



## **La sélection génomique : principe et perspectives d'utilisation pour l'amélioration des populations porcines**

Thierry Tribout, Jean Pierre Bidanel, Florence Phocas, Sandrine S. Schwob,  
Francois F. Guillaume, Catherine Larzul

### **► To cite this version:**

Thierry Tribout, Jean Pierre Bidanel, Florence Phocas, Sandrine S. Schwob, Francois F. Guillaume, et al.. La sélection génomique : principe et perspectives d'utilisation pour l'amélioration des populations porcines. 43. Journées de la Recherche Porcine, Feb 2011, Paris, France. pp.13-25. <hal-01000507>

**HAL Id: hal-01000507**

**<https://hal.science/hal-01000507v1>**

Submitted on 3 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# La sélection génomique : principe et perspectives d'utilisation pour l'amélioration des populations porcines

Thierry TRIBOUT (1,2), Jean-Pierre BIDANEL (1,2), Florence PHOCAS (1,2), Sandrine SCHWOB (1,2,3),

François GUILLAUME (1,2,4), Catherine LARZUL (1,2)

(1) INRA, UMR1313 Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France

(2) AgroParisTech, UMR1313 Génétique Animale et Biologie Intégrative, 75005 Paris, France

(3) IFIP-Institut du Porc, 35650 Le Rheu, France

(4) Institut de l'Elevage, 75012 Paris, France

thierry.tribout@jouy.inra.fr

## La sélection génomique : principe et perspectives d'utilisation pour l'amélioration des populations porcines

L'évaluation génomique est une nouvelle méthodologie d'estimation de la valeur génétique des animaux d'élevage. Son principe est de subdiviser le génome en un très grand nombre de segments chromosomiques (SC) déterminés par un ou plusieurs marqueurs SNP (Single Nucleotide Polymorphism), d'estimer l'effet de chacun de ces SC sur les caractères d'intérêt dans une population de référence (PR) phénotypée et génotypée pour ces marqueurs, puis de calculer la valeur génomique d'individus non phénotypés en sommant les effets estimés des SC portés par ces individus. Les méthodes d'estimation des effets des SC sont nombreuses et variables en termes de précision et de temps de calculs. L'efficacité d'une sélection fondée sur cette méthodologie augmente avec l'héritabilité du caractère considéré, la taille de la PR et le niveau de déséquilibre de liaison entre marqueurs adjacents, mais est plus faible dans les populations de grand effectif génétique ou lorsque la distance génétique entre la PR et les individus évalués est grande. Chez le porc, la sélection génomique pourrait permettre d'améliorer les populations pour des caractères non mesurables en routine, d'augmenter la précision des valeurs génétiques estimées, et de sélectionner les individus de l'étage de sélection pour leur valeur en croisement et en milieu de production. Les coûts de mise en œuvre d'une sélection génomique chez le porc seraient élevés, en raison principalement de la taille des PR nécessaires à une évaluation précise et du nombre élevé de candidats à génotyper. Le partage de PR par les différentes Organisations de Sélection et le recours aux techniques d'imputation génotypiques permettraient de maîtriser ces coûts.

## Genomic selection: principles and possible uses in pig genetic improvement schemes

Genomic evaluation is a new methodology to estimate the breeding value of livestock. Its principle is to subdivide the genome into a very large number of chromosomal segments (CS) determined by one or several SNP (Single Nucleotide Polymorphism) markers, to estimate the effect of each of these CS on the traits of interest in a reference population (RP) phenotyped and genotyped for these markers, and then to calculate the genomic value of unphenotyped animals by summing the estimated effects of the CS carried by these individuals. The methods to estimate the effects of the CS are numerous, and vary in terms of accuracy and computing time. The efficiency of a selection based on this methodology increases with the heritability of the considered trait, the size of RP and the level of linkage disequilibrium between adjacent markers, but is lower in populations with a large effective size or when the genetic distance between the RP and the evaluated animals is large. In pigs, genomic selection could be used to improve populations for traits too costly to be recorded routinely, to increase the accuracy of the estimated breeding values, and to select the nucleus populations for their crossbred value expressed in a commercial environment. The implementation of genomic selection in pigs would be costly, mainly because of the size of the RP required to estimate accurate breeding values and the high number of candidates that should be genotyped. The sharing of RP by several breeding organizations and the use of genotypes imputation techniques would help reduce the implementation costs.

## INTRODUCTION

Jusqu'à récemment, la grande majorité des programmes d'amélioration génétique des populations animales d'élevage s'appuyaient sur le « modèle infinitésimal », qui considère le génome comme une « boîte noire » contenant un très grand nombre de gènes non identifiés ayant chacun un effet très faible sur les caractères. Le principe sous-jacent à ce modèle est qu'un individu porteur d'allèles favorables aux gènes impliqués dans le déterminisme d'un caractère aura de bonnes performances, et qu'il va transmettre à ses descendants une partie de ses allèles et donc de sa supériorité. On peut ainsi estimer la valeur génétique des individus d'une population à partir de leurs performances propres et/ou de performances d'individus apparentés (les individus apparentés portant des allèles identiques pour un nombre de gènes proportionnel à leur degré de parenté) et d'informations généalogiques. La sélection des allèles favorables aux gènes dans la population se fait ainsi implicitement, sans avoir besoin de s'intéresser aux gènes eux mêmes.

Chez le porc, comme dans les autres espèces, les programmes d'amélioration génétique fondés sur le modèle infinitésimal ont été (et demeurent) très efficaces, comme en témoignent les évolutions génétiques réalisées au cours des précédentes décennies (voir par exemple Tribout *et al.*, 2003, 2004). Cependant, en dépit de sa simplicité et de son efficacité, il est évident que le « modèle infinitésimal » ne reflète qu'imparfaitement la réalité. On connaît en effet depuis longtemps des gènes ayant un effet majeur sur les performances, et on peut supposer qu'il existe un nombre non négligeable de gènes d'effet important. Il est alors tentant de sélectionner directement les individus de la population pour leur génotype à ces gènes, afin de fixer les allèles favorables plus rapidement qu'avec les programmes classiques.

Le génotype des individus à un gène d'intérêt n'est cependant en général pas accessible, le gène lui-même n'étant la plupart du temps même pas identifié (on sait seulement qu'il se trouve dans une région plus ou moins précise du chromosome, on parle alors de locus à effet quantitatif ou QTL) ; on est donc contraint de se contenter du génotype des individus à des « marqueurs » de ce gène (ou de ce QTL) pour distinguer les individus porteurs d'allèles favorables ou défavorables. Ces marqueurs peuvent être de nature diverse : visuels (coloration de la robe de l'animal), physiologiques (réaction de l'animal à un stimulus externe), biochimiques (protéines sanguines), ou encore génétiques (fragments d'ADN), et doivent présenter différentes formes dans la population, associées (dans l'idéal le plus étroitement possible) aux différents allèles du gène ou du QTL. La connaissance du variant d'un marqueur porté par un individu permet ainsi de savoir indirectement, avec toutefois parfois un risque de se tromper, si l'individu est génétiquement intéressant car porteur d'un allèle favorable au gène ou au QTL (et mérite donc d'être mis à la reproduction afin de transmettre sa supériorité à ses descendants) ou pas (et doit dans ce cas être réformé). Cette pratique est connue sous le nom de « sélection assistée par marqueurs » (SAM), dont l'une des toutes premières applications en sélection animale fut l'éradication de l'allèle de sensibilité à l'halothane dans la population porcine Landrace Français dans les années 1980. Ce programme est néanmoins longtemps resté la seule application de la SAM chez le porc, faute de marqueurs disponibles permettant de mettre en évidence des gènes d'effets moins marqués.

Dans les années 1990, la découverte des marqueurs microsatellites et la production dans les différentes espèces animales d'élevage de cartes génétiques (Archibald *et al.*, 1995) constituées grâce à ces marqueurs ont attisé l'enthousiasme de la communauté scientifique pour la SAM. De nombreux programmes de recherche ont été réalisés, conduisant à la mise en évidence d'un nombre important de QTL dans l'ensemble des espèces animales de rente. Des programmes de SAM ont été mis en place avec succès dans plusieurs espèces, en France par exemple chez les bovins laitiers (Fritz *et al.*, 2003) ou encore chez les ovins pour la résistance à la tremblante (Sidani *et al.*, 2010). Dans l'espèce porcine en revanche, force est de constater qu'à ce jour, les applications de la SAM ont été très peu nombreuses (Van der Steen *et al.*, 2005 ; Luther *et al.*, 2009 Schwob *et al.*, 2009) en dépit des centaines de QTL localisés (Hu *et al.*, 2010). Cette situation s'explique principalement par le fait que les marqueurs microsatellites, seuls marqueurs disponibles jusqu'à récemment, sont généralement en équilibre de liaison avec les QTL au niveau d'une population (obligeant à raisonner intra-famille ce qui complique les programmes de SAM), et éloignés des mutations causales (d'où une diminution rapide de l'efficacité de la sélection en raison des recombinaisons entre marqueurs et QTL).

Au cours des dernières années, les progrès du séquençage des génomes humains et des principales espèces domestiques ont permis de mettre en évidence un nouveau type de marqueurs répartis sur l'ensemble du génome et en nombre extrêmement élevé (plusieurs centaines de milliers à plusieurs millions de marqueurs sur l'ensemble d'un génome) : les marqueurs SNP, dont le polymorphisme consiste en la variation d'une seule base nucléotidique (« Single Nucleotide Polymorphism »). Grâce au développement de puces à SNP, il est aujourd'hui possible de génotyper très rapidement et pour un coût modéré (moins de 200€) un individu pour plusieurs dizaines de milliers de SNP. Avec une telle densité de marqueurs, les QTL sont obligatoirement en déséquilibre de liaison avec les SNP les plus proches, et le suivi de la transmission de l'ensemble du génome grâce aux marqueurs permet de suivre la transmission de tous les gènes intervenant sur un phénotype.

En 2001, Meuwissen *et al.* (2001) ont montré qu'en subdivisant le génome en un très grand nombre de segments déterminés par un ou plusieurs marqueurs, puis en estimant l'effet de chaque segment grâce à un échantillon d'individus phénotypés et génotypés pour ces marqueurs, on peut, en sommant les effets associés aux génotypes d'un individu pour tous ces segments, disposer d'une estimation très précise de sa valeur génétique, et ceci dès son plus jeune âge. C'est ce qu'on appelle l'évaluation génomique. Cette nouvelle méthodologie est utilisée dans les évaluations des populations bovines laitières de plusieurs pays depuis 2009, et la disponibilité d'une puce porcine de 60 000 SNP permet d'envisager sa mise en place dans l'espèce porcine.

