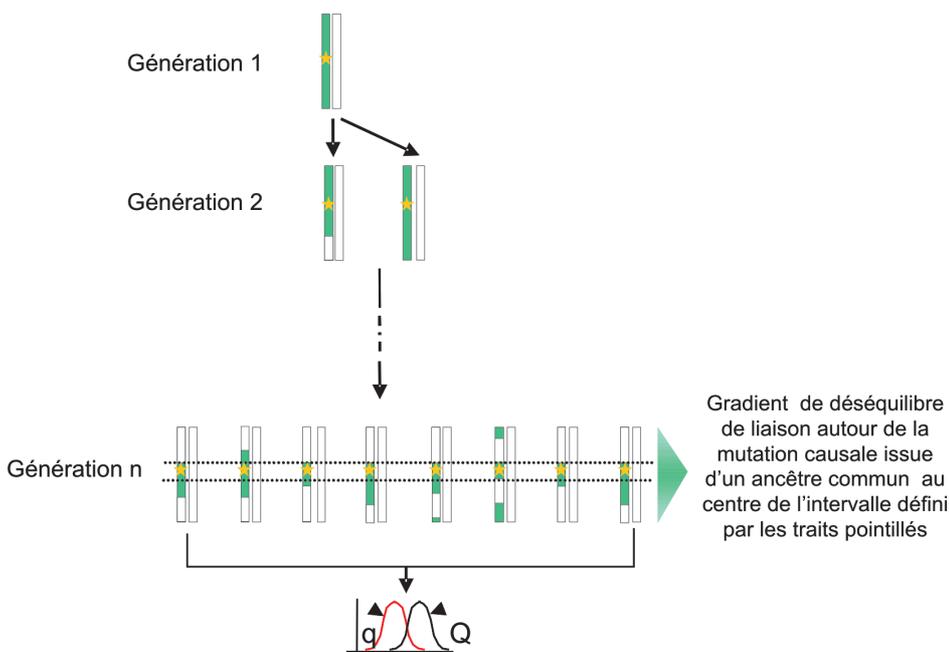
grande quantité doit ainsi permettre d'accroître la puissance et la résolution des dispositifs de détection et de cartographie fine de QTL.

### 3.3 / Identification du gène et du polymorphisme causal

L'identification du gène impliqué et du polymorphisme causal constitue l'étape ultime de la cartographie d'un gène. Lorsque l'intervalle de localisation est restreint, on commence en général par rechercher des gènes candidats positionnels à partir notamment de données de cartographie comparée. Les candidats sont choisis en tenant compte des données fonctionnelles disponibles pour ces gènes et de la localisation de ceux-ci dans l'intervalle de cartographie du QTL. Si aucun gène candidat n'est trouvé, une démarche de clonage positionnel est utilisée pour détecter les séquences codantes dans l'intervalle de localisation du gène (voir par exemple Eggen 2000). Des mutations dans les séquences codantes ou les régions régulatrices de ces gènes sont ensuite recherchées. L'étude du déséquilibre de liaison entre ces mutations et le caractère sur de grandes populations d'origines diverses pourra alors permettre de confirmer ou d'exclure ces polymorphismes candidats. Il est important de noter qu'un polymorphisme en total déséquilibre de liaison avec la mutation causale ne sera exclu avec certitude que par des études fonctionnelles de constructions transgéniques porteuses des différents variants alléliques identifiés.

Les différents gènes identifiés ces dernières années figurent dans le tableau 2. Cinq d'entre eux l'ont été dans le cadre de programmes que menait ou auquel participait l'INRA : le gène RN chez le porc, le gène «sans corne» chez la chèvre, les gènes responsables de l'hypermuscularité et de la syndactylie chez les bovins et les gènes d'hyperovulation «Booroola», et de conformation «Texel belge» chez les ovins. Il est important de noter que la majorité de ces gènes ont été identifiés par clonage positionnel, aucun argument fonctionnel à l'époque ne permettant de les considérer comme gènes candidats. Les polymorphismes causaux sont majoritairement des mutations ponctuelles de la séquence codante du gène, mais peuvent également correspondre à des situations plus complexes, comme une délétion intronique de plus de 11 kb dans le cas du gène sans corne chez la chèvre (figure 3), ou une mutation dans la région 3' non

**Figure 2.** Cartographie fine de QTL par la recherche de segments identiques par descendance (IBD).



Des échantillons d'ADN sont collectés sur des individus ayant hérité le même allèle au QTL d'un ancêtre commun. Le génotypage des individus pour un ensemble dense de marqueurs dans l'intervalle de localisation du QTL permettra de détecter la zone minimale IBD entre les individus génotypés.



ment utilisé, le modèle polygénique infinitésimal, faisait l'hypothèse que les caractères complexes étaient génétiquement déterminés par un très grand nombre de gènes à effets individuels faibles. Même s'il est désormais clair qu'il ne correspond pas à la réalité biologique, ce modèle s'est montré extrêmement efficace, avec une excellente capacité prédictive, et il a permis de développer depuis une quarantaine d'années des programmes performants d'amélioration génétique et de gestion des populations. L'information moléculaire peut néanmoins, comme nous le verrons, compléter voire remplacer dans certains cas, partiellement ou en totalité, l'information phénotypique.

#### 4.1 / Utilisation de l'information génomique en sélection

L'utilisation du génotype des animaux pour un ou plusieurs gènes ou marqueurs liés permet d'accroître plus rapidement la fréquence des allèles favorables. En contrepartie, la sélection est moins efficace pour les gènes ou QTL non génotypés et il reste nécessaire d'utiliser l'information phénotypique pour sélectionner cette variabilité génétique résiduelle. On parle de Sélection Génotypique (SAG) lorsque le gène cible est connu, de Sélection Assistée par Marqueurs lorsqu'elle est faite à partir de marqueurs génétiques liés. De nombreux travaux ont cherché à quantifier l'efficacité de la SAM ou de la SAG par rapport à une sélection phénotypique. Pour les caractères faisant déjà ou pouvant aisément faire l'objet d'un contrôle de performances, l'information moléculaire ne remplacera en général pas la sélection phénotypique, mais viendra la compléter en permettant d'accroître la précision de l'estimation de la valeur génétique des reproducteurs et/ou de réduire l'intervalle de génération. Lorsque la mesure du phénotype est difficile ou impossible à grande échelle, une sélection sur les seules informations moléculaires peut se révéler extrêmement efficace. Dans le cas d'une sélection utilisant à la fois les informations phénotypiques et moléculaires, l'apport des informations moléculaires dépend : 1/ du gain de précision qu'elles permettent ; 2/ de la proportion de la variance génétique expliquée par les QTL identifiés ; 3/ de la précision avec laquelle les effets des QTL sont estimés (Elsen 2000, Manfredi 2000). Le gain de précision sera faible si celle-ci est déjà élevée grâce aux seules informations phénotypiques. L'apport est ainsi beaucoup plus important pour des caractères fai-

blement héréditaires, ne s'exprimant que dans un sexe (reproduction, aptitudes maternelles, longévité...) ou mesurables uniquement sur apparentés (qualité de la viande) que pour les caractères de croissance et de composition corporelle. Il est également fonction de la fréquence des allèles favorables : il est plus important si les allèles favorables ont une fréquence initiale réduite et diminue lorsque les allèles favorables approchent la fixation. La précision de la SAM dépend également des caractéristiques des marqueurs et de la liaison entre les marqueurs et le locus cible. Si la densité de marqueurs est faible, ils sont en général relativement éloignés et en équilibre de liaison avec les QTL. On ne peut alors raisonner qu'intra-famille et les événements de recombinaison réduisent rapidement l'efficacité de la sélection. Avec une forte densité de marqueurs, il devient par contre possible de trouver des marqueurs en déséquilibre de liaison avec le QTL et ainsi de se rapprocher de la situation de la SAG.

