

# L'optimisation des programmes de sélection

F. PHOCAS

INRA, UMR1313 Génétique Animale et Biologie Intégrative, F-78352 Jouy-en-Josas, France  
AgroParisTech, Génétique Animale et Biologie Intégrative, 16 rue Claude Bernard, F-75231 Paris, France  
Courriel : [florence.phocas@jouy.inra.fr](mailto:florence.phocas@jouy.inra.fr)

Longtemps réalisée sur le seul phénotype, la sélection des animaux d'élevage a profondément évolué depuis le milieu du 20<sup>ème</sup> siècle grâce au développement de la génétique quantitative. La valeur génétique des animaux d'élevage est prédite par des index BLUP ou génomique. Dans un programme de sélection, ces index sont utilisés pour classer les candidats, les meilleurs étant retenus comme reproducteurs.

Dans des contextes de production différents associés à des contraintes biologiques, économiques, sociales et environnementales variées, l'organisation des programmes de sélection a pris diverses formes selon les espèces animales et les pays d'élevage. Des programmes de sélection se sont mis en place pour permettre l'amélioration génétique des animaux pour la production de viande, de lait, d'œuf et de poisson autour d'objectifs, critères et outils de sélection permettant de rationaliser et fiabiliser le choix des animaux reproducteurs. En France, les recherches sur les méthodes de sélection appliquées aux animaux d'élevage font partie des missions qui ont été confiées à l'INRA par la loi sur l'élevage du 28 décembre 1966 et confirmées par la loi d'orientation agricole du 5 janvier 2006. Ces recherches sont menées par le département de Génétique Animale en interaction constante avec les organisations professionnelles en charge des programmes d'amélioration génétique (les organismes et entreprises de sélection) ou de leur suivi (l'Institut de l'Élevage, l'Institut de la Filière Porcine, le Syndicat des Sélectionneurs Avicoles et Aquacoles Français...). Le but est d'effectuer toutes les recherches nécessaires au développement de schémas de sélection français efficaces et adaptés aux besoins tant des éleveurs que de l'ensemble de la filière pour contribuer à la pérennité de l'élevage en France, en intégrant les attentes sociétales en termes de Développement Durable.

D'un point de vue technique, un programme de sélection est efficace s'il permet de réaliser et de cumuler au cours des générations un certain progrès génétique, de limiter la prise de risque lors des choix de sélection et de mainte-

nir un certain niveau de variabilité génétique dans la population. D'un point de vue économique, un schéma de sélection est efficace si le bénéfice collectif issu de l'amélioration génétique est supérieur au coût de mise en place du programme. Selon Harris et Newman (1994), la mise en place d'un programme de sélection efficace repose sur quatre étapes fondamentales qui constituent le plan de cette présentation :

- la détermination d'objectifs et de critères de sélection,
- la réalisation d'évaluations génétiques permettant un choix objectif des reproducteurs,
- l'organisation et l'optimisation des différentes actions de création du progrès génétique,
- la diffusion de ce progrès génétique dans la population *via* diverses stratégies d'utilisation des reproducteurs.

## 1 / Les objectifs et critères de sélection

A partir des années 60, la sélection animale s'est organisée dans des programmes d'amélioration génétique des populations d'élevage visant une productivité maximale. Depuis les années 90, ces programmes se réorientent progressivement pour faire face à la nécessaire maîtrise des coûts de production dans un contexte de mondialisation des échanges et aux exigences sociétales en matière de sécurité et qualité des produits, de préservation de l'environnement (limitation des rejets, biodiversité, entretien et aménagement de l'espace rural) et d'éthique de production (conditions d'élevage, bien-être animal, biotechnologies). La réorientation effec-

tuée implique d'accroître le nombre de caractères inclus dans l'objectif de sélection d'une population.

### 1.1 / Les objectifs de sélection

La définition des objectifs est une étape préalable à la mise en place de tout programme de sélection. Les objectifs de sélection correspondent aux caractères héréditaires que l'on souhaite améliorer dans la population. Afin de choisir les reproducteurs parmi les candidats à la sélection, les généticiens quantitatifs cherchent à résumer cet ensemble de caractères dans une fonction – dite «objectif de sélection» – qui donne une valeur de synthèse des reproducteurs ou des races.

**L'objectif de sélection d'une population est une fonction des valeurs génétiques «vraies» (mais inconnues) des caractères à améliorer pour accroître le profit de l'éleveur, de la filière ou de tout autre niveau d'organisation.** L'intérêt de définir une telle fonction est de procurer un moyen rationnel de sélectionner les animaux, c'est-à-dire de choisir ceux qui offrent le meilleur compromis entre plusieurs caractères d'intérêt économique.

Une autre difficulté à lever pour une définition pertinente des objectifs de sélection est inhérente à la multiplicité des acteurs impliqués dans les filières et donc à la diversité voire la divergence de leurs intérêts économiques. Cette complexité a amené à raisonner les premiers programmes de sélection en termes d'objectifs techniques plutôt qu'économiques. A partir des années 90, compte tenu de l'évolution des programmes et outils de sélection, une approche globale plus rationnelle et

économique des objectifs de sélection s'imposait dans les espèces ayant des programmes d'amélioration collective, principalement les ruminants (Phocas *et al* 1997, Colleau et Régaldo 2001) et les porcs (Ducos 1995). En revanche, chez les espèces avicoles en situation très concurrentielle entre quelques grands sélectionneurs internationaux, la sélection s'effectue toujours sur des objectifs techniques, le but étant de pouvoir garantir un certain niveau de performances sur chacune des composantes de l'objectif plutôt que de maximiser le bénéfice attendu de la sélection pour l'éleveur ou le citoyen.

Il faut en effet distinguer trois raisonnements de nature bien différente pour déterminer une fonction des caractères importants à améliorer :

*i*) Les **objectifs techniques** consistent à fixer des niveaux de réponse souhaités sur les caractères. Un tel choix d'objectif reflète généralement une connaissance insuffisante du bénéfice économique espéré de la sélection, liée soit à la difficulté d'établir une valeur marchande pour certains caractères (qualités de race, aptitudes fonctionnelles, comportement, cas des objectifs de la durabilité), soit à une grande diversité d'objectifs entre acteurs de la sélection et des filières. Le choix d'un objectif technique peut également traduire des phénomènes économiques non linéaires (existence d'un optimum économique déjà atteint ou d'un seuil de rejet de la part des utilisateurs de la population sélectionnée). En théorie, un objectif technique peut être remplacé par un objectif économique équivalent, qui donne les mêmes réponses à la sélection. Les pondérations obtenues par une telle démarche mettent en évidence certaines aberrations économiques liées à l'utilisation d'un objectif technique établi sans connaissance suffisante du contexte économique (Gibson et Kennedy 1990). Aussi, la détermination d'un objectif économique doit-elle être considérée comme primordiale dans tout programme de sélection.

*ii*) Les **objectifs économiques** sont élaborés à partir des recettes et coûts associés à l'amélioration génétique des caractères. Cette fonction peut être très complexe à établir, notamment s'il faut intégrer les intérêts des divers acteurs de la filière concernée ou appréhender l'évolution future du contexte technique, économique et social. Un raisonnement prospectif est pourtant nécessaire car l'impact de la sélection des reproducteurs s'opère au niveau de leur descendance, donc par exemple cinq ans après l'action de sélection dans l'espèce bovine. Une synthèse des diverses méthodes de calcul proposées

et appliquées en races bovines laitières a été faite par Groen *et al* (1997). La manière la plus simple et la plus fréquente d'exprimer l'objectif de sélection consiste en une fonction linéaire des valeurs génétiques des caractères d'intérêt économique. Les coefficients de l'objectif sont appelés «pondérations économiques». La pondération économique d'un caractère est définie comme le gain marginal réalisé par l'amélioration d'une unité physique du caractère.

