

## **LA TRAÇABILITÉ GÉNÉTIQUE DE LA VIANDE ET DES BACTÉRIES**

**Bernard CHINA**

Département des Sciences des Denrées Alimentaires, Faculté de Médecine Vétérinaire,  
Université de Liège, Sart Tilman, Bât. B43bis, 4000 Liège

Correspondance : Dr. Bernard CHINA  
Tél. 04 366 4029 - Fax 04 366 47 53 - E-mail : bchina@ulg.ac.be

### **Résumé**

Afin d'assurer une transparence tout le long de la chaîne de transformation de la viande, il faut disposer d'un système de traçabilité performant. La traçabilité administrative montre des limites que l'utilisation de marqueurs génétiques pourrait surmonter. Le génome de chaque individu possède des différences de séquences, à la base du polymorphisme génétique, dont les marqueurs génétiques sont les témoins. Parmi ceux-ci, deux classes semblent s'imposer en matière de traçabilité : les microsatellites et les polymorphismes simple nucléotide. Les microsatellites sont caractérisés par un degré de polymorphisme important avec de nombreux allèles pour un même locus. De plus, leur détection se fait directement par amplification itérative. Techniquement, le problème se pose lors de la détection simultanée de plusieurs microsatellites où le patron obtenu est difficile à interpréter. Les polymorphismes simple nucléotide sont représentés par des mutations ponctuelles dans la séquence nucléotidique. Ils sont fréquents, stables, répartis de façon aléatoire et généralement bialléliques. La détection se fait soit par hybridation sur des biopuces, soit par spectrométrie de masse ou par d'autres techniques facilement automatisables. Cette automatisation est nécessaire pour pouvoir tester un grand nombre de polymorphismes simultanément sur un grand nombre d'échantillons, permettant ainsi de diminuer les coûts. Quant aux microorganismes contaminant la viande, il est également possible d'en établir l'empreinte génétique afin d'une part d'établir un lien objectif entre la souche isolée de patients et la souche isolée de l'aliment et d'autre part d'établir le caractère épidémique de la souche.

**MOTS-CLES :** Traçabilité-marqueurs génétiques- polymorphisme

### **Abstract**

In order to assure a tracability along the meat transformation chain, a powerful tracability system is required. The administrative tracability shows limits that the use of genetic markers could overcome. The individual genome contains sequence differences, basis of the genetic polymorphism of which the genetic markers are the witnesses. Among them, two

classes seem to dominate on the tracability field : the microsattellites and the single nucleotide polymorphisms. Microsattellites are characterized by a high level of polymorphism with a lot of alleles for the same locus. Moreover, their detection is directly possible by polymerase chain reaction. Technically, the problem raised when simultaneous detection of several microsattellites is achieved due to the complexity of the obtained pattern. Single nucleotide polymorphisms are single nucleotide mutations in the nucleotidic sequence. They are frequent, stable, randomly scattered and generally biallelic. Their detection is performed either by hybridization on biochips or by mass spectrometry or by other easily automatable technics. This automation is required to test simultaneously a great number of polymorphisms in a large number of samples allowing to decrease the cost. For microorganisms spoiling food, it is also possible to have DNA fingerprints in order first to establish the link between the strain isolated from patient and the strain isolated from food and second to establish the peidemic character of the strain

KEY-WORDS : Tracability-genetic markers-Polymorphism.

## Introduction

Les récentes crises de la “ vache folle ” ou encéphalopathie spongiforme bovine et de la “ dioxine ” ont accentué les besoins de disposer d’un système efficace de traçabilité de nos animaux d’élevage, de leur naissance à leur transformation en viande. Lors de la crise de la dioxine, des milliers d’animaux furent détruits entraînant des pertes énormes pour l’ensemble du secteur. Le nombre de ces pertes aurait pu être réduit si un système de traçabilité efficace avait existé. En effet, les systèmes de traçabilité existants (SANITEL et autres) permettent difficilement de remonter l’ensemble de la filière et les fraudes restent possibles (Evrard *et al.*, 2001). Récemment cependant, les progrès dans le domaine de la génétique moléculaire des eucaryotes sous l’impulsion de vastes programmes de séquençage ont permis, d’une part d’obtenir un nombre grandissant de séquences d’ADN disponibles dans des banques de données et d’autre part, le développement de technologies permettant l’analyse génétique.

Si on excepte le cas des vrais jumeaux, le génome de chaque individu est unique. En d’autres termes, au sein d’une espèce, il est possible de distinguer des individus différents car ils possèdent des différences dans la séquence de leur ADN. Ces variations sont générées par des mutations qui au niveau de la population sont le moteur de l’évolution. Bien que ces mutations puissent se traduire

facilité, les premières études portèrent sur les protéines plutôt que sur les gènes. Elles concernaient les groupes sanguins, les protéines sériques, ... (Juneja et Vogeli, 1998). Mais les techniques s’étant affinées depuis le début des années 1980, les marqueurs génétiques sont directement étudiés à partir de l’ADN (tableau I).

Essentiellement trois types majeurs de polymorphismes génétiques peuvent être utilisés comme marqueurs génétiques : l’ADN satellite, les mutations “ simple polymorphisme ” et l’amplification aléatoire d’ADN polymorphe.

Tableau I: Types de marqueurs génétiques

Type	nombre/génome	polymorphisme	Répartition	Détection	Automatisation
VNTR	>10 000	+++	localisés aux extrémités des chromosomes	Southern Blot	non
STR	>100 000	+++	aléatoire	PCR	oui
SNP	>1 000 000	+	aléatoire	Méthodes Diverses	oui
RFLP	>100 000	+	aléatoire	PCR	non
RAPD	?	++	aléatoire	PCR/restriction	non
AFLP	?	++	aléatoire	Restriction/ PCR	oui

VNTR: *Variable Number of Tandem Repeat*

STR: *Short Tandem Repeat*

SNP: *Single Nucleotide Polymorphism*

RFLP: *Restriction Fragment Length Polymorphism*

RAPD: *Random Amplification of Polymorphic DNA*

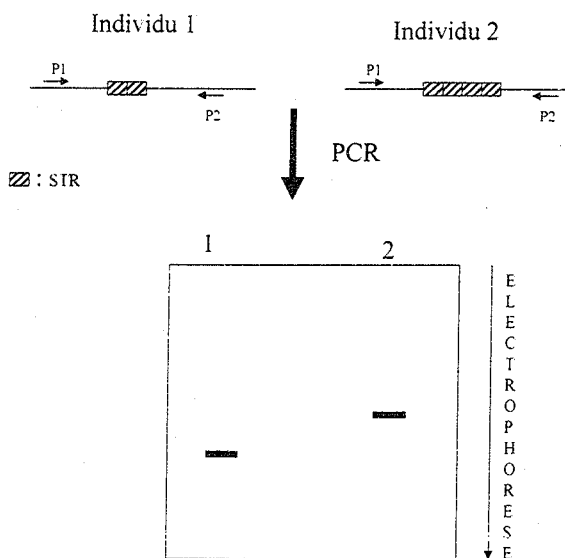
AFLP: *Amplified Fragment Length Polymorphism*