La sélection contre l'allèle de sensibilité à l'halothane chez les porcs Landrace Français a constitué une des premières utilisations de la SAM et de la SAG. Au démarrage du programme dans les années 80, les marqueurs utilisés étaient des marqueurs biochimiques (protéines sanguines), remplacés à partir de 1991 et la découverte de la mutation causale (Fuji *et al* 1991) par une sélection sur le gène lui-même. Par la suite, l'INRA a initié deux programmes de sélection utilisant à grande échelle des informations moléculaires. Un programme de SAG visant à accroître la résistance à la tremblante, une encéphalopathie spongiforme ovine, en sélectionnant les allèles favorables au locus PrP responsable d'une grande partie de la variation génétique du caractère, a été mis en place à partir de 2002 dans les races de moutons. Il a nécessité jusqu'à 100 000 génotypages annuels et dû gérer le compromis entre fixation rapide des allèles favorables, maintien d'un progrès génétique pour les autres caractères et maintien de la variabilité génétique. Un programme de sélection assistée par marqueurs de grande ampleur a été lancé en 2000 dans les trois principales races bovines laitières françaises, en partenariat avec les centres d'insémination artificielle. Il porte sur 14 régions chromosomiques, chacune d'entre elles affectant un à quatre caractères et étant suivie par 2 à 4 marqueurs. Les animaux génotypés sont d'une part des candidats à la sélection, des jeunes mâles ou des femelles candidates comme mères à taureau, d'autre

part des individus apparentés avec phénotype. Au total, environ 10 000 individus sont typés pour 45 marqueurs. La localisation des QTL étant au départ imprécise, le programme a fait dans un premier temps l'hypothèse d'absence de déséquilibre de liaison populationnel entre marqueurs et phénotype. Un premier bilan du programme a permis de mettre en évidence un gain d'efficacité de 5 à 8% selon le caractère par rapport à une sélection classique (Guillaume *et al* 2008). L'utilisation de marqueurs SNP proches et en déséquilibre de liaison avec les QTL, prévue en 2008, devrait permettre d'accroître sensiblement cette efficacité.

L'arrivée de puces pangénomiques à très haute densité bouleversera ici aussi les choses dans un proche avenir. Il deviendra en effet possible de réaliser une sélection génomique, c'est-à-dire de sélectionner les individus sur la base de leur génotype pour un nombre suffisant de marqueurs (plusieurs dizaines, voire plusieurs centaines de milliers de marqueurs) pour que l'ensemble des QTL affectant un caractère soient en déséquilibre de liaison avec au moins un marqueur. La sélection génomique (Meuwissen *et al* 2001) peut donc être vue comme une généralisation de la SAM à l'ensemble du génome et l'ensemble des QTL. Dans la mesure où elle peut décrire l'ensemble de la variabilité génétique, les phénotypes ne sont théoriquement plus nécessaires. En pratique, la situation n'est pas aussi simple, dans la mesure où, dans de nombreux cas, seul un sous-ensemble d'animaux est génotypé. Il est en outre nécessaire d'estimer de façon précise la liaison génotype-phénotype et son évolution sous l'effet de la sélection, et *in fine* de valider l'efficacité de la sélection. Ces différents points et, plus globalement, l'efficacité et les modalités de mise en œuvre de la sélection génomique font depuis quelques années l'objet de recherches extrêmement actives. L'INRA développe des travaux à la fois sur le plan méthodologique (Legarra et Misztal 2008) et sur le plan des applications dans les programmes de sélection, notamment chez les bovins laitiers.

#### 4.2 / Caractérisation et gestion de la variabilité génétique

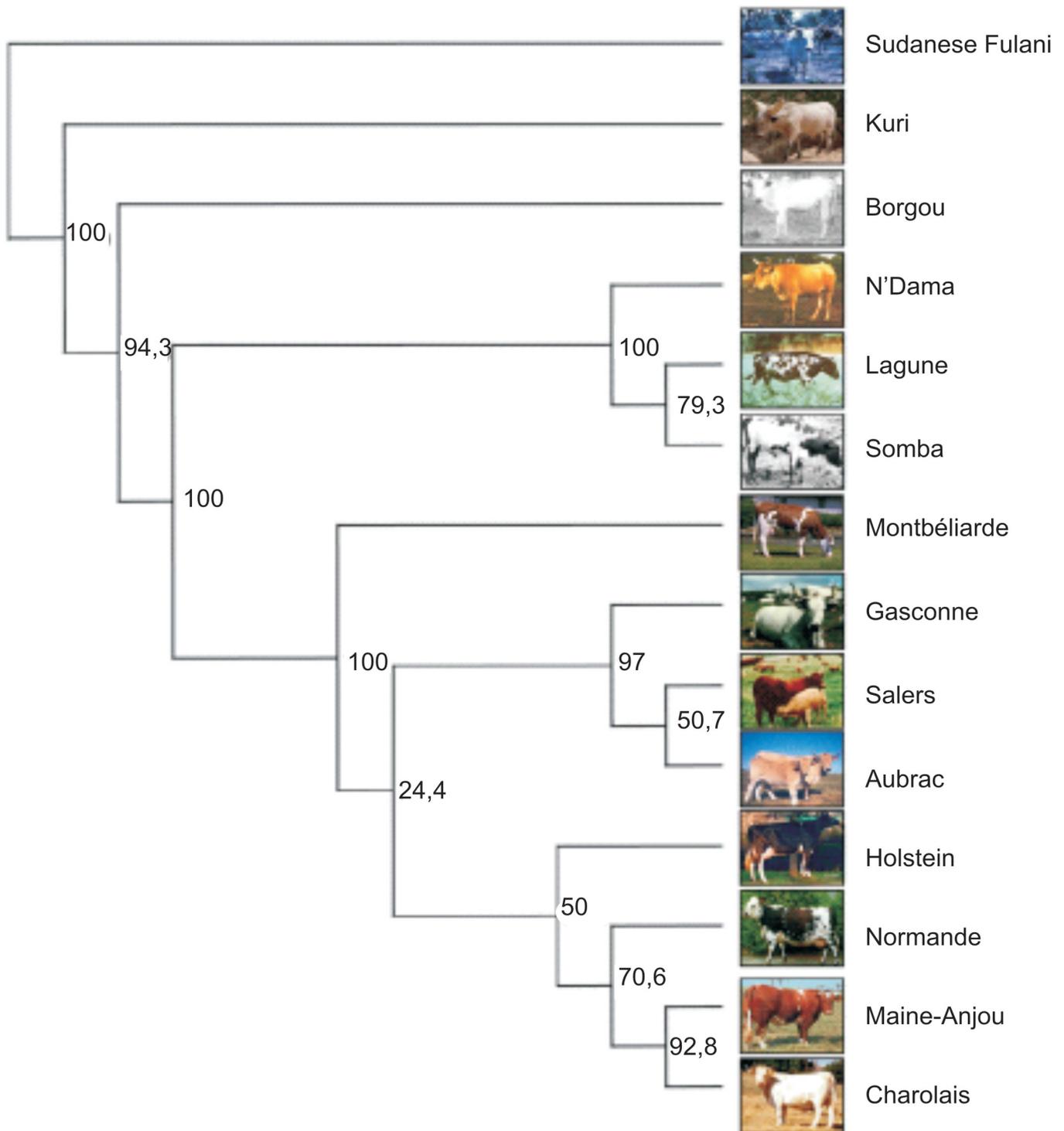
Les marqueurs génétiques ont révolutionné les recherches sur la caractérisation de la variabilité génétique, auparavant limitée à l'analyse des phénotypes et de quelques polymorphismes protéiques. L'utilisation d'un réseau de