*iii*) Les **objectifs «non marchands»**. Jusqu'à un passé récent, la démarche prônée par les généticiens pour rationaliser la construction de l'objectif de sélection repose sur la modélisation de systèmes de production dont l'efficacité est souvent considérée en termes strictement économiques. Cependant, certains caractères sans valeur économique quantifiable présentent un intérêt indéniable pour l'un des acteurs de la filière ou pour la société (morphologie, aplombs, comportement animal...). Par conséquent, la fonction «objectif de sélection» doit désormais intégrer diverses contraintes pour assurer conjointement viabilité économique (réduction des coûts de production), équité sociale (conditions de travail des éleveurs, qualité et sécurité alimentaire, bien-être animal) et durabilité écologique (limitation des rejets ou des intrants pharmaceutiques). Une telle approche nécessite l'acquisition de connaissances sur les mécanismes biologiques fins des aptitudes recherchées et des génomes impliqués, sur l'expression des aptitudes zootechniques en milieux avec contraintes, sur les caractères actuellement difficiles à mesurer, et sur les caractères non marchands requis par une valorisation propre au développement durable, en particulier à l'environnement. L'inclusion de ces derniers dépendra de l'intégration des notions de coût et recette écologiques dans notre société. Les concepts de développement durable ont été introduits en génétique animale dans l'article fondateur d'Olesen *et al* (2000). L'objectif de sélection est alors décrit comme une combinaison linéaire associant pondération économique et valeur non marchande pour chacune des aptitudes à améliorer. La partie non marchande peut être établie soit en raisonnant des gains génétiques désirés sur les caractères environnementaux (Kanis *et al* 2005) ou bien les réductions de gain acceptées sur les caractères de production (Nielsen *et al* 2005), soit par les méthodes d'évaluation contingente et les expériences de choix visant à estimer le consentement à payer du consommateur (Olesen *et al* 2006).

## 1.2 / Les critères et indices de sélection

Les critères de sélection correspondent aux caractères sur lesquels porte le choix des reproducteurs afin d'atteindre l'objectif. Ce sont les caractères mesurés pour lesquels sont calculés les index élémentaires de valeur génétique et dont une combinaison linéaire constitue le plus souvent l'index synthétique de sélection (Goddard 1983). Du fait des contraintes liées au coût ou à l'impossibilité de mesurer directement sur les candidats à la sélection certaines aptitudes à améliorer (mesures post mortem, production laitière des taureaux), les caractères inclus dans l'indice de sélection ne sont pas nécessairement les mêmes que ceux constituant l'objectif de sélection, mais ils doivent au moins en être de bons prédicteurs sur le plan génétique. Dans l'idéal, un bon prédicteur a une héritabilité plus élevée que le caractère objectif et peut être mesuré plus précocement dans la vie de l'animal afin de réduire l'intervalle de génération. Une illustration de cette distinction entre objectif et critère de sélection concerne l'efficacité alimentaire. Ce caractère est d'une grande importance économique pour la production de viande, mais il est difficilement mesurable sur un grand nombre d'animaux, car la mesure de la consommation alimentaire nécessite une conduite des animaux individualisée et donc des coûts conséquents. L'efficacité alimentaire (objectif de sélection) est donc rarement mesurée. Les mesures (critères de sélection) sont alors effectuées pour des caractères corrélés génétiquement à l'efficacité alimentaire, tels que la croissance et le développement musculaire.

**L'index (ou indice) de sélection est le critère sur lequel seront classés les animaux candidats à la reproduction. L'index de sélection synthétique est une fonction des index élémentaires, à savoir les valeurs génétiques «prédites» des caractères à sélectionner.** Sa détermination correspond donc à la première étape d'un programme de sélection.

Historiquement formalisée par Hazel (1943) en amélioration génétique des animaux, la théorie des index de sélection s'appuie sur le modèle génétique infinitésimal pour prédire, par régression linéaire multiple, les valeurs génétiques des animaux à partir de leur propre performance et/ou de performances d'animaux apparentés. Cette théorie suppose que les performances observées ( $y$ ) sont idéalement corrigées pour tous les effets de milieu. L'établissement d'un index de

sélection nécessite donc deux étapes préliminaires : l'estimation des effets de milieu que l'on peut identifier et enregistrer (effets fixés) et la correction des performances pour ces effets.

Si l'objectif économique de sélection s'écrit sous la forme  $H = \mathbf{a}'\mathbf{u}$ , où  $\mathbf{u}$  est le vecteur des valeurs génétiques à améliorer et  $\mathbf{a}$  le vecteur des pondérations économiques associées, l'indice de sélection optimal ( $I = \mathbf{b}'\mathbf{y}$ ) est construit de manière à maximiser la corrélation entre l'indice et l'objectif de sélection et donc maximise la réponse à la sélection pour l'objectif de sélection (Henderson 1963). Quelles que soient les distributions de  $\mathbf{y}$  et  $\mathbf{u}$ , la maximisation de la corrélation entre  $H$  et  $I$  aboutit aux relations suivantes :

$\text{cov}(H, I) = \text{var}(I)$  et  $E(H/I) = I$  avec  $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a}$ , où :

$\mathbf{P}$  est la matrice de variance-covariance phénotypique des performances  $\mathbf{y}$ ,

$\mathbf{G}$  la matrice de variance-covariance entre les composantes de  $H$  et de  $I$ ,

$\mathbf{b}$  le vecteur des pondérations des performances qui maximisent la corrélation entre  $H$  et  $I$ .

En outre, pour des distributions multivariées de  $\mathbf{y}$  et  $\mathbf{u}$ , l'indice de sélection maximise la probabilité de bon classement des candidats à la sélection et donc maximise la réponse à un cycle de sélection.

Les formules ci-dessus restent valables en remplaçant le vecteur  $\mathbf{y}$  par le vecteur  $\hat{\mathbf{u}}$ ,  $\mathbf{P}$  devenant alors la matrice de variance-covariance de  $\hat{\mathbf{u}}$ , dans le cas d'une évaluation basée sur la méthodologie BLUP (*Best Linear Unbiased Predictor*) explicitée au paragraphe 2.2 ci-dessous.

Par ailleurs, l'information apportée par des marqueurs moléculaires peut être facilement incluse par la théorie des index de sélection. Dans le cas d'une sélection sur un QTL et une composante polygénique correspondant au même caractère, Dekkers et van Arendonk (1998) et Meuwissen et Sonesson (2004) ont proposé une formulation de l'index de sélection permettant d'optimiser la sélection à plus ou moins long terme. Dans le cas d'une sélection génomique, c'est-à-dire grâce à des marqueurs répartis densément sur tout le génome, Goddard (2009) propose une formulation de l'index de sélection permettant de maximiser la réponse à long terme, les coefficients étant proportionnels à l'inverse de la racine carrée de  $p(1-p)$  où  $p$  est la fréquence de l'allèle favorable au marqueur.

## 2 / L'évaluation génétique et la sélection des reproducteurs

### 2.1 / Le contrôle des performances pour mesurer les critères de sélection

Les contrôles de performance peuvent soit s'opérer dans le milieu de production (l'élevage), soit dans des stations spécialisées où l'environnement est mieux contrôlé et les animaux directement comparables les uns aux autres. Le contrôle de performances peut concerner les candidats à la sélection et/ou des individus qui leur sont apparentés : classiquement on distingue le contrôle individuel (mesure sur le candidat à la sélection), le contrôle sur ascendance, sur descendance ou sur collatéraux (germains ou demi-germains). Toutefois, ces distinctions tendent à se perdre dans l'organisation actuelle des programmes de sélection dès lors que l'évaluation se fait à partir d'un BLUP-modèle animal de l'ensemble des performances et des parentés recueillies en élevages car à chaque évaluation l'ensemble des informations disponibles est pris en compte pour prédire les valeurs génétiques des candidats à la sélection.