## LES MARQUEURS GENETIQUES

### L'ADN SATELLITE

Dans le génome, on trouve des séquences répétées. Celles-ci sont classées en fonction du nombre de répétitions, de la taille du motif répété et de la dispersion de ces séquences répétées dans le génome. L'ultracentrifugation en gradient de chlorure de césium a été depuis des années la technique de prédilection pour isoler de l'ADN génomique. On constate que l'ADN génomique forme une bande majeure dans le gradient et une bande plus fine appelée ADN satellite qui possède une densité différente due à des différences dans la composition en bases. L'analyse de cet ADN satellite révèle de petites séquences répétées plusieurs centaines ou milliers de fois. Cet ADN répété peut être localisé à des endroits précis du chromosome (centromère, télomère) ou dispersé de façon plus homogène dans le génome. Ces répétitions peuvent être groupées en tandem à un endroit précis du génome. Les minisatellites ont des unités répétées de 9 à 64 nucléotides, voire d'avantage (Moran, 1998). Le nombre de répétitions varie très fort à un endroit particulier entre individus ce qui en fait potentiellement une source de marqueurs génétiques. Le polymorphisme est dû à des différences dans le nombre de répétitions. En raison de l'hypervariabilité dans le nombre de copies répétées en tandem, les minisatellites sont aussi appelés VNTRs (*Variable Number Tandem Repeats*) (Nakamura *et al.*, 1987). Bien qu'ils soient très polymorphes, les minisatellites ne sont guère utilisés comme marqueur génétique principalement car ils ne sont pas de bonnes cibles pour l'amplification itérative (*Polymerase Chain Reaction* ou PCR) en raison de la taille importante des répétitions et du nombre important de répétitions. De plus, leur distribution ne semble pas aléatoire. On leur préfère les microsatellites qui comme les minisatellites sont des VNTRs. Cependant, ils ont l'avantage d'avoir des unités répétées plus petites (de 1 à 10 nucléotides), d'où leur appellation de STR (*Short Tandem Repeat*), pour une longueur totale de 10 à 100 pb et un nombre de copies raisonnablement bas ce qui en fait des cibles idéales pour l'amplification par PCR (figure 1). Les microsatellites sont répartis de façon relativement aléatoire le long du génome. On en compte en moyenne 50 000 par génome. Ils constituent donc de bons

candidats comme marqueurs génétiques. De plus, de nombreux microsatellites ont été identifiés chez les principaux animaux de rente. Ainsi, si on recherche dans la banque de séquences d'ADN (Genbank), on obtient les résultats suivants pour la recherche conjointe du terme "microsatellite" et de l'animal : *Bos taurus* (1345 références), *Gallus gallus* (749 références), *Sus scrofa* (348 références), *Ovis aries* (191 références) et *Capra hircus* (54 références). Une fois, un microsatellite identifié, il faut encore déterminer son degré de polymorphisme, ainsi que la fréquence des différents allèles dans la population étudiée et ce à partir d'un échantillon de celle-ci. En terme de traçabilité, on choisira les marqueurs les plus polymorphes. Ainsi, on constate que plus le nombre d'allèles pour un marqueur est important, plus le contenu en information est important et plus il sera facile de distinguer deux individus sur base de ce marqueur (Peelman *et al.*, 1998). En pratique, on est amené à combiner différents microsatellites pour obtenir un test suffisamment discriminant. Mais ceci n'est pas sans poser des problèmes d'ordre technique. En effet, il faut choisir des amplimères pour la PCR tel que la température optimale d'*annealing* soit proche pour chaque paire, de plus, il faut que la taille des produits de PCR générés ne se chevauche pas d'un microsatellite à l'autre. Il en résulte un patron de produits d'amplification assez complexe et difficile à interpréter. Cependant, la demande aidant, les appareils de séquençage automatique sont maintenant équipés notamment au niveau des logiciels pour pouvoir analyser ce genre de patron. Le degré de multiplexation ne dépasse pas 4 STRs pour la PCR (Heyen *et al.*, 1997, Peelman *et al.*, 1998). Mais des produits d'amplification marqués par le même fluorochrome peuvent être analysés simultanément par un appareil de séquençage automatique si les tailles des fragments sont suffisamment différentes. En utilisant cette stratégie, Heyen et collaborateurs (1997) ont pu analyser simultanément 22 microsatellites bovins.



**Figure 1.** Détection des Microsatellites. L'individu 1 et 2 présentent un nombre de répétitions différentes au même locus. La région comprenant le microsatellite est amplifiée par PCR avec les mêmes amorces. La taille des produits d'amplification est analysée par électrophorèse.

## LES “SINGLE NUCLEOTIDE POLYMORPHISMS”.

### Définition

Les SNPs pour “*single nucleotide polymorphism*” sont définis par la coexistence d’au moins deux bases différentes à la même position, chacun avec une fréquence > 1% (Krawczak, 1999). Ils affectent une seule paire de base soit par substitution (remplacement d’un nucléotide par un autre) soit par délétion ou insertion d’une base. Ces mutations ponctuelles sont fréquentes dans le génome (1/1000 pb). On estime leur nombre à environ  $3 \times 10^6$  pour le génome humain (Wang *et al.*, 1998 ; Sauer *et al.*, 2000). Les SNPs sont répartis uniformément tout le long du génome et sont, au même titre que les microsatellites, utilisés pour construire des cartes de linkage. Ces cartes fournissent des outils puissants pour la localisation et la caractérisation de facteurs génétiques (Collins *et al.*, 1997). Les SNPs sont relativement stables de génération en génération (Krawczak, 1999), mais ils ont tendance à être spécifiques d’une population. Le risque étant que les marqueurs développés pour une race au sein d’une espèce animale ne soient pas utilisables pour une autre race. Un des avantages majeurs des SNPs est leur capacité à être typés de façon rapide et semi-automatique (Landergrén *et al.*, 1998). Les stratégies utilisées sont nombreuses (China *et al.*, 2001), elles regroupent d’une part les techniques basées sur des différences de mobilités électrophorétiques (DGGE, SSCP, ...) ou de temps de résolution (DHPLC) qui ne permettent que de tester une ou deux mutations à la fois et qui sont peu automatisables et d’autre part, les techniques plus facilement automatisables et qui permettent de tester un grand nombre de SNPs en même temps : les puces à ADN, la spectrométrie de masse et l’électrophorèse capillaire.

