

Apport des nouveaux outils de la post-génomique aux recherches en physiologie chez le porc

Marie DAMON (1), Isabelle P. OSWALD (2), Isabelle LURON (1) et François HATEY (3).

(1) I.N.R.A. Unité Mixte de Recherches sur le veau et le Porc, 35590 Saint-Gilles

(2) I.N.R.A. Laboratoire de Pharmacologie-Toxicologie, 31931 Toulouse

(3) I.N.R.A. Laboratoire de Génétique Cellulaire, 31326 Castanet Tolosan

Apport des nouveaux outils de la post-génomique aux recherches en physiologie chez le porc

Les recherches en génomique ont permis de réaliser des progrès considérables dans l'élaboration de cartes physiques et génétiques et de déterminer la séquence de nombreux génomes. Elles ont ouvert la voie et grandement contribué à l'acquisition de nouvelles connaissances en physiologie. Parallèlement, le développement de nouvelles technologies d'analyses moléculaires à grande échelle (transcriptomique et protéomique) et l'essor de la bio-informatique rendent maintenant possible l'exploitation de ce flux de données et permettent par une étude globale de l'expression des gènes d'avoir une vision moins réductionniste du vivant. Ces progrès sont à l'origine d'une nouvelle discipline de la génomique appelée la post-génomique (ou génomique fonctionnelle). La post-génomique est une approche globale de fonctions biologiques complexes qui impliquent de nombreux gènes. Elle vise à comprendre la fonction et la régulation de tous les gènes et protéines qui conditionnent le fonctionnement de l'organisme. L'utilisation de cette approche dans le domaine médical a d'ores et déjà porté ses fruits. En revanche dans le domaine des productions animales, les programmes qui utilisent ce type d'approche sont plus récents. Après une présentation des nouvelles technologies mises en œuvre, cette synthèse exposera les intérêts d'une telle démarche et les perspectives offertes en production porcine pour l'étude de la qualité de la viande, de la nutrition, de la reproduction et de la santé.

Contribution of new post-genomic technologies to research in pig physiology

Genomic research has enabled considerable progress to be made in the establishment of physical and genetic maps and well as in the determination of numerous genome sequences. Consequently, it has helped in the acquisition of new knowledge concerning biological systems. At the same time, the development of large scale, high throughput molecular analysis systems (transcriptome and proteome analyses) and the improvement in the bioinformatics area will allow further exploitation of this data, resulting in a more comprehensive knowledge of living systems. This new field of genomics, which is called post-genomics (or functional genomics) is currently developing. It is a global approach used to look at complex biological functions concerning many genes. It aims to help in the understanding of the function and regulation of all genes and proteins that play important roles in how an organism works. Although in medical research, the post-genomic approach has already provided some potential applications, in livestock production this strategy is rarely exploited. In this paper, we present the different technologies which can be used in post-genomic research. We then go on to describe their potential application in the study of different aspects of pig meat quality, nutrition, reproduction and health.

INTRODUCTION

L'amélioration de la compétitivité de l'élevage, de la conduite des troupeaux et de la qualité des produits d'origine animale nécessite notamment une maîtrise accrue de la croissance, de la reproduction, de l'alimentation, de la santé et du bien-être animal. Celle-ci repose en partie sur une connaissance de plus en plus approfondie des grandes fonctions physiologiques des espèces domestiques et des interactions entre ces fonctions.

Par ailleurs, le séquençage entier de génomes modèles (levure, nématode, drosophile) et du génome humain (International Human Genome Sequencing Consortium, 2001) est sans doute le fait le plus médiatique de la génomique. Ce terme introduit en 1986 lors de l'apparition d'une revue scientifique « Genomics », se définit comme la science qui étudie les génomes. Les apports de la génomique sont considérables. Ils résident principalement dans le caractère systématique des études, le recours à des systèmes d'analyse à haut débit, le travail en réseaux /consortium et l'établissement de liens de plus en plus étroits entre la génétique moléculaire et l'informatique. La génomique a ainsi fait naître de nouveaux objets (transcriptome, protéome, analyse des réseaux d'interactions...) qui vont progressivement changer notre manière de concevoir l'étude de la physiologie. En effet, on parle maintenant de « post-génomique » (ou génomique fonctionnelle, les deux termes étant équivalents) pour désigner les recherches de biologie fonctionnelle où les produits des gènes (ARNm et protéines) sont au cœur des analyses et non plus seulement les gènes et leurs séquences. Cette nouvelle ère de la biologie nécessite le développement de nouvelles technologies d'investigation massives (d'où le suffixe -ome) de la fonction des gènes. Ces technologies doivent réduire le décalage entre la vitesse et la relative facilité avec laquelle des milliers de séquences sont obtenues et les efforts nécessaires pour connaître le rôle exact d'une seule protéine. Ainsi, la démarche génomique peut être considérée comme le point de départ de démarches post-génomiques plus globales qui intègrent différents niveaux d'organisation, depuis la fonction des gènes dans la cellule jusqu'aux fonctions complexes de l'organisme en interaction avec son environnement.

L'objectif de cette synthèse est d'éclaircir la notion de post-génomique, de présenter les nouvelles technologies d'investigation moléculaire ainsi que quelques perspectives offertes par ce type d'approche en production porcine pour la maîtrise de la croissance musculaire, de la reproduction, de la nutrition et de la santé. En raison d'un relatif retard de ces recherches chez les animaux d'élevage, ces applications seront illustrées par des données d'ores et déjà disponibles dans d'autres espèces (principalement l'homme et la souris) ou par des projets et des études en cours chez le porc.

1. LA POST-GÉNOMIQUE : ENJEUX ET TERMINOLOGIES

L'adage « 1 gène, 1 protéine, 1 fonction » (MORANGE, 2000) est remis en cause par les données issues du séquençage du génome humain : l'estimation de 30 000 à 40 000 gènes semble indiquer que la complexité de l'organism

me résulte plus de la régulation des gènes, de leurs interactions et de celles de leurs produits que de leur nombre (JORDAN, 2000). De plus le phénomène bien connu d'épissage alternatif, qui produit plusieurs transcrits et donc plusieurs protéines à partir d'un même gène, joue un rôle essentiel dans les processus biologiques. Des données récentes, issues du séquençage des génomes, suggèrent en effet que la fréquence de ce phénomène est loin d'être négligeable (de l'ordre de 40 %, BRETT et al., 2000). Les études de post-génomique visent donc, grâce au développement de nouvelles technologies (cf 2), à comprendre comment l'activité intégrée des gènes de l'organisme peut rendre compte des fonctions physiologiques et comportementales complexes de celui-ci.

De la même façon que l'ensemble des gènes d'un organisme constitue son génome, l'ensemble des ARNm issus de la transcription des gènes et des protéines exprimées sont appelées respectivement le transcriptome et le protéome (fig.1). Cependant, alors que le génome (l'ensemble des gènes) est constant dans toutes les cellules d'un organisme le protéome et le transcriptome peuvent varier considérablement dans le temps (en fonction du stade de développement), dans l'espace (selon le type cellulaire, tissulaire ou organique) ou/et selon le statut physiopathologique. Ce sont donc des entités dynamiques dont l'étude est complémentaire et constitue le cœur d'une démarche de post-génomique.

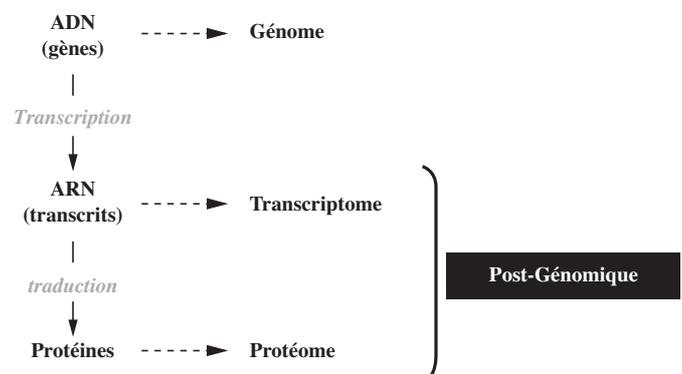


Figure 1 - Les différents niveaux d'approche de la génomique

2. PRÉSENTATION DES DIFFÉRENTES TECHNIQUES

Jusqu'à présent et en raison de contraintes techniques, les expériences de biologie moléculaire étaient limitées à l'analyse d'un ou quelques gènes ou protéines, donnant peu d'informations compte tenu de la complexité des systèmes étudiés. Cette approche ne peut suffire pour la compréhension de systèmes biologiques contenant plusieurs milliers de gènes ; une vision globale aux niveaux de l'ADN, de l'ARN et des protéines est indispensable. Toute une batterie de nouvelles technologies d'analyses moléculaires, robotisées et à grande échelle a donc vu le jour : « puces » à ADN pour l'étude du transcriptome, électrophorèse bidimensionnelle associée à la spectrométrie de masse pour le protéome, technique du double hybride pour la recherche d'interactions protéines-protéines...

2.1 Le transcriptome et les « puces » à ADN

L'application la plus importante de l'analyse du transcriptome est la mesure de l'expression des gènes. Le transcriptome est en effet un déterminant majeur du phénotype et de la fonction cellulaire. La transcription d'un gène en ARN est la première étape de la synthèse des protéines et des variations dès cette étape peuvent être responsables de changements morphologiques (développement) ou indiquer une réponse cellulaire à des perturbations environnementales. Le transcriptome est en effet très dynamique et change rapidement et de façon importante en réponse à des perturbations ou lors d'évènements cellulaires tels que la réplication de l'ADN et la division cellulaire. Si l'on veut comprendre la fonction d'un gène il est important de connaître quand, où et à quel niveau ce gène est exprimé. De plus, l'analyse simultanée de milliers de gènes peut conduire aux mécanismes de régulation permettant de mieux appréhender le fonctionnement cellulaire et les processus biochimiques, matérialisant ainsi le caractère « intégratif » de l'approche.