marqueurs neutres judicieusement répartis sur le génome permet en effet de caractériser de façon directe et efficace la variabilité génétique entre- et intra- population, indépendamment des effets de milieu. De nombreux projets internationaux, le plus souvent européens, ont été mis en place pour caractériser la diversité des principales espèces d'animaux d'élevage : porc,

bovins, poules, lapins, etc. Ils sont en général basés sur le génotypage pour quelques dizaines de marqueurs d'échantillons représentatifs de chacune des populations étudiées. Dans la plupart des races, la diversité génétique est ensuite appréciée à partir de différentes approches basées sur l'hétérozygotie ou le calcul de distances génétiques entre populations, utilisées dans

le cadre de modèles de génétique des populations pour la construction de pseudo-arbres phylogénétiques, de méthodes de classification ou d'analyses multivariées (figure 4). L'INRA a apporté une contribution majeure à plusieurs de ces projets, que ce soit en assurant leur coordination scientifique, ou en réalisant l'analyse des données à l'aide des différentes méthodes évoquées

**Figure 4.** Relations phylogénétiques entre 14 races bovines européennes et africaines (Gautier et al 2007). L'arbre a été établi en utilisant la distance de Nei et la méthode de jonction des voisins. Les chiffres (en %) figurant au niveau de chaque nœud indiquent la fiabilité de chaque jonction.



ci-dessus. Les résultats obtenus sont particulièrement utiles pour proposer des recommandations sur les stratégies de conservation des populations animales. Par la suite, des études basées sur le même principe, mais avec un objectif de cartographie de gènes, ont été réalisées pour les régions dans lesquelles des QTL avaient été identifiés (par exemple, projet DIVQTL chez le porc, Iannuccelli *et al* 2006).

Comme précédemment, l'arrivée des puces SNP à haut débit bouleverse les travaux sur la diversité. Il devient en effet possible d'étudier de façon beaucoup plus fine la diversité des populations animales et aller jusqu'à la caractérisation de la structure haplotypique des populations comme cela a été réalisé chez l'Homme dans le cadre du projet HAPMAP. Une première étude effectuée chez les bovins (Gautier *et al* 2007) a permis de décrire la structure haplotypique dans un ensemble de 14 populations bovines avec, entre races, des traces de déséquilibre de liaison ancestral sur de courtes distances (< 10 kb). Les blocs ancestraux ainsi définis sont regroupés intra-race dans des blocs haplotypiques de plus grande taille (quelques centaines de kb). Ces nouvelles approches, qui permettent en pratique de combiner dans une même étude, l'analyse fine de la diversité génétique et la mise en évidence des zones du génome (haplotypes) associées aux variations phénotypiques, sont sans nul doute amenées à se développer de manière importante au cours des prochaines années.

### 4.3 / Contrôle de généalogie – traçabilité – certification

L'inventaire des applications des marqueurs moléculaires en élevage ne saurait être complet sans évoquer les applications développées en matière de contrôle ou d'assignation de filiation, de traçabilité et de certification. L'INRA a été pionnier en matière de contrôle de filiation en mettant en place dès les années 50 un outil basé sur le polymorphisme des groupes sanguins dans le cadre du laboratoire d'Analyse des Groupes Sanguins (Grosclaude *et al* 1996). A partir de la fin des années 90, les groupes sanguins ont été progressivement remplacés par des marqueurs microsatellites, alors que le laboratoire INRA a évolué en 1994 vers une structure de type GIE sous le nom de LABOGENA. A l'heure actuelle,

l'ensemble des contrôles de filiation est réalisé à partir de jeux de 10 à 20 marqueurs microsatellites selon les espèces et les populations, qui permettent de détecter toujours plus de 95% et souvent plus de 99% des incompatibilités parentales dans les populations animales. Les marqueurs microsatellites seront vraisemblablement remplacés à l'avenir par des marqueurs SNP. Bien que moins informatifs individuellement, leur faible coût devrait permettre d'en génotyper un grand nombre, pour approcher un taux de détection de 100%.

Ces mêmes marqueurs, complétés le cas échéant par quelques marqueurs supplémentaires, peuvent être utilisés à des fins de traçabilité individuelle des animaux ou de produits dérivés comme la viande. La faisabilité de ce type approche, qui s'apparente aux tests utilisés en médecine légale, a été établie dans le cas des bovins par San Cristobal-Gaudy *et al* (2000). Elle nécessite de disposer d'un échantillon d'ADN de l'animal dont l'origine est irréfutable. Ceci pourrait par exemple être obtenu grâce au prélèvement, puis au stockage par l'abattoir de l'ensemble oreille/boucle d'identification de tout animal abattu. Les marqueurs génétiques peuvent également être employés pour déterminer l'origine raciale d'un animal ou d'une pièce de viande. Elle peut également, dans une certaine mesure, être utilisée pour vérifier l'utilisation ou non d'une race ou d'une lignée donnée dans un plan de croisement. L'approche utilisée est de type probabiliste, mais peut être extrêmement puissante si les marqueurs permettent de bien caractériser les particularités génétiques de chaque population.

Les vérifications *a posteriori* décrites ci-dessus peuvent être envisagées dans le cadre d'actions en justice, mais également dans le cadre d'une vérification par échantillonnage du respect d'un cahier des charges. Plus qu'une éventuelle origine raciale, les cahiers des charges comportent souvent des clauses relatives à l'absence d'allèles défavorables (par exemple l'allèle de sensibilité à l'halothane chez le porc). Si l'on peut envisager le génotypage systématique des mâles, il apparaît beaucoup plus difficile de l'envisager chez les femelles et plus encore sur l'ensemble des animaux. Une alternative peut alors être de garantir une fréquence allélique

inférieure à un seuil donné en mettant en place une stratégie d'échantillonnage appropriée des animaux à typer.