### Détection

#### a. Les Puces à ADN.

De nombreuses analyses en biologie moléculaire reposent sur l’hybridation d’une sonde d’ADN sur des fragments d’ADN immobilisés sur un support solide (nylon, ...). Ces séquences cibles ont été déposées ou transférées sur ces membranes, appelées maintenant “*macroarray*”. Ceux-ci doivent être distingués des “*microarrays*”, terme réservé aux “*arrays*” miniaturisés appelés aussi biopuces ou puces à ADN. Les biopuces répondent, entre-autres, à la nécessité de pouvoir analyser les polymorphismes d’un génome (Wilgenbus et Lichter, 1999). L’hybridation couplée à l’utilisation de puces à ADN (Chee *et al.*, 1996) permet de tester un grand nombre (plusieurs milliers) de SNPs à la fois (Wang *et al.*, 1998). Les technologies actuelles permettent la fabrication de biopuces de 1,28 cm de côté pouvant contenir 400 000 oligonucléotides, chacun occupant 20 sur 20  $\mu\text{m}$  et contenant plus de  $10^7$  copies de la sonde (Chee *et al.*, 1996). Parmi les diverses solutions techniques disponibles sur le marché ou en développement (Marshall et Hodgson, 1998), la technique de choix pour la détection des SNPs consiste en des oligonucléotides courts synthétisés directement sur un support solide en utilisant la chimie des nucléotides photolabiles masqués séquentiellement (Fodor *et al.*, 1991 ; Pease *et al.*, 1994). L’ADN cible marqué par un fluorochrome et constitué de produits de PCR ou d’ADN génomique est appliqué et la détection des fluorochromes est effectuée. L’avantage des fluorochromes est que l’on peut utiliser simultanément différents marqueurs (figure 2A). Le dépôt des oligonucléotides et la détection du signal peuvent être automatisés. Les biopuces ont été utilisées pour la détection de SNPs (Wang *et al.*, 1998, Winzeler *et al.*, 1998 ; Cargill *et al.*, 1999 ; Halushka *et al.*, 1999). Pour éviter les problèmes de bruit de fonds dus au fait qu’il s’agit de distinguer des cibles différant par un seul nucléotide, deux types d’approches

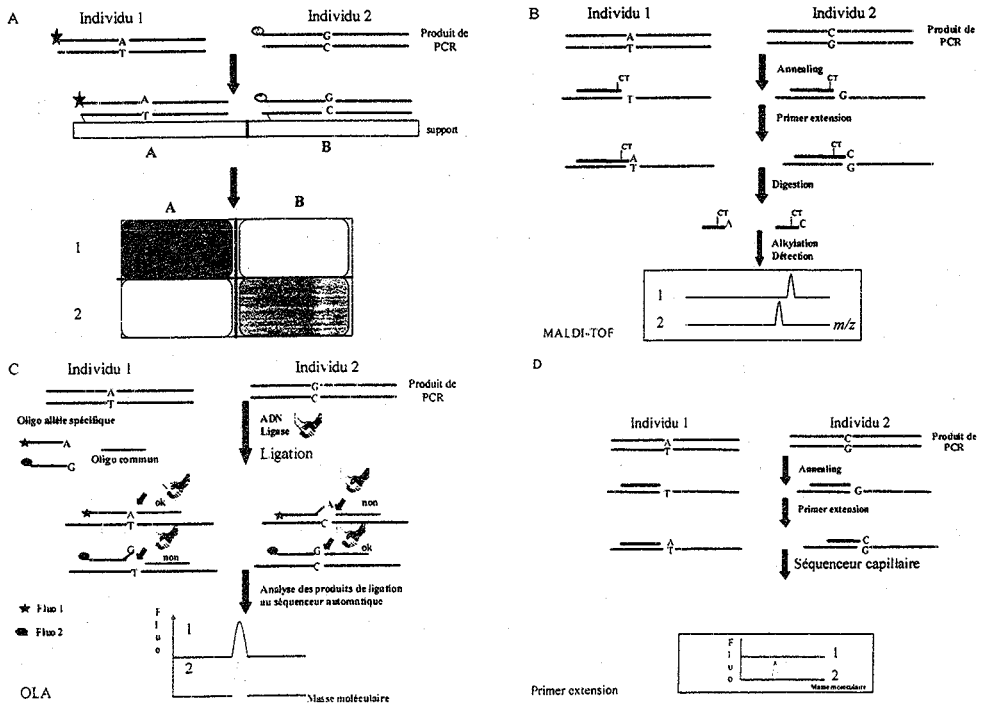
semblent prometteuses. D'une part, coupler l'hybridation à une réaction d'extension d'amorce (Dubiley *et al.*, 1999, Pastinen *et al.*, 2000). D'autre part, cette technologie peut être couplée à celles des " balises " moléculaires (" *molecular beacons* ") (Tyagi *et al.*, 1998) ce qui permettrait de diminuer les problèmes de bruit de fond en rendant l'hybridation plus spécifique. Les balises moléculaires sont des oligonucléotides qui ont la forme d'une épingle à cheveux avec un fluorophore masqué dont la fluorescence est restaurée quand elle lie un ADN cible. Ces molécules sont particulièrement utiles dans des situations où on travaille en excès de sonde. Pour la détection des SNPs, on pourrait charger le *microarray* avec plusieurs beacons marqués avec un fluorochrome différent ce qui permet de détecter plusieurs SNPs à la fois.

#### b. La spectrométrie de masse

En invoquant la présence d'un bruit de fond important si on se contente d'appliquer de l'ADN génomique et la nécessité donc de passer par une étape de purification avant l'analyse en puce à ADN qui limite la capacité de la méthode, Sauer et collaborateurs (2000) proposent une nouvelle approche basée sur la spectrométrie de masse MALDI-TOF (*Matrix-assisted laser desorption/ionization time-of-flight*) (Karas *et al.*, 1988) pour détecter les SNPs. La spectrométrie de masse (MS) *time-of-flight* (TOF) est simple, peu coûteuse, avec une grande sensibilité et une gamme pratiquement illimitée de masse. Cette technique révèle un intérêt tout particulier pour l'analyse des macromolécules biologiques comme les protéines, les hydrates de carbone et les oligonucléotides. La spectrométrie de masse TOF (TOFMS) n'est pas une technique nouvelle. Cependant, un intérêt nouveau dans cette technique a vu le jour grâce à l'introduction de la " *matrix assisted laser desorption* " (MALD) en 1988 par Karas et Hillenkamp d'une part et Tanaka et collaborateurs, d'autre part. En TOFMS, les ions sont formés dans une courte région de résonance en présence d'un champ électrique qui accélère les ions dans une plus longue région de dérive. Idéalement, tous les ions entrant dans la région de dérive ont la même énergie cinétique et ils ont une vitesse qui dépend de leur masse. Le temps nécessaire pour traverser la région de dérive dépend aussi de la masse de l'ion de sorte que le spectre de temps peut être converti en un spectre de masse. Les ions sont générés par des impulsions électriques à intervalles réguliers qui ouvrent ou ferment un faisceau d'ions sur l'échantillon ou qui contrôlent un laser pulsé (Cotter, 1992). Cette technique a été appliquée pour la détection des SNPs (Ross *et al.*, 1997; Li *et al.*, 1999). Sauer et collaborateurs (2000) proposent une technique en un seul tube complètement automatisée (figure 2B). En bref, (i) la région contenant le SNP est amplifiée par PCR, (ii) les nucléotides restant sont inactivés par une phosphatase, (iii) l'étape suivante est une expérience d'extension d'amorce (*primer extension*) sur le produit de PCR. L'amorce utilisée contient, près de l'extrémité 3', un nucléotide chargé. Des didéoxynucléotides modifiés ( $\alpha$ -S-ddNTP) sont utilisés lors de cette expérience. Les produits d'extension sont clivés par une endonucléase libérant les trois derniers nucléotides. Les bases sont alkylées et analysées en spectrométrie de masse.