En termes de génétique, la sélection s'est pendant longtemps focalisée sur les seuls caractères de production et de standard de race. Cependant, depuis plus de 20 ans, les qualités d'élevage (aptitudes fonctionnelles) et les qualités des produits font de plus en plus l'objet d'une évaluation génétique, et sont donc introduits dans les contrôles de performances au côté des caractères de production, pour être améliorés dans les populations d'animaux d'élevage. Mais le choix des reproducteurs s'effectue le plus souvent en systèmes de conduite à l'optimum où l'environnement est très contrôlé, cela est particulièrement vrai chez les monogastriques, mais l'est beaucoup moins chez les ruminants en particulier laitiers (contrôle en ferme des filles des reproducteurs mâles). Or, le développement de systèmes de production à haute performance économique et environnementale demande de développer des ressources génétiques capables d'assurer une production relativement stable dans des environnements variés et/ou moins contrôlés. Mais à un même génotype peut correspondre des phénotypes variables dans des environnements différents, ce qui correspond à l'existence d'interactions génotype-environnement qui n'est généralement pas considérée dans l'évaluation génétique classique (Laloë 2011). Deux phénomènes biologiques différents peuvent être à l'origine de telles interactions. La

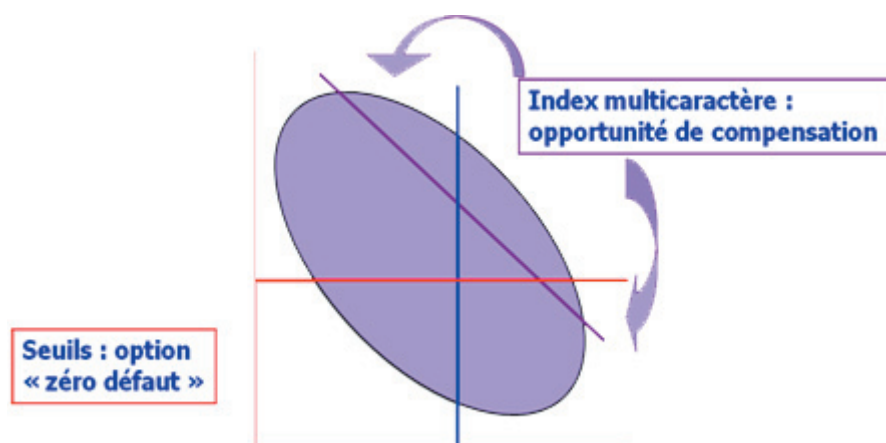
première source engendre seulement une hétérogénéité de variance entre environnements, ce qui cause des effets d'échelle et des réponses à la sélection plus ou moins fortes par rapport à la réponse attendue. La seconde source d'interaction entraîne quant à elle un reclassement des génotypes entre environnements. Il s'agit alors d'inclure dans l'évaluation, la sélection et l'utilisation des reproducteurs les interactions génotype-environnement dont l'importance ne peut que s'accroître avec la variabilité des milieux de production. Cette prise en compte peut être faite de diverses manières selon que l'on s'intéresse à une meilleure estimation de la valeur génétique globale ou bien à l'estimation d'une valeur génétique dans un environnement particulier.

L'arrivée de la sélection génomique ouvre des perspectives nouvelles en termes de développement d'évaluation des reproducteurs : associés aux performances, les profils génomiques des animaux permettent d'établir les relations phénotype-génotype sur une population d'animaux dite «de référence». Les «équations» qui en découlent peuvent ensuite être utilisées dans la prédiction des valeurs génétiques de candidats à la sélection pour lesquels on ne dispose que des génotypes sans phénotypes associés. Il devient donc possible d'envisager la valorisation génétique du recueil dans un nombre limité d'élevages ou de fermes expérimentales, de nouveaux phénotypes difficiles ou coûteux à mesurer sur un grand nombre d'animaux. La sélection génomique requiert toutefois de grandes populations de référence constituées de quelques milliers ou dizaines de milliers d'individus à la fois génotypés et phénotypés pour les caractères à évaluer. Cet objectif est inatteignable pour de nombreuses populations en sélection. Une mutualisation du recueil des performances et de leur valorisation entre populations est alors la solution à privilégier pour maximiser l'efficacité de la sélection génomique et diminuer le coût global de la sélection (Hayes *et al* 2009, de Roos *et al* 2009, Harris et Johnson 2010, Kizilkaya *et al* 2010, Toosi *et al* 2010).

### 2.2 / L'évaluation génétique et la sélection des reproducteurs

Les concepts de base de l'évaluation génétique des reproducteurs ont été présentés dans d'autres articles (voir notamment Mulsant *et al* 2011). Rappelons simplement ici que l'évaluation génétique consiste à obtenir un index pour chaque candidat à la sélection, c'est-à-dire une prédiction de sa valeur génétique additive (part transmissible à la descendance). Les index sont calculés à partir des performances recueillies sur l'animal à évaluer et ses

**Figure 1.** Comparaison d'une sélection par seuils indépendants et par index de synthèse.



apparentés. La réalisation d'une évaluation génétique suppose donc l'enregistrement des généalogies et des performances des animaux.

L'évaluation des animaux s'est tout d'abord faite sur des index calculés selon la théorie des index de sélection. Depuis la fin des années 80, la méthode statistique de référence utilisée pour prédire les valeurs génétiques est le *Best Linear Unbiased Prediction* ou BLUP (Henderson 1963) appliqué à un modèle animal. La prise en compte des relations de parenté entre tous les animaux permet de comparer les valeurs génétiques d'animaux dans l'espace et le temps dès lors qu'il existe des animaux apparentés qui ont été contrôlés dans différents environnements et à diverses périodes (Foulley *et al* 1992, Laloë 1993). La seule propriété statistique qui différencie le BLUP de l'index de sélection est l'élimination du biais lié à une mauvaise connaissance a priori des effets de milieu. Ceci est permis par l'estimation simultanée des effets de milieu (fixés) et des effets génétiques (aléatoires), grâce à un modèle statistique mixte d'analyse des performances.

Dans le cas d'un objectif multicaractère  $H = \mathbf{a}'\mathbf{u}$  dont les composantes sont corrélées, il est souhaitable que l'évaluation génétique soit multicaractère afin d'augmenter la précision de l'évaluation et d'éviter les biais dus à une sélection séquentielle des animaux (Ducrocq 1994). Une telle sélection s'opère par exemple sur les données concernant des poids à des âges successifs chez les ruminants allaitants (naissance, sevrage, un an). L'index de sélection BLUP sur lequel les reproducteurs sont choisis pour maximiser la réponse à la sélection sur  $H$  est alors  $I = \mathbf{a}'\hat{\mathbf{u}}$ , qui contient les valeurs génétiques prédites  $\hat{\mathbf{u}}$  de tous les caractères inclus dans l'objectif de sélection. Afin de maximiser la réponse à un cycle de

sélection directionnelle, il faut sélectionner les reproducteurs dont l'index dépasse un seuil unique commun à tous les candidats à la sélection (Henderson 1977, Goffinet et Elsen 1984). Toutefois, si le souhait n'est pas une réponse maximale sur un objectif global mais l'élimination systématique d'animaux avec des défauts génétiques rédhibitoires (par exemple difficulté de vêlage, agressivité), une sélection par seuils est nécessaire comme l'illustre la figure 1. L'index synthétique favorise en effet le maintien d'animaux extrêmes sur chacune des composantes de l'index de synthèse.