## Conclusion

L'essor de la génomique depuis 20 ans a constitué un changement fondamental dans la recherche en biologie avec l'apparition d'une dimension industrielle (grands nombres d'échantillons, automatisation...) qui en était quasiment absente. Elle se traduit par la mise en place de structures adaptées aux analyses à grande échelle et au traitement de masses d'information importantes et une spécialisation des tâches. Le coût de ces structures et la durée de vie courte des équipements qu'elles gèrent ont conduit à les développer sous forme de plates-formes communes à plusieurs organismes (génopoles), gérées de façon très formalisée (mise en place de démarche qualité, certification...). L'INRA participe à plusieurs de ces plates-formes et y tient parfois un rôle de leader (génopoles de Toulouse ou de Rennes).

Les perspectives et les enjeux de la génomique sont considérables et dépassent largement le seul secteur de l'amélioration génétique animale, avec des applications potentielles dans les domaines de la santé animale et humaine, de la traçabilité, du diagnostic, etc. Les évolutions actuelles devraient se poursuivre et s'intensifier à l'avenir, compte tenu de l'accélération du rythme des innovations. Ces évolutions exacerbent la concurrence entre équipes de recherche et entre pôles d'excellence dotés ou cherchant à se doter des outils les plus performants et des meilleurs chercheurs. Elles touchent aussi les professionnels de l'élevage, qui doivent accompagner les évolutions scientifiques et technologiques, l'exploitation des droits de propriété intellectuelle et répondre au mieux aux attentes des éleveurs, des consommateurs et des citoyens. Ces enjeux ont été fort bien compris par le secteur de l'élevage en France, qui n'a pas hésité depuis le début des années 2000 à se regrouper entre organisations concurrentes pour pouvoir apporter un soutien financier substantiel et participer avec l'INRA à une co-construction des programmes de recherche, dans le cadre notamment du programme AGENAE.

## Références

- Archibald A.L., Brown J.F., Couperwhite S., Mc Queen H.A., Nicholson D., Haley C.S., Coppieters W., Van de Weghe A., Stratil A., Winterö A.K., Fredholm M., Larsen N.J., Nielsen V.H., Milan D., Woloszyn N., Robic A., Dalens M., Riquet J., Gellin J., Caritez J.C., Hue D., Burgaud G., Ollivier L., Bidanel J.P., Vaiman M., Renard C., Geldermann H., Davoli R., Ruyter D., Versteger E.J.M., Groenen M.A.M., Davies W., Höyheim B., Keiserud A., Andersson L., Ellegren H., Johansson M., Marklund L., Miller J.R., Anderson Dear D.V., Signer E., Jeffreys A.J., Moran C., Le Tissier P., Rothschild M.F., Tuggle C.K., Vaske D., Helm J., Liu H.C., Rahman A., Yu T.P., Larson R.G., Schmitz C.B., 1995. The PiGMap consortium linkage map of the pig (*Sus scrofa*). *Mammalian Genome*, 6, 157-175.
- Barillet F., Arranz J.J., Carta A., 2005. Mapping quantitative trait loci for milk production and genetic polymorphisms of milk proteins in dairy sheep. *Genet. Sel. Evol.*, 37 (suppl. 1), S109-S123.
- Basso B., Fritz S., Druet T., Guillaume F., Rossignol M.N., Amigues Y., Gabriel R., Sellem E., Salas-Cortes L., Humblot P., Druet X., 2005. Estimation de paramètres génétiques et détection de QTL liés à des caractères de fertilité mâle, de production de semence et de qualité de la semence chez le taureau laitiers. *Renc. Rech. Rum.*, 12, 145-148.
- Beaumont C., Roussot O., Fève K., Vignoles F., Leroux S., Pitel F., Faure J.M., Mills A.D., Guéméné D., Sellier N., Mignon-Grasteau S., Le Roy P., Vignal A., 2005. A genome scan with AFLP<sup>TM</sup> markers to detect fearfulness-related QTLs in Japanese quail. *Anim. Genet.*, 36, 401-407.
- 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.
- Bidanel J.P., Milan D., Iannuccelli N., Amigues Y., Boscher M.Y., Bourgeois F., Caritez J.C., Gruand J., Le Roy P., Lagant H., Quintanilla R., Renard C., Gellin J., Ollivier L., Chevalet C., 2001. Detection of quantitative trait loci for growth and fatness in pigs. *Genet. Sel. Evol.*, 33, 289-309.
- Bidanel J.P., Rosendo A., Iannuccelli N., Gilbert H., Caritez J.C., Billon Y., Prunier A., Milan D., 2008. Detection of quantitative trait loci for teat number and female reproductive traits in Meishan x Large White F2 pigs. *Animal*, 2, accepté.
- Blott S., Kim J.J., Moiso S., Küntzel A.S., Cornet A., Berzi P., Cambisano N., Ford C., Grisart B., Johnson D., Karim L., Simon P., Snell R., Spelman R., Wong J., Vilkki J., Georges M., Farnir F., Coppieters W., 2003. Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics*, 163, 253-266.
- Bodin L., Di Pasquale E., Fabre S., Bontoux M., Monget P., Persani L., Mulsant P., 2007. A novel mutation in the bone morphogenetic protein 15 gene causing defective protein secretion is associated with both increased ovulation rate and sterility in Lacaune sheep. *Endocrinology*, 148, 393-400.
- Boichard D., Grohs C., Bourgeois F., Cerqueira F., Faugeras R., Neau A., Rupp R., Amigues Y., Boscher M.Y., Levéziel H., 2003. Detection of genes influencing economic traits in three French dairy cattle breeds. *Genet. Sel. Evol.*, 35, 77-101.
- Botstein D., White R.L., Skolnick M., Davis R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, 32, 314-331.