#### c. L'électrophorèse capillaire

Le développement de l'électrophorèse capillaire a été rapide ces dernières années pour la séparation de l'ADN (Mitchelson, 2001). Elle offre un grand débit et une grande résolution ainsi qu'une automatisation des opérations : chargement des échantillons, électrophorèse, détection du signal et analyse des résultats. Cette technologie est particulièrement efficace



**Figure 2.** Méthode de détection des SNPs. A. Les biopuces. La région contenant le SNP est amplifiée par PCR. Le produit de PCR est marqué spécifiquement avec un fluorochrome différent selon la mutation. Le fragment marqué est hybridé sur un support sur lequel est fixé un oligonucléotide spécifique de l'allèle. L'hybridation est analysée par la présence ou l'absence de la fluorescence de la longueur d'onde correspondant aux fluorochromes utilisés. B. La spectrométrie de masse "Matrix-Assisted Laser Desorption Ionization Time of Flight" ou MALDI-TOF. La région contenant la mutation est amplifiée par PCR. Sur le produit de PCR, une expérience d'extension d'amorce est menée, en utilisant une amorce chargée et modifiée. Après extension, la partie modifiée résiste à la digestion par la phosphodiesterase et le produit est alkylé avant d'être analysé par spectrométrie de masse. Le spectre détermine la différence de masse en fonction de l'allèle présent chez l'individu 1 ou 2. C. *Oligonucleotide Ligation Assay* (OLA). Les régions contenant la mutation sont amplifiées par PCR. On réalise une expérience de ligation à l'aide d'une ADN ligase et de trois oligonucléotides, l'un commun pour les deux allèles et les deux autres spécifiques de chaque allèle et marqués par un fluorochrome spécifique (fluo 1 ou fluo 2). Si l'hybridation des oligonucléotides est parfaite, la ligation a lieu et un produit de ligation est détecté. Dans le cas contraire, la ligation n'a pas lieu et aucun produit de ligation n'est observé. D. L'extension d'amorce. Un oligonucléotide est hybridé en amont de la mutation. Cet oligonucléotide sert d'amorce pour l'ADN polymérase qui va ajouter un didéoxynucléotide correspondant à l'allèle présent. Chaque didéoxynucléotide est marqué par un fluorochrome différent. Les produits d'extension d'amorce sont analysés par électrophorèse capillaire. La couleur détectée détermine l'allèle présent.

dans la détection des SNPs (Syvanen, 1999). Deux méthodes existent : la OLA (Oligonucleotide Ligation Assay) et l'extension d'amorce (Primer Extension). Les tests de ligation tirent profit de la capacité d'une ADN ligase à joindre deux oligonucléotides adjacents quand ils sont hybridés sur un ADN cible (Landergrén *et al.*, 1988 ; Nickerson *et al.*, 1990). La ligation se déroule seulement quand les oligonucléotides sont parfaitement complémentaires de l'ADN cible. Dans le cas où on teste les deux allèles associés avec un SNP biallélique, trois oligonucléotides sont utilisés : deux oligonucléotides allèle-spécifique et un oligonucléotide commun (figure 2C). La nature exacte des modifications pour les trois oligonucléotides dépend du mode de détection et du format de la réaction. La stratégie de base est de déterminer quel oligonucléotide allèle-spécifique se lie à l'oligonucléotide commun afin de révéler l'allèle présent dans un échantillon d'ADN donné. En choisissant les amplimères de PCR et les oligonucléotides de ligation qui s'hybrident à des températures différentes, on peut combiner tous les réactifs pour réaliser la PCR et la ligation dans un seul tube (Eggerding, 1995). De plus, les tests de ligation peuvent être multiplexés (Grossman *et al.*, 1994). Dans l'extension d'amorce, un oligonucléotide s'hybride juste en amont de la mutation et il sert d'amorce à l'élongation par une ADN polymérase qui ajoute le didéoxynucléotide (ddNTP) correspondant à l'allèle présent. Chaque ddNTP est marqué par un fluorochrome différent. La multiplexation est possible soit dès l'étape de PCR en choisissant des amorces pour l'extension de taille différente pour chaque SNP soit lors de l'électrophorèse par injection rapprochée dans le même capillaire des produits d'extension. Dans les deux cas, le taux de multiplexation est de l'ordre de 10 SNPs/capillaire (Figure 2D).

#### d. Autres méthodes

D'autres alternatives comme la PCR en temps réel (Livak *et al.*, 1995) sont aussi prometteuses. La PCR en temps réel permet de quantifier l'ADN cible mais aussi de détecter des mutations ponctuelles. Cette technique permet de réaliser la PCR et l'hybridation dans le même tube. Cette technique (i) diminue les risques de contamination (ii) ne nécessite pas de manipulation post-PCR (iii) est une méthode rapide permettant une discrimination facile. L'analyse de deux allèles simultanément est possible. Cependant, cette technique est actuellement limitée par la faible possibilité de multiplexation. Quoique de constants progrès soient effectués (Germer *et al.*, 2000) et le "pyrosequencing" (Ronaghi *et al.*, 1996) s'avère également une technique de choix pour la détection des SNPs (Nyren *et al.*, 1997). Elle est basée sur un séquençage avec une détection fluorescente des bases incorporées. Le pyrosequencing est très intéressant pour évaluer la fréquence des allèles dans une population (Kruglyak, 1999). La détection simultanée de plusieurs SNPs restent problématique, même si la technique peut être automatisée.

En résumé, les trois techniques d'avenir pour la détection automatisée des SNPs sont : les biopuces, la spectrométrie de masse et l'électrophorèse capillaire.