La première partie de cette communication sera consacrée à la présentation du principe de l'évaluation génomique et des différentes méthodes d'estimation des segments chromosomiques. Nous présenterons ensuite les facteurs influençant l'efficacité de cette méthodologie d'évaluation. Enfin, les perspectives d'applications dans les populations porcines seront évoquées.

## 1. LA SÉLECTION GENOMIQUE : PRINCIPE

La mise en œuvre de la sélection génomique comporte deux étapes distinctes : (1) l'estimation des effets des segments chromosomiques dans une population de référence, et (2) la prédiction, à partir des effets précédemment estimés, des valeurs génomiques des animaux qui ne font pas partie de cette population de référence (des candidats à la sélection par exemple).

### 1.1. Constitution de la population de référence

Concrètement, l'étape d'estimation des effets chromosomiques nécessite la constitution d'un groupe d'individus - la « population de référence » (PR) - pour lesquels on dispose à la fois du génotype pour un très grand nombre de marqueurs et du phénotype pour les caractères pour lesquels on souhaite réaliser une sélection. La densité de marqueurs doit être suffisamment grande pour que tous les loci contribuant à la variabilité génétique d'un caractère présentent un déséquilibre de liaison suffisant avec les marqueurs qui leur sont proches (ou avec une combinaison de ces marqueurs). En pratique, plusieurs dizaines de milliers de marqueurs répartis sur l'ensemble du génome semblent nécessaires. Cette PR doit être représentative de la population que l'on souhaite sélectionner (dans l'idéal, l'ensemble des allèles des différents loci ségrégeant dans la population doivent y être représentés), et d'effectif suffisamment grand pour permettre une estimation précise de l'effet des fragments chromosomiques, car de celle-ci dépendra la précision de la prédiction des valeurs génomiques des candidats.

### 1.2. Estimation des effets des segments chromosomiques

Une fois la PR constituée, la principale difficulté consiste à estimer un très grand nombre d'effets de segments chromosomiques ( $M$  = plusieurs dizaines de milliers) à partir d'un nombre en général beaucoup plus faible ( $N$  = quelques centaines à quelques milliers) de phénotypes, ce qui revient à vouloir résoudre le système ci-dessous contenant plus d'inconnues que d'équations :

$$\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \end{bmatrix} = \begin{bmatrix} \mu + \sum_m X_{i,m} g_m + e_i \\ \vdots \\ \mu + \sum_m X_{i,m} g_m + e_i \\ \vdots \end{bmatrix}$$

où  $y_i$  est le phénotype de l'individu  $i$  de la PR,  $\mu$  est la moyenne générale,  $X_{i,m}$  est le génotype (ou haplotype) de l'individu  $i$  au  $m^{\text{ème}}$  segment chromosomique,  $g_m$  est l'effet du  $m^{\text{ème}}$  segment chromosomique sur le caractère, et  $e_i$  est une résiduelle non expliquée.

Pour cela, plusieurs méthodologies statistiques ont été proposées, présentant des coûts de calcul et des efficacités très variables. Ces méthodes peuvent s'appliquer à l'estimation des effets de marqueurs individuels ou d'haplotypes, selon que chaque segment de chromosome est déterminé par un seul marqueur ou par plusieurs marqueurs contigus. Dans le premier cas, on estimera un seul effet par marqueur, alors que dans le second cas, plusieurs effets devront être estimés pour chaque segment.

#### 1.2.1. Méthode des moindres carrés

La méthode des moindres carrés est fréquemment utilisée en statistique pour trouver les valeurs des paramètres décrivant

le mieux les données (minimisant l'écart entre performances observées et performances prédites à l'aide de ces paramètres). Cependant, elle ne peut s'appliquer que lorsque le nombre de paramètres à estimer est inférieur au nombre d'observations. Dans le cas de l'évaluation génomique, on est donc amené dans un premier temps à sélectionner un sous ensemble de  $M'$  segments chromosomiques parmi  $M$  (avec  $M' < N$ ) en estimant l'effet de chacun des  $M$  segments par régression dans un modèle uni-varié, et en ne conservant que les  $M'$  segments dont l'effet est le plus significatif.

Dans un second temps, les effets  $g_m$  de ces  $M'$  segments sont ré-estimés dans un modèle de régression multiple comme ci-dessous, les segments non conservés étant supposés avoir un effet nul :

$\hat{g} = X'X^{-1}X'y$ , où  $X$  est la matrice d'incidence reliant les  $M'$  génotypes (ou haplotypes) aux  $N$  phénotypes.

Cette méthode, très simple d'un point de vue calculatoire, pose cependant les problèmes du choix du seuil de significativité qui ne doit pas être trop laxiste afin de ne pas conserver plus de segments que de phénotypes dans la PR, et de la forte surestimation des effets estimés de certains segments. Pour ces raisons, cette méthode n'est en pratique pas recommandée.

#### 1.2.2. Méthode de la régression pénalisée

Afin de contourner ce problème de surestimation et de pré-sélection des segments à conserver, il est possible d'estimer les effets de tous les segments chromosomiques simultanément (ils sont alors considérés comme des effets aléatoires dans le modèle), en régressant vers 0 (en « pénalisant ») les estimations des effets. Ainsi, seuls les segments ayant un effet important auront un effet dans l'estimation de la valeur génomique, l'influence des segments de plus faible effet étant minimisée. Diverses méthodes de pénalisation ont été proposées pour déterminer la valeur de la pénalité à appliquer dans le modèle : « Ridge Regression » (Whittaker *et al.*, 2000 ; Xu, 2003), « LASSO » (Least Absolute Shrinkage and Selection Operator, Tibshirani, 1996), ...

Dans le cas particulier de la « Ridge Regression » où la pénalité appliquée est égale au rapport de la variance résiduelle sur la variance génétique, on retrouve le BLUP utilisé par Meuwissen *et al.* (2001), qui fait l'hypothèse que la variance de l'effet de chaque segment chromosomique est identique :

$$\hat{g} = \left( X'X + \frac{\sigma_e^2}{\sigma_g^2} I \right)^{-1} X'y$$

#### 1.2.3. Les méthodes bayésiennes

Les QTL et gènes mis en évidence dans les programmes expérimentaux de détection ou de validation de gènes candidats n'expliquent qu'une partie de la variabilité génétique des caractères, et il est communément admis que celle-ci résulte d'un petit nombre de loci d'effet important, de quelques loci d'effet moyen à faible, et d'un grand nombre de loci ayant un effet très faible ou nul (Hayes et Goddard, 2001).

L'intérêt des méthodes bayésiennes est de combiner de manière optimale cette information a priori avec les données réellement observées sur la PR (phénotypes et génotypes / haplotypes aux segments chromosomiques), afin d'obtenir de meilleures estimations des effets des segments chromosomiques que celles qui pourraient être obtenues en considérant uniquement les observations.

Cependant, l'obtention de la distribution conjointe a posteriori de tous les paramètres est trop complexe pour être obtenue

directement, et il est nécessaire de recourir à des techniques d'échantillonnages successifs et répétés de chaque paramètre, connaissant les observations et l'état des autres paramètres à cet instant. Le gain de précision dans l'estimation des effets des segments chromosomiques est donc obtenu au prix d'un accroissement considérable des temps de calcul.

Plusieurs méthodes bayésiennes d'estimation des effets des segments chromosomiques ont été développées, variant entre elles au niveau de l'information considérée a priori.

Chronologiquement, la première méthode développée, appelée « Bayes A », fait l'hypothèse qu'un grand nombre de QTL ont un effet faible et que quelques QTL ont un effet important (Meuwissen *et al.*, 2001). Ceci se traduit par l'utilisation d'une distribution de Chi2 inversée comme distribution a priori des variances des effets des segments chromosomiques. En réalité cependant, beaucoup des segments chromosomiques ne contiennent aucun QTL et n'ont donc strictement aucun effet sur le caractère. La distribution a priori de la méthode « Bayes A », qui ne permet pas à un nombre suffisamment grand de segments d'avoir un effet nul, n'est donc pas réellement adéquate. Meuwissen *et al.* (2001) ont alors proposé une autre méthode, nommée « Bayes B », avec laquelle une proportion  $\pi$  de segments n'a aucun effet, une distribution de Chi2 inversée étant à nouveau utilisée comme distribution a priori des variances des effets des  $(1-\pi)$  autres segments chromosomiques.

De nombreuses variantes à ces méthodes ont été proposées, afin de remédier aux imperfections des approches précédentes, mais globalement leur efficacité reste assez similaire dans la plupart des cas.

Bien que les méthodes bayésiennes requièrent une information a priori sur la distribution des effets des QTL, et que cette information est nécessairement biaisée car les expérimentations de recherche de QTL ne peuvent en général détecter que les QTL à effet fort, des études montrent leur efficacité et leur robustesse (Verbyla *et al.*, 2009). Cependant, le gain de précision dans l'estimation des valeurs génétiques par rapport aux méthodes de régression pénalisée (BLUP) peut cependant sembler modeste au regard de leur coût en temps de calcul. Hayes *et al.* (2008) rapportent en effet des gains de précision de seulement 1% à 7% dans les évaluations génétiques des populations bovines laitières australiennes, américaines, néo-zélandaises et hollandaises comparativement aux méthodes « BLUP ».

### 1.3. Estimation des valeurs génomiques

Une fois les effets des segments chromosomiques estimés, on peut calculer la valeur génomique de n'importe quel individu de la population pour lequel on dispose du génotype aux marqueurs considérés dans la PR, en sommant simplement les effets estimés des segments portés par l'individu :

$$\widehat{VG}_i = \sum_m X_{i,m} \hat{g}_m$$

où  $X_{i,m}$  est le génotype (ou l'haplotype) de l'individu  $i$  au segment chromosomique  $m$  et  $\hat{g}_m$  est l'effet estimé du segment chromosomique  $m$ .