2.1.1. Méthodologies

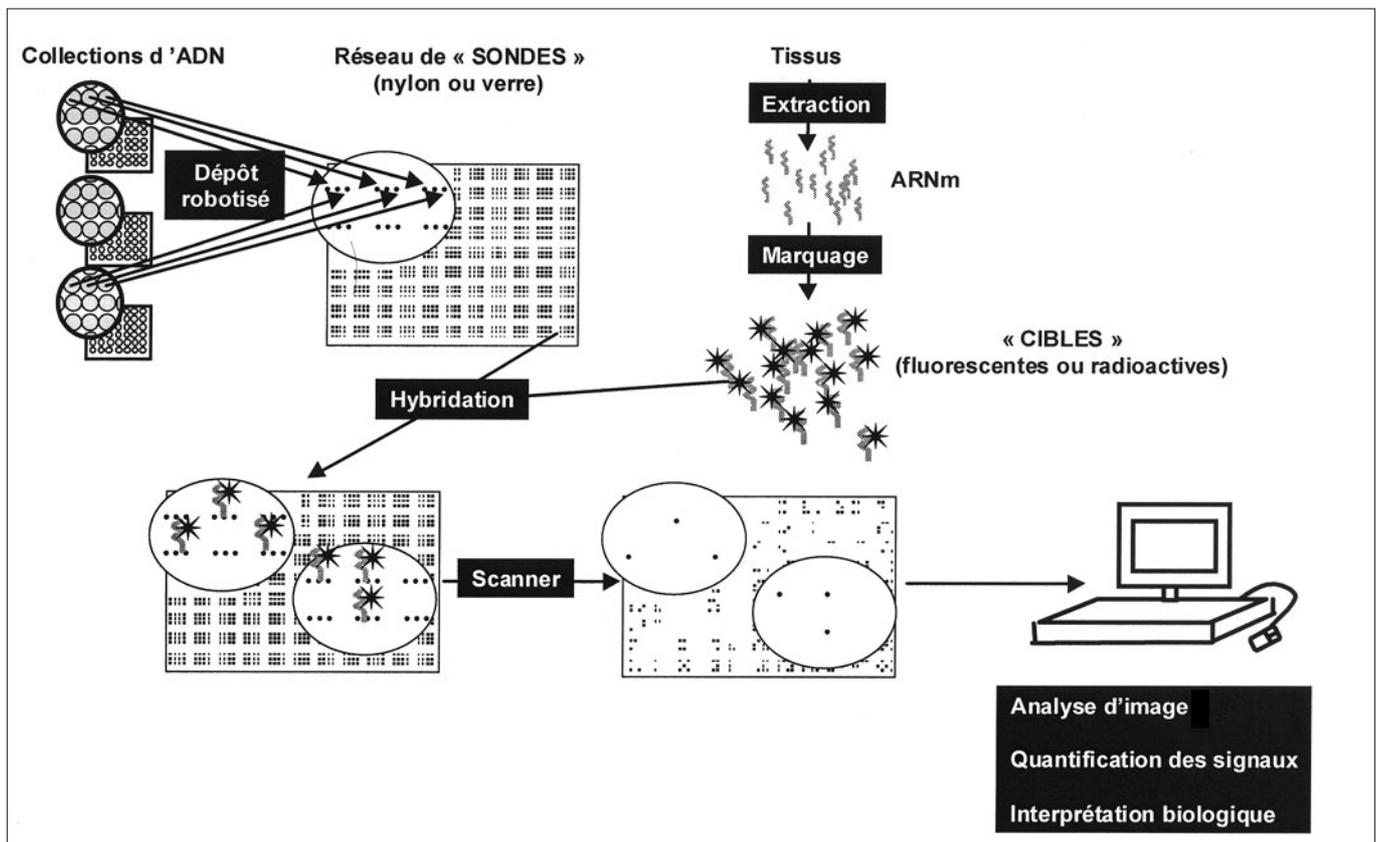
Les techniques appropriées à la mesure des niveaux d'expression existent depuis de nombreuses années (Northern blot, RT-PCR). Parmi les plus récentes, le mRNA differential display (LIANG et PARDEE, 1992) est une technique basée sur la PCR dans laquelle des sous-populations de transcrits sont comparées entre elle pour identifier ceux dont le niveau d'expression varie en fonction du traitement ou de la situation expérimentale. Une approche plus efficace est l'hybridation soustractive suppressive (HSS) qui combine soustraction

et normalisation pour permettre l'identification des transcrits exprimés différemment tout en réduisant la redondance du produit soustrait (DIATCHENKO et al., 1996). Cependant, ces méthodes sont souvent sujettes aux faux positifs et sont souvent très coûteuses et laborieuses (WATSON et al., 2000). Deux stratégies dominantes se sont donc développées pour effectuer une analyse quantitative et qualitative en série de l'expression des gènes : les méthodes de séquençage d'étiquettes (méthode SAGE) et d'hybridation (les « puces » à ADN).

La technique SAGE ou analyse en série de l'expression de gènes, est basée sur l'isolement à partir d'une population d'ARNm, de courtes séquences ou étiquettes d'une dizaine de paires de bases (Tag) représentatives de chacun des messages (VELCULESCU et al., 1995). Après une étape de séquençage massif, l'analyse informatique permet de donner le niveau d'expression de chaque gène en fonction du nombre d'occurrence des étiquettes provenant de ce gène, par rapport au nombre total d'étiquettes analysées. Toutefois, cette technique n'est pas adaptée à l'étude de nombreux échantillons et requiert de grandes bases de données de séquences génomiques pour les espèces étudiées.

Grâce à la robotisation, la miniaturisation des techniques, les « puces » à ADN apportent pour la première fois la possibilité de détecter des milliers de molécules d'acides nucléiques de façon simultanée sur des matrices solides mesurant quelques centimètres carrés. Cette technologie repose sur l'hybridation (SOUTHERN, 1975) c'est-à-dire l'appariement de séquences nucléiques complémentaires (WATSON et CRICK, 1957). Les « puces à ADN » (fig.2)

Figure 2 - Principe des réseaux et puces à ADN



sont constituées d'un support solide (lame de verre identique à celles utilisées en microscopie traditionnelle ou membrane de nylon) sur lequel des milliers de fragments d'ADN sont déposés de façon géométrique à l'aide d'un système robotisé (spotteur). Grâce à cette technique, chacun des fragments d'ADN est représenté par un point sur le support (ou puce). Ils servent d'appâts (ou sondes) pour fixer de façon très spécifique les fragments de gènes complémentaires (ou cibles) présents dans les échantillons biologiques à tester : l'« hybridation » peut être mise en évidence grâce au « marquage » préalable de l'échantillon au moyen de traceurs fluorescents ou radioactifs. La quantification des signaux obtenus et l'identification des fragments de gènes reconnus sont ensuite réalisées par des systèmes informatiques d'acquisition d'image et d'analyse de données. Les résultats obtenus sont ensuite validés par une analyse statistique et interprétés dans leur contexte biologique (LOCKARD et WINZELER, 2000 ; BERTUCCI et al., 2001, FREEMAN et al., 2000).

Le terme « puces à ADN » est générique. On distingue actuellement deux grands types différents de puces qui se distinguent essentiellement par leur procédé de fabrication et leur densité : (1) les macro et microréseaux avec un dépôt direct de molécules d'ADN sur leur support et (2) les puces avec la synthèse in situ des sondes oligonucléotidiques sur une surface solide. Dans le premier cas, il est nécessaire d'établir une collection de sondes spécifiques d'un grand nombre de gènes, dans le second, il est nécessaire d'en connaître la séquence.

Pour les macroréseaux, les dépôts (sondes) sont des clones d'ADNc ou des produits de PCR déposés à haute densité sur une membrane de nylon (8 x 12 cm). Le marquage est le plus souvent radioactif et le criblage est réalisé en excès de cible, on obtient ainsi une mesure de l'abondance relative de chacun des ARNm présents dans l'échantillon de départ. On parle de macroréseau jusqu'à une densité d'environ 2500 fragments d'ADN déposés sur une telle membrane (ZHAO et al., 1995 et NGUYEN et al., 1995).

Mis au point par l'équipe de P. Brown à Stanford (SCHENA et al., 1995), les microréseaux permettent la miniaturisation des dépôts d'ADN. Les sondes sont généralement des fragments d'ADN amplifiés par PCR ou des oligonucléotides longs (50-70 mères) greffés sur la puce. 12 000 sondes peuvent être déposées sur une lame, préalablement traitée chimiquement. Les cibles sont marquées par des fluorophores différents (Cy-3 vert et Cy-5 rouge) et sont hybridées simultanément. Les signaux d'hybridation sont analysés grâce à un lecteur capable de discriminer les 2 fluorochromes et de générer des images en fausses couleurs avec des spots allant du vert au rouge en passant par le jaune selon les rapports d'expression des différents gènes dans les deux échantillons. Le rapport des intensités de fluorescence rouge/fluorescence verte pour chacun des spots donne une mesure de l'expression différentielle des gènes dans les 2 échantillons biologiques étudiés. Par ailleurs, on peut également essayer de regrouper des gènes ayant le même profil d'expression sur plusieurs expériences ayant un rapport biologique. Cette comparaison permet d'inférer que des gènes sont sous la dépendance d'un même mécanisme de régulation.

Les puces à ADN proprement dites sont bien représentées par les puces de la société Affymetrix. Dans ce cas, les sondes sont des oligonucléotides de 20 ou 25 mères synthétisés in situ par une technique de photolithographie. On peut ainsi disposer d'un jeu d'oligonucléotides représentant 30 000 gènes sur une surface d'environ 1 cm². A chaque ARNm correspondent deux jeux de sondes, l'un parfaitement complémentaire de la cible et l'autre comprenant en position centrale un mésappariement, ce qui permet de définir un bruit de fond non spécifique. La différence des signaux relatifs aux hybridations de ces deux types de sondes est directement proportionnelle à l'abondance de l'ARNm dans la cible. Cette technique est d'une spécificité supérieure aux autres méthodes mais elle ne permet pas d'étudier l'expression de gènes inconnus (LOCKHART et al., 1996, MAHADEVAPPA et WARRINGTON 1999)

Ils existent d'autres technologies qui utilisent des électrodes de carbone, des fibres optiques, des métaux précieux... mais ces techniques sont encore le plus souvent en développement (GELI, 2000 ; JORDAN, 2002).

2.1.2 Intérêts et limites des « puces » à ADN

La très grande souplesse des techniques de fabrication et d'hybridation des réseaux d'ADN en fait une technique utilisable dans beaucoup de laboratoires et aux champs d'applications très vastes (cf 2.4). Les principales limites, dans l'utilisation de cette technique, résident dans la disponibilité des clones des séquences appâts (sondes) déposées sur le support, dans la qualité de l'extraction des ARN des échantillons à analyser et dans la sensibilité de cette technique. Cette sensibilité est cependant tout à fait raisonnable car à l'heure actuelle quelle que soit la plate-forme utilisée, on peut atteindre des niveaux de sensibilité de l'ordre de 1/100 000 à partir de quelques microgrammes d'ARNm (GRANJEAUD et al., 1999). Le problème est donc plus lié à la quantité d'échantillon disponible nécessaire pour obtenir une quantité suffisante d'ARNm de bonne qualité. Des techniques d'amplification des ARNm avant marquage ont donc été mises au point mais elles pourraient fausser la distribution réelle des ARNm (abondance relative) dans l'échantillon de départ. Bien qu'il soit possible de déposer des séquences non caractérisées (correspondant à des gènes inconnus), la démarche la plus utilisée est celle d'un choix pertinent de sondes à déposer en limitant la redondance et en choisissant un répertoire de gènes le plus large possible (DUGGAN et al., 1999). Les puces Affymetrix sous couvert d'un brevet, sont en revanche moins souples d'utilisation, très coûteuses et couvrent un nombre d'espèces encore très limité.

L'avantage clé des puces réside dans le fait qu'il n'est pas nécessaire de prédire quels seront les gènes importants ou les mécanismes mis en avant dans le processus biologique étudié. Elles donnent une vision plus complète et objective du fonctionnement cellulaire (LOCKHART et WINZELER, 2000). Cette approche, horizontale et massive ne fait aucune hypothèse sur les acteurs d'une fonction. Elle n'est pas naturelle pour un physiopathologiste conventionnel, qui trouve plus cartésienne une recherche verticale, fondée sur des hypothèses que l'acceptation ou le rejet font progresser dans telle

ou telle direction. Cette analyse apportera cependant des données totalement nouvelles qui serviront de support à des hypothèses originales pour mener des recherches plus conventionnelles.

Outre la mesure des profils d'expression, l'autre grand intérêt des puces à ADN consiste à étudier les variations affectant la séquence d'ADN, ou polymorphisme génétique (WANG et al., 1998). Cet aspect est traité dans la synthèse de BIDANEL et al., 2003.

2.2. Le protéome et la spectrométrie de masse

Si l'analyse de l'expression des gènes au niveau du transcriptome est extrêmement profitable et largement pratiquée, la corrélation entre les taux d'ARNm et de protéines qui sont les véritables « ouvrières » de la cellule n'est malheureusement pas très bonne en raison principalement de l'existence des régulations post-traductionnelles (méthylation, glycosylation, phosphorylations...). De plus, l'activité des protéines dépend beaucoup de leur environnement physico-chimique (pH, potentiel d'oxydo-réduction, conformation...), dimension qui n'est évidemment pas prise en compte par la seule analyse du transcriptome (CANTOR, 2000). Les techniques de l'analyse du protéome se sont donc également développées.