## LE POLYMORPHISME ALEATOIRE

Un troisième type de polymorphisme correspond à la détection dans le génome de différences dont on ne connaît ni l'emplacement, ni la séquence. Parmi ceux-ci, on trouve principalement la RAPD (*Random Amplification of Polymorphic DNA*) et l'AFLP (*Amplified fragment length polymorphism*). Dans la RAPD, un oligonucléotide de séquence aléatoire sert d'amorce pour la PCR. Il va s'accrocher sur le génome à des endroits présentant une séquence

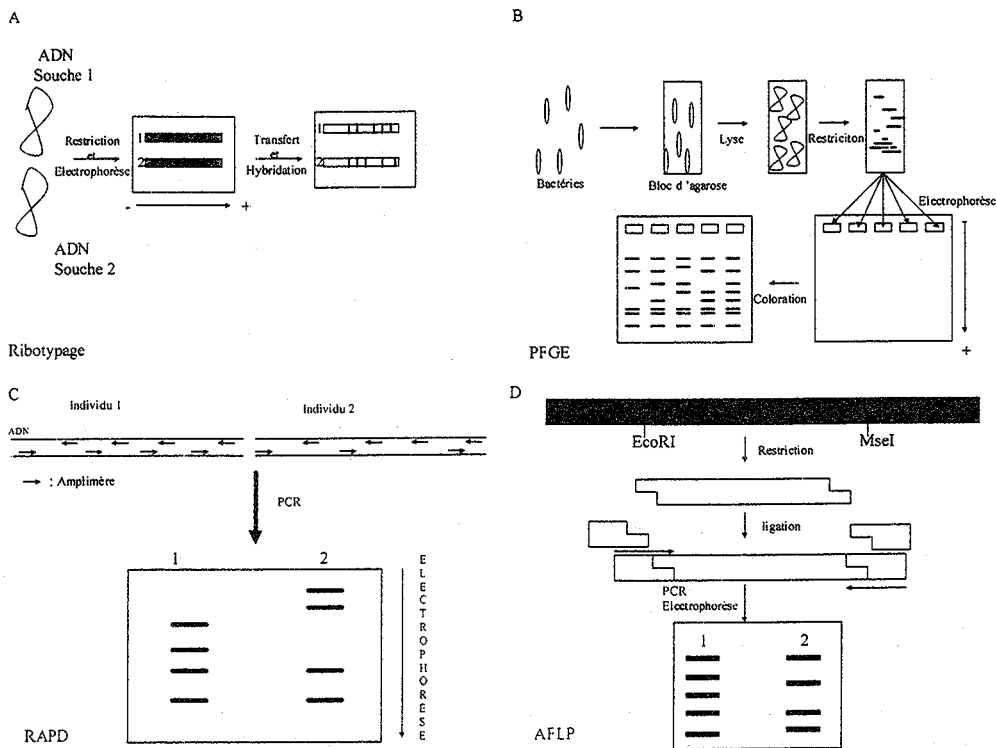


complémentaire à la sienne. L'amplification par PCR produit une série de fragments de tailles différentes (figure 3C). Le patron obtenu est caractéristique d'un individu. Le problème de la RAPD réside dans la reproductibilité des résultats qui varient en fonction des conditions de PCR, de l'enzyme ou de l'appareil utilisé (Meunier et Grimont, 1993), mais aussi d'une fois à l'autre toute chose étant égale (Perez *et al.*, 1998). De plus, les résultats observés et les résultats théoriques déduits de la séquence du phage lambda ne correspondaient pas (Perez *et al.*, 1998). Par conséquent, la nature de la variation génétique mise en évidence par la RAPD n'est pas claire. Compte tenu de ces considérations, la prudence devrait guider les conclusions basées sur la seule RAPD. L'AFLP (Vos *et al.*, 1995) se déroule en trois étapes (figure 3D). L'ADN génomique est digéré par deux enzymes et les fragments générés sont ligaturés à des adaptateurs oligonucléotidiques. La deuxième étape est une étape de PCR en deux phases : d'abord une amplification non sélective de chaque fragment de restriction en utilisant des primers complémentaires aux adaptateurs et ensuite une amplification sélective en utilisant des primers complémentaires des adaptateurs, du site de restriction et de un ou deux nucléotides en aval du site, afin de diminuer la complexité du patron obtenu. La troisième étape est une analyse électrophorétique des fragments générés avec obtention d'un patron spécifique de l'ADN étudié. Cette technique a été utilisée avec succès dans la caractérisation d'ADN porcins notamment (Ovilo *et al.*, 2000).

### **La traçabilité des microorganismes.**

Lorsqu'on isole une souche de bactérie pathogène d'un aliment deux types de questions viennent à l'esprit du scientifique : la souche isolée de l'aliment est-elle la même que celle isolée des patients ayant ingéré l'aliment et cette souche est-elle semblable ou apparentée aux autres souches isolées auparavant? Pour répondre à ces questions, des études comparatives s'imposent. Celles-ci se classent en deux catégories : les tests phénotypiques basés sur l'étude des caractères exprimés par la bactérie (antigéniques ou biochimiques) et les tests génotypiques basés sur l'étude de la séquence en ADN (Milch, 1998 ; Van, 2002). Nous allons focaliser notre attention sur ces derniers. Quand on veut comparer le génome de deux souches bactériennes, la chose la plus évidente est d'en établir la séquence nucléotidique complète et de les comparer. Quoique les progrès en matière de séquençage aient été spectaculaires ces dernières années en raison des programmes de séquence entrepris, il n'est pas encore possible d'obtenir cette information facilement et rapidement. C'est pourquoi, les chercheurs ont développé des méthodes alternatives. Elles sont basées sur l'utilisation de méthodes indirectes permettant d'avoir des informations sur la séquence de l'ADN. Historiquement, la première technique utilisée fut le ribotypage (Hofstra *et al.*, 1994). Elle est basée sur la présence dans les souches de bactéries de plusieurs copies de l'opéron codant pour les ARN ribosomaux 16S et 23S. Ce nombre de copie varie d'une bactérie à l'autre (7 pour *E. coli*, 10 pour *Clostridium perfringens*). Le principe est le suivant (figure 3A), l'ADN de la bactérie est restreint par une endonucléase de restriction reconnaissant un site à 6 paires de bases (un site toute les 4096 pb). Après restriction, les nombreux (mille environ) fragments d'ADN sont séparés sur une électrophorèse en gel d'agarose. Le nombre élevé de fragments ne permet pas de visualiser le polymorphisme, c'est pourquoi les fragments sont transférés sur une membrane de nylon. Celle-ci est hybridée avec une sonde correspondant à l'ADN ribosomal 16S ou 23S. Après révélation, il apparaît des fragments dont la taille est fonction de la conservation des sites de