Au cours des dix dernières années, quelques programmes de Sélection Assistée par Marqueurs (SAM) ont été mis en place avec succès dans plusieurs espèces, par exemple chez les bovins laitiers pour des caractères de production et de fertilité (Fritz *et al* 2003) ou les porcs pour des caractères de qualité technologique des viandes (Schwob *et al* 2009). Au cours des dernières années, le développement de puces à ADN avec des marqueurs SNP (*Single Nucleotide Polymorphism*) répartis densément sur l'ensemble du génome permet de mettre en œuvre une nouvelle méthodologie d'évaluation, appelée «sélection génomique» (Meuwissen *et al* 2001), sous réserve de disposer de populations de référence où les animaux soient à la fois phénotypés et génotypés, les candidats à la sélection pouvant être quant à eux seulement génotypés. Cette innovation a en premier lieu révolutionné les programmes de sélection bovins laitiers car elle permet de sélectionner avec une bonne précision les jeunes mâles dès leur naissance, sans avoir besoin de les soumettre à un testage sur descendance long et coûteux (Boichard *et al* 2010).

### 3 / La modélisation des programmes de sélection

La modélisation et l'optimisation des programmes de sélection sont des outils d'aide à la décision pour les sélectionneurs. Ces outils permettent de déterminer le plan de sélection optimal (en termes de progrès génétique, de gain économique...) à mettre en place dans une situation particulière, parmi une très grande diversité de schémas possibles. Ils sont également susceptibles d'aider à repérer les pertes de charge dans la réalisation des plans de sélection, par comparaison des réponses à la sélection attendue et réalisée, et donc de concourir à la mise en œuvre d'indicateurs de pertes de charge des programmes de sélection complexes.

La diversité des programmes d'amélioration génétique des animaux d'élevage dépend non seulement de la variabilité des contraintes biologiques (niveau de fécondité) mais aussi des types de production (importance des effets d'hétérosis sur les caractères ayant le plus d'incidence économique) et objectifs de sélection (degré d'antagonisme génétique entre performances de production et de reproduction). Deux grandes méthodes sont classiquement distinguées : le croisement et la sélection. Il y a sur certains points un recouvrement : la sélection de lignées spécialisées pour le croisement (cas des monogastriques), la sélection pour l'appétitude au croisement et le croisement d'absorption qui s'apparente à un processus de sélection. L'objectif du croisement est de tirer parti de la variabilité génétique entre populations (races ou lignées) disponibles à un moment donné. Dans cet exposé, le croisement est considéré comme une stratégie d'utilisation des reproducteurs.

#### 3.1 / Organisation de la population sélectionnée

Il convient d'abord de décrire l'ensemble des règles d'évaluation, de sélection et d'accouplements des animaux. Ces trois étapes successives constituent un cycle de sélection et se répètent au cours du temps puisque les accouplements faits vont générer une nouvelle génération de descendants, et donc un retour à la première étape du cycle. La description des règles correspond à modéliser le programme de sélection en établissant d'une part un modèle conceptuel, puis un modèle mathématique pour prévoir l'évolution de l'état de la population soumise à sélection. Ces règles sont ensuite optimisées dans le but d'améliorer le niveau génétique de la population sur une fonction objectif.

La population est décrite par des paramètres qui la caractérisent sur les plans :

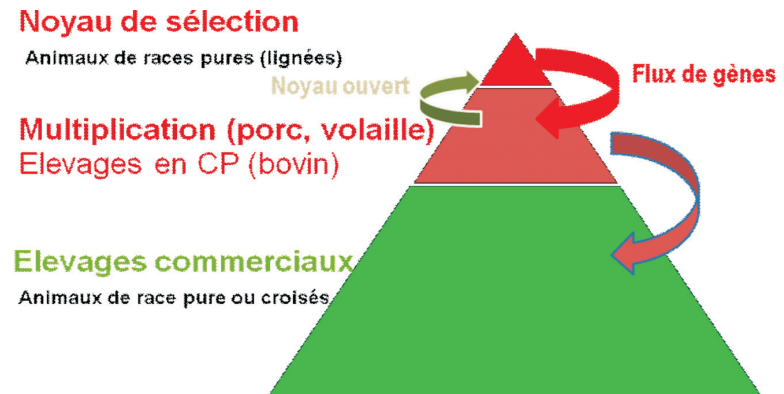
- démographique : taille de la population, fertilité, sex-ratio, mortalité, intervalle de génération ou âge à la reproduction, etc.,
- génétique : héritabilités, corrélations génétiques entre les caractères d'intérêt,
- zootechnique : moyennes et variances des caractères, taille des troupeaux, etc.,
- économique : coût d'entretien et de mesure des animaux, coût de calcul des indices, marge par unité de caractère, etc.

Les programmes de sélection des populations d'animaux de ferme s'organisent souvent au sein d'une sous-population d'effectif limité, appelée noyau de sélection (figure 2), sur laquelle les efforts de sélection sont concentrés. Le progrès génétique créé dans le noyau est diffusé dans le reste de la population (base de production) par l'intermédiaire principalement des mâles sélectionnés dans le noyau. Il existe un écart de niveau génétique entre les animaux nés au même moment dans le noyau de sélection et ceux nés dans la base de production (Bichard 1971). De même des écarts de niveaux génétiques se créent sous l'effet de la sélection dans des structures de noyau ouvert, c'est-à-dire dans des populations où une partie des animaux nés en base de production (avec un certain degré d'information, notamment l'enregistrement de l'état civil) sont mis à la reproduction dans le noyau.

Les taux de sélection des femelles sont généralement plus faibles que les taux de sélection des mâles, principalement en raison de leur plus faible prolificité (Ollivier 1974). En outre, les femelles et les mâles ne sont pas forcément évalués avec la même quantité d'information. Ces phénomènes conduisent à distinguer les mâles et les femelles dans la modélisation. Dans le noyau, des intensités de sélection différentes peuvent être appliquées aux reproducteurs d'un sexe donné selon l'utilisation prévue de leur descendance (production ou reproduction) ainsi que selon le sexe des descendants candidats à la sélection : chez les bovins, les pères et les mères à taureaux sont généralement plus sélectionnés que les pères et les mères à vaches.

Les populations d'animaux de ferme sont le plus souvent à générations chevauchantes : les candidats à la sélection à un moment donné sont d'âges différents, et ont donc en espérance des valeurs génétiques différentes. L'écart génétique entre deux classes de candidats nés deux années successives et

Figure 2. Organisation pyramidale de la sélection.



considérés au même âge (par exemple à la naissance) est le progrès génétique annuel (Bichard *et al* 1973). Afin de prédire correctement ce progrès, il faut décrire les différentes règles de sélection et de reproduction des animaux. A cette fin, une population animale est considérée comme un ensemble de groupes d'animaux soumis aux mêmes règles d'évaluation, sélection et reproduction (Elsen 1992). Ces groupes sont donc généralement définis en fonction de l'âge, du sexe, du mode de reproduction et du mode de sélection des animaux. En générations chevauchantes, la sélection des reproducteurs peut intervenir de deux manières différentes :

i) Chaque candidat est sélectionné une seule fois dans sa vie ; il est ensuite mis à la reproduction pendant un certain nombre d'années, identique en espérance pour tous les candidats sélectionnés en même temps que lui. La réforme intervient pour des causes non liées à sa valeur génétique. Une telle sélection concerne par exemple les animaux uniquement évalués en station ;

ii) Chaque année, les reproducteurs disponibles sont mis en compétition les uns par rapport aux autres. Un animal est donc sélectionné plusieurs fois au cours de sa vie, sur sa valeur génétique prédite. Cette valeur peut être supposée constante, connue avec une même précision d'une année sur l'autre, ou non : dans ce dernier cas, l'arrivée d'information en continu au cours de la carrière de l'animal est prise en compte (performances individuelles répétées dans le temps ou performances de générations successives de descendants).