Le terme de protéome est apparu en 1995 (WASINGER et al., 1995) et désigne l'ensemble des produits fonctionnels du génome afin de mieux appréhender la complexité du fonctionnement cellulaire. ANDERSON et ANDERSON (1998) définissent l'analyse protéomique comme « l'utilisation de la quantification au niveau de la protéine comme mesure objective de l'expression génique caractérisant un processus biologique donné (processus physiologique ou pathologique, effets de médicaments, de l'environnement...) et comme moyen de décodage des mécanismes contrôlant cette expression ».

2.2.1. Méthodologies

A l'heure actuelle, il n'existe pas d'équivalent des puces à ADN pour les protéines, en raison de contraintes technologiques importantes (JORDAN, 2002 ; ABBOTT, 2002). Cette technologie en est donc encore au stade initial de développement et pour l'instant les puces à protéines disponibles sur le marché ne permettent d'étudier que quelques dizaines de protéines. Les éléments essentiels à la construction de telles puces sont la production et l'isolement d'un large répertoire de molécules de reconnaissance avec lesquelles vont interagir spécifiquement et de façon extrêmement sensibles les protéines cibles, les techniques d'immobilisation sur la puce de ces molécules de reconnaissance et la technique de détection de la liaison de la protéine cible sur la molécule de reconnaissance. Actuellement les molécules de reconnaissance le plus souvent utilisées sont des anticorps mais ceux-ci sont très sensibles à la dénaturation, l'existence de brevets limite leur utilisation et il semble difficilement concevable de produire des anticorps spécifiques de toutes les protéines d'un organisme. Parmi les technologies en développement, celle de la société Cyphergen semble très prometteuse. Cette puce se présente comme une barrette d'aluminium dont la

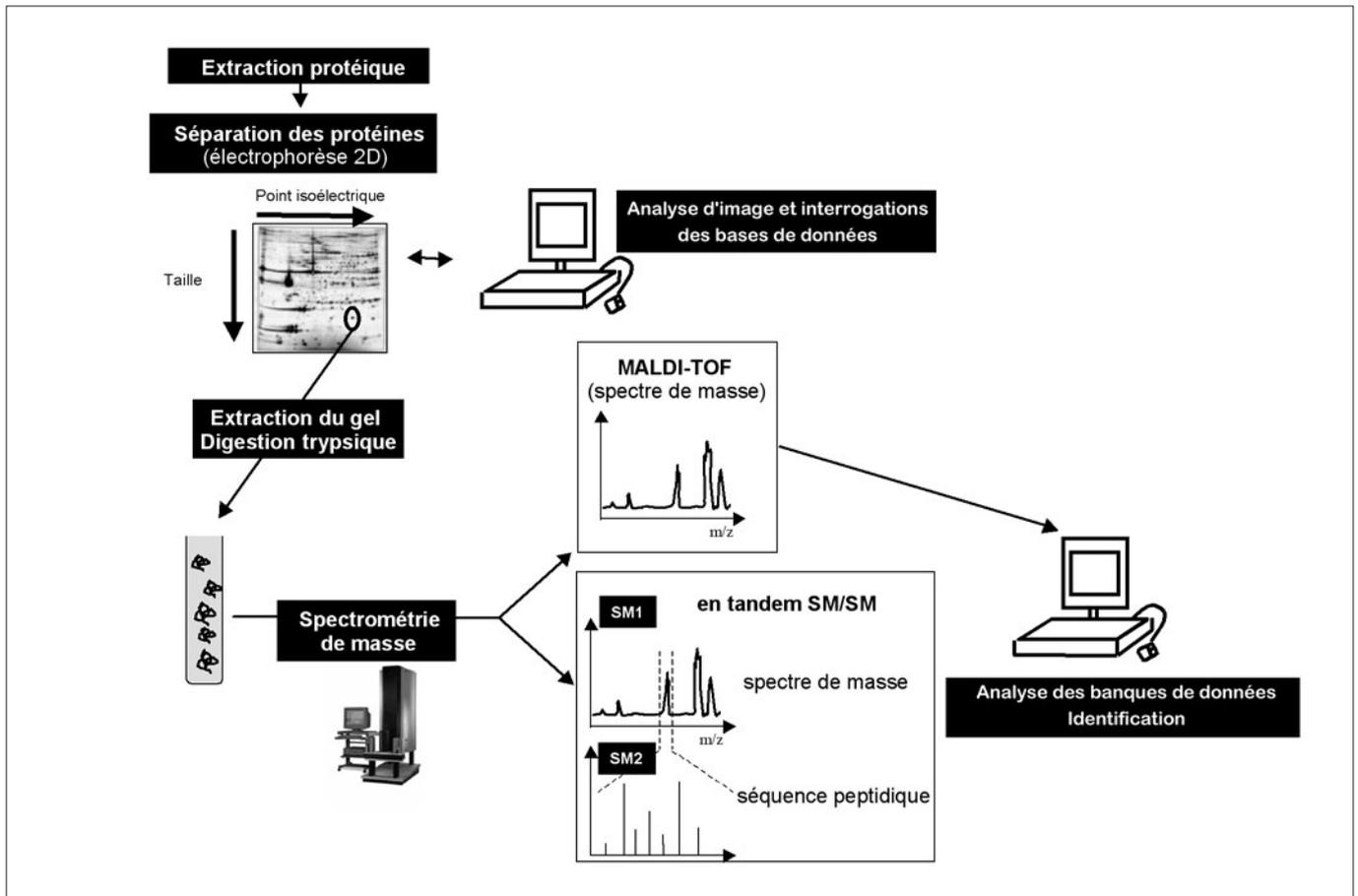
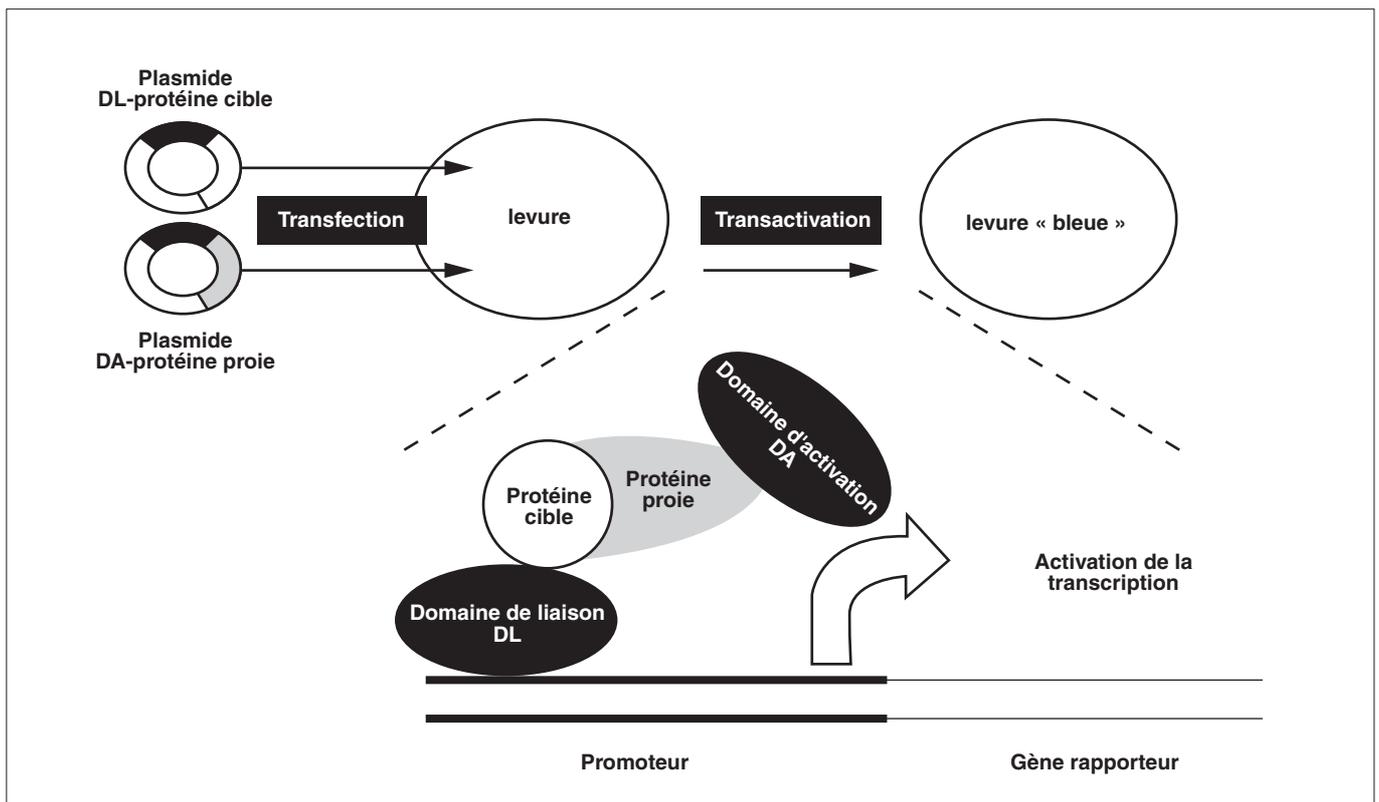
surface a été modifiée afin de capter les protéines présentes dans l'échantillon. Les supports disponibles actuellement utilisent le même principe que la chromatographie (échangeurs d'ions, résidus hydrophobes ou hydrophiles...) ou des ligands biologiques. L'échantillon biologique est déposé sans traitement préalable puis la barrette est introduite dans un spectromètre de masse (SELDI, Surface Enhanced Laser Desorption Ionisation). Après ionisation, le spectromètre de masse mesure les masses molaires de ces protéines, ce qui permet leur identification.

Outre la puce à protéines qui est sans doute une voie d'avenir, les deux grandes techniques d'analyse protéomique utilisées couramment aujourd'hui sont la comparaison ou la quantification de profils d'expression de protéines par électrophorèse bidimensionnelle couplée à la spectrométrie de masse et l'identification de complexes protéiques, grâce notamment à l'utilisation de la technique du double-hybride.

L'analyse des profils d'expression protéique comporte trois étapes principales qui font appel à différentes technologies (fig.3). L'analyse débute par une électrophorèse bidimensionnelle (gel 2D) qui combine deux types de séparations protéiques réalisées selon le point isoélectrique puis la masse, ce qui lui confère un pouvoir de résolution élevé (O'FARRELL, 1975). La deuxième étape est la mise en image et le traitement des cartes protéiques que constituent les gels 2D. La dernière étape est celle de l'identification et de la caractérisation des protéines, le plus souvent réalisée par spectrométrie de masse. Deux grands types de spectromètre sont utilisés en protéomique : MALDI-TOF (Matrix Assisted Laser Desorption Ionisation Time Of Flight) et MS/MS (spectromètre de masse en tandem). L'appareil MALDI-TOF sépare les peptides selon leur masse et permet d'obtenir une carte peptidique massique (CPM). La protéine est alors identifiée après comparaison de la CPM aux CPMs théoriques des séquences présentes dans les bases de données protéiques. Si la séquence protéique est inconnue et donc non répertoriée dans les bases de données, la CPM est inopérante. On a alors recours au spectromètre MS/MS, qui permet le séquençage de peptide (THEBAULT et al., 2001).

L'identification des protéines par les techniques précédentes est automatisable et applicable à une grande échelle. Cependant comme l'efficacité d'ionisation est variable selon les protéines, la spectrométrie de masse ne peut fournir directement des données réellement quantitatives tandis que le caractère quantitatif de l'analyse d'image des cartes protéiques issues des gels 2D est encore critiqué. On a alors recours à des méthodes d'analyse protéomique quantitative qui font le plus souvent appel à l'incorporation dans les échantillons d'isotopes lourds stables comme étalon interne. Dans cette méthode l'électrophorèse 2D et l'analyse d'image des cartes protéiques sont supprimées. Cependant, toutes les modifications post-traductionnelles ne peuvent être ainsi identifiées (THEBAULT et al., 2001).