restriction pour l'enzyme utilisé et donc de la séquence nucléotidique de la séquence. Puisque l'opéron est présent en plusieurs copies, on observe plusieurs bandes. La comparaison est d'autant plus puissante que le nombre de bandes est important. Les problèmes de cette technique étaient initialement son manque d'automatisation, son caractère long et fastidieux mais aussi la possibilité d'obtenir de fausses images liées à la présence de restrictions partielles de l'ADN. Récemment cependant, cette technique a été complètement automatisée grâce au Riboprinter de la firme Qualicon (Grif et al., 1998). Cette machine honnêteuse (150 000 ₮) est utilisée actuellement dans les laboratoires de référence spécialisés dans le typage des souches bactériennes. Plus récemment, une alternative à cette technique a été proposée, il s'agit de l'utilisation de l'électrophorèse en champs pulsé (Butsmante et al., 1993 ; Mendez-Alvarez et al., 1995). Cette technique permet de séparer des fragments d'ADN de grandes tailles. Dans une électrophorèse classique, l'ADN d'une taille supérieure à 50 kb se trouve bloqué dans les puits du gel car il est incapable de s'insinuer entre ces mailles. Dans l'électrophorèse en champs pulsé, la direction du champs électrique est réorienté régulièrement, ce qui permet à l'ADN de ramper entre les mailles du gel. Ainsi, les bactéries sont emprisonnées dans des bloc d'agarose. La lyse de la bactérie ainsi que la restriction se font par diffusion à travers ces blocs afin de préserver l'intégrité de l'ADN. Les blocs contenant l'ADN restreint sont placés dans les puits d'un gel d'agarose qui subit l'électrophorèse en champs pulsé. Cette fois, l'enzyme a été choisie de manière à couper rarement l'ADN de la bactérie (une vingtaine de fragments). Après électrophorèse, ces fragments sont directement visualisés après coloration de l'ADN. Le polymorphisme est de nouveau basé sur la conservation ou non des sites de restriction. L'avantage de cette technique est qu'on évite les étapes de transfert et d'hybridation présentes dans le ribotypage. L'inconvénient de la PFGE (Pulsed Field Gel Electrophoresis) est la lenteur de la manipulation (environ 3 jours) due au temps nécessaire pour la lyse, la restriction et l'électrophorèse. Les techniques d'amplification aléatoire RAPD et AFLP sont également d'application pour les bactéries. Des kits commerciaux d'AFLP sont d'ailleurs disponibles pour les bactéries (Applied Biosystems). De plus, l'AFLP peut être analysé par électrophorèse capillaire (Desler-Nurmi et al., 2000) Enfin, même si ce n'est pas encore répandu, il semble que les chercheurs s'intéressent de près à la détection de SNPs chez les bactéries. Quelque soit la technique utilisée, on se trouve toujours face à des profils de bandes de taille différentes à analyser. Cette analyse se fait grâce à des logiciels basés sur le principe suivant : on calcule d'abord une matrice de similitude basé sur le calcul d'un indice par exemple l'indice de Jaccard (Priest et Austin, 1993). Si on considère a : le nombre de bandes présente dans les deux souches, b : le nombre de bandes présentes dans la souche 1 et pas dans la souche 2 et c : le nombre de bandes présentes dans la souche 2 et pas dans la souche 1 alors,  $S_j = a/a+b+c$ . L'indice est compris entre 0 et 1. Deux souches sont jugées identiques si  $S_j=1$  c'est-à-dire si toutes les bandes sont identiques. A partir des indices, on construit une matrice de similitudes qui donne le pourcentage de similitude entre les diverses souches testées. Cette matrice permet enfin de générer des dendrogrammes très utiles pour une meilleure visualisation des résultats (Priest et Austin, 1993).



**Figure 3. Typage moléculaire des bactéries.** A : le ribotypage. L'ADN des souches bactériennes est digéré par une endonucléase de restriction. Les produits de la restriction sont séparés par électrophorèse en gel d'agarose. Les fragments sont transférés sur membrane de nylon et hybridés avec une sonde correspondant à l'ADN ribosomal. Le profil d'hybridation est analysé. B : L'électrophorèse en champs pulsé (PFGE). Les bactéries sont emprisonnées dans des blocs d'agarose. La lyse et la restriction se font à travers ces blocs. Après restriction, les blocs sont chargés dans les puits d'un gel d'agarose et les fragments d'ADN sont séparés par électrophorèse en champs pulsé et visualisés par coloration (Bromure d'éthidium, Sybr Green,.....) C : *Random Amplification of Polymorphic DNA* (RAPD). Le même amplimère est utilisé pour réaliser une réaction de PCR sur le génome de l'individu 1 et 2. Selon les endroits où l'amorce vient s'hybrider, la taille des fragments d'amplification varie. D : *Amplification of Fragment Length Polymorphism* (AFLP). L'ADN est restreint par deux enzymes différentes (EcoRI et MseI). Les fragments de restriction générés sont ligués avec des adaptateurs de séquence connue. Les produits de la ligation sont amplifiés par PCR avec des amorces complémentaires aux adaptateurs et des deux premières bases du fragment de restriction. Les produits d'amplification sont analysés par électrophorèse.

## Discussion

Puisque le patrimoine génétique contenu dans chacune des cellules des individus est identique et puisque deux individus présentent des variations dans leur séquence en ADN, un individu peut être caractérisé par la séquence de son ADN. Ainsi, l'utilisation des marqueurs génétiques permet de suivre, à la trace, l'individu ou n'importe quelle cellule nucléée de celui-ci. C'est pourquoi cette approche devrait permettre de suivre un animal de rente de l'étable à la table tout au long de la filière de transformation. Quelles sont les qualités requises pour une plate-forme de génotypage ? Il faut que les analyses soient rapides et peu coûteuses et que les marqueurs génétiques choisis soient discriminants (entre espèces, entre races et entre individus). Les résultats doivent être facilement interprétables et transformables en un génotype tenant en une formulation simple éventuellement à ajouter aux systèmes de traçabilité administratives déjà existants.