- Chang C.M., Coville J.L., Coquerelle G., Gourichon D., Oulmouden A., Tixier-Boichard M., 2006. Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC Genomics*, 7, 19.
- Chantry-Darmon C., Urien C., de Rochambeau H., Allain D., Pena B., Hayes H., Bertaud M., Grohs C., Cribiu E.P., Deretz-Picoulet S., Larzul C., Save J.C., Neau A., Chardon P., Rogel-Gaillard C., 2006. A first-generation microsatellite-based integrated genetic and cytogenetic map for the European rabbit (*Oryctolagus cuniculus*): application to the localization of angora and albino characters. *Anim. Genet.*, 37, 335-341.
- Ciobanu D.C., Bastiaansen J., Malek M., Helm J., Woollard J., Plastow G.S., Rothschild M.F., 2001. Evidence for new alleles in the protein kinase AMP-activated, subunit gene associated with low glycogen content in pig skeletal muscle and improved meat quality. *Genetics*, 159, 1151-1162.
- Clop A., Marcq F., Takeda H., Pirottin D., Tordo X., Bibé B., Bouix J., Caiment F., Elsen J.M., Eychenne F., Larzul C., Laville E., Meish F., Milenkovic D., Tobin J., Charlier C., Georges M., 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genet.*, 38, 813-818.
- Corpet F., Chevalet C., 2000. Analyse informatique des données moléculaires. In : *Génétique moléculaire : principe et application aux populations animales*, numéro hors série, INRA Prod. Anim., 191-195.
- Cosseddu G.M., Oustry-Vaiman A., Jego B., Moreno C., Taourit S., Cribiu E.P., Elsen J.M., Vaiman D., 2002. Sheep/human comparative map in a chromosome region involved in scrapie incubation time shows multiple breakpoints between human chromosomes 14 and 15 and sheep chromosomes 7 and 18. *Chromosome Res.*, 10, 369-378.
- Courcelle E., Beausse Y., Letort S., Stahl O., Fremez R., Ngom-Bru C., Gouzy J., Faraut T., 2008. Narcisse: a mirror view of conserved syntenties. *Nucleic Acids Res.*, 36, D485-D490.
- De Givry S., Bouchez M., Chabrier P., Milan D., Schiex T., 2005. CARTHA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics*, 21, 1703-1704.
- Demeure O., Sanchez M.P., Riquet J., Iannuccelli N., Demars J., Fève K., Kernaléguen L., Gogué J., Billon Y., Caritez J.C., Milan D., Bidanel J.P., 2005. Exclusion of the swine leukocyte antigens as candidate region and reduction of the position interval for the *Sus scrofa* chromosome 7 QTL affecting growth and fatness. *J. Anim. Sci.*, 83, 1979-1987.
- Désautés C., Bidanel J.P., Milan D., Iannuccelli N., Amigues Y., Bourgeois F., Caritez J.C., Renard C., Chevalet C., Mormède P., 2002. Genetic mapping of quantitative trait loci for behavioural and neuroendocrine traits in pigs. *J. Anim. Sci.*, 80, 2276-2285.
- Dib C., Fauré S., Fizames C., Samson D., Drouot N., Vignal A., Millasseau P., Marc S., Hazan J., Seboun E., Lathrop M., Gyapay G., Morissette J., Weissenbach J., 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*, 380, 152-154.
- Druet T., Fritz S., Boussaha M., Ben Jemaa S., Derbala D., Zele D., Lechner D., Charon C., Boichard D., Gut I.G., Eggen A., Gautier M., 2008. Fine-mapping of QTLs affecting female fertility in dairy cattle on BTA03 using a dense SNP map. *Genetics*, sous presse.
- Du Z., Vincent-Naulleau S., Gilbert H., Vignoles F., Crechet F., Shimogiri T., Yasue H., Leplat J.J., Bouet S., Gruand J., Horak V., Milan D., Le Roy P., Geffrotin C., 2007. Detection of novel quantitative trait loci for cutaneous melanoma by genome-wide scan in the MeLiM swine model. *Int. J. Cancer*, 120, 303-320.
- Duchesne A., Gautier M., Chadi S., Grohs C., Floriot S., Gallard Y., Caste G., Ducos A., Eggen A., 2006. Identification of a doublet missense substitution in the bovine LRP4 gene as a candidate causal mutation for syndactyly in Holstein cattle. *Genomics*, 88, 610-621.
- Eggen A., 2000. Cartographie fine d'un gène et clonage positionnel. In : *Génétique moléculaire : principe et application aux populations animales*, numéro hors série, INRA Prod. Anim., 133-136.
- Elsen J.M., 2000. Sélection et introgression assistées par marqueurs. In : *Génétique moléculaire : principe et application aux populations animales*, numéro hors série, INRA Prod. Anim., 233-237.
- Eriksson J., Larson G., Gunnarsson U., Bed'hom B., Tixier-Boichard M., Strömstedt L., Wright D., Jungerius A., Vereijken A., Randi E., Jensen P., Andersson L., 2008. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet.*, 4, e1000010.
- Foissac S., Bardou P., Moisan A., Cros M.J., Schiex T., 2003. EUGENE/HOM: A generic similarity-based gene finder using multiple homologous sequences. *Nucleic Acids Res.*, 31, 3742-3745.
- Frazer K.A., Ballinger D.G., Cox D.R., Hinds D.A., Stuve L.L. *et al.*, 2007. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*, 449, 851-861.
- Freking B.A., Murphy S.K., Wylie A.A., Rhodes S.J., Keele J.W., Leymaster K.A., Jirtle R.L., Smith P.L., 2002. Identification of the single base change causing the Callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. *Genome Res.*, 12, 1496-1506.
- Fujii J., Otsu K., Zorzato F., de Leon S., Khanna V.K., Weiler J.E., O'Brien P.J., MacLennan D.H., 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science*, 253, 448-451.
- Galloway S.M., McNatty K.P., Cambridge L.M., Laitinen M.P.E., Juengel J.L., Jokiranta T.S., McLaren R.J., Luro K., Dodds K.G., Montgomery G.W., Beattie A.E., Davis G.H.,