Une autre stratégie de la protéomique, souvent oubliée, vise à établir des cartes d'interactions protéiques de manière à reconstituer des voies métaboliques. De telles cartes permettent de guider sur le rôle d'une protéine totalement inconnue, pour peu qu'elle interagisse avec une autre dont on connaît

Figure 3 - Stratégie d'une analyse protéomique (d'après THEBAULT et al., 2001)**Figure 4** - Présentation de la technique du double hybride

la fonction. Pour cela, des plate-formes technologiques se sont développées en se fondant sur la technique du double hybride (FREDERICKSON, 1998 ; fig.4). Celle-ci consiste à étudier, dans un organisme (levure ou bactérie), la capacité d'une protéine connue (la cible) à interagir avec une protéine inconnue (la proie) *in vivo*. L'idée repose sur l'utilisation d'un facteur de transcription issue de la levure : la protéine Gal4p. Ce facteur comporte 2 domaines fonctionnels : un domaine de liaison à l'ADN (DL) et un domaine d'activation de la transcription (DA). Pour déterminer si la proie interagit avec la cible, on transforme des levures avec 2 plasmides. L'un est capable de produire la protéine de fusion DL-cible et l'autre la protéine proie-DA. Dans ce système, la levure pourra dégrader le galactose (synthèse d'un pigment bleu) si et seulement si il y a formation d'un complexe moléculaire qui active l'expression du gène rapporteur.

2.2.2. Intérêts et limites

Comparées aux analyses du transcriptome, les techniques basées sur l'électrophorèse 2D ne permettent pas une étude aussi exhaustive dans un organisme ou une cellule. L'hétérogénéité chimique des protéines, dont les points iso-électriques peuvent varier de 1 à 12 et les PM de moins de 5 kDa à plus de 300 kDa, pose un réel problème de résolution des gels 2D. Une autre limite réside dans les différences de solubilité des protéines. Ainsi, les protéines hydrophobes, membranaires sont difficilement analysables. De ce fait les protéines régulatrices de la cellule (récepteurs membranaires, protéines nucléaires...) qui sont présentes en faible quantité échappent à ces analyses. De plus, les conditions dénaturantes nécessaires à la solubilisation peuvent faire obstacle à leur caractérisation fonctionnelle ultérieure. Enfin, le dernier problème est lié à la dynamique d'expression des protéines pouvant couvrir 6 ordres de grandeur (10^6). De ce fait, il existe un biais important car les protéines très fortement exprimées sont plus facilement détectables que les protéines faiblement exprimées. Or le niveau d'expression n'est pas lié à l'importance de la protéine dans la cellule et de nombreuses protéines essentielles sont faiblement exprimées.

Dans les limites indiquées ci-dessus, ces techniques sont adaptées à l'étude des modifications post-traductionnelles essentielles pour la fonction des protéines puisqu'elles peuvent en déterminer l'activité, la stabilité, la localisation et le *turn-over*. Cependant, seules les phosphorylations sont facilement décelables par spectrométrie de masse (PANDEY et MANN, 2000). De plus, cette approche peut également discriminer les variants d'épissage dont nous avons parlé plus haut.

Un intérêt majeur des analyses protéomiques réside dans le fait qu'elles permettent d'appréhender le fonctionnement des organites (mitochondrie, Golgi...) par fractionnement subcellulaire, une caractérisation qui n'est pas prise en compte par les analyses du transcriptome. La localisation subcellulaire d'une protéine inconnue est une information importante pour la recherche de sa fonction. Ces études ciblées sont considérées comme prometteuses pour accéder aux protéines mineures.

A ce jour, les méthodologies disponibles n'accèdent pas encore au haut débit atteint par la transcriptomique. Le développement de la spectrométrie de masse et des puces à protéines devraient cependant à l'avenir atteindre cet objectif tout en augmentant encore la sensibilité des analyses.

2.3. Analyse de données, bio-informatique et modélisation des réseaux de gènes

Si les puces à ADN ont maintenant presque une dizaine d'années et mettent en évidence de longues listes de gènes aux profils d'expression altérés, l'analyse des données reste quant à elle assez subjective et encore en développement (BUSTIN et DORUDI, 2002). Les analyses ont longtemps été basées sur l'étude des ratios d'expression des gènes entre 2 conditions physiologiques. Aujourd'hui, les méthodes statistiques sont souvent des analyses de variance (ANOVA) et/ou un test bootstrap de ré-échantillonnage des résidus qui semblent les meilleurs moyens d'obtenir pour chacun des gènes des mesures d'expression différentielle et un intervalle de confiance correct (KERR et al., 2000). Pour identifier des associations de profils d'expression la plupart des données sont analysées par des méthodes dites de « clustering » : classification hiérarchique, regroupement à partir de la moyenne K, la méthode SOM (Self Organizing Maps)... La méthode la plus connue est celle de EISEN et al., (1998) qui est une classification hiérarchique de gènes co-régulés ou fonctionnellement proches s'apparentant aux méthodes utilisées lors des analyses phylogénétiques. En amont de ces analyses, le biologiste a évidemment recours aux moyens informatiques (ordinateur et logiciels) pour gérer et stocker d'énormes quantités de données d'expression issues de l'analyse d'image.

La bio-informatique quant à elle est une discipline de la biologie dont la définition porte souvent à confusion. En effet, la bio-informatique est constituée par l'ensemble des concepts et des techniques, qui en l'absence d'expérimentation sont nécessaires à l'interprétation de l'information biologique (bioinformation) contenue dans les séquences et les structures des macromolécules biologiques (CLAVERIE, 1999). Elle n'est donc pas un dérivé de l'informatique mais une discipline biologique qui utilise l'ordinateur et ses méthodes de calculs. On parle aussi de biologie *in silico*. Le but de la bio-informatique tout comme la transcriptomique ou la protéomique est de comprendre le fonctionnement de la cellule et elle participe à la démarche de biologie intégrative de la post-génomique. Elle est absolument nécessaire pour identifier des biomarqueurs (gènes ou protéines) à partir d'autres critères que les similarités de séquence ou de structure comme les profils d'expression et le polymorphisme.

Dans le cadre de la post-génomique, l'enjeu central est de comprendre comment les macromolécules (gènes, ARN et protéines) coopèrent pour l'accomplissement des fonctions biologiques dans l'organisme (JACQ et THIEFFRY, 2000). L'exploitation des données bibliographiques, l'expérimentation (cf 2.1 et 2.2) et la bio-informatique sont nécessaires pour mettre en évidence et modéliser le fonctionnement des réseaux génétiques formés par les nombreuses interactions entre toutes ces macromolécules. L'étude de ces réseaux d'in-

teraction n'en est encore qu'à ses premiers pas. Cependant, la modélisation du fonctionnement des gènes semble d'ores et déjà capitale pour comprendre comment l'information contenue dans les gènes d'un individu (génotype) s'exprime et se combine à d'autres facteurs pour déterminer ses caractéristiques réelles (phénotype).

2.4. Principaux champs d'applications de ces nouvelles technologies

Les champs d'application des outils de la post-génomique sont nombreux et ils concernent toutes les sciences du vivant, en recherche fondamentale ou appliquée (agronomie, médecine, pharmacologie). Les applications citées ici ne sont donc pas exhaustives. Au niveau fondamental, ces démarches peuvent par exemple permettre d'identifier de nouveaux gènes contrôlant certains processus biologiques, de mieux comprendre les mécanismes d'action mis en jeu, d'identifier des réseaux de gènes co-régulés, de préciser ou définir la fonction du produit des gènes... En pharmacologie, de grands espoirs sont placés dans la post-génomique pour le développement de nouvelles molécules thérapeutiques et pour étudier les effets secondaires des médicaments. Les liens entre la génomique et la recherche médicale sont très étroits et c'est dans ce domaine que les puces ont été les plus productives. Les nouveaux outils peuvent permettre d'identifier des gènes cibles pour les mutations à l'origine des maladies mono et polygéniques, de réaliser un diagnostic grâce à la mise en évidence de gènes marqueurs ou de profils d'expression typiques d'une maladie ou même d'établir un pronostic... Les puces à ADN permettent en effet de déceler rapidement les mutations génétiques indiquant qu'un individu a de fortes probabilités d'être atteint d'une maladie. Dans ce cadre, un test de dépistage pour les femmes risquant d'être atteintes d'un cancer du sein est d'ores et déjà envisagé. Cette pratique de dépistage soulève de nombreuses questions d'ordre éthique. Le principe du diagnostic repose sur la détection du gène BRCA codant une protéine intervenant dans la réparation de l'ADN. La jonction commence tout juste à s'établir entre la post-génomique et les productions animales. Cependant ces outils seront utiles dans de nombreux domaines : la préservation des ressources génétiques, l'optimisation des techniques d'élevage, l'amélioration des produits par une valorisation des caractères d'intérêts économiques comme la qualité de la viande, la prophylaxie sanitaire ou encore la protection de l'environnement.

3. APPLICATIONS À L'ÉTUDE DE GRANDES FONCTIONS PHYSIOLOGIQUES

3.1. La physiologie musculaire

3.1.1 Recherches fondamentales ou médicales

Les travaux concernant la fonction musculaire sont essentiellement destinés à mieux comprendre le fonctionnement musculaire au cours du développement (différenciation, vieillissement), lors d'une activité physique, suite à des modifications nutritionnelles (cf 3. 2) ou lors de pathologies (atrophie musculaire, diabète). Au travers de quelques exemples, nous

pourrons découvrir certaines potentialités des études de post-génomique. Les études publiées ne concernent pour l'instant que l'analyse du transcriptome.

Un répertoire de gènes exprimés spécifiquement dans le muscle a été établi chez l'homme et 48 nouveaux gènes exprimés spécifiquement dans le muscle ont ainsi été découverts (PIETU et al., 1996 ; PIETU et al., 1999). Chez la souris, l'analyse comparative d'un muscle rouge oxydatif (soleus) et d'un muscle blanc glycolytique (quadriceps) a été réalisée afin de mieux comprendre l'origine des différences métaboliques et contractiles observées entre ces 2 muscles (CAMPBELL et al., 2001). Parmi les 3000 gènes analysés, 49 montrent une expression différentielle entre ces deux muscles de types métabolique et contractile contrastés. Pour une majorité, ils sont impliqués dans la contraction musculaire ou le métabolisme mais 9 nouveaux gènes, impliqués dans la signalisation intracellulaire ou le contrôle de la transcription sont également mis en évidence dans cette étude. En utilisant des microréseaux d'ADN comportant 4000 gènes humains, ROTH et al., (2002) ont identifié différents facteurs contrôlant fortement l'expression des gènes musculaires comme le sexe (200 gènes modulés), l'âge (50 gènes modulés) et l'activité physique (70 gènes modulés).

L'étude des profils d'expression des gènes au cours de la myogénèse a été effectuée chez la souris (MORAN et al., 2002). Sur les 11 000 séquences criblées, 400 gènes montrent des variations d'expression au cours de ce processus. Le regroupement de ces gènes en fonction de leur profil d'expression révèle que la plupart des modifications d'expression ont lieu lors de la transition prolifération/différenciation ou au cours de la différenciation. Sans surprise, cette analyse montre que des groupes de gènes impliqués dans la contraction musculaire, l'adhésion cellulaire, le métabolisme, le transport mitochondrial ou le cycle cellulaire sont modifiés au cours de la myogénèse. En revanche, cette analyse suggère l'implication de gènes jouant un rôle dans la fonction immunitaire. Ceci n'est pas trop étonnant dans la mesure où une régénération musculaire s'accompagne souvent d'un processus inflammatoire. Cette étude révèle également l'implication de gènes connus comme régulant la différenciation osseuse. Récemment dans une étude menée chez l'homme sur le vieillissement, WEELE et al., (2002) ont montré que seul 1/3 des gènes impliqués dans ce processus chez la souris (LEE et al., 1999) présentaient une modification de leur profils d'expression chez l'homme.