Il résulte de cette première analyse que les marqueurs génétiques les plus efficaces semblent être les microsatellites et les SNPs. L'utilité des SNPs en tant que marqueurs génétiques reste un sujet de discussion. Il est important d'en apprécier les avantages et d'en connaître les limites. L'intérêt des SNPs s'est focalisé sur les technologies potentielles pour tester cette classe de marqueurs génétiques qui engendrent des tests simples du type oui/non pouvant être réalisés en parallèle et automatisés. Les SNPs sont généralement bialléliques, il y a deux allèles qu'un individu peut avoir pour chaque marqueur. Cela signifie que l'information contenue par SNP est relativement faible si on la compare à l'information contenue dans les marqueurs de type microsatellites qui peuvent avoir 10 allèles ou plus (selon le nombre de fois que l'unité de base est répétée). On estime qu'il faut en moyenne 5 marqueurs SNPs pour égaler l'information contenue dans un microsatellite (Chakraborty *et al.*, 1999). Il faut donc un nombre suffisant de marqueurs (au moins 20) pour distinguer des individus. Ceci pose le problème de la multiplexation de la PCR qui est souvent l'étape limitante (Pastinen *et al.*, 2000). Les SNPs ont tendance à être spécifiques d'une population ; un marqueur qui est polymorphe dans une population peut ne pas l'être dans une autre. Cela signifie que les marqueurs SNPs devront être générés spécifiquement pour la population étudiée. En comparaison, les marqueurs microsatellites sont polymorphes à travers diverses populations ce qui signifie qu'une fois générés, ils sont d'usage universel. Mais si on étudie une population où les croisements consanguins sont fréquents comme c'est le cas des élevages porcins en Belgique, les mêmes SNPs peuvent être utilisés. De plus, si on sélectionne un nombre suffisant de marqueurs (40-50) le pouvoir discriminant en sera augmenté. De plus, les SNPs sont fréquents (tous les 1000 bases) dans la plupart des génomes offrant le potentiel pour générer des cartes génétiques très denses qui peuvent être très utiles pour le repérage de gènes ou de régions d'intérêt. De part la nature des SNPs, ils peuvent être associés avec les phénotypes pathologiques étudiés. Le taux de mutation bas des SNPs en font également de bons marqueurs pour étudier des traits génétiques complexes. Le défi futur sera d'intégrer au mieux ces deux types de marqueurs génétiques (SNPs et microsatellites) pour développer des outils de génotypages à la fois puissants et flexibles.

En ce qui concerne les microorganismes, les diverses méthodes utilisées (Ribotypage, PFGE, RAPD, AFLP, ...) visent toutes à obtenir des renseignements indirects sur la conservation de la séquence d'ADN. L'automatisation de ces techniques est satisfaisante pour le ribotypage et l'AFLP. Mais le développement du génotypage par détection de SNPs via des puces à ADN risque de venir bousculer ces techniques et de s'imposer dans le génotypage à grande échelle des souches bactériennes.

## Références

- BUSTAMANTE C., GURRIERI S., SMITH S. B. Towards a molecular description of pulsed-field gel electrophoresis. *Trends Biotechnol.*, 1993, 11, 23-30.
- CARGILL M., ALTSHUER D., IRELAND J., SKLAR P., ARDLIE K., PATIL N., LANE C. R., LIM E. P., KALAYANARAMAN N., NEMESH J., ZIAUGRA L., FRIEDLAND L., ROLFE A., WARRINGTON J., LIPSHUTZ R., DALEY G. Q., LANDER E. S. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.*, 1999, 22, 231-238.
- CHAKRABORTY R., STIVERS DN, SU B., ZHONG Y., BUDWOLE B. The utility of short tandem repeat loci beyond human identification : implications for development of new DNA typing systems. *Electrophoresis*, 1999, 20, 1682-1696.
- CHEE M., YANK R., HUBBELL E., BERNO A., HUANG X. C., STERN D., WINKLER J., LOCKHART D. J., MORRIS M. S., FODOR S. P. A. Accessing Genetic information with high-density arrays. *Science*, 1996, 274, 610-613
- CHINA B., EVRARD V., NOIRFALISE R., CLINQUART A., DAUBE G. La traçabilité dans la filière viande. II. Les marqueurs génétiques. *Annales de Médecine Vétérinaire*, 2001, 145, 15-24.
- COLLINS F. S., GUYER M. S., CHAKRAVARTI A. Variations on a theme : cataloging human DNA sequence variations. *Science*, 1997, 278, 1580-1581.
- COTTER R.J. Time-of-flight mass spectrometry for the structural analysis of biological molecules. *Anal. Chem.*, 1992, 64, 1027A-1039A.
- DRESLER-NURMI A., TEREFWORK Z., KAIJALAINEN S., LINDSTROM K., HATAKKA A. Silver stained polyacrylamide gels and fluorescence-based automated capillary electrophoresis for detection of amplified fragment length polymorphism patterns obtained from white-rot fungi in the genus *Trametes*. *J. Microbiol. Methods*, 2000, 41, 161-172.
- DUBILEY S., KIRILOV E., MIRZABEKOV A. Polymorphism analysis and gene detection by minisequencing on an array of gel-immobilized primers. *Nucl. Ac. Res.*, 1999, 27, e19
- EGGERDING F.A. A one-step coupled amplification and oligonucleotide ligation procedure for multiplex genetic typing. *PCR Methods Appl.*, 1995, 4 : 337-345.
- EVRARD, V., CHINA, B., NOIRFALISE, R., DAUBE G., CLINQUART, A. La traçabilité dans la filière viande. I. La traçabilité administrative. *Annales de Médecine Vétérinaire*, 2001, 145, 5-12.
- FODOR S. P. A., READ J. L., PIRRUNG M. C., STRYER L., LU A. T., SOLAS D. Light-directed spatially addressable parallel chemical synthesis. *Science*, 1991, 251, 767-773.
- GERMER S., HOLLAND M. J., HIGUCHI R. High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. *Genome Res.*, 2000, 10, 258-266.
- GRIF K., Karch H., Schneider C., Daschner F. D., Beutin L., Cheasty T., Smith H., Rowe B., Dierich M. P., Allerberger F. Comparative study of five different techniques for epidemiological typing of *Escherichia coli* O157. *Diagn. Microbiol. Infect. Dis.*, 1998, 32, 165-176.