Les modèles supposent souvent que les reproducteurs mâles et femelles sélectionnés sont accouplés au hasard. Une situation d'hétérogamie évidente concerne le croisement entre races et l'utilisation de lignées mâles et de lignées femelles spécialisées. Certains modèles décrivent des phénomènes d'homogamie entre les mâles et les femelles retenus pour la reproduction,

par l'intermédiaire de classes de parents de niveaux génétiques différents. Afin de prédire correctement le progrès génétique, il faut alors prendre en compte l'hétérogénéité de niveaux génétiques des descendants liée à ces accouplements non panmixtiques.

### 3.2 / Principes généraux du choix d'un programme de sélection

Afin d'optimiser la structure d'un programme de sélection, il faut comprendre les facteurs qui déterminent la prise de décision dans leur conception.

#### a) Valeurs des paramètres génétiques

Le premier paramètre clé est la variabilité génétique additive  $\sigma_a^2$  du caractère. Sans variabilité génétique, aucune évolution génétique de la population n'est possible.

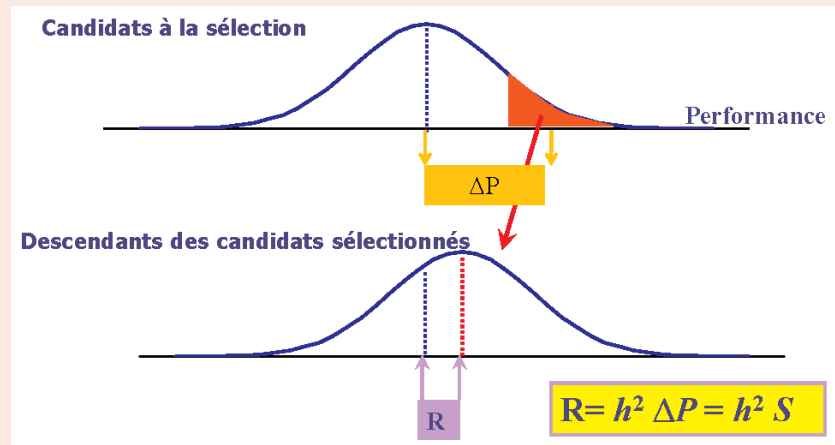
Un paramètre clé de la théorie de la génétique quantitative est l'héritabilité  $h^2$  du phénotype étudié, défini ici comme la performance corrigée des effets identifiés de milieu. Il s'agit de quantifier la part de variation des phénotypes qui est expliquée par la variation des valeurs génétiques additives :  $h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_e^2)$ . Ce paramètre, variant de 0 (absence de variabilité génétique additive) à 1 (absence de variabilité résiduelle), permet d'évaluer le progrès génétique que l'on peut espérer en sélectionnant le caractère. Classiquement, on parle de caractère à héritabilité faible en-dessous de 0,2 (très faible, si  $h^2 < 0,1$ ), à héritabilité moyenne entre 0,2 et 0,4 et à héritabilité forte au-dessus de 0,4 (très forte si  $h^2 > 0,6$ ). Les caractères à héritabilité moyenne à forte sont les caractères de production sur lesquels les premiers efforts de sélection se sont naturellement portés (croissance et morphologie, production laitière, rendement à l'abattage...). Les caractères à faible héritabilité concernent en particulier les capacités adaptatives et

### Encadré 1. Réponse $R$ à un cycle de sélection phénotypique

La réponse à un cycle de sélection par troncature sur les distributions normales des index indépendants et identiquement distribués (figure 3) est :  $R = i \rho_{IA} \sigma_a$

Où  $i = S/\sigma_l$  est l'intensité de sélection, avec  $S$  la différentielle de sélection observée, c'est-à-dire la valeur moyenne des indices des animaux retenus exprimée en écart à la moyenne des candidats à la sélection ;  $\sigma_l$  est l'écart type de l'indice de sélection ;  $\sigma_a$  est l'écart type génétique additif ;  $\rho_{IA}$  est la corrélation entre l'indice de sélection  $I$  et la valeur génétique  $A$ .

Pour une distribution normale des performances,  $i$  est égal à  $f(x)/p$ , où  $f(x)$  est la valeur de la fonction de densité d'une loi normale centrée réduite au point  $x$  qui est le seuil de troncature et où  $p$  est la proportion d'animaux sélectionnés.



reproductives des animaux, non par absence de variabilité génétique mais par absence d'identification suffisante des facteurs de variation dus au milieu conduisant à des modèles d'analyse à forte variabilité résiduelle.

L'héritabilité joue un rôle important dans la réponse  $R$  à une sélection phénotypique (encadré 1). Toutefois, Houle (1992) démontre que, dans un certain nombre de cas, il n'est pas pertinent d'utiliser  $h^2$  comme mesure de l'«évolabilité» d'un caractère, c'est-à-dire sa capacité à répondre à la sélection. Cela aboutit à de fausses conclusions quand  $h^2$  faible n'est pas associée à  $\sigma_a^2$  faible. Houle (1992) propose comme mesure de l'«évolabilité» d'un caractère, le coefficient de variation génétique du caractère :  $CV_a = 100 \times \sigma_a / \mu$ , où  $\mu$  est la moyenne phénotypique. Contrairement à  $h^2$ , ce paramètre est souvent plus important pour les caractères associés à la capacité adaptative de l'animal que pour les caractères de production.

La valeur des paramètres génétiques détermine la quantité d'information nécessaire à contrôler sur le candidat et ses apparentés pour garantir une précision suffisante de la sélection. A son tour, cette quantité d'information souhaitée impacte les choix en termes de contrôle des performances (performance individuelle, descendance, collatéraux) et d'intervalle de génération selon une stratégie qui est fonction de la bio-

logie de l'espèce et en particulier de sa capacité reproductive.

#### b) Capacité reproductive

La fécondité naturelle des espèces et des races ou lignées intra-espèces peut être un critère de choix pour la mise en place de stratégies de croisement ou de dissémination du progrès génétique. Ollivier (1974) a montré que le taux de renouvellement qui maximise le progrès génétique est d'autant plus élevé que l'espèce est prolifique, et que le renouvellement des mâles doit être plus rapide que celui des femelles, en particulier dans les espèces peu prolifiques.

La fécondité naturelle d'une espèce peut être modifiée par les techniques de reproduction artificielle (Insémination Artificielle (IA), superovulation et transfert d'embryon, clonage). Ceci peut concourir à accroître la capacité reproductive des espèces à faible fécondité et permettre des accroissements de progrès génétique par accroissement de la précision ou de l'intensité de sélection ou la réduction de l'intervalle de génération (Colleau *et al* 1998).

#### c) Infrastructures sociales

La répartition géographique et la dimension des élevages du noyau de sélection ainsi que leurs conduites d'élevage (degré d'utilisation de l'insémination artificielle, type de production) peuvent avoir des répercussions importantes sur l'organisation de la

sélection tant sur le plan des outils et méthodes de sélection que sur le plan politique. Ainsi, il peut y avoir nécessité à mutualiser les efforts de sélection au-delà des seuls moyens disponibles pour une entreprise privée, notamment pour les espèces à intervalle de génération long. Une organisation collective des programmes de sélection a des conséquences importantes en termes d'objectif de sélection et capacités des citoyens à influencer sur l'amélioration génétique des animaux d'élevage (Dockès *et al* 2011). La qualité de l'identification des animaux et du contrôle des performances (automatisation, centralisation des données) est également dépendante de l'ampleur du dispositif génétique (national ou centralisé sur quelques élevages).