Au cours de certaines pathologies (diabète, dystrophie musculaire), on observe également des modifications importantes du transcriptome. Grâce à l'utilisation de puces Affymetrix, YECHOOR et al., (2002) ont mis en évidence chez la souris de nouveaux mécanismes de régulation par l'insuline pouvant même conduire à la découverte de nouvelles cibles thérapeutiques. La myodystrophie de Duchenne est une maladie monogénique identifiée il y a une dizaine d'années mais la fonction du gène responsable est encore méconnue ce qui ne permet pas, à l'heure actuelle, de proposer un traitement efficace. En utilisant des microréseaux d'ADN, TKATCHENKO et al., (2001) ont identifié 9 gènes dont l'expression est modifiée chez des patients souffrant de

cette maladie. La fonction de 3 de ces gènes est inconnue à l'heure actuelle.

3.1.2 Etude de la qualité de la viande

Pour les espèces animales d'élevage, les travaux utilisant les technologies de la génomique et appliqués à la physiologie musculaire sont peu nombreux. Il s'agit d'études préliminaires et de projets annoncés visant à améliorer la qualité de la viande. Dans le domaine de la production porcine, la maîtrise et l'amélioration de la qualité de la viande sont un enjeu majeur pour toute la filière. Cette qualité dépend de l'association entre diverses caractéristiques complexes (caractéristiques des fibres, teneur en lipides du muscle, nature du collagène, intensité de la protéolyse, etc...) qui peuvent être affectées par de nombreux facteurs génétiques, nutritionnels ou environnementaux. Les approches développées jusqu'à présent, comme l'analyse de l'expression et de la fonction d'un nombre restreint de gènes 'candidats' connus (facteurs de différenciation myogénique et adipocytaire, myosines, récepteurs hormonaux, activités enzymatiques, protéines de transport), ne permettent pas de résoudre cette complexité. De plus, si des gènes majeurs ont été identifiés pour 2 défauts de qualité de la viande (viandes PSE et acides ; OTSU et al., 1991 et MILAN et al., 2000), la plupart des problèmes rencontrés sont encore mal maîtrisés. Ainsi, le gène (MI) à effet majeur sur le taux de lipides intramusculaire (LIM) dont l'allèle récessif *imf* (pour "intramuscular fat") augmente fortement la teneur en LIM, n'est pas identifié (SANCHEZ et al., 2002). De plus, si de nombreux QTLs influençant la teneur en LIM ont été détectés (BIDANEL et al., 2000), les gènes contrôlant ce caractère et l'apparition de la déstructuration demeurent inconnus. Pourtant, le taux de LIM, caractéristique déterminante pour les qualités organoleptiques de la viande (goût, flaveur, jutosité), est en moyenne de 1.5 % dans le muscle *Longissimus Dorsi* (LD, noix de côtelette), alors qu'une étude d'un jury de consommateurs donne une qualité organoleptique optimale pour un taux de 2.5 à 3 % (FERNANDEZ et al., 1999). Il est donc aujourd'hui nécessaire de mieux maîtriser la teneur en LIM, et ceci indépendamment du dépôt de lipides externes. De même, le défaut de déstructuration touche environ 17% des jambons cuits et entraîne une perte économique importante pour les industriels. Ce phénomène n'est décelable qu'après le désossage du jambon, notamment au moment de la découpe. Les muscles touchés ont alors perdu leur aspect fibreux au profit d'une masse musculaire apparemment sans structure organisée et molle. Cette altération se rapprocherait de lésions de types PSE (MINVIELLE et al., 2001).

Pour étudier le déterminisme de la teneur en LIM chez le porc, plusieurs approches sont envisagées. En utilisant les premiers filtres haute densité porcins issus du programme AGENAE (HATEY et al., 2000) comportant un millier de séquences d'une banque multi-tissus normalisée d'ADNc de porc, nous avons étudié et comparé les transcriptomes de muscle de porcs F2 Duroc X LW à teneur haute (LIM-H, 3.4 %) ou basse (LIM-B, 1.9 %) en LIM (DAMON et al., 2002). Entre ces deux lots d'animaux, 8 gènes semblent être différentiellement exprimés. Six de ces 8 gènes codent pour des protéines connues et deux pour des protéines inconnues.

Actuellement, la relation entre la fonction de ces protéines et le déterminisme de la teneur en LIM n'est pas clairement établie. Tout un travail de recherches bibliographiques est actuellement nécessaire pour formuler des hypothèses concernant le rôle de ces gènes et leurs mécanismes d'action dans le contrôle de la teneur en LIM.

D'autre part, une meilleure maîtrise des processus de prolifération et de différenciation des précurseurs adipocytaires issus du muscle semble essentielle pour la teneur finale en LIM ; en effet, ces deux processus déterminent respectivement le nombre et la taille des adipocytes. Cependant si les mécanismes contrôlant l'hypertrophie des tissus adipeux sous-cutanés sont relativement bien connus, le métabolisme de l'adipocyte intramusculaire et sa régulation restent à caractériser chez le porc. Des travaux portant sur la comparaison d'adipocytes du tissu adipeux sous-cutané et du tissu adipeux interne chez le rat ou l'homme montrent que l'expression des gènes et leur régulation hormonale diffèrent suivant la localisation de ces tissus (WAJCHENBERG et al., 2000). Ainsi, deux types d'approches sont actuellement envisagées :

- l'identification d'un ensemble de gènes dont l'expression différerait fortement en fonction du type d'adipocyte (adipocyte mature isolé d'un tissu adipeux sous-cutané ou du tissu adipeux intramusculaire) permettrait de définir un profil d'expression spécifique des adipocytes intramusculaires. Les gènes ainsi identifiés pourraient ensuite être utilisés pour étudier spécifiquement le développement et le métabolisme du tissu adipeux intramusculaire.
- la recherche de marqueurs de la différenciation du tissu adipeux intramusculaire par comparaison des profils d'expression des transcrits exprimés au cours de la différenciation adipocytaire (préadipocytes lors de la mise en culture, préadipocytes lors de la confluence cellulaire, préadipocytes en différenciation terminale) dans le but d'obtenir des marqueurs précoces du développement des préadipocytes.

La recherche de prédictors précoces et de QTL associés aux défauts de qualité de la viande de porc à travers l'analyse du transcriptome est un programme en cours qui associe 4 laboratoires INRA, l'ENV de Lyon et FRANCE HYBRIDES (DAMON et al., 2002). Deux problèmes majeurs de qualité de viande seront analysés : la déstructuration du jambon cuit et l'hétérogénéité de la texture de la viande fraîche. Des facteurs tels que la sélection génétique, la nutrition, l'environnement en élevage et avant l'abattage vont moduler l'expression de certains gènes dont la présence est indispensable à l'obtention d'une viande avec des caractéristiques qualitatives optimales. L'objectif de ce programme est de mettre en relation le profil d'expression des gènes et protéines musculaires, les caractéristiques physico-chimiques du muscle et la qualité finale des viandes. Les niveaux d'expression seront appréciés à partir de l'intensité des signaux d'hybridation entre une collection d'ADNc d'une banque multi-tissus de porc (programme AGENAE) disposée sur des membranes et des cibles synthétisées à partir de muscles prélevés à l'abattage sur 400 animaux. L'analyse du protéome, techniquement plus délicate permettra de tenir compte des modifications post-traductionnelles et donc de toute la dynamique du

système. L'association originale des analyses génétique, transcriptomique et protéomique permettront, d'une part d'augmenter la puissance du dispositif de détection des QTL par une définition plus fine des caractères, d'autre part de faciliter l'identification des gènes et polymorphismes sous jacents aux QTL.

Un projet européen Quality Pork Genes (PLASTOW et al., 2000, www.qualityporkgenes.com) s'attache à identifier des gènes responsables de la variation de la qualité de la viande de porc en relation avec le stress et le bien-être animal. Il repose sur l'utilisation des technologies de la post-génomique en associant des analyses transcriptomique et protéomique. Les objectifs principaux sont la construction d'une base de données de mesures phénotypiques du muscle et de la viande, la production d'une banque d'ADNc musculaire et de puces à ADN et l'identification de marqueurs de la qualité de la viande.

Chez le bovin, une démarche de génomique fonctionnelle est également engagée pour mettre en évidence de nouvelles caractéristiques musculaires permettant de rendre compte de la variabilité de la qualité de la viande (SUDRE et al., 2002). La comparaison d'animaux aux caractéristiques divergentes en terme de type génétique (différence de potentiel de croissance) ou de conditions d'élevage (à base d'herbe ou d'ensilage de maïs) est actuellement en cours. Une analyse protéomique récente a permis de mettre en évidence de nouveaux marqueurs potentiels de l'hypertrophie musculaire en comparant des bovins culards (Blanc Bleu Belge) ou non (BOULEY et al., 2002). L'analyse protéomique de la protéolyse post-mortem permet de mettre en évidence la dégradation de protéines, telles que l'actine et certaines enzymes du métabolisme comme la créatine kinase et la glycogène phosphorylase (LAMETSCH et al., 2002). Leur protéolyse ne peut être mise en relation directe avec la tendreté de la viande mais les peptides issus de cette dégradation pourraient être de bons marqueurs de la tendreté de la viande. En effet, par des analyses de double hybride, PURINTRAPIBAN et al., (2001) ont montré que ces enzymes étaient des substrats de la calpaine dont l'activité est fortement corrélée à la tendreté de la viande (UYTTERHAEGEN et al., 1994).

A terme, toutes ces études devraient conduire à l'identification de nouveaux gènes-clés impliqués dans le déterminisme des caractéristiques d'intérêt (défauts de qualité de viande et mise en place du tissu adipeux intramusculaire) et ainsi fournir de nouvelles cibles d'action. Ces études sont d'autant plus nécessaires que la qualité de viande est une thématique spécifique aux animaux d'élevage qui ne sera donc pas prise en compte dans des études réalisées sur des espèces modèles comme la souris.

3.2. La nutrition

Le champ de la nutrition est particulièrement vaste, depuis le métabolisme jusqu'à la santé des populations, et nécessite des approches très diverses, depuis la biochimie et la biologie moléculaire jusqu'à l'épidémiologie. L'utilisation des outils de la biologie moléculaire classique a montré que de nombreux gènes, enzymes et protéines de structure sont

régulés au niveau transcriptionnel par les aliments. L'utilisation de la transcriptomique et de la protéomique en nutrition nous donne la possibilité d'élargir notre connaissance sur la façon dont les aliments complexes affectent un tissu ou un compartiment cellulaire. Ce type de démarche pourra permettre de révéler des gènes dont l'importance fonctionnelle est totalement inconnue à l'heure actuelle. Elle va pouvoir également montrer comment des variations génétiques individuelles peuvent affecter l'expression phénotypique en fonction du régime. Ces aspects changeront notre compréhension des phénomènes de santé chez les animaux et l'homme à l'échelle de l'individu ou d'une population. En conséquence, des changements à venir sont à prévoir en agriculture (production des matières premières), dans l'industrie agro-alimentaire (formulation d'aliments et traçabilité) et des critères nouveaux d'évaluation des matières premières et des aliments devraient apparaître.

Les premières applications en nutrition disponibles à l'heure actuelle concernent à notre connaissance, l'homme et des modèles rongeurs. Cependant, outre l'intérêt de ces nouvelles approches pour les productions animales, les animaux d'élevage et, en particulier le porc, pourraient constituer des modèles pour la nutrition humaine, de part notamment les possibilités expérimentales plus étendues.