- GROSSMAN P. D., BLOCH W., BRINSON E., CHANG C. C., EGGERDING A., FUNG S., IOVANNISCI D. A., WOO S., WINN-DEEN S. High-density multiplex detection of nucleic acid sequences : oligonucleotide ligation assay and sequence-coded separation. *Nucl. Ac. Res.*, 1994, 22, 4527-4534.
- HALUSHKA M. K., FAN J. B., BENTLEY K., HSIE L., SHEN N., WEDER A., COOPER R., LIPSHUTZ R., CHAKRAVIATI A. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat. Genet.*, 1999, 22, 239-247.
- HEYEN D. W., BEEVER J. E., DAY Y., EVERT R. E., GREEN C., BATES S. R., ZIGLE J. S., LEWIN H. A. Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semi-automated parentage testing. *Anim. Genet.*, 1997, 28, 21-27.
- HOFSTRA H., VAN DER VOSSEN J. M., VAN DER PLAS J. Microbes in food processing technology. *FEMS Microbiol. Rev.*, 1994, 15, 175-183.
- JUNEJA R. K., VOGELI P. Biochemical genetics In : Rotschild M. F., Ruvinsky A. (Eds ), *The genetics of the pig*. CAB International : Wallingford, 1998, 105-134.
- KARAS M., HILLENKAMP F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.*, 1988, 60, 2299.
- KRAWCZAK M. Informatively assessment for biallelic single nucleotide polymorphisms. *Electrophoresis*, 1999, 20, 1676-1681.
- KRUGLYAK L. Prospects for the whole genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.*, 1999, 22, 139-144.
- LANDERGREN U., KAISER R., SANDERS J., HOOD L. *Science*, 1988, 241, 1077-1080.
- LANDERGREN U., NILSSON M., KWOK P. Y. *Genome Res.*, 1998, 8, 769-776.
- LI J., BUTLER J. M., LIN H., ROYER S., OHLER L., SHALER T. A., HUNTER J. M., POLLART D. J., MONFORTE J. A., BECKER C. H. Single nucleotide polymorphism determination using primer extension and time-of-flight mass spectrometry. *Electrophoresis*, 1999, 20, 1258-1285.
- MARSHALL A., HODGSON J. DNA chips : an array of possibilities. *Nat. Biotechnol.*, 1998, 16, 27-31.
- MENDEZ-ALVAREZ S., PAVON V., ESTEVE I., GUERRERO R., GAJU N. Analysis of bacterial genomes by pulsed field gel electrophoresis. *Microbiologica*, 1995, 11, 323-336.
- MEUNIER J. R., GRIMONT P. A. Factors affecting reproductibility of random amplified polymorphic DNA. *Res. Microbiol.*, 1993, 144, 373-379.
- MILCH H. Advance in bacterial typing methods (a review). *Acta Microbiol. Immunol. Hung.*, 1998, 45, 401-408.
- MITCHELSON, K. R. The application of capillary electrophoresis for DNA polymorphism analysis. *Methods in molecular biology*, 2001. 162, 3-26
- MORAN C. Molecular genetics. In : Rotschild M. F., Ruvinsky A. (Eds.), *The genetics of the pig*. CAB International : Wallingford, 1998, 135-162 .
- NAKAMURA Y., LEPPERT M., O'CONNELL P., WOLFF R., HOLM T., CULVER M., FUJIMOTO E., HOFF M., KUMLIN E., WHITE R. Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, 1987. 235, 1616-1622.
- NEI M., ROYCHOUDHURY A. K. Sampling variances of heterozygosity and genetic distance. *Genetics*, 1974, 76, 379-390.

NICKERSON D.A., KAISER R., LAPPIN S., STEWART J., HOOD L., LANDERGRÉN U. Automated DNA diagnostics using ELISA-based oligonucleotide ligation assay. *Proc. Natl. Acad. Sci. USA*, 1990, 87, 8923-8927

NYREN P., KARAMOHAMED S., RONAGHI M. Detection of Single-base changes using bioluminometric primer extension assay. *Anal. Biochem.*, 1997, 244, 366-373.

OVILO C., CERVERA MT, CASTELLANOS C., MARTÍNEZ-ZAPATER JM. Characterisation of iberian pig genotypes using AFLP markers. *Animal Genetics*, 2000, 31, 117-122.

PASTINEN T., RAITIO M., LINDROSS K., TAINOLA P., PELTONEN L., SYVANEN A.-C. A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Research*, 2000, 10, 1031-1042.

PEASE A. C., SOLAS D., SULLIVAN E. J., CRONIN M. T., HOLMES C. P., FODOR S. P. A. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. USA*, 1994, 91, 5022-5026.

PEELMAN L. J., MORTIAUX F., VAN ZEVEREN A., DANSERCOER A., MOMMENS G., COOPMAN F., BOUQUET Y., BURNY A., RENAUVILLE R., PORTETELLE D. Evaluation of the genetic variability of 23 bovine microsatellite markers in four belgian cattle breeds. *Animal Genetics*, 1998, 29, 161-167.

PEREZ T., ALBORNOZ J., DOMINGUEZ A. An evaluation of RAPD fragment reproductibility and nature. *Mol. Ecol.*, 1998, 7, 1345-1357

PRIEST F. and AUSTIN B. *Modern bacterial taxonomy*. Chapman et Hall, London, 1993, 228 p.

RONAGHI M., KARAMOHAMED S., PETERSSON B., UHLEN M., NYREN P. Real-Time DNA sequencing using detection of pyrophosphate release. *Anal. Biochem.*, 1996, 242, 84-89.

ROSS P. L., LEE K., BELGRADER P. Discrimination of single-nucleotide polymorphisms of human DNA using peptide nucleic acid probes detected by MALDI-TOF mass spectrometry. *Anal. Chem.*, 1997, 69, 4197-4202.

SAUER S., LECHNER D., BERLIN K., LEHRACH H., ESCARY J. L., FOX N., GUT I. G. A novel procedure for efficient genotyping of single nucleotide polymorphisms. *Nucl. Acid Res.*, 2000, 28, e13.

SYVANEN, A.-C., From gels to chips : "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum. Mutat.*, 1999, 13, 1-10.

TANAKA K., WAIKI H., IDO Y., AKIHA S., YOSHIDA Y., YOSHIDA T. Rapid Commun. Mass Spectrom., 1988, 2, 151

TYAGI S., BRATU D., KRAMER F.R. Multicolor beacons for allele discrimination. *Nature Biotechnol.*, 1998, 16, 49-53.

VAN B.A. Molecular typing of micro-organisms : at the centre of diagnostics, genomics and pathogenesis of infectious diseases ? *J. Med. Microbiol.*, 2002, 51, 7-10.

VOS P., HOGERS R., BLEEKER M. et al. AFLP : a new technique for DNA fingerprinting. *Nucl. Ac. Res.*, 1995, 23, 4407-4414.

WANG D. G., FAN J. B., SIAO C. J., BERNO A., YOUNG P., SAPOLSKY R., GHANDOUR G., PERKINS N., WINCHESTER E., SPENCER J., KRUGLYAK L., STEIN L., HSIE L., TOPALOGLOU I., HUBBELL E., ROBINSON E., MITTMANN M., MORRIS M. S., SHEN N., KILBURN D., RIOUX J., NUSBAUM C., ROZEN S., HUDSON T. J., LIPSHUTZ R., CHEE M., LANDER E. S. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, 1998, 280, 1077-1082.

WILGENBUS K. K., LICHTER P. DNA chip technology ante portas. *J. Mol. Med.*, 1999, 77, 761-768.

WINZELER E. A., RICHARDS D. R., CONWAY A. R., GOLDSTEIN A. L., KALMAN S., McCULLOUGH M. J., McCUSKER J. H., STEVENS D. A., WODICKA L., LOCKHART D. J., DAVIS R. W. Direct allelic variation scanning of the yeast genome. *Science*, 1998, 281, 1194-1196.