### 1.4. Le BLUP génomique

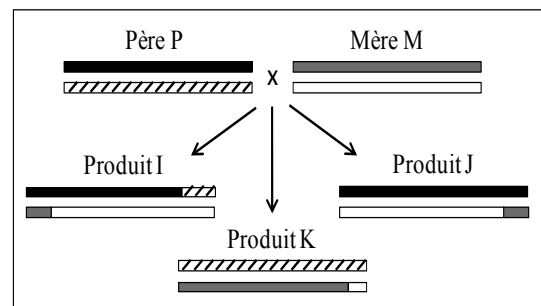
Dans les méthodes décrites précédemment, la valeur génomique d'un animal était définie comme la somme d'effets de segments chromosomiques. Une autre approche

équivalente (Van Raden, 2009) consiste à estimer directement la valeur génomique d'un individu (pour lequel on dispose du génotype) sans estimer les effets des segments chromosomiques, mais en résolvant un système analogue au système des équations du modèle mixte obtenu lorsque les performances sont décrites avec un modèle animal individuel. Dans ce système, la matrice de parenté entre individus ( $A$ ) classiquement utilisée est remplacée par une « matrice de parenté génomique » ( $G$ ) décrivant la parenté « réelle » entre les individus observée à travers leurs génotypes aux marqueurs :

$$[Z'Z + \alpha G^{-1}][\alpha] = [Z'y]$$

où  $\alpha$  est le vecteur des valeurs génomiques des individus génotypés,  $y$  est le vecteur des phénotypes des individus de la PR,  $Z$  est la matrice d'incidence reliant les performances aux valeurs génomiques, et  $\alpha$  est égal au rapport de la variance résiduelle sur la variance génétique. Cette méthode est connue sous le nom de « BLUP Génomique ».

Alors que la matrice de parenté  $A$  habituellement utilisée pour estimer les valeurs génétiques contient pour chaque paire d'individus évalués leur parenté espérée compte tenu de leur pedigree, les coefficients de parenté entre les paires d'individus contenus dans la matrice de parenté génomique  $G$  sont calculés en fonction du nombre d'allèles communs entre individus sur l'ensemble des marqueurs. Ainsi, si on considère 3 germains  $I$ ,  $J$  et  $K$  (issus des mêmes père et mère), les 3 coefficients de parenté entre  $I$  et  $J$ , entre  $I$  et  $K$  et entre  $J$  et  $K$  seront identiques dans la matrice  $A$  (par exemple 0,5), mais pourront différer dans la matrice  $G$ , par exemple si les individus  $I$  et  $J$  ont reçu plus d'allèles identiques de leurs parents que l'individu  $K$ , comme illustré sur la figure 1. Dans ce cas,  $I$  et  $J$  seront considérés plus apparentés entre eux que  $I$  et  $K$  et que  $J$  et  $K$ .



**Figure 1** - Génotype à un chromosome de 3 individus  $I$ ,  $J$  et  $K$  issus des mêmes parents  $P$  et  $M$ .  $I$  et  $J$  partagent une plus grande quantité de matériel génétique entre eux qu'avec  $K$ .

### 1.5. Intérêts de l'évaluation génomique par rapport aux méthodes de sélection classiques

L'évaluation génomique présente potentiellement de nombreux avantages par rapport aux méthodes d'évaluation classiques :

- Il est possible de disposer d'une estimation de la valeur génétique de tout individu de la population pour lequel on dispose d'un échantillon d'ADN, sans qu'il soit nécessaire de le phénotyper, ou de phénotyper un individu apparenté (collatéral ou descendant).
- Cette méthode peut s'appliquer à n'importe quel caractère, y compris aux caractères s'exprimant tardivement (longévité) et/ou dans un seul sexe (prolificité, fertilité, production de semence) ou dont le phénotypage

nécessite l'abattage de l'animal (qualité des tissus, ...), du moment que la constitution d'une PR pour le(s) caractère(s) considéré(s) est possible.

- Sous réserve de disposer d'une PR de taille suffisante et de marqueurs suffisamment nombreux pour assurer un déséquilibre de liaison adéquat avec tous les QTL, la précision d'estimation des valeurs génomiques peut être très élevée.
- La valeur génomique d'un individu peut être estimée dès sa naissance, ce qui permet de réduire fortement l'intervalle entre générations dans la population par rapport à un dispositif s'appuyant sur du testage sur descendance qui reporte nécessairement la mise en service d'un reproducteur.
- Dans le cas où le coût de génotypage d'un candidat est inférieur à celui de son testage (performances propres, contrôle de collatéraux ou testage sur descendance), l'évaluation génomique permet d'augmenter le nombre de candidats à la sélection et l'intensité de sélection des reproducteurs de renouvellement, sur la voie mâle comme sur la voie femelle.

Le progrès génétique réalisé annuellement dans une population en sélection étant proportionnel à la précision d'estimation des valeurs génétiques des candidats et à l'intensité de sélection pratiquée, et inversement proportionnel à l'intervalle entre générations (Falconer, 1981), la sélection génomique est donc potentiellement source d'un gain d'efficacité important dans les schémas de sélection.

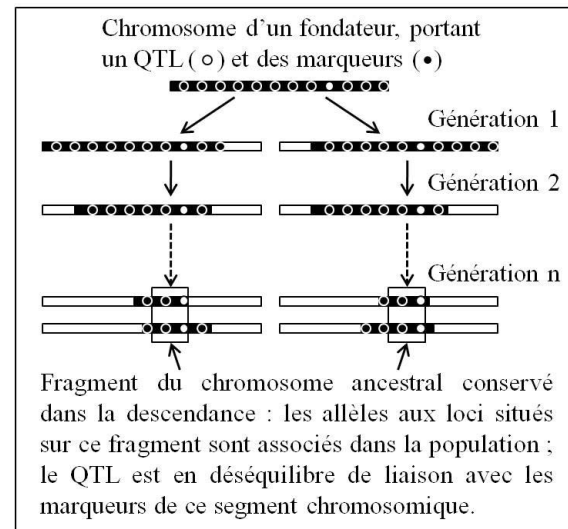
C'est par exemple le cas des schémas d'amélioration des 3 principales races bovines laitières françaises dans lesquelles le traditionnel testage sur descendance se voit aujourd'hui remplacé par une évaluation génomique (Boichard *et al.*, 2010). S'appuyant sur une population de testage de plus de 3600 taureaux testés sur descendance et sur une puce de 54 000 marqueurs SNP, cette nouvelle évaluation permet de disposer pour les jeunes taureaux d'index presque aussi précis que dans le dispositif classique, mais sans nécessiter un testage sur descendance long (5 ans) et coûteux (environ 40 000 euros par mâle testé). Les candidats sélectionnés peuvent donc être mis à la reproduction dès leur maturité sexuelle, soit une diminution d'1/3 de l'intervalle entre générations. De plus, les schémas n'étant plus limités par les capacités de testage sur descendance, un plus grand nombre de candidats mâles peuvent être évalués, augmentant ainsi l'intensité de sélection sur la voie mâle. Enfin, la possibilité de génotyper des génisses pour disposer de valeurs génomiques estimées avec la même précision que celles des candidats mâles est à l'étude, ce qui permettrait d'améliorer l'efficacité des schémas sur la voie femelle.

## 2. FACTEURS INFLUENÇANT L'EFFICACITE DE LA SELECTION GENOMIQUE

### 2.1. Niveau du déséquilibre de liaison, taille efficace de la population et densité de marqueurs

L'évaluation génomique ne peut fonctionner que si l'association entre les allèles aux marqueurs (ou les haplotypes) et les allèles aux QTL est suffisamment forte, afin que les effets estimés des segments chromosomiques reflètent bien les effets des QTL dans la population. La force de l'association entre un marqueur et un QTL, ou entre deux marqueurs, s'appelle le déséquilibre de liaison (DL). Au sein d'une population, les associations entre les allèles à plusieurs

loci d'une région chromosomique sont le résultat de la transmission sans recombinaison de cette région au travers des générations, comme illustré sur la figure 2.



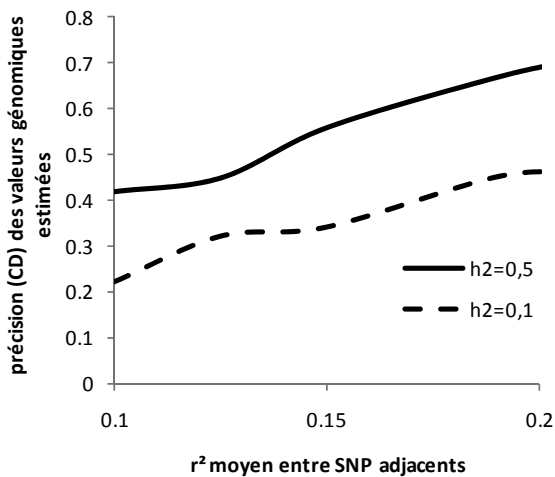
**Figure 2** - Origine du déséquilibre de liaison dans une population. Un fondateur transmet à ses produits des gamètes créés à partir de ses chromosomes, modelés par des événements de recombinaison a priori aléatoires.

Au fur et à mesure des générations et au gré des événements de recombinaisons, des segments de taille décroissante du chromosome ancestral sont conservés et sont portés par un grand nombre d'individus de la population ; les allèles aux loci présents sur ces segments ancestraux sont statistiquement associés, c'est à dire en déséquilibre de liaison.

Le génome des individus d'une population est donc constitué d'un ensemble de segments chromosomiques issus de reproducteurs influents et conservés au cours des générations, la longueur de ces segments étant d'autant plus faible que le nombre de générations séparant des individus considérés de leur ancêtre commun est grand.

En considérant que les différentes races porcines sont toutes issues d'une même population ancestrale ayant divergé un grand nombre de générations plus tôt, on peut donc trouver des associations entre loci communes aux différentes races, mais ces associations se limitent à des loci beaucoup plus proches qu'intra population.

Le déséquilibre de liaison entre deux loci peut être quantifié par divers indicateurs, par exemple au travers du paramètre  $r^2$  (Hill, 1981) qui varie entre 0 (indépendance entre les allèles aux deux loci) et 1 (les allèles aux deux loci sont totalement liés). Des études réalisées par simulations (Calus *et al.*, 2008, Meuwissen, 2009) ont montré que la précision d'estimation des valeurs génomiques est fortement dépendante du niveau de  $r^2$  entre marqueurs adjacents. En effet, la position des QTL étant en général inconnue, le  $r^2$  moyen entre marqueurs répartis sur l'ensemble du génome est un indicateur du niveau de  $r^2$  attendu entre un QTL et les marqueurs l'entourant. La figure 3, réalisée d'après les résultats de Calus *et al.* (2008), présente la précision d'estimation des valeurs génomiques d'individus non phénotypés (CD = Coefficient de Détermination = carré de la corrélation entre valeur génomique estimée et valeur génétique vraie de l'individu) pour deux caractères d'héritabilité ( $h^2$ ) 0,1 et 0,5, en fonction du  $r^2$  moyen entre marqueurs contigus.