La nutrition s'intéresse aux interactions moléculaires entre bactéries et cellules de l'hôte, aux mécanismes de digestion-absorption, à la régulation du métabolisme énergétique et à diverses situations physiopathologiques (vieillesse, diabète, cancers, ...). Quelques exemples d'application de la post-génomique en nutrition sont rapportés.

Le tractus gastro-intestinal est le lieu d'interactions entre les muqueuses, les nutriments, divers composés exogènes et la flore microbienne. Les interactions flore bactérienne-hôte ont récemment été abordées par HOOPER et al., (2001) qui ont examiné l'effet de la colonisation de l'intestin de souris par une souche bactérienne particulière (*Bacteroides thetaiotaomicron*). Ainsi, la prolifération de cette souche bactérienne modifie l'expression de nombreux gènes impliqués dans diverses fonctions telles que l'absorption des nutriments, la fonction barrière, la maturation postnatale de l'intestin. Par ailleurs, le type et la densité des bactéries colonisant l'intestin affecteraient diversement les fonctions intestinales, renforçant la nécessité de développer des recherches sur les interactions entre l'hôte et la flore commensale ou introduite par l'alimentation. La complexité de ces interactions, illustrée par la modification de l'expression de nombreux gènes, nécessite le recours à des techniques d'analyses systématiques comme celles que propose la post-génomique. L'analyse simultanée du transcriptome bactérien et celui du compartiment cellulaire-cible de l'hôte permettra de plus d'explorer les interactions entre ces organismes et apportera un éclairage nouveau pour le développement notamment de substituts alternatifs aux antibiotiques.

Dans un but à la fois pronostic et thérapeutique, l'analyse du transcriptome réalisée chez des sujets sains et des patients souffrant de différentes formes de maladies inflammatoires de l'intestin (la colite ulcéreuse et la maladie de Crohn ;

DIECKGRAEFFE et al., 2000) révèle des gènes communs et des gènes spécifiques à ces deux maladies. Sur environ 6500 gènes représentés sur des puces à ADN Affymetrix, 70 gènes transcrits dans la muqueuse intestinale sont surexprimés dans la colite ulcéreuse. Ces gènes comprennent les gènes de la NO synthase inductible, des constituants de la matrice extracellulaire et du système immunitaire (IL-1, IL-8, chemokines,...). Par ailleurs, 60 gènes exprimés uniquement chez les patients atteints de la colite ulcéreuse ou de la maladie de Crohn sont révélés par clustering, ouvrant de nouvelles perspectives thérapeutiques.

Les insuffisances nutritionnelles affectent de nombreux organes et tissus. Cependant, seul un nombre limité de gènes affectés par le statut nutritionnel est caractérisé. La technique du Differential Display, qui permet d'identifier les gènes différemment exprimés, a été appliquée à la muqueuse intestinale de rats recevant un régime déficient ou supplémenté en zinc. Elle a permis d'identifier 9 gènes dont l'expression était modulée par la teneur en zinc du régime alimentaire. Parmi ceux-ci, les gènes de la phosphatase alcaline II, de la FABP-I, de deux hormones (la cholécystokinine et l'uroguanyline), de quatre gènes de ménage et de la chaîne J des immunoglobulines. Cette dernière est nécessaire à la polymérisation et la sécrétion des IgM et des IgA, indiquant une modification du statut immunitaire lors d'une déficience en zinc (BLANCHARD et COUSINS 2000 ; BLANCHARD et al., 2001). La surexpression de l'uroguanylin impliquée dans la régulation des échanges d'eau et d'électrolytes au niveau de la muqueuse intestinale, favoriserait la sécrétion d'eau et de chlore dans l'intestin et ainsi l'apparition de diarrhées lors d'une déficience en zinc.

L'effet bénéfique d'une restriction énergétique sur le vieillissement était démontré, mais les mécanismes par lesquels la restriction énergétique contribue à accroître la longévité restaient inconnus. L'analyse de l'expression de 6347 gènes chez des souris jeunes et âgées, soumises ou non à une restriction énergétique, a permis de montrer que l'expression d'environ 1 % des gènes était modifiée par le vieillissement (LEE et al., 1999). Ceci était assez surprenant car chez la levure environ 30 % des gènes modifient leur expression quand l'apport de glucose est diminué (DE RISI et al., 1997). La plupart de ces modifications sont généralement abolies par la restriction calorique. Il semblerait que la restriction calorique retarde le vieillissement en augmentant le turn-over protéique, la gluconéogenèse et le cycle des pentoses phosphates tandis qu'elle diminue les gènes liés au stress oxydatif (protéines de choc thermique, de réparation de l'ADN et du système de détoxification).

Dans le domaine de l'obésité, l'utilisation de puces à ADN appliquées au tissu adipeux blanc de souris (SOUKAS et al., 2000) a permis d'identifier par clustering les gènes dont l'expression était modulée par la leptine. Ces analyses mettent en avant les relations entre la leptine, le métabolisme et le contrôle expressionnel de gènes impliqués dans les voies de synthèse du cholestérol et des acides gras.

Ainsi, les premiers résultats concernent essentiellement les recherches en nutrition humaine. De la même manière, un

grand nombre de questions posées en production animale pourront être explorées en utilisant ces outils de la post-génomique.

3.3. La fonction de reproduction

La fonction de reproduction illustre bien la complexité du vivant puisqu'elle implique plusieurs organes, comme l'hypothalamus, l'hypophyse, l'ovaire et l'utérus chez la femelle. Leur activité coordonnée résulte d'une régulation hormonale complexe, associant des mécanismes endocrines et auto ou paracrines. Cette caractéristique pose le problème du choix des organes ou tissus qui feront l'objet de l'étude, même si l'ovaire et le testicule constituent des cibles expérimentales privilégiées.

Les approches moléculaires « réductionnistes », seules disponibles jusqu'ici, sont insuffisantes pour permettre le développement de nouvelles stratégies en reproduction. Les progrès spectaculaires réalisés dans ce domaine l'ont été essentiellement sur la base des données physiologiques et sont illustrés dans les deux autres exposés de cette session. Le développement des approches de génomique, de caractère global et systématique, permettent d'entrevoir de nouvelles possibilités à la fois dans la compréhension fine des mécanismes de régulation et dans l'intégration de ces données : par exemple l'identification d'un ensemble de gènes présentant un même profil de régulation, signature possible d'une éventuelle cascade de signalisation. Nous n'en sommes toutefois qu'aux prémices mais, dans ce domaine, grâce aux possibilités d'expérimentation, les recherches sur l'animal pourraient progresser plus vite que celles conduites dans l'espèce humaine.

Une première série d'approches ciblées permet d'accéder directement à des gènes a priori importants dans la fonction. Le tri différentiel dans laquelle deux répliques d'une banque d'ADNc ovariens sont hybridées avec des ADNc représentatifs des transcrits de deux populations de cellules de la granulosa de truie traitées ou non par la FSH a été mis en œuvre par TOSSER et al., (1997). Parmi les 238 clones isolés, les auteurs ont identifié 136 clones uniques dont 60 % ne présentent aucune similarité avec les séquences de la base de données GenBank. Pour les clones identifiés, 4 correspondent à des séquences porcines (ferritine, inhibiteur du complément, protéine de liaison à l'héparine, polyubiquitine), 35 correspondent à des gènes humains et 15 à des gènes d'autres mammifères. Une autre stratégie permettant d'identifier les gènes différemment exprimés, le Differential Display, a été appliquée aux cellules de la granulosa de truie par CLOUSCARD-MARTINATO et al., (1998), aux cellules de Sertoli murines par HELLANI et al., (2000) et aux oocytes bovins par ROBERT et al., (2001). Des gènes intéressants ont ainsi été identifiés, par exemple l'OSP dans les cellules de Sertoli murines. Cet « Oligodendrocyte Specific Protein » ou Claudin 11 joue un rôle dans la formation de la barrière hémato-testiculaire. Elle est également exprimée de manière spécifique au cours du développement fœtal, après le pic d'expression du facteur SRY et avant celui de l'hormone anti-müllérienne. Cependant, le très grand nombre de faux positifs obtenus avec l'une et l'autre métho-

de en limite l'application. L'hybridation soustractive suppressive (HSS) est une approche différentielle plus performante. Elle a été utilisée par le groupe d'Adashi chez la souris pour identifier les gènes exprimés spécifiquement ou préférentiellement dans l'ovaire, avec l'idée que les gènes dont l'expression est spécifique d'un tissu jouent certainement un rôle important dans ce tissu. Ils ont ainsi identifié 21 et 15 gènes exprimés spécifiquement ou préférentiellement dans l'ovaire, dont certains encore inconnus (HENNEBOLD et al., 2000). En utilisant la même méthode, ROBERT et al., (2001) ont isolés plus de 50 transcrits exprimés différemment entre des oocytes bovins compétents ou non. Chez la truie enfin, des travaux en cours ont permis d'identifier quelques dizaines de transcrits différemment exprimés dans les cellules de la granulosa (A. BONNET, communication personnelle). En particulier, des gènes connus pour être impliqués dans le développement folliculaire tels que le cytochrome P450_{sc}, la 3- β -hydroxystéroïde déshydrogénase, l'actine ou la GST ont été retrouvés, validant ainsi la pertinence de l'approche HSS.

Au delà de ces approches ciblées, des approches à plus long terme, systématiques, se développent, soit à travers des projets généraux comme le projet AGENAE initié par l'INRA, soit à travers des projets ciblés reproduction comme celui développé par un consortium américain (TUGGLE et al., 2002). Ce projet regroupe 5 laboratoires différents, centrés sur les universités de l'Iowa, et a construit 18 banques d'ADNc normalisées ou non, à partir d'embryons, hypophysés, hypothalamus, ovaires, utérus et placenta à différents stades du cycle ou de la gestation. Cette diversité de tissus illustre bien la complexité de la fonction de reproduction que nous évoquions au début de cet exposé. Près de 15 000 séquences ont été obtenues, correspondant à 9 000 gènes différents. Le travail se poursuit actuellement par la localisation de ces gènes, soit directement en utilisant les panels d'hybrides somatiques classique ou irradié, soit en utilisant les ressources de la carte comparée avec l'homme.

Un deuxième projet de génomique fonctionnelle, développé par l'université de Nebraska-Lincoln (CAETANO et al., 2002), utilise des lignées de porc divergentes sur la base d'un index de sélection des composants de la taille de portée (taux d'ovulation, mortalité embryonnaire, etc...). La lignée indexée présente un taux d'ovulation supérieur de 7,4 par rapport à la lignée témoin et un QTL significatif a été identifié dans une population F2 (CASSADY et al., 2001) ; cet avantage résulte d'une modification de la dynamique de maturation folliculaire pendant la phase folliculaire du cycle. Une banque normalisée d'ADNc a été construite à partir de follicules de différentes tailles isolés à partir d'ovaires de truies des 2 lignées et séquencée partiellement (POMP et al., 2001). Les produits d'amplification de 3636 clones uniques représentatifs des clusters de séquences, dont certains ne correspondent à aucun gène identifié jusqu'ici, ont été déposés sur lame de verre. L'hybridation de ces microréseaux avec les transcrits ovariens de chacune des 2 lignées a permis d'identifier 92 gènes présentant une différence d'expression significative entre les 2 lignées, en particulier le TNF- α , la 3- β -hydroxystéroïde déshydrogénase ou le P450 17 α -hydroxylase. Une approche de clusterisation de ces

données met en évidence différents groupes de gènes dont les profils d'expression sont comparables. Ce travail se poursuit par la recherche de la fonction de ces gènes sur la base d'homologies fonctionnelles avec l'homme ou la souris. Sans perdre de vue que les données du transcriptome doivent être complétées par celle du protéome et des approches physiopathologiques, ces données de génomique sont susceptibles de servir de base à des hypothèses concernant la régulation du taux d'ovulation chez le porc. Cette connaissance peut conduire au développement de stratégies innovantes permettant d'améliorer les performances de reproduction chez le porc.