- Ritvos O., 2000. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nature Genet.*, 25, 279-283.
- Gautier M., Roy Barcelona R., Fritz S., Grohs C., Druet T., Boichard D., Eggen A., Meuwissen T.H.E., 2006. Fine mapping and physical characterization of two linked Quantitative Trait Loci affecting milk fat yield in dairy cattle on BTA26. *Genetics*, 172, 425-436
- Gautier M., Faraut T., Moazami-Goudarzi K., Navratil V., Foglio M., Grohs C., Boland A., Garnier J.G., Boichard D., Lathrop G.M., Gut I.G., Eggen A., 2007. Genetic and haplotypic structure in 14 European and African cattle breeds. *Genetics*, 177, 1059-1070.
- Gharbi K., Gautier A., Danzmann R.G., Gharbi S., Sakamoto T., Høyheim B., Taggart J.B., Cairney M., Richard Powell R., Krieg F., Okamoto N., Ferguson M.M., Holm L.E., Guyomard R., 2006. A linkage map for Brown Trout (*Salmo trutta*): chromosome homeologies and comparative genome organization with other salmonid fish. *Genetics*, 172, 2405-2419.
- Gilbert H., Riquet J., Gruand J., Billon Y., Iannuccelli N., Sellier P., Noblet J., Bidanel J.P., 2008. Détection de QTL pour la consommation alimentaire dans un croisement de type back-cross Large White x Piétrain. *Journ. Rech. Porcine Fr.*, 40, 149-150.
- Girardot M., Guibert S., Laforet M.P., Levéziel H., Julien R., Oulmouden A., 2003. Exploitation des gènes de la coloration de la robe pour une traçabilité raciale des produits d'origine bovine. *Renc. Rech. Rum.*, 10, 33-36.
- Girardot M., Guibert S., Laforet M.P., Gallard Y., Larroque H., Oulmouden A., 2006. The insertion of a full-length Bos Taurus LINE element is responsible for at transcriptional deregulation of the Normande Agouti gene. *Pigment Cell Res.*, 19, 346-355.
- Goureau A., Yerle M., Schmitz A., Riquet J., Milan D., Pinton P., Frelat G., Gellin J., 1996. Human and porcine correspondence of chromosome segments using bidirectional chromosome painting. *Genomics*, 36, 252-262.
- Grisart B., Coppeters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M., Snell R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res.*, 12, 222-231.
- Grobet L., Martin L.J., Poncelet D., Pirottin D., Brouwers B., Riquet J., Schoeberlein A., Dunner S., Ménéssier F., Massabanda J., Fries R., Hanset R., Georges M., 1997. A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nature Genet.*, 17, 71-74.
- Grobet L., Poncelet D., Royo L.J., Brouwers B., Pirottin D., Michaux C., Ménéssier F., Zanotti M., Dunner S., Georges M., 1998. Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle. *Mammalian Genome*, 9, 210-213.
- Grosclaude F., Mercier J.C., Vaiman M., Levéziel H., Gellin J., 1996. La génétique moléculaire des espèces d'élevage : des groupes sanguins à la cartographie du génome. In : 50 ans de recherche en productions animales, numéro hors série, INRA Prod. Anim., 57-69.
- Guérin G., Eggen A., Vaiman D., Nocart M., Laurent P., Bechet D., Ferrara M., 1994. Further characterization of a somatic cell hybrid panel: ten assignments to bovine genome. *Anim. Genet.*, 25, 31-35.
- Guérin G., Bailey E., Bernoco D., Anderson L., Antczak D.F. *et al.*, 1999. Report of the international equine gene mapping workshop: male linkage map. *Anim. Genet.*, 30, 341-354.
- Guillaume F., Fritz S., Boichard D., Druet T., 2008. Estimation by simulation of the efficiency of the french Marker Assisted Selection program in dairy cattle. *Genet. Sel. Evol.*, 40, 91-102.
- Gunnarsson U., Hellström A.R., Tixier-Boichard M., Minvielle F., Bed'hom B., Ito S., Jensen P., Rattink A., Vereijken A., Andersson L., 2007. Mutations in SLC45A2 cause plumage color variation in chicken and japanese quail. *Genetics*, 175, 867-877.
- Hatey F., 2000. Identification de gènes associés à des fonctions : l'approche fonctionnelle. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 153-160.
- Hawken R.J., Murtaugh J., Flickinger G.H., Yerle M., Robic A., Milan D., Gellin J., Beattie C.W., Schook L.B., Alexander L.J., 1999. A first-generation porcine whole-genome radiation hybrid map. *Mammalian Genome*, 10, 824-830.
- Hayes H., 1995. Chromosome painting with human chromosome-specific DNA libraries reveals the extent and distribution of conserved segments in bovine chromosomes. *Cytogenet. Cell Genet.*, 71, 168-174.
- Hayes H., Elduque C., Gautier M., Schibler L., Cribiu E., Eggen A., 2003. Mapping of 195 genes in cattle and updated comparative map with man, mouse, rat and pig. *Cytogenet. Genome Res.*, 102, 16-24.
- Hiragaki T., Inoue-Murayama M., Miwa M., Fujiwara A., Mizutani M., Minvielle F., Ito S., 2008. Recessive black is allelic to the yellow plumage locus in japanese quail and associated with a frameshift deletion in the ASIP gene. *Genetics*, 178, 771-775.
- Hu Z. L., Fritz E.R., Reecy J.M., 2007. AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. *Nucleic Acids Res.*, 35, D604-D609.
- Humphray S.J., Scott C., Clark R., Marron B., Bender C., Camm N., Davis J., Jenks A., Noon A., Patel M., Sehra H., Yang F., Rogatcheva M., Milan D., Chardon P., Rohrer G., Nonneman D., de Jong P., Meyers S., Archibald A., Beever J., Schook L., Rogers J., 2007. A high utility integrated map of the pig genome. *Genome Biol.*, 8, R139.
- Iannuccelli N., Riquet J., Mercat M.J., Legros H., San Cristobal M., Lacoste A., Bidanel J.P., Milan D., 2006. Diversité génétique des populations porcines françaises dans les régions chromosomiques soumises à la sélection : projet DIVQTL. Actes du 6<sup>ème</sup> Colloque National Ressources Génétiques, 2-4 octobre, La Rochelle, France, 111-128 (<http://www.brg-prd.fr/brg/pdf/LaRocheMilan.pdf>).
- International Human Genome Sequencing Consortium, 2001. Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.
- Jennen D.G., Crooijmans R.P., Morisson M., Grootemaat A.E., Van Der Poel J.J., Vignal A., Groenen M.A., 2004. A radiation hybrid map of chicken chromosome 15. *Anim. Genet.*, 35, 63-65.
- Kambadur R., Sharma M., Smith T. P., Bass J. J., 1997. Mutations in myostatin (GDF8) in double-muscling Belgian Blue and Piedmontese cattle. *Genome Res.*, 7, 910-916.
- 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.
- Larroque H., Gallard Y., Delacroix-Buchet A., Bach C., Ogier J.C., Mahé M.F., Miranda G., Gautier M., Grohs C., Legros H., Druet T., Colleau J.J., Boichard D., 2007. Détection de QTL influençant l'aptitude à la transformation fromagère des laits dans un croisement Holstein Normand. *Renc. Rech. Rum.*, 14, 125-128.
- Laurent P., Schibler L., Vaiman A., Laubier J., Delcros C., Cosseddu G., Vaiman D., Cribiu E.P., Yerle M., 2007. A 12 000-rad whole-genome radiation hybrid panel in sheep: application to the study of the ovine chromosome 18 region containing a QTL for scrapie susceptibility. *Anim. Genet.*, 38, 358-363.
- Legarra A., Misztal I., 2008. Computing strategies in genome-wide selection. *J. Dairy Sci.*, 91, 360-366.
- Le Roy P., Elsen J.M., 2000. Principes de l'utilisation des marqueurs génétiques pour la détection des gènes influençant les caractères quantitatifs. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 211-215.
- Levéziel H., Cribiu E., 2000. Etat des lieux de la cartographie du génome des ruminants. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 103-107.
- Manfredi E., 2000. Intérêt et limites de la sélection intra-race assistée par marqueurs. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 239-242.
- McPherron A.C., Lee, S. J., 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proc. Nat. Acad. Sci. USA*, 94, 12457-12461.
- Meijerink E., Neuenschwander S., Fries R., Dinter A., Bertschinger H.U., Stranzinger G., Vögeli P., 2000. A DNA polymorphism influencing alpha(1,2)fucosyltransferase activity of the pig FUT1 enzyme determines susceptibility of small intestinal epithelium to *Escherichia coli* F18 adhesion. *Immunogenetics*, 52, 129-136.
- Meuwissen T. H. E., Hayes B. J., Goddard M. E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819-1829.
- Meyers S.N., Rogatcheva M.B., Larkin D.M., Yerle M., Milan D., Hawken R.J., Schook L.B., Beever J.E., 2005. Piggy-BACing the human genome II. A high-resolution, physically anchored, comparative map of the porcine autosomes. *Genomics*, 86, 739-752.
- Milan D., Le Roy P., Woloszyn N., Caritez J.C., Elsen J.M., Sellier P., Gellin J., 1995. The RN locus for meat quality maps to pig chromosome 15. *Genet. Sel. Evol.*, 27, 195-199.
- Milan D., Yerle M., Robic A., Lahbib-Mansais Y., Riquet J., Iannuccelli N., 2000a. Etat des lieux de la cartographie du génome du porc. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 109-111.