**Figure 3** - Effet du niveau moyen de déséquilibre de liaison entre marqueurs contigus ( $r^2$ ) sur la précision d'estimation des valeurs génomiques. D'après Calus *et al.*, 2008.

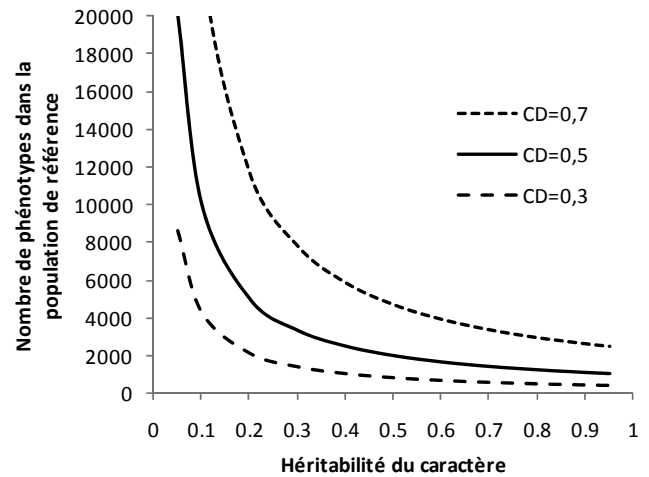
On constate que la précision est positivement liée au niveau de DL entre marqueurs. Ainsi, pour un caractère d'héritabilité 0,1, le CD est doublé lorsque le  $r^2$  moyen entre marqueurs contigus passe de 0,1 à 0,2.

Dans les populations de taille finie, ce qui est le cas des populations animales d'élevage, le  $r^2$  attendu entre 2 loci dépend principalement de 2 paramètres. Tout d'abord, la probabilité de recombinaison entre deux loci étant inversement liée à la distance les séparant, le  $r^2$  sera en moyenne d'autant plus élevé que les loci considérés sont proches. En conséquence, un moyen d'améliorer l'efficacité de l'évaluation génomique consiste à disposer d'un marquage du génome très dense, afin de maximiser la probabilité que les QTL soient en fort déséquilibre de liaison avec les marqueurs les entourant. Ainsi, dans les populations bovines laitières, environ 30 000 SNP répartis sur le génome sont nécessaires pour disposer en moyenne d'un  $r^2$  de 0,2 entre deux marqueurs contigus, soit un marqueur tous les 100kb environ. Outre l'augmentation de la densité de marqueurs, un autre moyen d'améliorer la précision de l'évaluation génomique consiste à définir les segments chromosomiques dont on cherche à estimer l'effet à l'aide de plusieurs marqueurs contigus plutôt qu'avec un seul marqueur. Les haplotypes ainsi déterminés présentent en effet un déséquilibre de liaison plus fort avec les allèles des QTL, ce qui permet de mieux estimer leur effet. Cependant, des simulations semblent indiquer que l'utilisation d'haplotypes ne présente un avantage que pour de faibles densités de SNP, les deux approches ayant des efficacités très proches pour des  $r^2$  moyens entre marqueurs adjacents supérieurs à 0,2 (Calus *et al.*, 2008).

Le second paramètre dont dépend le  $r^2$  est la taille efficace  $N_e$  de la population considérée. Dans une population de faible effectif génétique, un reproducteur va transmettre en espérance des fragments de son génome à une proportion importante de la population, ce qui résulte en une faible variété d'haplotypes. En conséquence, le  $r^2$  moyen entre loci sera plus élevé dans une population de faible effectif génétique que dans une population de taille efficace importante.

## 2.2. Taille de la population de référence et héritabilité des caractères

La figure 4, établie d'après Meuwissen (2009), présente la taille approximative de la PR nécessaire pour atteindre des précisions (CD) de 0,3, 0,5 ou 0,7 dans la prédiction des valeurs génomiques de candidats, en fonction de l'héritabilité du caractère considéré.



**Figure 4** - Taille de la population de référence requise pour atteindre une précision d'estimation (coefficient de détermination = CD) donnée des valeurs génomiques (0,3, 0,5 ou 0,7) pour des individus sans phénotype, en fonction de l'héritabilité du caractère considéré. La taille efficace de la population ( $N_e$ ) considérée est 100.

On constate, d'après cette figure, que la taille de la PR influence fortement la précision des valeurs génomiques prédites, un plus grand nombre de phénotypes permettant une meilleure estimation des effets des segments chromosomiques. Ainsi, pour un caractère d'héritabilité 0,2, environ 2000 individus sont nécessaires pour obtenir des CD de 0,3 alors qu'il en faut six fois plus pour disposer de CD de 0,7. Par ailleurs, la taille de la PR requise pour prédire des valeurs génomiques avec une précision souhaitée est inversement liée à l'héritabilité des caractères. Autrement dit, pour une taille de PR fixée, la précision de prédiction de la valeur génomique d'un individu pour un caractère peu héritable sera plus faible que pour un caractère plus héritable. Par exemple, d'après ces résultats, une PR de 5000 individus permet de prédire des valeurs génomiques avec des CD de 0,3 et 0,7 pour des caractères d'héritabilités 0,10 et 0,50, respectivement. Ceci peut s'expliquer par la plus faible part de variance phénotypique expliquée par les loci impliqués dans le déterminisme des caractères peu hértables, nécessitant plus d'observations pour pouvoir être mis en évidence.

Dans ce contexte, les schémas bovins laitiers sont dans une position avantageuse. En effet, leurs PR sont constituées de taureaux génotypés testés sur descendance dont le « phénotype » est une moyenne ajustée des performances de leurs filles. Tout se passe alors comme si l'héritabilité du caractère était égale à la précision de la valeur du taureau sur descendance. C'est par exemple le cas de la fertilité, extrêmement peu héritable ( $h^2=0,02$ ) mais dont l'héritabilité « apparente » dans la PR est proche de 0,60.

### 2.3. « Distance génétique » entre population de référence et population sélectionnée

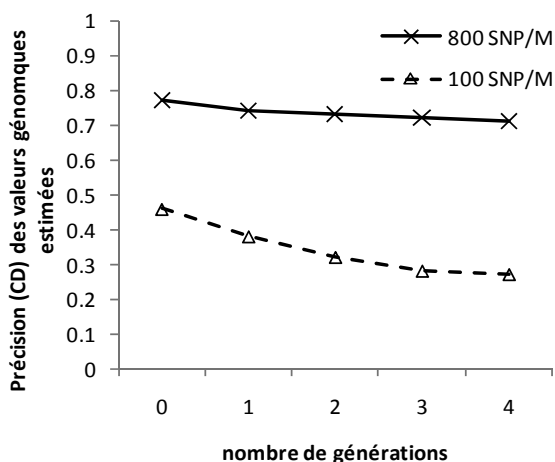
L'évaluation génomique repose sur l'hypothèse que les associations entre allèles aux marqueurs et allèles aux QTL sont les mêmes dans la PR et dans la population des candidats à la sélection. Cependant, plus ces deux populations sont génétiquement différentes, moins cette hypothèse est valide, en particulier si la densité de marqueurs est faible, c'est-à-dire si la distance entre marqueurs et QTL est grande.

En pratique, la sélection génomique est toujours mise en œuvre dans une population différente de la PR utilisée pour estimer les effets des fragments chromosomiques : les candidats à la sélection pour lesquels on estime des valeurs génomiques à l'aide de ces estimations peuvent être de la même race que la PR mais nés dans des générations plus récentes, où appartenir à des races ou des lignées différentes.

#### 2.3.1. La population de référence et les candidats à la sélection sont de la même race mais ne sont pas contemporains

Au cours des générations, des recombinaisons chromosomiques se produisent. Lorsqu'une recombinaison a lieu à proximité d'un QTL, l'association entre les allèles à ce QTL et les allèles aux marqueurs situés au-delà de la recombinaison se réduit, rendant caduques les effets des segments chromosomiques de cette région estimés dans la population d'origine.

La précision d'estimation de la valeur génomique des nouveaux candidats s'en trouve donc réduite. Ce phénomène est d'autant plus rapide que la densité de marqueurs est faible, comme illustré sur la figure 5, réalisée d'après les résultats de Solberg *et al.* (2009) obtenus par simulation.



**Figure 5** - Diminution de la précision d'estimation des valeurs génomiques estimées en fonction du nombre de générations séparant les individus constituant la population de référence des candidats à la reproduction, pour deux valeurs de densité de marqueurs sur le génome. Taille efficace  $N_e$  de la population considérée = 100.

Habier *et al.* (2007a) ont montré que la perte d'efficacité de l'évaluation génomique au cours des générations varie selon les méthodes d'estimation des effets des segments chromosomiques, la dégradation étant moins rapide avec les méthodes bayésiennes qu'avec celles reposant sur le BLUP. Dans tous les cas cependant, le maintien d'un niveau de précision d'estimation des valeurs génomiques des candidats nécessite de ré-estimer périodiquement les effets des segments chromosomiques à l'aide d'une PR renouvelée.

Diverses stratégies peuvent être envisagées, allant du renouvellement complet de la PR après un nombre fixé de générations au remplacement en continu des individus les plus anciens de la PR par des individus contemporains des candidats.

#### 2.3.2. La population de référence et les candidats à la sélection appartiennent à des races différentes

On peut envisager d'estimer les valeurs génomiques de candidats d'une race à partir des effets de segments chromosomiques estimés grâce à une PR d'une autre race. Cependant, la précision d'estimation des valeurs génomiques des candidats est dans ce cas en général assez faible.

Ainsi, Harris *et al.* (2008) ont obtenu des corrélations entre les valeurs génétiques vraies de taureaux Holstein et leurs valeurs génomiques estimées à partir d'effets de SNP obtenus sur une population PR Jersiaises (et inversement) variant entre -0,1 et 0,3.

Outre le fait que les associations entre marqueurs et QTL peuvent varier entre deux populations qui ont divergé depuis de nombreuses générations en raison des recombinaisons évoquées ci-dessus, d'autres facteurs peuvent également expliquer ces résultats. Tout d'abord, certains gènes polymorphes dans une population peuvent être fixés dans une autre. De plus, des QTL peuvent avoir des effets différents dans diverses populations.

La manière optimale de composer une PR destinée à estimer la valeur génomique d'individus appartenant à différentes races n'est pas encore clairement établie, mais une densité de marqueurs plus forte que pour une application intra-race, ainsi qu'une représentation de tous les types génétiques que l'on souhaite évaluer semblent requises (De Roos *et al.*, 2008 ; Harris *et al.*, 2008).