L'état de la technologie engendre aussi de fortes répercussions en termes d'enregistrement de nouvelles données, de systèmes d'évaluation génétique fiable et de possibilités d'évolution des programmes de sélection. Ainsi, depuis 2009, la sélection génomique est une profonde innovation de rupture pour la sélection animale, en particulier pour les espèces à faible fécondité comme les bovins. Elle a un intérêt évident pour réduire l'intervalle de génération puisqu'elle permet une évaluation génétique précise des animaux dès leur naissance (Schaeffer 2006). Par ailleurs, le coût réduit du typage par rapport à la valeur du reproducteur bovin permet d'envisager une pression de sélection élevée tout en augmentant le nombre de reproducteurs sélectionnés afin d'éviter l'augmentation de consanguinité dans les populations (Colleau *et al* 2009). En revanche, chez les espèces avicole et porcine, l'intérêt économique de la sélection génomique est moins évident et demande une optimisation sérieuse car d'une part, il n'y a guère d'amélioration possible de l'intervalle de génération proche du minimum biologique, d'autre part le coût élevé du génotypage par rapport à la valeur du reproducteur est une contrainte forte (Tribout *et al* 2011).

#### d) Finalité du programme de sélection

Les facteurs qui jouent sur la réponse à la sélection dépendent de l'horizon d'intérêt défini, c'est-à-dire du nombre de cycles de sélection considérés. Selon Hill (1985), le court terme concerne les 5 premières générations de sélection, même si cela correspond à un long terme pour un éleveur de bovins (25 ans). Ce court terme génétique correspond à la période pendant laquelle l'impact de la consanguinité et des mutations est mineur sur l'évolution de la variabilité génétique et sur les progrès

génétiques annuels en l'absence de phénomènes de dépression de consanguinité ou de sélection d'intensité très forte (Hill 1985, Meuwissen 1989, Wray et Hill 1989). Selon l'objectif de sélection (nature et nombre de caractères concernés) et la finalité du programme (réponse à la sélection à court *versus* long terme), l'importance à accorder au maintien de la variabilité génétique dans l'optimisation du programme de sélection est variable. Le maintien de la variabilité génétique au-delà des 4-5 premières générations de sélection devient un problème majeur dès lors que l'on s'intéresse à des populations d'effectifs génétiques limités en raison des problèmes de dérive génique et de consanguinité.

Dans une population à générations chevauchantes, Dickerson et Hazel (1944) calculent le progrès génétique annuel comme étant :  $\Delta G = R/L$ , où **R** est la réponse à la sélection par génération (encadré 1) et **L** l'intervalle de génération, c'est-à-dire l'âge moyen des parents à la naissance de leurs descendants. Cette équation suppose une relation linéaire entre le niveau génétique de la population et la durée pendant laquelle elle est soumise à sélection. Ceci n'est pas vrai d'une année sur l'autre au démarrage d'une sélection en générations chevauchantes, cependant cette formule est une bonne approximation du progrès génétique annuel attendu en moyenne lors des 5 premières générations de sélection dans un noyau homogène et fermé (Elsen et Mocquot 1974, Hill 1974).

La réponse à un cycle de sélection augmente avec les paramètres de base suivants : **i**, l'intensité de sélection ; **p**, la précision de l'indice de sélection ; et **1/L**, l'inverse de l'intervalle de génération. Etant donné que ces paramètres ne sont pas indépendants, il faut choisir leur meilleure combinaison possible, c'est-à-dire celle qui maximise l'espérance de la réponse à la sélection. Ainsi, l'augmentation de la précision de la sélection associée à l'utilisation de mesures répétées dans le temps (lactations ou portées successives) ou celle des performances de la descendance est souvent contrebalancée par l'allongement de l'intervalle de génération. De même, à intervalle de génération constant, la précision augmente avec le nombre de collatéraux pris en compte dans l'évaluation, mais si l'abattage des collatéraux est nécessaire pour effectuer les mesures (qualités des carcasses et de la viande), l'intensité de sélection est alors réduite. A taille de population constante, l'intensité de sélection possible est d'autant plus forte que la durée d'utilisation des reproducteurs est importante, car le nombre de reproducteurs à renouveler chaque année est alors d'autant plus fai-

ble. Mais une plus longue durée d'utilisation des reproducteurs correspond à un intervalle de génération plus important et donc à un moindre rapport  $R/L$ .

Ce choix de la meilleure combinaison possible entre intensité de sélection, précision de l'indice de sélection et intervalle de génération dépend de la quantité et de la nature des informations utilisées pour évaluer les reproducteurs et donc de la méthode d'évaluation génétique. En générations chevauchantes et/ou en populations hétérogènes ou de faible taille, se posent en outre des problèmes de tri et d'utilisation comme reproducteurs des meilleurs animaux parmi un ensemble de groupes d'animaux de niveaux génétiques a priori différents ainsi que des problèmes de dérive génique et de consanguinité. Par ailleurs, bien que les méthodes de prédiction de l'efficacité des schémas de sélection intégrant l'information des marqueurs moléculaires aient été étudiées depuis les années 60 (Neimann-Sorensen et Robertson 1961, Ruane et Colleau 1995), l'arrivée des puces pangenomiques à haute densité renouvelle complètement les questions autour de la modélisation et l'optimisation des programmes de sélection dans toutes les espèces d'élevage (cf. § 4.4).

### 3.3 / Formalisation mathématique

Un modèle mathématique est une représentation simplifiée de la structure et du comportement d'un système donné par un jeu d'équations établies pour répondre à un objectif donné. Le système considéré ici est le programme de sélection dans le but de prédire l'ef-

ficacité de la sélection sur un critère donné (encadré 2).

Il faut déterminer les variables et les relations mathématiques entre ces variables qui permettent de calculer la valeur du critère d'efficacité choisi. Les principes qui vont être présentés ci-dessous sont décrits par Elsen (1992).

#### a) Variables et relations entre variables

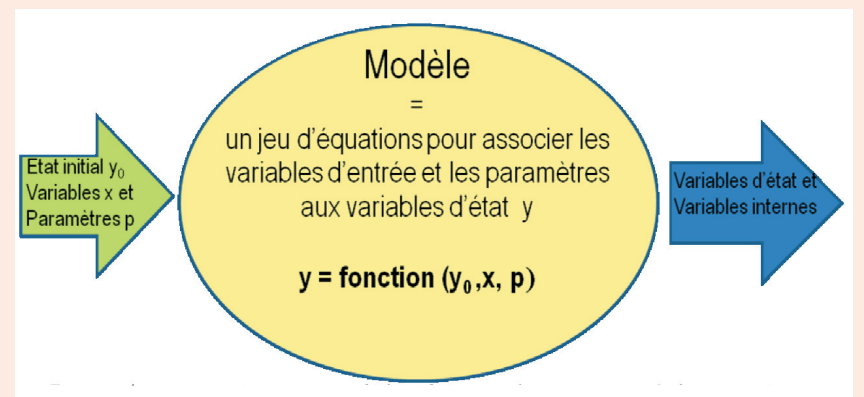
La population se modélise comme un ensemble de classes d'animaux soumis aux mêmes décisions de sélection et de reproduction. Ces classes sont donc fréquemment définies en fonction de l'âge, du sexe, du mode de sélection, du mode de reproduction, ainsi que de l'origine des animaux (noyau de sélection, base de production). Chacune de ces classes est caractérisée sur le plan génétique par des variables d'état : espérance de la valeur génétique moyenne, variance génétique, niveau de consanguinité. La nature des variables d'état choisies dépend du critère d'efficacité étudié. Le modèle doit rendre compte de l'évolution des valeurs prises par ces variables d'état. A cette fin, des variables élémentaires sont nécessaires.

Parmi les variables élémentaires, trois groupes peuvent être distingués :

- les paramètres qui caractérisent la population étudiée sur les plans démographique (mortalité, fertilité, sex-ratio, taille de la population...), génétique (héritabilité, corrélations), zootechnique (performances moyennes, variances phénotypiques) et économique (coût des mesures, du calcul, marge brute de l'élevage...),

#### Encadré 2. Concepts associés à une modélisation d'un programme de sélection.

Le modèle doit prédire l'état du système (par des variables d'état  $y$ ) en fonction d'un état initial donné ( $y_0$ ) et de variables élémentaires que sont des paramètres techniques ( $p$ ) et des variables de décision ( $x$ ) dont les valeurs sont fournies en entrée.