3.4. La réponse immunitaire

Chez les animaux comme chez l'homme, la réponse immunitaire est un mécanisme de défense essentiel contre les agressions par des agents pathogènes ou contre toute atteinte à l'intégrité de l'organisme (brûlure, coupure...). Deux mécanismes différents sont impliqués dans la réponse immunitaire : la réponse inflammatoire et la réponse immunitaire associée à la mise en place d'une mémoire (aussi appelée immunité acquise). L'inflammation est une réponse non spécifique qui se met en place très rapidement et conduit à une activation des cellules phagocytaires (macrophages et neutrophiles). Les cellules activées sécrètent de nombreuses molécules telles que des cytokines inflammatoires qui permettent le recrutement et l'activation de nouvelles cellules afin de se débarrasser de l'agent agresseur, des métabolites de l'acide arachidonique (prostaglandines et leucotriènes) et des composés réactifs de l'azote ou de l'oxygène. La réponse immunitaire acquise résulte d'une sensibilisation contre l'agent étranger par stimulation des lymphocytes. Lors d'une stimulation ultérieure, l'organisme qui a gardé en mémoire la stimulation précédente, élabore une réponse rapide et spécifique. Deux types de cellules participent à cette réponse : (i) les lymphocytes B qui sécrètent les anticorps et établissent la réponse à médiation humorale, (ii) les lymphocytes T qui participent à la réponse à médiation cellulaire en développant une activité cytotoxique et en sécrétant des cytokines. On peut distinguer les cytokines de type Th1 et Th2 qui orientent l'immunité vers une réponse à médiation cellulaire ou humorale respectivement. La réponse immunitaire est donc un phénomène complexe qui mobilise de nombreux gènes. Les puces à ADN fournissent un outil très puissant d'analyse de cette réponse car elles permettent d'obtenir une vue globale de l'expression des ARNm des gènes d'intérêt (réponse inflammatoire, réponse immune...) avec un gain de temps et d'échelle par rapport aux techniques classiques.

Dans le cadre de l'étude des réponses immunitaires, des travaux chez l'homme et dans les modèles rongeurs ont utilisé soit des puces « généralistes » (c'est-à-dire des puces comportant un très grand nombre de gènes) soit des puces « dédiées » qui regroupent un nombre plus restreint de gènes mais tous impliqués dans la réponse immunitaire. L'utilisation de ces puces en immunologie a permis pour la première fois d'analyser simultanément l'expression de centaines voire de milliers de gènes. Ceci a conduit à définir des « signatures » caractéristiques d'un type de cellule immuni-

taire, d'un état de différenciation ou d'activation, d'une voie de signalisation... (SHAFFER et al., 2001). Dans le cadre de la réponse immunitaire élaborée contre un agent pathogène, les puces ont permis d'analyser la chorégraphie moléculaire mise en œuvre lors de la réponse, certains gènes étant exprimés de façon systématique quel que soit l'agent pathogène alors que d'autres gènes sont au contraire très spécifiques de l'agent infectieux (BOLDRICK et al., 2002). Ces études servent à mieux comprendre le fonctionnement des cellules immunitaires et le dialogue moléculaire qui s'instaure au cours d'une infection. Dans l'avenir elles devraient permettre de mieux définir les cibles cellulaires à atteindre dans un cadre thérapeutique de mise au point de nouvelles molécules. Elles aideront également à identifier les voies cellulaires qu'il est nécessaire de stimuler au cours de la vaccination pour orienter de manière optimale la réponse protectrice.

Les puces à ADN ont permis de mettre en évidence de nouvelles régulations au cours d'infections par des agents pathogènes. Cette technologie a, par exemple, montré le recrutement de neutrophiles dans certaines lignées de souris lors d'une infection par *Schistosoma mansoni* (HOFFMANN et al., 2001). Elle a aussi révélé une régulation du métabolisme du glucose et du mevalonate lors de l'infection des fibroblastes humains par des toxoplasmes (BLADER et al., 2001). L'utilisation croissante de cette technologie devrait également conduire à identifier les fonctions de gènes orphelins et peut-être à assigner de nouvelles fonctions à des gènes déjà connus. Comme déjà cité ci-dessus, cette technologie a montré l'implication de gènes du système immunitaire dans des fonctions physiologiques « non immunitaires » telle que la physiologie musculaire.

Si des puces ADN sont déjà commercialisées pour analyser la réponse immunitaire de l'homme et des rongeurs de laboratoire (Atlas System de Clontech, GEM Microarray de Genome Systems...) il n'en est pas de même chez les animaux domestiques. Chez les bovins, le groupe de P. COUSSENS a développé plusieurs outils post-génomique pour analyser la réponse immunitaire (YAO et al., 2001 ; COUSSENS et NOBIS 2002). En effet, cette équipe a fabriqué une banque normalisée de cDNA de leucocytes bovins, de laquelle 721 EST spécifiques ont été extraits afin de confectionner une puce à ADN (YAO et al., 2001). Cette puce a été utilisée pour comparer les profils d'expression des cellules sanguines mononuclées de bovins contrôles et d'animaux infectés de façon clinique et sub-clinique par *Mycobacterium paratuberculosis* (COUSSENS et al., 2002). Les résultats montrent une répression génique importante (16 gènes) chez les animaux infectés cliniquement tandis que les bovins ne présentant pas de signe clinique montrent des régulations à la fois positives (11 gènes) et négatives (16 gènes) de l'expression des EST cibles. Dans le domaine aviaire, une puce contenant plus d'un millier de cDNA a été construite à partir d'une banque d'EST provenant de lymphocytes T de poulets (TIRUNAGARU et al., 2000). Utilisant cet outil MORGAN et coll (2001) ont montré que l'infection de fibroblastes embryonnaires de poulet par le virus de la maladie de Marek induit l'expression d'une dizaine de gènes spécifiques.

Chez le porc, il n'y a pas de travaux publiés utilisant des approches génomiques dans le cadre de la réponse immunitaire. Deux approches complémentaires sont actuellement développées à l'INRA dans le cadre du programme AGENAE pour créer des puces ADN utilisables pour étudier les réponses immunitaires. D'une part, une banque multi-tissus normalisée d'ADNc de porc a été construite pour réaliser des filtres haute densité représentant une large part du génome porcin exprimé (HATEY et al., 2000). Ces filtres « généralistes » déjà utilisés dans le cadre de l'étude de la physiologie musculaire (cf. 3.1.) sont également utilisables pour étudier les changements d'expression de gènes dans le cadre de la réponse contre un agent infectieux. Des filtres moyenne densité, spécifiquement dédiés à l'étude de la réponse immunitaire du porc sont aussi réalisés dans le cadre du programme AGENAE. La fabrication de ces filtres a été facilitée par le fait que de nombreuses cytokines porcines sont aujourd'hui clonées et séquencées mais également par la forte homologie entre les protéines humaines et porcines. Cette collection comprend à ce jour 70 gènes :

- 30 gènes codant pour des cytokines et des chimiokines
- 5 gènes codant pour des récepteurs de cytokines
- 24 gènes codant pour des marqueurs d'activation et autres gènes impliqués dans la réponse immunitaire
- 10 gènes de ménage et contrôles négatifs.

Dans les prochains mois, cette collection génique sera complétée afin d'atteindre 150 à 200 gènes spécifiques de la réponse immunitaire du porc. Cette puce est utilisée pour analyser la réponse immunitaire des porcelets lors d'infections colibacillaires dans le cadre du projet européen AEECInfection (<http://www.inra.fr/Internet/Projets/aeec/>). Elle permet également d'étudier les effets délétères de l'ingestion de mycotoxines présentes dans les aliments sur les réponses immunitaires des porcs (transversalité INRA-mycotoxines <http://www.inra.fr/Internet/Projets/mycotoxines/>).

Enfin, il est important de souligner que la réponse immunitaire contre un agent infectieux met face à face deux génomes : celui de l'hôte mais aussi celui de l'agent infectieux. L'existence d'un nombre de plus en plus important de génomes complètement séquencés chez des bactéries d'intérêt médical autorise la mise au point de puce « génome total » (notamment si elles sont associées à des systèmes de quantification du signal), domaine qui est particulièrement prometteur pour l'analyse des interactions entre bio-agresseur microbien et cellules ou entre microorganisme et organisme supérieur. L'utilisation de puces permettra d'analyser les messagers microbiens et cellulaires et de discerner très finement les composantes du dialogue moléculaire à différentes étapes de l'interaction. La compréhension de ce dialogue est porteuse d'applications novatrices, par exemple en termes de nouveaux agents thérapeutiques alternatifs aux antibiotiques.

En conclusion les outils de la post-génomique ont ouvert un champ nouveau dans l'étude des réponses immunitaires. D'un point de vue fondamental, les puces à ADN permettront de décrypter le dialogue moléculaire qui s'instaure lors d'une infection. Cette meilleure compréhension de la physiologie cellulaire pourrait conduire à la découverte de nou-

velles cibles thérapeutiques (pharmaco-génomique) mais aussi vers le développement de vaccins plus ciblés et plus efficaces (vaccino-génomique). Les puces à ADN permettront également de prendre en compte la variabilité individuelle et d'adapter, comme c'est déjà le cas en cancérologie, les « traitements » thérapeutiques ou vaccinaux de façon plus ciblée en fonction des populations animales (POLLACK et al., 2002). Enfin ces outils, en participant à l'identification des gènes de résistance, pourront conduire au développement de lignées génétiquement résistantes aux infections.

CONCLUSION

Déterminer la fonction des gènes d'un organisme consiste à identifier l'ensemble des gènes portés par son génome (programmes de séquençage et d'annotations), établir les conditions qui déterminent leur expression (étude de la régulation des gènes et des mécanismes d'action), définir la fonction de chacune des protéines, étudier leurs interactions en tenant compte de toute la dynamique du système (temps, espace, conditions physiopathologiques...). Les deux dernières étapes sont la base même des démarches de la post-génomique ou génomique fonctionnelle. Depuis une dizaine d'an-

nées, les outils de biologie moléculaire à haut débit se sont considérablement développés et associés à la bio-informatique, ils ont dynamisé les recherches qui s'effectuent maintenant de façon globale et systématique. Cependant, l'approche post-génomique n'a de sens et sera constructive que si elle est replacée dans le contexte et s'appuie sur la physiologie. L'ensemble des données générées constitueront à terme les bases de ce que l'on pourra appeler une physiologie moléculaire intégrée. Ces nouvelles approches devraient permettre de mieux comprendre le fonctionnement des organismes et donner lieu à des développements dans de nombreux domaines et notamment la médecine, la pharmacologie et l'agronomie. Dans le domaine des productions animales et notamment pour la filière porcine, ces nouveaux outils sont encore en construction et s'il est possible de décrire toute la potentialité des études de post-génomique en physiologie, il est en revanche difficile d'en prédire toutes les applications. Cependant, cette démarche devrait permettre à terme la mise à disposition de la filière, de nouveaux outils de sélection, des marqueurs ou prédicteurs de la qualité de la viande, de nouvelles stratégies alimentaires ou prophylactiques (agents alternatifs aux antibiotiques), des méthodes de traçabilité...