## 3. PERSPECTIVES D'UTILISATION DE LA SÉLECTION GÉNOMIQUE DANS LES POPULATIONS PORCINES

### 3.1. Forces et faiblesses des schémas d'amélioration génétiques porcins actuels

Chez le porc, les dispositifs d'amélioration génétique ont une structure pyramidale, composée schématiquement de trois étages : sélection, multiplication et production (pour plus de détails, voir par exemple Ollivier, 1998).

L'ensemble du travail d'amélioration génétique est réalisé à l'étage de sélection sur les populations de race pure et les lignées synthétiques ; l'objectif final est cependant d'améliorer les performances de reproduction et d'aptitudes maternelles des truies parentales croisées et les performances d'engraissement et de qualité de la carcasse des porcs charcutiers terminaux des élevages de production. Cette concentration de l'effort de sélection au sommet de la pyramide est fondée sur l'hypothèse que le progrès génétique qui y est réalisé se transmet jusqu'à l'étage de production.

A l'exception du schéma bavarois qui s'appuie en partie sur un testage sur descendance (Habier *et al.*, 2007b), les dispositifs d'évaluation génétique porcins actuels reposent sur un contrôle en élevage des performances des jeunes candidats mâles et femelles pour quelques caractères d'engraissement et de composition de la carcasse. Ces mesures sont parfois complétées par le contrôle en station puis l'abattage d'un nombre limité de collatéraux des candidats, permettant de disposer de performances pour des caractères trop complexes ou trop coûteux à mesurer en élevage (efficacité alimentaire)



ou enregistrés post mortem (caractères de qualité de la viande, composition de la carcasse, ...). Dans les populations maternelles, les performances de reproduction des truies (prolificité, aptitudes maternelles) sont également considérées. Les valeurs génétiques sont estimées grâce à la méthodologie du BLUP appliquée à un modèle animal. Chaque individu (jeune candidat mâle ou femelle, reproducteur) dispose ainsi d'une valeur génétique estimée pour tous les caractères de l'objectif de sélection de sa population, y compris pour les caractères pour lesquels il ne dispose pas de phénotype propre. En général, seules les performances des individus de la population sélectionnée sont considérées dans l'évaluation génétique, mais certains schémas prennent également en compte des performances d'apparentés croisés.

Les décisions de sélection ou de réforme des candidats mâles et femelles sont prises dès la fin de leur contrôle de performances en ferme (un peu avant 6 mois d'âge), et les intensités de sélection sont fortes. Ainsi, dans la population collective Large White type femelle française par exemple, 1 candidate sur 11 est conservée pour la reproduction en sélection, et 1 candidat sur 65 entre en CIA et sera utilisé en sélection ; dans la population collective Piétrain française, ces taux sont de 1/16 et 1/50, respectivement.

Les individus retenus sont mis à la reproduction dès 8 mois d'âge, et leur carrière de reproducteur en sélection est courte, environ 1 an pour les mâles et 2 ans pour les femelles.

Ce taux de renouvellement rapide s'explique d'une part par le progrès génétique réalisé dans les populations, qui fait que la supériorité génétique des reproducteurs en activité par rapport au niveau des meilleurs candidats s'amenuise rapidement, ce qui impose leur réforme et leur remplacement, et d'autre part par la volonté de limiter la descendance des verrats dans les élevages de sélection afin de préserver la variabilité génétique dans la population.

En conséquence directe de ces pratiques, les intervalles de génération sont très faibles, tant sur la voie mâle que sur la voie femelle, et seules des carrières encore plus courtes pourraient les réduire davantage.

*Faiblesses des dispositifs actuels.* Cette organisation des schémas d'amélioration porcins, malgré des intervalles de génération faibles et des intensités de sélection fortes tous deux propices à un progrès génétique important, présente cependant plusieurs faiblesses.

Tout d'abord, la faible durée de carrière des reproducteurs limite le nombre de leurs produits phénotypés ; la précision de leurs valeurs génétiques reste donc globalement faible. Ainsi, le CD moyen des verrats de CIA actifs dans les populations collectives françaises pour les caractères mesurés sur les individus contrôlés en station varie entre 0,18 pour le rendement technologique estimé (caractère le moins héritable,  $h^2 \approx 0,28$ ) et 0,50 pour le taux de muscle des pièces (caractère le plus héritable,  $h^2 \approx 0,75$ ), et se limite à 0,30 pour le nombre de porcelets nés vivants par portée ( $h^2 \approx 0,10$ ) dans les populations maternelles. Il en va de même pour le CD des valeurs génétiques des jeunes candidats en fin de contrôle de performances, en moyenne très modeste pour ces mêmes caractères (0,10, 0,40 et 0,16, respectivement). Cette précision limitée des valeurs génétiques estimées au moment de la sélection des reproducteurs de renouvellement pénalise le progrès génétique, les risques de sélectionner un individu qui se révélera finalement mauvais étant non négligeables. Des schémas fondés sur un testage sur descendance des verrats (comme le schéma bavarois) permettraient certes de

disposer de valeurs génétiques estimées plus précises et donc de diminuer les risques lors de la sélection des candidats, mais au prix cependant d'un allongement de l'intervalle entre générations et d'une réduction de l'intensité de sélection, venant réduire le progrès génétique réalisé.

Comme évoqué précédemment, le travail de sélection est réalisé dans les noyaux de race pure en faisant l'hypothèse que le progrès génétique réalisé se transmet (avec un retard d'une ou deux générations) aux individus croisés de l'étage de production. Cependant, les valeurs génétiques estimées en race pure peuvent parfois se révéler être des prédicteurs imparfaits de celles des descendants croisés élevés dans des conditions de production, pour deux principales raisons.

Tout d'abord, les individus de l'étage de sélection sont élevés dans des environnements extrêmement maîtrisés et d'excellent statut sanitaire, et ces conditions d'élevage favorables peuvent différer sensiblement des conditions de production. Or, les animaux les plus performants dans un milieu peuvent se révéler moyens voire médiocres dans un environnement différent, et inversement. De telles interactions génotype x milieu ont par exemple été mises en évidence chez le porc par Merks (1988). Par ailleurs, des effets de dominance et d'épistasie inexistantes dans deux races pures peuvent exister chez les individus issus de leur croisement. Ainsi, plusieurs auteurs rapportent des corrélations entre les performances d'individus de race pure d'élevages de sélection et celles de leurs descendants croisés comprises entre 0,3 et 1 selon les caractères et les populations considérées (Merks, 1988 ; Merks et Hanenberg, 1998 ; Lutaaya *et al.*, 2001). En conséquence, le progrès génétique effectivement réalisé peut être sensiblement plus faible à la base de la pyramide qu'à son sommet.

Il est possible d'augmenter le progrès génétique réalisé en élevage de production, en combinant les performances des individus de race pure à celles d'apparentés croisés lors de l'évaluation génétique des populations en sélection (Wei et Van der Steen, 1991). Plusieurs schémas de sélection porcins se sont engagés dans cette voie. Cependant, la remontée d'informations en provenance des élevages de production s'avère en général difficile faute de généalogies de qualité suffisante et en l'absence de contrôle de performances. De plus, 1 ou 2 générations (parfois plus dans le cas de croisements complexes) séparent les populations sélectionnées des porcs charcutiers et, sauf à vouloir organiser un testage sur descendance (qui augmenterait considérablement l'intervalle entre générations et pénaliserait le progrès génétique réalisé annuellement), le gain de précision obtenu pour les jeunes candidats grâce à leurs apparentés croisés serait minime.

### 3.2. Opportunités offertes par la sélection génomique dans l'espèce porcine

Une puce pangénomique de 60 000 marqueurs SNP est disponible chez le porc depuis 2009. Les résultats de Du *et al.* (2007) et de Huisman *et al.* (2010) dans diverses populations porcines semblent indiquer qu'avec une telle densité de marqueurs, le déséquilibre de liaison moyen ( $r^2$ ) entre SNP adjacents serait supérieur à 0,20, ce qui, au vu des exemples des schémas bovins laitiers français et étrangers, permet d'envisager la mise en place d'une sélection génomique dans l'espèce porcine. De plus, il est probable que des puces beaucoup plus denses seront disponibles dans les prochaines années, qui permettront de tirer parti d'associations encore

plus fortes entre marqueurs et QTL. Cependant, avant de se lancer dans une telle évolution des schémas d'amélioration génétique porcins, il est impératif de se demander en quoi la sélection génomique pourrait améliorer leur efficacité.

### 3.2.1. Intervalles entre générations

Les intervalles entre générations dans les dispositifs porcins actuels sont très faibles, de l'ordre de 2 ans dans les populations collectives françaises (Maignel *et al.*, 1998). Les verrats et les cochettes sont en effet mis à la reproduction peu après leur puberté, leurs carrières sont courtes, et un changement des modalités d'évaluation génétique n'aurait qu'un impact limité sur ces paramètres.

### 3.2.2. Intensités de sélection

Dans un dispositif fondé sur l'évaluation génomique, chaque individu génotypé peut se voir calculer une valeur génomique, même s'il n'a pas de performance propre et ne dispose d'aucun apparenté phénotypé.

Le choix des reproducteurs de renouvellement peut donc se faire en théorie parmi l'ensemble des porcelets nés, permettant de maximiser l'intensité de sélection et donc d'augmenter le progrès génétique réalisé.

En pratique, le nombre de candidats dans un tel dispositif, et par conséquent l'intensité de sélection effective, seraient fonction des éventuels surcoûts liés au génotypage que les sélectionneurs seraient prêts à supporter compte tenu des gains génétiques supplémentaires attendus.

Les intensités de sélection actuellement réalisées dans les schémas porcins sont cependant déjà élevées (le coût des mesures réalisées en ferme sur les candidats est faible, permettant de contrôler une forte proportion des porcelets nés), et le gain attendu en jouant sur ce paramètre pourrait donc ne pas être majeur.