En sélection,  $y$  peut être par exemple la valeur génétique moyenne de la population ou de diverses sous-populations, sa variance génétique, le niveau de consanguinité.

- les variables de décision qui sont pour l'essentiel des effectifs d'animaux (contrôlés ou sélectionnés) et des durées d'utilisation des reproducteurs,

- les variables internes au modèle, qui sont des quantités communément utilisées en génétique quantitative (intervalle de génération, différentielle de sélection, coefficient de détermination d'un indice de sélection, coefficient de consanguinité...), et qui permettent, d'une part, de résumer l'information des paramètres et des variables de décision, et, d'autre part, de lier l'évolution des variables d'état au critère d'efficacité.

Les relations entre variables sont de deux types :

- les contraintes dues à des disponibilités en animaux limitées ; ces contraintes sont de trois natures : biologique (prolificité...), démographique (renouvellement de la population...), organisationnelle (capacité de contrôle en station...);

- les relations fonctionnelles qui décrivent comment variables de décision et paramètres sont liés pour former les variables internes et au-delà pour définir l'évolution des variables d'état.

#### b) Typologie des modèles de prédiction de l'efficacité du plan de sélection

Tous les modèles décrits ici reposent sur les hypothèses du modèle génétique infinitésimal. En outre, seule la sélection directionnelle est considérée dans les modèles décrits. Toutefois, d'autres types de sélection (canalisante, divergente) sont envisageables, mais concernent généralement les expériences de sélection plutôt que les populations d'animaux de ferme pour lesquelles le but est le plus souvent de sélectionner les meilleurs animaux sur les caractères d'intérêt économique.

**Les modèles déterministes** sont des modèles où le critère d'efficacité est complètement décrit et calculé à l'aide d'équations mathématiques reliant variables d'état et variables élémentaires. Le calcul s'opère sur les espérances et covariances des distributions des valeurs génétiques des animaux.

Les modèles de base ont déjà été introduits au § 3.2 pour donner l'espérance de la réponse à un cycle de sélection.

Il existe deux types de modèles déterministes de prédiction de la réponse à plusieurs cycles de sélection, cités dans l'ordre chronologique de leur apparition :

i) **les modèles asymptotiques** calculent directement le résultat de la sélection

(progrès génétique asymptotique, variance génétique asymptotique...) obtenu lorsqu'un régime d'évolution génétique constante de la population soumise à sélection peut être atteint à moyen terme (après 5 à 6 générations de sélection). Les modèles asymptotiques sont utilisés pour décrire l'impact au bout de quelques générations de sélection de phénomènes opérant dès le démarrage d'une sélection et amenant à un état d'équilibre génétique (stabilisation des écarts génétiques entre groupes d'animaux d'une population hétérogène, réduction des variances sous l'effet de la sélection).

Ainsi, Rendel et Robertson (1950) donnent l'expression du progrès génétique annuel pour une population où l'intervalle de génération moyen sur les 4 voies de transmission des gènes (père-fils, père-fille, mère-fils et mère-fille) est noté L :

$$\Delta G = [(i_1 \rho_1 + i_2 \rho_2 + i_3 \rho_3 + i_4 \rho_4) \sigma_a^2 / 4] / L.$$

Une telle population est dite fermée et homogène car tous les descendants d'un sexe donné sont nés d'une seule catégorie de pères et d'une seule catégorie de mères. Les seules sources d'hétérogénéité génétique de la population que ce modèle peut décrire, sont le sexe et l'âge des reproducteurs.

Au démarrage d'une sélection en générations chevauchantes, le progrès génétique annuel n'est pas constant d'une année à l'autre (Elsen et Mocquot 1974, Hill 1974). Toutefois, le progrès génétique annuel asymptotique est une bonne approximation du progrès génétique annuel réalisé en moyenne sur les 5 à 6 premières générations de sélection. La prédiction du progrès génétique asymptotique par la formule de Rendel et Robertson (1950) n'est valable que pour une population fermée et homogène. Sinon, les différences a priori de niveaux génétiques entre groupes de candidats à la sélection doivent être identifiées et intégrées dans le calcul du progrès génétique. Les premiers modèles intégrant de telles différences sont associés d'une part à la prise en compte des générations chevauchantes (Hill 1974, Elsen et Mocquot 1974), d'autre part à l'étude des populations à plusieurs étages (James 1977, Elsen 1993, Phocas *et al* 1995). Ces modèles asymptotiques supposent que les flux d'échanges de gènes entre étages sont constants au cours des générations de sélection.

ii) **les modèles dynamiques** décrivent l'évolution de la sélection à chaque unité de temps et permettent donc de rendre compte de tous les phénomènes de la sélection, y compris de phénomènes

de long terme qui n'aboutissent pas forcément à l'atteinte d'un équilibre génétique (consanguinité et mutations). Elsen et Mocquot (1974) et Hill (1974) ont établi indépendamment le premier modèle permettant de décrire la dynamique de l'évolution génétique d'une population. Ce modèle correspond à une relation matricielle de récurrence qui s'inspire des processus de Markov pour décrire le vecteur des niveaux génétiques de la population à chaque unité de temps  $A(t)$  en fonction du niveau génétique à l'instant précédent  $A(t-1)$  et des différentielles de sélection  $S(t)$  :

$$A(t) = P(t) A(t-1) + S(t)$$

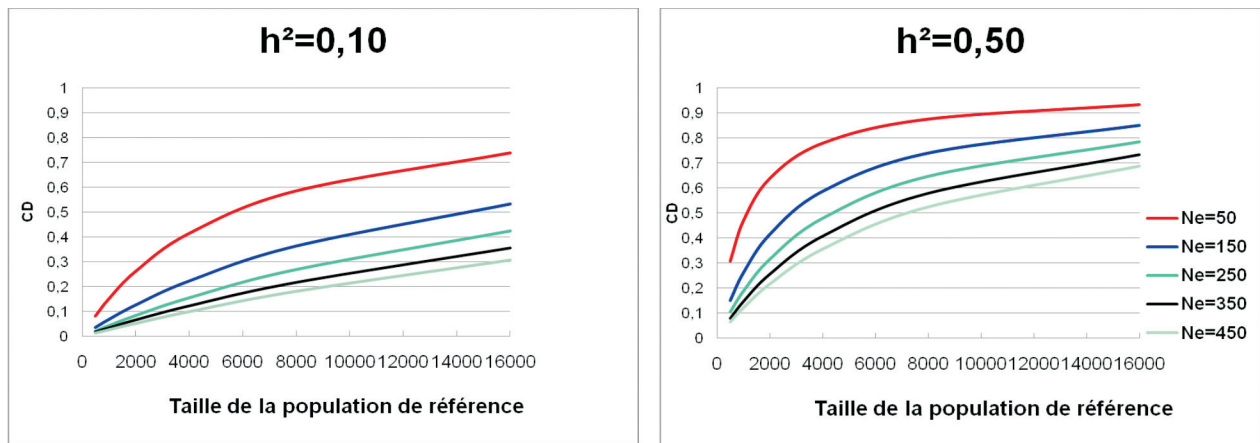
Chaque élément des vecteurs  $A(t)$  et  $S(t)$  correspond à une classe d'animaux définie en fonction de la population décrite.  $P(t)$  est une matrice de transition dont les éléments rendent compte du vieillissement et du renouvellement de la population. L'hypothèse est souvent faite que les caractéristiques démographiques de la population sont constantes, c'est-à-dire  $P(t) = P$ . Sous cette hypothèse, ce modèle est asymptotiquement équivalent à celui de Rendel et Robertson (1950).

iii) **Les modèles stochastiques** permettent de pouvoir considérer des phénomènes complexes (effets de milieu, effectif limité, consanguinité, variabilité de la réponse...) qui sont difficiles à décrire et/ou à calculer dans les modèles déterministes. Ces modèles sont de ce fait utilisés pour valider les hypothèses faites dans les modèles déterministes. L'inconvénient des modèles stochastiques est la lourdeur et la lenteur des calculs qu'ils génèrent, ainsi que l'absence de compréhension claire des relations entre phénomènes qu'ils traduisent. Dans les modèles stochastiques, les individus de la population sont distingués. Les distributions des valeurs génétiques vraies et des performances des animaux sont simulées par tirage aléatoire dans des distributions multinormales. Les données ainsi générées sont traitées comme des données réelles et utilisées dans l'évaluation des animaux. Le critère d'efficacité est alors calculé à partir de statistiques simples (moyenne, variance) des valeurs génétiques prédites et réelles des animaux sélectionnés.