REFERENCES BIBLIOGRAPHIQUES

- BERTUCCI F., LORIOD B., TAGETT R., GRANJEAUD S., BIRNBAUM D., NGUYEN C., HOULGATTE R. 2001. *Bull Cancer*, 88, 243-252.
- BIDANEL J.P., CHARDON P., HATEY F., RIQUET J., LE ROY P., MILAN D. 2003. *Journées Rech. Porcine*, 35, (sous presse).
- BLADER I.J., MANGER I.D., BOOTHROYD J.C., 2001. *J. Biol. Chem.* 276, 24223-31.
- BLANCHARD R.K., COUSINS R.J., 2000. *J. Nutr.*, 130, 1393S-1398S.
- BLANCHARD R.K., MOORE J.B., GREEN C.L., COUSINS R.J., 2001. *Proc. Natl. Acad. Sci. U.S.A.*, 98, 13507-13513.
- BOLDRICK J.C., ALIZADEH A.A., DIEHN M., DUDOIT S., LIU C.L., BELCHER C.E., BOTSTEIN D., STAUDT L.M., BROWN P.O., RELMAN D.A. 2002. *Proc Natl Acad Sci U S A.* 99, 972-7.
- BOULEY J., CHAMBON C. ET PICARD B., 2002. *Viandes Prod. Carnés, Hors Série*, 131-132.
- BUSTIN S.A., DORUDI S. 2002. *Trends Mol. Med.*, 8, 269-272.
- CAETANO A.R., JOHNSON R.K., POMP D. 2002. 7th World Congress on Genetic Applied to Livestock Production, Montpellier, France.
- CAMPBELL W.G., GORDON S.E., CARLSON C.J., PATTISON J.S., HAMILTON M.T., BOOTH F.W. 2001. *Am J Physiol Cell Physiol.*, 280, C763-C768.
- CANTOR C.R. 2000. *Biofutur*, 206, 58-61.
- CASSADY J.P., JOHNSON R.K., POMP D., ROHRER G.A., VAN VLECK L.D., SPIEGEL E.K., GILSON K.M. 2001. *J. Anim. Sci.*, 79, 623-633.
- CLAVERIE J.M. 1999. In « Développement et applications de la génomique - L'après-génome ». 136-143. Tec & Doc Lavoisier éd., Paris, 231p.
- CLOUSCARD-MARTINATO C., MULSANT P., ROBIC A., BONNET A., GASSER F., HATEY F. 1998. *Anim. Genet.*, 29, 98-106.
- COUSSENS P.M., COLVIN C.J., WIERSMA K., ABOUZIED A., SIPKOVSKY S. 2002. *Infect Immun.* 70, 5494-502.
- COUSSENS P.M., NOBIS W. 2002. *Vet Immunol Immunopathol.* 86, 229-44.
- DAMON M., LIAUBET L., VINCENT A., HERPIN P. ET HATEY F, 2002. *Viandes Prod. Carnés, Hors Série*, 139-140.
- DERISI J.L., IYER V.R., BROWN P.O. 1997. *Science*, 278, 680-686
- DIECKGRAEFE B.K., STENSON W.F., KORZENIK J.R., SWANSON P.E., HARRINGTON C.A., 2000. *Physiol. Genomics*, 4, 1-11.
- EISEN M.B., SPELLMAN P.T., BROWN P.O., BOTSTEIN D. 1998. *Proc Natl Acad Sci U S A.*, 95, 14863-148688
- FERNANDEZ X., MONIN G., TALMANT A., MOUROT J., LEBRET B., 1999. *Meat Sci.*, 53, 59-65.
- FREDERICKSON R.M. 1998. *Curr. Opin. Biotechnol.*, 9, 90-96.
- FREEMAN W.M., ROBERTSON D.J., VRANA K.E. 2000. *Biotechniques*, 29, 1042-1055.
- GELI F. 2000. *Biofutur, Technoscope n°128*, 3-14.
- GRANJEAUD S., BERTUCCI F., JORDAN B.R. 1999. *Bioessays*, 21, 781-790.
- HATEY F., MARTIN P., DOUAIRE M., LE GAC F., DAMBRINE G., HERPIN P., MONGET P., 2000. *INRA Prod. Anim.*, 175-180
- HELLANI A., JI J., MAUDUIT C., DESCHILDRE C., TABONE E., BENAHMED M. 2000. *Endocrinology*, 141, 3012-9.
- HENNEBOLD J.D., TANAKA M., SAITO J., HANSON B.R. ADASHI E.Y. 2000. *Endocrinology*, 141, 2725-2734.
- HOFFMANN K.F., McCARTY T.C., SEGAL D.H., CHIRAMONTE M., HESSE M., DAVIS E.M., CHEEVER A.W., MELTZER P.S., MORSE H.C., WYNN T.A., 2001. *FASEB J.* 15, 2545-7
- HOOPER L.V., WONG M.H., THELIN A., HANSSON L., FALK P.G., GORDON J.I., 2001. *Science*, 291, 881-884.
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM, 2001. *Nature*. 409, 860-921.
- JACQ B., THIEFFRY D. 2000. *Biofutur*, 206, 72-75.
- JORDAN B. 2002. *Médecine/Sciences*, 18, 297-301.

- JORDAN B., 2000. *Biofutur*, 206, 38-43.
- KERR M.K., MARTIN M., CHURCHILL G.A. 2000. *J Comput Biol.*, 7, 819-837.
- LAMETSCH R., ROEPSTORFF P., BENDIXEN E., 2002. *J Agric Food Chem.*, 50, 5508-5512.
- LEE C.K., KLOPP R.G., WEINDRUCH R., PROLLA T.A. 1999. *Science*, 285, 1390-1393.
- LEE C.K., KLOPP R.G., WEINDRUCH R., PROLLA T.A., 1999. *Science*, 285, 1390-1393.
- LIANG P., PARDEE A.B. 1992. *Science*. 257, 967-971.
- LOCKHART D.J., DONG H., BYRNE M.C., FOLLETTIE M.T., GALLO M.V., CHEE M.S., MITTMANN M., WANG C., KOBAYASHI M., HORTON H., BROWN E.L. 1996. *Nat Biotechnol.*, 14, 1675-1680.
- LOCKHART D.J., WINZELER E.A. 2000. *Nature*, 405, 827-836.
- MAHADEVAPPA M., WARRINGTON J.A. 1999. *Nat Biotechnol.*, 17,1134-1136.
- MILAN D., JEON J.T., LOOFT C., et al., 2000. *Science*, 288, 1248-1251.
- MINVIELLE B., LE STRAT P., LEBRET B., HOUIX Y., BOULARD J., CLOCHEFERT N., 2001. *Journées Rech. Porcine*, 33, 95-101.
- MORAN J.L., LI Y., HILL A.A., MOUNTS W.M., MILLER C.P. 2002. *Physiol Genomics*, 10, 103-111.
- MORANGE M., 2000. *Biofutur*, 206, 24-25.
- MORGAN R.W., SOFER L., ANDERSON A.S., BERNBERG EL, CUI J., BURNSIDE J. 2001. *J Virol*. 75:533-9.
- NGUYEN C., ROCHA D., GRANJEAUD S., BALDIT M., BERNARD K., NAQUET P., JORDAN B.R. 1995. *Genomics*, 29, 207-216.
- O'FARRELL P.H. 1975. *J. Biol. Chem.*, 250, 4007-4021.
- OTSU K., KHANNA V.K., ARCHIBALD A.L. et MACLENNAN D.H., 1991. *Genomics*, 11, 744-750.
- PANDEY A. et MANN M. 2000. *Nature*, 405, 837-846.
- PIETU G., ALIBERT O., GUICHARD V., LAMY B., BOIS F., et al., 1996. *Genome Res.*, 6, 492-503
- PIETU G., EVENO E., SOURY-SEGURENS B., FAYEIN N.A., MARIAGE-SAMSON R., et al., 1999. *Genome Res.*, 9, 1313-1320.
- PLASTOW G., SONICKI A., MALTIN C., et al., 2002. In *Proceedings of 48th ICOMST Italy*, 2, 638.
- POLLACK J.R., SORLIE T., PEROU C.M., REES C.A., JEFFREY S.S., LONNING P.E., TIBSHIRANI R., BOTSTEIN D., BORRESEN-DALE A.L., BROWN P.O. 2002. *Proc Natl Acad Sci U S A*. 99, 12963-8.
- POMP D., CAETANO A.R., BERTANI G.R., GLADNEY C.D., JOHNSON R.K. 2001 *Reprod Suppl.*, 58, 277-92.
- PURINTRAPIBAN, J.; WANG, M.; FORSBERG, N. E. 2001. *Int. J. Biochem. Cell Biol.*, 33, 531-540.
- ROBERT C., GAGNE D. BOUSQUET D., BARNES F.L., SIRARD M.A. 2001. *Biol Reprod* 64, 1812-1820.
- ROTH S.M., FERRELL R.E., PETERS D.G., METTER E.J., HURLEY B.F., ROGERS M.A. 2002. *Physiol Genomics*, 10, 181-190.
- SANCHEZ M.P., LE ROY P., GRIFFON H., CARITEZ J.C., FERNANDEZ X., LEGAULT C., GANDEMER G., 2002. *Journées Rech. Porcine*, 34, 39-43.
- SCHENA M., SHALON D., DAVIS R.W., BROWN P.O. 1995. *Science*, 270, 467-470.
- SHAFFER A.L., ROSENWALD A., HURT E.M., GILTANANE J.M., LAM L.T., PICKERAL O.K., STAUDT L.M. 2001. *Immunity*. 15, 375-85.
- SOUKAS A., COHEN P., SOCCI N.D., FRIEDMAN J.M., 2000. *Genes Dev*. 14, 963-980.
- SOUTHERN EM. 1975. *J Mol Biol.*, 98, 503-517.
- SUDRE K., CASSAR-MALEK I., LEROUX C., LISTRAT A., JURIE C., RENAND G., MARTIN P. et HOCQUETTE J.F., 2002. *Viandes Prod. Carnés, Hors Série*, 111-112.
- THEBAULT S., MACHOUR N., PERROT F., JOUENNE T., LANGE C., HUBERT M., FONTAINE M., TRON F. et CHARLIONET R. 2001. *Médecine/Sciences*, 17, 609-618.
- TIRUNAGARU V.G., SOFER L., CUI J., BURNSIDE J. 2000. *Genomics*. 66, 144-51.
- TKATCHENKO A.V., PIETU G., CROS N., GANNOUN-ZAKI L., AUFRAY C., LEGER J.J., DECHESSNE C.A. 2001. *Neuromuscul Disord.*, 11, 269-277.
- TOSSER-KLOPP G., BENNE F., BONNET A., MULSANT P., GASSER F., HATEY F. 1997. *Mamm Genome*, 8, 250-254.
- TUGGLE C.K., GREEN J.A., FITZSIMMONS C., WOODS R., PRATHER R.S., et al., 2002. *Plant, Animal & Microbes Genomes X*, San Diego.
- UYTTERHAEGEN L.; CLAEYS E.; DEMEYER, D. 1994. *J. Anim. Sci.*, 72, 1209-1223.
- VELCULESCU V.E., ZHANG L., VOGELSTEIN B., and KINZLER K.W. 1995. *Science* 270, 484-487
- WAJCHENBERG B.L., 2000. *Endocr. Rev.*, 21, 697-738.
- WANG D.G., FAN J.B., SIAO C.J., BERNO A., YOUNG P., et al., 1998. *Science*, 280, 1077-1082.
- WASINGER V.C., CORDWELL S.J., CERPA-POLJAK A., YAN J.X., GOOLEY A.A., et al., 1995. *Electrophoresis*, 16, 1090-1094.
- WATSON J. and CRICK F. 1953. *Nature*, 171, 737-738.
- WATSON S.J., MENG.F., THOMPSON R.C., AKIL H. 2000. *Biol Psychiatry*, 48, 1147-1156.
- WELLE S, BROOKS A, THORNTON CA. 2001. *Physiol Genomics*, 5, 67-73.
- YAO J., BURTON J.L., SAAMA P., SIPKOVSKY S., COUSSENS P.M. 2001. *Acta Vet Scand*. 42, 391-405.
- ZHAO N., HASHIDA H., TAKAHASHI N., MISUMI Y., SAKAKI Y. 1995. *Gene*, 156, 207-213.