### 3.2.3. Précision des valeurs génétiques estimées

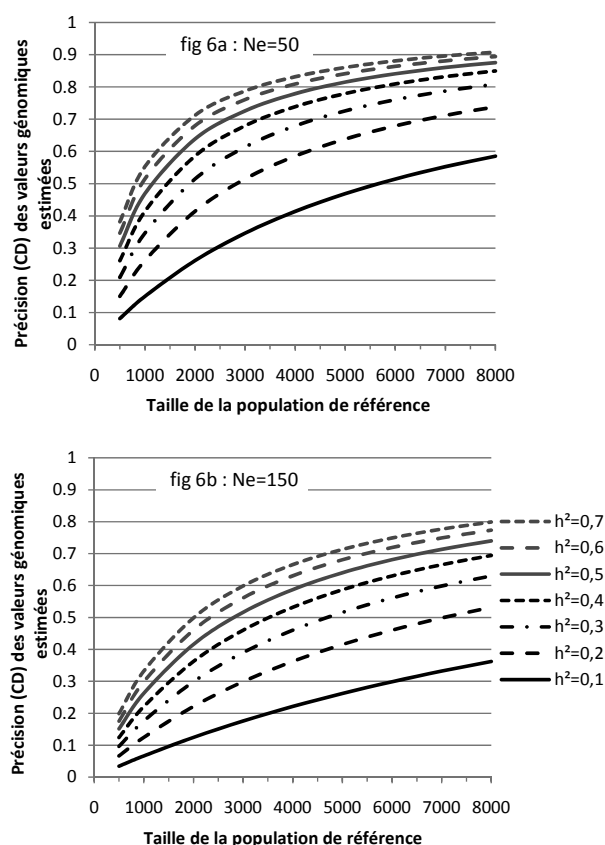
L'évaluation génomique peut en revanche avoir un impact beaucoup plus important sur l'efficacité des schémas par le biais d'une augmentation de la précision des valeurs génétiques estimées.

Comme indiqué précédemment, les VG des candidats pour les caractères de reproduction (dans les lignées femelles) et pour les caractères mesurés sur collatéraux (caractères de qualité, d'efficacité alimentaire) sont estimées avec des précisions souvent modestes (CD moyens compris entre 0,10 et 0,40 selon le caractère).

Le retour d'expérience des schémas de sélection génomique bovins laitiers et d'études réalisées par simulations (figure 6, établie d'après Meuwissen, 2009) montrent qu'il serait possible d'estimer pour de jeunes candidats des valeurs génomiques plus précises que leurs valeurs génétiques actuelles.

Les effectifs génétiques considérés sur la figure 6 ( $N_e=50$  et  $N_e=150$ ) correspondent approximativement à la plage de variation des tailles efficaces des lignées et races porcines en sélection (Maignel *et al.*, 1998 ; Habier *et al.*, 2009 ; Welsh *et al.*, 2009).

On constate que l'obtention de CD supérieurs à 0,3 pour des caractères faiblement héritables, ce qui est le cas de la qualité de la viande, de la prolificité et des aptitudes maternelles par exemple, requerrait une PR d'environ 2000 à 5000 individus selon la taille efficace de la population évaluée.



**Figure 6** - Précision (CD) des valeurs génomiques estimées en fonction de la taille de la population de référence et de l'héritabilité ( $h^2$ ) du caractère considéré, pour une population de taille efficace  $N_e=50$  (figure 6a) ou 150 (figure 6b). D'après Meuwissen, 2009.

### 3.2.4. Sélection sur des caractères non mesurables en routine

L'évaluation génomique pourrait également permettre de sélectionner les populations pour des caractères nouveaux, trop complexes ou trop coûteux à mesurer en routine (y compris dans le cadre d'un contrôle sur collatéraux) ou nécessitant la mise en place de protocoles expérimentaux spécifiques : mesures comportementales, résistance immunitaire, composition des rejets, maturité des porcelets à la naissance, homogénéité des porcelets intra-portée, odeurs sexuelles, qualité des aplombs ... On pourrait par exemple imaginer d'enregistrer dans quelques élevages équipés d'automates de pesée les poids des porcelets à la naissance et au sevrage pour un échantillon de truies génotypées ; ces truies constitueraient une PR permettant d'estimer les valeurs génomiques des candidats à la sélection pour l'homogénéité du poids des porcelets intra-portée, la croissance des porcelets en période d'allaitement et la production laitière des truies, ... Les coûts de génotypage et de phénotypage liés à la constitution initiale et au renouvellement périodique de telles PR seraient cependant très élevés, et ce type d'approche se limiterait vraisemblablement à quelques caractères parmi les plus importants pour la filière.

### 3.2.5. Augmentation du progrès génétique en production

Comme évoqué précédemment, le fait de sélectionner des individus de race pure élevés dans un milieu très contrôlé dans le but d'améliorer les performances de leurs descendants

croisés en conditions de production constitue l'une des principales faiblesses des dispositifs d'amélioration génétique porcins. C'est probablement sur ce point que la sélection génomique pourrait présenter son plus grand intérêt chez le porc. Il serait en effet possible, en s'appuyant sur une PR constituée d'animaux croisés, idéalement phénotypés dans des conditions de milieu semblables à celles des élevages de production, d'estimer la valeur génomique de candidats de race pure pour des caractères exprimés par des animaux commerciaux (troues parentales ou porcs charcutiers).

Dekkers (2007) et Kinghorn *et al.* (2010) ont montré par simulation qu'une telle approche permettrait un progrès génétique plus important dans la population croisée qu'une sélection génomique s'appuyant sur une PR constituée d'animaux de race pure. Elle serait également plus efficace qu'une sélection classique (BLUP-modèle animal) combinant les informations des individus de race pure et de leurs apparentés croisés (Dekkers, 2007), sans nécessiter l'organisation d'un enregistrement en routine des généalogies et des performances des descendants de production et sans augmenter l'intervalle entre générations.

Deux modèles ont été proposés pour estimer les effets des segments chromosomiques à partir d'une PR constituée d'animaux croisés (Ibanz-Escriche *et al.*, 2009 ; Kinghorn *et al.*, 2010). Pour décrire ces modèles, considérons par exemple que les individus de la PR sont issus du croisement entre une truie parentale croisée ayant des parents des races pures A et B et un verrat terminal de race pure C, et que chaque segment chromosomique est défini par un seul marqueur biallélique.

Le premier modèle, le plus simple, fait l'hypothèse que l'effet du génotype à un marqueur sur la performance d'un animal croisé est identique quelles que soient les populations d'origine (A, B ou C) des allèles reçus par l'animal pour ce marqueur :

$$y_i = \mu + \sum_j X_{ij} \beta_j + e_i$$

, où  $y_i$  est la performance de l'individu croisé  $i$ ,  $\mu$  est la performance moyenne,  $X_{ij}$  est le nombre d'allèles 1 (0, 1 ou 2) au locus du  $j^{\text{ème}}$  marqueur de l'individu  $i$ ,  $\beta_j$  est l'effet de substitution de l'allèle 1 au locus du  $j^{\text{ème}}$  marqueur dans la PR, et  $e_i$  est le résidu pour l'individu  $i$ .

La valeur génomique estimée VGE d'un candidat  $q$  d'une quelconque des 3 races pures A, B ou C se calcule alors simplement comme la somme sur l'ensemble des marqueurs considérés du produit du nombre d'allèles 1 portés par l'individu  $q$  au locus  $j$  ( $X_{qj}$ ) par l'effet de substitution estimé

$$VGE_q = \sum_j X_{qj} \beta_j$$

pour cet allèle :

Le second modèle considère au contraire que l'effet du génotype à un marqueur peut être différent en fonction de la population d'origine des allèles au marqueur :

$$y_i = \mu + \sum_j (A_{ijk}^P \beta_{jk}^P + A_{ijl}^M \beta_{jl}^M) + e_i$$

, où  $y_i$  est la performance de l'individu croisé  $i$ ,  $\mu$  est la performance moyenne,  $A_{ijk}^P$  est l'allèle porté par l'individu au  $j^{\text{ème}}$  SNP provenant de la race  $k$  transmis par le père de l'individu (ici, l'allèle provient forcément de la race C puisque le père est de race pure), et  $\beta_{jk}^P$  est l'effet de substitution de l'allèle  $A_{ijk}^P$ . De même,  $A_{ijl}^M$  est l'allèle porté par l'individu au  $j^{\text{ème}}$  SNP

provenant de la race  $l$  transmis par la mère de l'individu (ici, l'allèle provient de la race A ou de la race B puisque la mère de l'individu est croisée), et  $\beta_{jl}^M$  est l'effet de substitution de

l'allèle  $A_{ijl}^M$ . Dans le cas d'une population croisée constituée d'individus (AxB)xC, on estimera donc pour chaque SNP

3 effets : un effet  $\beta_{jc}^P$  pour l'allèle transmis par les pères, et

deux effets  $\beta_{jA}^M$  et  $\beta_{jB}^M$  pour l'allèle transmis par les mères selon qu'il provient à l'origine de la population A ou B, respectivement.

La valeur génomique des candidats de chacune des 3 races pures constituant le croisement est ensuite calculée à partir des effets des marqueurs estimés pour la race correspondante ; par exemple, pour un individu  $q$  de la race pure A :

$$VGE_q = \sum_j X_{qjA} \beta_{jA}$$

La sélection génomique s'appuie sur le déséquilibre de liaison existant entre marqueurs et QTL. Cependant, les associations entre marqueurs et QTL, et par conséquent les effets des génotypes aux marqueurs, peuvent varier entre populations. Le modèle estimant les effets des marqueurs intra-race semble donc plus pertinent, en particulier si les populations de race pure considérées ont divergé depuis de nombreuses générations (par exemple, populations paternelles vs populations maternelles) ou si la densité de marqueurs est faible (augmentant la probabilité de recombinaisons entre marqueurs et QTL).

Cependant, le nombre de paramètres à estimer avec ce modèle est plus important que lorsqu'on estime un seul effet par SNP commun à toutes les races, et une PR de plus grande taille est donc nécessaire pour conserver une bonne précision d'estimation des effets des marqueurs (Ibanz-Escriche *et al.*, 2009). De plus, le modèle « intra-race » nécessite de connaître l'origine raciale de chaque allèle porté par les individus de la PR croisée, mais des erreurs d'attribution seraient en pratique inévitables, réduisant peut-être son efficacité par rapport à un modèle plus simple.

Ces deux modèles, appliqués à une PR croisée, permettraient d'estimer l'effet de substitution des allèles présents dans les races pures dans le fond génétique de la population croisée que l'on cherche à améliorer, en incluant implicitement dans ces estimations une partie des effets de dominance et d'épistasie s'exprimant en croisement (Ibanz-Escriche *et al.*, 2009 ; Kinghorn *et al.*, 2010). De plus, l'enregistrement des performances de la PR dans un environnement représentatif du milieu de production permettrait d'estimer pour les individus des races pures des valeurs génomiques pour les caractères exprimés en conditions commerciales, et de s'affranchir ainsi du problème d'éventuelles interactions génotype x milieu.