- Milan D., Jeon J.T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundström K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P., Andersson L., 2000b. A Mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science*, 288, 1248-1251.
- Milan D., Bidanel J.P., Iannuccelli N., Riquet J., Amigues Y., Boscher M.Y., Bourgeois F., Caritez J.C., Gruand J., Le Roy P., Lagant H., Quintanilla R., Renard C., Gellin J., Ollivier L., Chevalet C., 2002. Detection of quantitative trait loci for carcass composition traits in pigs. *Genet. Sel. Evol.*, 34, 705-728.
- Milan D., Mercat M.J., Dechamp N., Iannuccelli N., Gilbert H., Billon Y., Bouffaud M., Schwob S., Sanchez M.P., Riquet J., Bidanel J.P., 2008. Estimation des effets de 6 QTL affectant les caractères de production dans des populations porcines commerciales françaises : premiers résultats du projet BIOMARK. *Journ. Rech. Porcine. Fr.*, 40, 145-148.
- Minvielle F., Kayang B.B., Inoue-Murayama M., Miwa M., Vignal A., Gourichon D., Neau A., Monvoisin J.L., Ito S., 2005. Microsatellite mapping of QTL affecting growth, feed consumption, egg production, tonic immobility and body temperature of Japanese quail. *BMC Genomics*, 6, 87.
- Minvielle F., Kayang B.B., Inoue-Murayama M., Miwa M., Vignal A., Gourichon D., Neau A., Monvoisin J.L., Ito S., 2006. Search for QTL affecting the shape of the egg laying curve of the Japanese quail. *BMC Genetics*, 7, 26.
- Morisson M., Denis M., Milan D., Klopp C., Leroux S., Bardes S., Pitel F., Vignoles F., Gêrus M., Fillon V., Douaud M., Vignal A., 2007. The chicken RH map: current state of progress and microchromosome mapping. *Cytogenet. Genome Res.*, 117, 14-21.
- Muller C., Denis M., Gentzbittel L., Faraut T., 2004. The Iccare web server: an attempt to merge sequence and mapping information for plant and animal species. *Nucleic Acids Res.*, 32, W429-W434.
- Mulsant P., Lecerf F., Fabre S., Schibler L., Monget P., Lanneluc I., Pisselet C., Riquet J., Monniaux D., Callebaut I., Cribiu E.P., Thimonier J., Teyssier J., Bodin L., Cognié Y., Chitour N., Elsen J.M., 2001. Mutation in bone morphogenetic protein receptor-1B is associated with increased ovulation rate in Booroola Merino ewes. *Proc. Nat. Acad. Sci. USA*, 98, 5104-5109.
- Nadaf J., Gilbert H., Pitel F., Berri C.M., Fève 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.
- Nadeau, N.J., Minvielle F., Ito S., Inoue-Murayama M., Gourichon D., Johns S.A., Burke T., Mundy N.I., 2008. Characterization of Japanese quail dominant yellow as a genomic deletion upstream of the avian homologue of the mammalian ASIP (agouti) gene. *Genetics*, 178, 777-786.
- 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. *Nature Genet.*, 29, 453-458.
- Pitel F., Riquet J. 2000. Les marqueurs anonymes et la détection de leur polymorphisme. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 43-53.
- Pitel F., Bergé R., Coquerelle G., Crooijmans R.P.M.A., Groenen M.A.M., Vignal A., Tixier-Boichard M., 2000. Mapping the Naked Neck (NA) and Polydactyly (PO) mutants of the chicken with microsatellite molecular markers. *Genet. Sel. Evol.*, 32, 73-86.
- Quintanilla R., Bidanel J.P., Milan D., Iannuccelli N., Amigues Y., Gruand J., Renard C., Chevalet C., Bonneau M., 2003. Genetic mapping of quantitative trait loci for fat androstenone level in pigs. *J. Anim. Sci.*, 81, 385-394.
- Renand G., Bonnot A., Levéziel H., Payet N., Malafosse A., Hocquette J.F., Lepetit J., Rousset S., Denoyelle C., Dodelin V., 2007. Teneur en lipides du muscle et adiposité des carcasses en races bovines à viande : variabilité génétique et effets de marqueurs dans les gènes DGAT1 et Thyroglobuline. *Renc. Rech. Rum.*, 14, 113-116.
- Renard C., Hart E., Sehra H., Beasley H., Coggin P., Howe K., Harrow J., Gilbert J., Sims S., Rogers J., Ando A., Shinegari A., Shiina T., Inoko H., Chardon P., Beck S., 2006. The genomic sequence and analysis of the swine major histocompatibility complex. *Genomics*, 88, 96-110.
- Rieder S., Taourit S., Mariat D., Langlois B., Guérin G., 2001. Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotype in horses (*Equus caballus*). *Mammalian Genome*, 12, 450-455.
- Rogel-Gaillard C., 2000. Les banques de grands fragments d'ADN. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 79-85.
- Rouzaud F., Martin J., Gallet P.F., Delourme D., Goulemot-Léger V., Amigues Y., Ménissier F., Levéziel H., Julien R., Oulmouden A., 2000. A first genotyping assay of French cattle breeds based on a new allele of the extension gene encoding the melanocortin-1 receptor (MC1R). *Genet. Sel., Evol.*, 32, 511-520.
- Sabeti P.C., Varilly P., Fry B., Lohmueller J., Hostetter E., Cotsapas C., Xie X.H., Byrne E.H., McCarroll S.A., Gaudet R., Schaffner S.F., Lander E.S., 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449, 913-918.
- Sanchez M.P., Riquet J., Iannuccelli N., Gogué J., Billon Y., Demeure O., Caritez J.C., Burgaud G., Fève K., Péry C., Lagant H., Le Roy P., Bidanel J.P., Milan D., 2006. Effects of quantitative trait loci on chromosomes 1, 2, 4, and 7 on growth, carcass, and meat quality traits in backcross Meishan x Large White pigs. *J. Anim. Sci.*, 84, 526-537.
- Sanchez M.P., Basso B., Bidanel J.P., Billon Y., Gandemer G., Gilbert H., Iannuccelli N., Larzul C., Legault C., Milan D., Riquet J., Le Roy P., 2007. Detection and localization of a major gene for intramuscular fat in pig. *BMC Genetics*, 8, 55.
- San Cristobal-Gaudy M., Renand G., Amigues Y., Boscher M.Y., Levéziel H., Bibé B., 2000. Traçabilité individuelle des viandes bovines à l'aide des marqueurs génétiques. *INRA Prod. Anim.*, 13, 269-276.
- Schibler L., Roig A., Mahé M.F., Save J.C., Gautier M., Taourit S., Boichard D., Eggen A., Cribiu E.P., 2004. A first generation bovine BAC-based physical map. *Genet. Sel. Evol.*, 36, 105-122.
- Schibler L., Roig A., Mahé M.F., Laurent P., Hayes H., Rodolphe F., Cribiu E.P., 2006. High-resolution comparative mapping among man, cattle and mouse suggests a role for repeat sequences in mammalian genome evolution. *BMC Genomics*, 7, 194.
- Snelling W.M., Chiu R., Schein J.E., Hobbs M., Abbey C.A., Adelson D.L., Aerts J., Bennett G.L., Bosdet I.E., Boussaha M., Brauning R., Caetano A.R., Costa M.M., Crawford A.M., Dalrymple B.P., Eggen A., Everts-van der Wind A., Floriot S., Gautier M., Gill C.A., Green R.D., Holt R., Jann O., Jones S.J.M., de Jong P.J., Kappes S.M., Keele J.W., Larkin D.M., Lewin A.H., McEwan J.C., McKay S., Marra M.A., Mathewson C.A., Matukumalli L.K., Moore S.S., Murdoch B., Nicholas F., Osoegawa K., Roy A., Salih H., Schibler L., Schnabel R., Silveri L., Skow L.C., Smith T.P.L., Sonstegard T.S., Taylor J., Tellam R., Van Tassell C.P., Williams J.L., Womack J.E., Wye N.H., Yang G., Zhao S., 2007. The International Bovine BAC Mapping Consortium. 2007. A physical map of the bovine genome. *Genome Biology*, 8, R165.
- Tilquin P., Barrow P.A., Marly J., Pitel F., Plisson-Petit F., Velge P., Vignal A., Baret P.V., Bumstead N., Beaumont C., 2005. A genome scan for quantitative trait loci affecting the Salmonella carrier-state in the chicken. *Genet. Sel. Evol.*, 37, 539-561.
- Tribout T., Iannuccelli N., Druet T., Gilbert H., Riquet J., Guéblez R., Mercat M.J., Bidanel J.P., Milan D., Le Roy P., 2008. Detection of quantitative trait loci for reproduction and production traits in Large White and French Landrace pig populations. *Genet. Sel. Evol.*, 40, 61-78.
- Vaez M., Follett S.A., Bed'hom B., Gourichon D., Tixier-Boichard M., Burke T., 2008. A single point-mutation within the melanophilin gene causes the lavender plumage colour dilution phenotype in the chicken. *BMC Genetics*, 9, 7.
- Vaiman D., Schibler L., Bourgeois F., Oustry A., Amigues Y., Cribiu E.P., 1996a. A genetic linkage map of the male goat genome. *Genetics*, 144, 279-305.
- Vaiman D., Koutita O., Oustry A., Elsen J.M., Manfredi E., Fellous M., Cribiu E.P., 1996b. Genetic mapping of the autosomal region involved in XX sex-reversal and horn development in goats. *Mammalian Genome*, 7, 133-137.
- Van Laere A.S., Minh N., Braunschweig M., Nezer C., Collette C. *et al.*, 2003. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*, 425, 832-836.
- Venter C.J., Adams M.D., Myers E.W., Li P.W., Mural R.J. *et al.*, 2001. The sequence of the human genome. *Science*, 291, 1304-1351.
- Vignal A., 2000. Etat de la carte de la poule. In : Génétique moléculaire : principe et application aux populations animales, numéro hors série, INRA Prod. Anim., 113-114.
- Vitezica Z.G., Moreno C.R., Lantier F., Lantier I., Schibler L., Roig A., Francois D., Bouix J., Allain D., Brunel J.C., Barillet F., Elsen J.M., 2007. Quantitative Trait Loci linked to PRNP gene controlling health and production traits in INRA 401 sheep. *Genet. Sel. Evol.*, 39, 421-430.
- Weissenbach J., Gyapay G., Dib C., Vignal A., Morissette J., Millasseau P., Vaysseix G., Lathrop M., 1992. A second-generation linkage map of the human genome. *Nature*, 359, 794-801.
- Yerle M., 2000. Etablissement des cartes cytogénétiques et physiques. In : Génétique moléculaire : principe et application aux populations