Le progrès génétique réalisé dans la population croisée à l'étage de production serait donc plus important grâce à un tel dispositif ; en revanche, le progrès réalisé à l'étage de sélection serait nécessairement plus faible qu'avec des objectifs de sélection définis en race pure.

Bien que ces modèles n'aient pour l'instant pas été testés sur de vraies populations mais aient uniquement fait l'objet d'études par simulation, la possibilité de sélectionner les

individus de races pures en se fondant sur des valeurs génomiques estimées à l'aide d'une PR croisée semble prometteuse pour les schémas porcins.

### 3.3. Mise en place d'une évaluation génomique chez le porc : de l'idéal à la réalité

Compte tenu de l'ensemble des éléments précédents, on pourrait imaginer faire évoluer les actuels dispositifs d'amélioration génétique porcins organisés autour d'un contrôle de performances des candidats et de collatéraux et d'évaluations génétiques BLUP-modèle animal vers des dispositifs fondés sur des évaluations génomiques.

Dans l'idéal, des PR constituées de porcs charcutiers croisés élevés dans des conditions de production et phénotypés pour l'ensemble des caractères d'intérêt de production et de qualité, permettraient d'établir les équations de prédiction de la valeur génomique des populations en sélection.

Ce dispositif serait complété dans les populations maternelles par une évaluation génomique pour la prolificité et les aptitudes maternelles, s'appuyant sur des PR constituées de truies parentales croisées d'élevages de production.

Cependant, la mise en place effective d'un tel scénario nécessite de résoudre un certain nombre de difficultés pratiques.

#### 3.3.1. Coût de la constitution des populations de référence et du génotypage des candidats

On ne dispose pour l'instant pas des éléments nécessaires au dimensionnement précis des PR à mettre en place en vue d'une sélection génomique chez le porc. Cependant, compte tenu des résultats de simulations, du retour d'expérience des schémas génomiques bovins laitiers, et du niveau présumé du déséquilibre de liaison entre les marqueurs de la puce porcine 60 K actuellement disponible, une PR de plusieurs milliers d'individus semble nécessaire pour atteindre des précisions satisfaisantes (c'est à dire au moins égales aux précisions obtenues dans les dispositifs actuels) pour une population sélectionnée dont la taille efficace est proche de 100, en particulier si on s'intéresse à des caractères peu héréditaires comme la qualité de la viande ou les aptitudes maternelles. Il est probable que les coûts de génotypage diminueront au cours des prochaines années, mais en considérant un prix approximatif de 150 € par individu, la constitution initiale d'une PR s'élèverait donc à plusieurs centaines de milliers d'euros, auxquels viendraient s'ajouter les coûts de phénotypage, évidemment dépendants de la liste des caractères considérés. De plus, une PR doit être renouvelée régulièrement afin de préserver la qualité de prédiction des valeurs génomiques des candidats (Solberg *et al.*, 2009). La vitesse de perte d'efficacité de la prédiction est inversement liée au niveau du déséquilibre de liaison et à la densité des marqueurs, et même si des références précises manquent encore pour l'espèce porcine, on peut néanmoins faire l'hypothèse de la validité d'une PR pendant 3 générations, ce qui impliquerait de renouveler la dépense initiale environ tous les 5 ans. De plus, ces coûts seraient approximativement doublés pour les populations maternelles, puisque les phénotypes pour les caractères de production et de reproduction ne pourraient pas être mesurés dans une même PR. Enfin, l'évaluation génomique d'éventuels nouveaux caractères dont le phénotypage serait incompatible avec des conditions d'élevage classiques pourrait nécessiter la constitution de PR spécifiques supplémentaires.

A ces charges viendraient s'ajouter celle du génotypage des candidats à la sélection, directement liées au nombre de candidats à évaluer, et donc à l'intensité de sélection et au progrès génétique réalisés. Des études (voir par exemple Druet et Georges, 2010) montrent que, sous certaines conditions (par exemple, génotypage d'une partie des reproducteurs de la population pour l'ensemble des marqueurs considérés dans la PR), il serait possible de ne génotyper les candidats que pour quelques milliers de marqueurs informatifs répartis de manière homogène sur le génome, et « d'imputer » avec un très faible taux d'erreur (quelques pourcents) leur génotype aux marqueurs manquants. Ces techniques d'imputation permettraient de diminuer sensiblement les coûts de génotypage des candidats, tout en préservant l'efficacité de l'évaluation génomique. Néanmoins, même sous l'hypothèse favorable d'un coût de 30 euros par individu pour une puce de 3 000 SNP, le génotypage des 9 200 candidats mâles et des 35 000 candidates femelles contrôlés en ferme en 2008 dans la population collective française Large White type femelle s'élèverait par exemple à plus de 1 300 K€ chaque année.

Les sommes à engager pour mettre en place une sélection génomique dans les dispositifs porcins sont donc importantes. Contrairement à la situation des schémas bovins laitiers pour lesquels la mise en place de la sélection génomique, en se substituant à un testage sur descendance très coûteux (40 000 euros par candidat mâle testé en France), a permis de réaliser des économies considérables couvrant les frais de génotypage, un mode de financement reste à trouver chez le porc. Le coût des mesures actuellement réalisées sur les candidats en ferme, qui se limitent en général à des pesées et des mesures par ultrasons, est faible, et un éventuel arrêt du contrôle en ferme permis par une évaluation génomique ne conduirait pas à des économies substantielles. Le contrôle des collatéraux est en revanche plus coûteux, nécessitant l'équipement des élevages ou des stations en automates d'alimentation et le suivi des animaux à l'abattoir pour les mesures de qualité, avec parfois une diminution de la valeur de la carcasse en cas de prélèvement de tissus. Cependant, on peut dans une première approche considérer que le phénotypage des individus des PR remplacerait celui des collatéraux, et là encore aucun report de financement vers des génotypes n'est à espérer.

Toutefois, si on tient compte du volume annuel de la production porcine (25 millions de porcs charcutiers ont été abattus en France en 2008), l'ensemble des coûts liés à la constitution des PR, au génotypage des candidats et à l'encadrement des dispositifs serait a priori faible comparé au profit supplémentaire réalisé à l'étape de production, grâce à un accroissement du progrès génétique permis par la sélection génomique. On pourrait en conséquence imaginer de financer l'augmentation du coût de la sélection par une hausse du prix de vente des reproducteurs et de la semence.

Les valeurs individuelles d'une cochette et d'un verrat sont certes faibles (entre 260 et 300 euros pour une cochette, environ 650 euros pour un verrat (IFIP, 2009)), mais le nombre de reproducteurs vendus annuellement est très élevé (environ 400 000 cochettes et 9 400 verrats ont été vendus en France en 2008, tous types génétiques confondus (IFIP, 2009)), si bien que cette augmentation serait a priori modérée.

Des études économiques devront cependant être conduites afin de quantifier de manière plus précise les sommes concernées et trouver le mode de financement le plus adapté à une évolution des dispositifs porcins.

### 3.3.2. Cas des petites populations

De nombreuses lignées et variétés de races en sélection comptent moins de 200 femelles reproductrices, et même si la taille de la PR nécessaire à l'estimation de valeurs génomiques avec une précision donnée est plus faible dans les populations de petite taille efficace ( $N_e$ ), la constitution de dispositifs suffisamment puissants au sein de chacune de ces populations ne sera pas possible. Il semble néanmoins envisageable dans de tels cas d'utiliser des équations de prédiction des valeurs génomiques établies dans une autre race. L'efficacité d'une telle solution nécessite cependant une densité de marqueurs beaucoup plus forte que lorsque la PR et la population sélectionnée sont du même type génétique, afin de disposer d'associations entre marqueurs et QTL valides dans l'ensemble des populations. De Roos *et al.* (2008) ont ainsi estimé à 300 000 le nombre de SNP nécessaires à une évaluation efficace de la population bovine laitière Jersiaise à l'aide d'effets de marqueurs estimés en Holstein. Il est possible que la puce porcine de 60 000 SNP actuellement sur le marché soit insuffisante pour de telles applications, mais des puces plus denses seront certainement bientôt disponibles.

Par ailleurs, des résultats suggèrent qu'une représentation, même faible, d'individus de la population à évaluer au sein de la PR permettrait d'améliorer sensiblement la précision d'évaluation (Harris *et al.*, 2008 ; de Roos *et al.*, 2008) ; les organisations de sélection détentrices de petites populations seront donc vraisemblablement amenées à collaborer afin de constituer des dispositifs multiraciaux.

### 3.3.3. Partage des ressources

Cette mise en commun de ressources ne se limitera vraisemblablement pas aux seules petites populations. En effet, compte tenu du nombre élevé de populations porcines sélectionnées (par exemple, en France, 5 grandes races, une dizaine de variétés de ces races et une quinzaine de lignées composites), de la variété des truies parentales et des types de porcs charcutiers produits, des capacités de phénotypage limitées pour certains caractères, ainsi que des budgets à engager, il semble peu probable de pouvoir constituer et entretenir des PR spécifiques pour un grand nombre de populations sélectionnées ou de croisements pratiqués. Une mise en commun des phénotypes et génotypes de populations génétiquement proches détenues par plusieurs Organisations de Sélection semble donc nécessaire, permettant d'augmenter la puissance des dispositifs tout en limitant les coûts engagés par chaque partenaire. Des collaborations internationales sont également possibles, à l'image du projet Eurogenomics destiné

à mettre en commun les phénotypes et génotypes de 16 000 taureaux Holstein de 6 pays européens (Lund *et al.*, 2010).

## CONCLUSIONS

La sélection génomique offre plusieurs pistes d'amélioration de l'efficacité des schémas d'amélioration génétique porcins, bien qu'il soit peu probable que sa mise en place entraîne des mutations aussi profondes que les changements auxquels nous assistons dans les schémas bovins laitiers, principalement avec l'arrêt du testage sur descendance.

Les perspectives les plus intéressantes semblent être la possibilité d'estimer plus précisément la valeur génétique des individus pour les caractères peu héréditaires et non observables sur les candidats (qualité de la viande, efficacité alimentaire, aptitudes maternelles) et de sélectionner plus efficacement les individus de race pure pour les caractères exprimés en croisement et dans des conditions de production.

Bien que cette méthodologie permette d'estimer la valeur génétique de candidats sans performance propre et sans performance d'apparenté, l'enregistrement de phénotypes pour les caractères que l'on souhaite améliorer reste tout de même primordial afin d'établir les équations de prédiction des valeurs génomiques. Les stations publiques porcines, bien que les conditions d'environnement y soient sensiblement différentes de celles des élevages commerciaux, seraient naturellement un outil de choix pour la constitution des PR. Elles permettent en effet de mesurer de manière précise les performances des individus pour un grand nombre de caractères, et la nécessité de leur maintien dans des schémas d'amélioration génétique s'appuyant sur la sélection génomique est donc primordiale. La création d'un réseau d'élevages de référence dont la fonction serait analogue à celles des stations pour les caractères de reproduction et d'aptitudes maternelles semble également nécessaire.

De nombreuses incertitudes demeurent cependant quant aux modalités de mise en œuvre de la sélection génomique dans les schémas porcins, qui devraient être levées par les prochaines avancées technologiques (puces plus denses et à moindre coût) et méthodologiques (meilleure prise en compte des effets génétiques non additifs, mélanges de populations). Des études par simulations devront également être conduites afin de déterminer la manière optimale de réorganiser les schémas d'amélioration génétique porcins autour de cette nouvelle méthodologie.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- Archibald A. L., Haley C. S., Brown J. F., Couperwhite S. et al., 1995. The PIGMaP consortium linkage map of the pig (*Sus scrofa*), *Mamm. Genome*, 6, 157-175.
- Boichard D., Guillaume F., Baur A., Croiseau P., Rossignol M.N., Boscher M.Y., Druet T., Genestout L., Eggen A., Journaux L., Ducrocq V., Fritz S., 2010. Genomic selection in french dairy cattle. *Proc 9th WCGALP*, Leipzig, Germany August 1-6, 2010. *Comm.* 716.
- Calus M.P.L., Meuwissen T.H.E., de Roos A.P.W., Veerkamp R.F., 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics*, 178, 553-561.
- Dekkers J.C.M., 2007. Marker assisted selection for commercial crossbred performance. *J. Anim. Sci.*, 85(9), 2104-2114.
- De Roos A.P.W., Hayes B.J., Spelman R., Goddard M.E., 2008. Linkage disequilibrium and persistence of phase in Holstein Friesian, Jersey and Angus cattle. *Genetics*, 179, 1503-1512.
- Druet T., Georges M., 2010. A Hidden Markov Model Combining Linkage and Linkage Disequilibrium Information for Haplotype Reconstruction and Quantitative Trait Locus Fine Mapping. *Genetics*, 184, 789-798.
- Du F.X., Clutter A.C., Lohuis M.M., 2007. Characterizing Linkage Disequilibrium in Pig Populations. *Int. J. Biol. Sci.*, 3(3), 166-178.
- Falconer, D. S. 1981. *Introduction to Quantitative Genetics*. 2nd ed., Longman Inc., New York, NY.
- Fritz S., Colleau J.J., Druet T., Boscher M.Y., Rossignol M.N., Malafosse A., Boichard D., 2003. Mise en place d'une sélection assistée par marqueurs dans les trios principales races bovines laitières françaises. *Renc. Rech. Ruminants*, 10, 53-56.

- Habier D., Fernando R.L., Dekkers J.C.M., 2007a. The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, 177, 2389-2397.
- Habier D., Götz K.U., Dempfle L., 2007b. Estimation of genetic parameters on test stations using purebred and crossbred progeny of sires of the Bavarian Piétrain. *Livest. Sci.*, 107, 142-151.
- Habier D., Götz K.U., Dempfle L., 2009. Breeding programme for Piétrain pigs in Bavaria with an estimation of genetic trends and effective population size. *Livest. Sci.*, 123, 187-192.
- Harris B.L., Johnson D.L., Spelman R.J., 2008. Genomic selection in New Zealand and the implications for national genetic evaluation. *Proc. Interbull Meeting*, Niagara Falls, Canada.
- Hayes B., Goddard M.E., 2001. The distribution of the effects of genes affecting quantitative traits in livestock. *Genet. Sel. Evol.*, 33, 209-229.
- Hayes B.J., Bowman P.J., Chamberlain A.J., Goddard M.E., 2008. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.*, 92, 433-443.
- Hill W.G., 1981. Estimation of effective population size from data on linkage disequilibrium. *Genet. Res.*, 38, 209-216.
- Hu Z.L., Park C.A., Fritz E.R., Reecy J.M., 2010. QTLdb: A Comprehensive Database Tool Building Bridges between Genotypes and Phenotypes. *Proc. 9th WCGALP*, Leipzig, Germany August 1-6, 2010. Comm. 17.
- Huisman A.E., Cherel P., van Haandel B., 2010. Linkage disequilibrium and signatures of selection on chromosome 1 in a commercial sire line and dam line. *Proc. 9th WCGALP*, Leipzig, Germany August 1-6, 2010. Comm. 840.
- Ibanz-Escriche N., Fernando R.L., Toosi A., Dekkers J.C.M., 2009. Genomic selection of purebreds for crossbred performance. *Genet. Sel. Evol.*, 41, 12-21.
- IFIP - Institut du Porc, 2009. Le porc par les chiffres. Ed. IFIP, 149, rue de Bercy, 75595 Paris cedex 12.
- Kinghorn B.P., Hickey J.M., van der Werf J.H.J., 2010. Reciprocal recurrent genomic selection for total genetic merit in crossbred individuals. *Proc. 9th WCGALP*, Leipzig, Germany August 1-6, 2010. Comm. 36.
- Lund M. S., de Roos A.P.W., de Vries A.G., Druet T., Ducrocq V., Fritz S., Guillaume F., Guldbbrandtsen B., Liu Z., Reents R., Schrooten C., Seefried M., Su G., 2010. Improving genomic prediction by EuroGenomics collaboration. *Proc. 9th WCGALP*, Leipzig, Germany, Comm. 0880.
- Lutaaya E., Misztal I., Mabry J.W., Short T., Timm H.H., Holzbauer R., 2001. Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model. *J. Anim. Sci.*, 79, 3002-3007.
- Luther H., Vögeli P., Hofer A., 2009. Increasing genetic E. coli F18 resistance in Swiss pigs. *Proc. 60th EAAP*, Barcelona, Spain, 24-27 August 2009. Comm. 18-03.
- Maignel L., Tribout T., Boichard D., Bidanel J.P., Guéblez R., 1998. Analyse de la variabilité génétique des races porcines Large White, Landrace Français et Piétrain, sur la base de l'information généalogique. *Journées Rech. Porcine en France*, 30, 109-116.
- Merks J.W.M. 1988. Genotype x environment interactions in pig breeding programs. Thèse de doctorat. Wageningen Agricultural University, Wageningen, The Netherlands, 137 pp.
- Merks J.W.M., Hanenberg E.H.A.T., 1998. Optimal selection strategy for crossbred performance in commercial pig breeding programs. *Proc. 6th WCGALP*, Armidale, Australia. 23, 575-579.
- Meuwissen T.H.E., 2009. Accuracy of breeding values of unrelated individuals predicted by dense SNP genotyping. *Genet. Sel. Evol.*, 41, 35.
- Meuwissen T.H.E., Hayes B.J., Goddard M.E., 2001. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, 157, 1819-1829.
- Ollivier L., 1998. Improvement. In M.F. Rothschild & A. Ruvinsky (Eds), *The genetics of the pig*, 511-540. CAB International, Wallingford, Oxon OX10 8DE, UK.
- Sidani C., Astruc J.-M., Baelden M., Barillet F., Bibé B. *et al.*, 2010. The French ovine scrapie plan: results and prospects. *Proc. 9th WCGALP*, Leipzig, Germany August 1-6, 2010. Comm. 412.
- Schwob S., Riquet J., Bellec T., Kernalguen L., Tribout T., Bidanel J.P., 2009. Mise en place d'un programme de sélection assistée par marqueurs dans la population sino-européenne Duochan. *Journées Rech. Porcine*, 41, 29-30.
- Solberg T.R., Sonesson A.K., Woolliams J.A., Odegard J., Meuwissen T.H.E., 2009. Persistence of accuracy of genome-wide breeding values over generations when including a polygenic effect. *Genet. Sel. Evol.*, 41, 53.
- Tibshirani R., 1996. Regression shrinkage and selection via the Lasso. *J. Royal Stat. Soc.*, 58, 267-288.
- Tribout T., Caritez J.C., Gogué J., Gruand J., Billon Y., Bouffaud M., Lagant H., Le Dividich J., Thomas F., Quesnel H., Guéblez R., Bidanel J.P., 2003. Estimation, par utilisation de semence congelée, du progrès génétique réalisé en France entre 1977 et 1998 dans la race porcine Large White : résultats pour quelques caractères de reproduction femelle. *Journées Rech. Porcine*, 35, 285-292.
- Tribout T., Caritez J.C., Gogué J., Gruand J., Bouffaud M., Billon Y., Péry C., Griffon H., Brenot S., Le Tiran M.H., Bussièrès F., Le Roy P., Bidanel J.P., 2004. Estimation, par utilisation de semence congelée, du progrès génétique réalisé en France entre 1977 et 1998 dans la race porcine Large White : résultats pour quelques caractères de production et de qualité des tissus gras et maigres. *Journées Rech. Porcine*, 36, 275-282.
- Van der Steen H.A.M., Prall G.F.W., Plastow G.S., 2005. Application of genomics to the pork industry. *J. Anim. Sci.*, 83(E. Suppl.), E1-E8.
- Van Raden P.M., 2009. Efficient methods to compute genomic predictions. *J. Dairy Sci.*, 91, 4414-4423.
- Verbyla K., Bowman P., Hayes B., Raadsma H., Khatkar M., Goddard M.E., 2009. Comparison of Bayesian methods for genomic selection using real dairy data. *Proc. 60th EAAP*, Barcelona, Spain, 24-27 August 2009. Comm. 28-2.
- Wei M., Van der Steen H.A.M., 1991. Comparison of reciprocal recurrent selection with pure-line selection systems in animal breeding (a review). *Anim. Breed. Abstr.*, 59, 281-298.
- Welsh C.S., Blackburn H.D., Schwab C., 2009. Population status of major U.S. swine breeds. *Proceedings of American Society of Animal Science Western Section*, June 16-18, 2009 Fort Collins, CO, 60, 42-45.
- Whittaker J.C., Thompson R., Denham M.C., 2000. Marker-assisted selection using ridge regression. *Genet. Res.*, 75, 249-252.
- Xu S., 2003. Estimating polygenic effects using markers of the entire genome. *Genetics*, 163, 789-801.