animales, numéro hors série, INRA Prod. Anim., 87-93.

Yerle M., Echard G., Robic A., Mairal A., Dubut-Fontana C., Riquet J., Pinton P., Milan D., Lahbib-Mansais Y., Gellin J., 1996. A somatic

cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. *Cytogenet. Cell Genet.*, 73, 194-202.

Yerle M., Pinton P., Delcros C., Arnal N., Milan D., Robic A., 2002. Generation and

characterization of a 12,000-rad radiation hybrid panel for fine mapping in pig. *Cytogenet. Cell Genet.*, 97, 219-228.

## Résumé

Cet article retrace les principales étapes et la contribution de l'INRA au développement de la génomique, qui révolutionne depuis deux décennies les connaissances sur la structure et le fonctionnement du génome des animaux d'élevage. L'élaboration, dans les années 90, des premières cartes de marqueurs microsatellites a rapidement permis de mettre en évidence de nombreux locus à effets quantitatifs et de localiser les premiers gènes majeurs. En parallèle, les travaux de cytogénétique et de génomique comparative permettaient de tirer parti de l'avancée des connaissances sur le génome de l'homme et de la souris. A la fin des années 90, la construction d'outils de cartographie à haute densité, cartes d'hybrides d'irradiation et banques de grands fragments, a permis l'essor des travaux de cartographie fine et l'identification des premières mutations causales. Le début des années 2000 a été marqué par le développement des outils d'étude systématique de l'expression des gènes, micro-réseaux et puces à ADN, le démarrage du séquençage des premiers génomes d'animaux d'élevage, et l'essor des bases de données génomiques et de la bioinformatique. La connaissance des séquences permet ensuite de détecter *in silico* leurs variations, en particulier les très nombreuses variations ponctuelles (SNP) et de caractériser finement la structure des génomes des populations animales. La génomique a également renouvelé les méthodes d'amélioration génétique des populations, avec la mise en place de programmes de sélection assistée par marqueurs, de caractérisation et de gestion de la variabilité génétique et le développement d'applications en matière de contrôle de généalogie ou de traçabilité.

## Abstract

### *From genetics to genomics*

This article outlines the main steps and the contribution of INRA to the development of genomics, which has revolutionised over the last two decades the knowledge of livestock genome structure and function. The development in the 1990's of the first microsatellite marker maps quickly allowed the first major genes to be localised and numerous quantitative trait loci to be detected. Meanwhile, cytogenetics and comparative genomics studies allowed livestock genomics to benefit from advances in the knowledge of human and mouse genomes. At the end of the 1990's, the construction of high-density mapping tools such as radiation hybrid maps and large insert libraries made it possible to develop fine mapping programmes and to identify the first causative mutations. The early 2000's have been marked by the development of systematic gene expression tools, *i.e.* micro-arrays and DNA chips, the start of the first animal genome sequencing projects, and the rise of genomic databases and bioinformatics. Available sequence data then made it possible to detect *in silico* their variations, in particular the numerous single point polymorphisms (SNP) and to finely characterise the structure of livestock population genomes. Genomics has also renewed breeding programmes, with the development of programmes for marker-assisted selection, for the characterisation and the management of genetic variability and of parentage control or traceability applications.

BIDANEL J.-P., BOICHARD D., CHEVALET C., 2008. De la génétique à la génomique. *INRA Prod. Anim.*, 21, 15-32.