### 3.4 / Modélisation de la précision de l'évaluation génétique

A paramètres génétiques fixés, intensité de sélection et intervalle de génération donnés, reste à prédire la précision (ou le carré de la précision noté CD) de la sélection en modélisant correctement la quantité d'informations disponibles pour les candidats à la sélection. Si cette



**Figure 3.** CD génomique sous les hypothèses du modèle génétique infinitésimal (Meuwissen 2009).

prédiction déterministe du CD est aisée dans le cadre de la théorie des index de sélection (Walsh et Lynch 2009), elle l'est moins dans le cas d'une sélection intégrant de plus en plus d'informations (effets de milieu fixés, diversité des apparentements, multiplicité des marqueurs) telles que la sélection par BLUP-Modèle Animal (BLUP-MA) ou la Sélection Génomique (SG).

Wray et Hill (1989) qualifient d'index pseudo-BLUP un index de sélection calculé à partir de la performance de l'animal à évaluer, des performances de ses collatéraux proches et des index BLUP de ses parents, car ils montrent que la précision d'un tel index est proche de celle d'une évaluation BLUP-MA. Pour une sélection faite en cumulant toutes les données et la généalogie depuis une population de base non sélectionnée et non consanguine, Harris et Johnson (1998) permettent de fournir une précision approchée d'une évaluation par BLUP-MA en utilisant la décomposition de la valeur génétique prédite d'un candidat à la sélection entre sa performance individuelle, les performances de ses descendants et les index de ses parents :

$$CD = \rho^2 = CD_{1+2+3} = \frac{CD_{1+2} + CD_3 - 2CD_1 + 2CD_8}{1 - CD_{1+2}CD_8}$$

où

$CD_1 = h^2$  est la précision d'une sélection massive

$CD_2 = n h^2 / [4 + (n-1) h^2]$  est la précision d'un contrôle sur  $n$  descendants

$CD_3 = (CD_{\text{père}} + CD_{\text{mère}}) / 4$  est la précision d'une sélection sur ascendance

$$CD_{1+2} = \frac{CD_1 + CD_2 - 2CD_1CD_2}{1 - CD_1CD_2}$$

Contrairement à un index BLUP sur ascendance, un index SG permet de quantifier l'importance de l'aléa de

méiose sur la valeur génétique du candidat sans avoir à contrôler sa descendance. La SG permet donc de fournir des valeurs génétiques prédites des candidats à la sélection avec une bonne précision dès la naissance. Sous l'hypothèse d'un modèle génétique infinitésimal et d'une population de référence constituée d'animaux non apparentés, Meuwissen (2009) donne la formule permettant de prédire la précision d'une évaluation génomique en fonction de l'héritabilité du caractère, de la taille de population de référence et de l'effectif génétique (cf. définition au § 3.5c). Selon l'objectif de précision souhaité pour la SG, il s'agit de définir la taille et de la structure de la population de référence à mettre en place selon l'héritabilité du caractère à améliorer et l'effectif génétique  $N_e$  de la population (figure 3).

### 3.5 / Phénomènes à considérer dans la prédiction de l'efficacité de la sélection

#### a) Espérance de la réponse à la sélection à court terme

Même à court terme, la prédiction de l'efficacité de sélection est compliquée si l'on souhaite s'écarter des hypothèses du modèle génétique infinitésimal en considérant par exemple des interactions génotype-environnement, des phénomènes de dominance et d'épistasie, des effets environnementaux corrélés (Walsh et Lynch 2009).

Dans le cadre restrictif du modèle génétique infinitésimal, la prédiction de l'efficacité de la sélection sur des caractères gouvernés par un grand nombre de gènes (fréquences alléliques constantes) et sans interaction de dominance (absence de dépression de consanguinité) doit intégrer les deux phénomènes suivants :

i) la réduction de l'intensité de sélection due à l'apparentement des candidats en population finie ;

ii) la réduction des variances génétiques due à la sélection.

La sélection crée un déséquilibre de liaison entre loci (Bulmer 1971). Ce phénomène est le seul effet direct de la sélection sur la réduction des variances sous l'hypothèse d'un modèle génétique infinitésimal. Pour une sélection directionnelle, le déséquilibre correspond à un déficit de combinaisons extrêmes des gènes, ce qui induit des covariances négatives entre les effets des gènes, et donc une réduction de la variance génétique. Du fait de la recombinaison des gènes lors de la méiose, la variance génétique se stabilise à une valeur d'équilibre (ou valeur asymptotique) après quelques générations de sélection. Si la sélection est interrompue, la variance génétique retrouve sa valeur initiale.

Des formules de calcul direct des variances et covariances génétiques asymptotiques ont été établies en fonction des seuls paramètres de la population de base dans le cas de générations discrètes et de populations fermées et homogènes. Ainsi Dekkers (1992) présente une formulation asymptotique de la variance génétique pour un schéma à quatre voies de sélection et une sélection sur indice BLUP-MA. Cette formulation repose sur le fait que la variance d'erreur de prédiction n'est pas affectée par une sélection sur indice BLUP-MA prenant en compte toute l'information (performances et généalogies) depuis la population de base non sélectionnée, non consanguine et non apparentée. Phocas et Colleau (1996) décrivent l'évolution possible de la corrélation génétique entre deux caractères dans le cas d'une sélection par BLUP.

Pour une structure de population complexe et/ou une situation de générations chevauchantes, les variances et covariances génétiques asymptotiques ne peuvent pas être exprimées analyti-

quement en fonction des paramètres de la population de base.

#### b) Variance de la réponse à la sélection en population de taille finie

Dans une population de taille finie, de nombreux phénomènes génèrent une variabilité de la réponse à la sélection. Les moyennes génétiques des individus sélectionnés vont dépendre du nombre de candidats à la sélection, de la précision des indices des candidats ainsi que de leur structure d'apparement qui est elle-même conditionnée par l'état antérieur de la population et par les fluctuations aléatoires qui s'y sont produites. Les variances génétiques des individus sélectionnés sont soumises aux mêmes fluctuations que leurs moyennes, auquel s'ajoute une fluctuation liée au déséquilibre de liaison entre loci créé par la sélection (Avery et Hill 1977). L'aléa d'échantillonnage de la méiose joue sur les moyennes et les variances génétiques à la naissance ainsi que sur les déséquilibres de liaison entre gènes.

Pour une distribution multinormale de variables identiquement et indépendamment distribuées, la variance de la différentielle de sélection  $S$  est inversement proportionnelle au nombre  $n$  de candidats à la sélection retenus (Finney 1956) :

$$\text{Var}(S) = n^{-1} [1 - i(i-x) + (1-p)(i-x)^2] \sigma_I^2$$

pour  $E(S) = i \sigma_I$

Sous l'hypothèse que les paramètres génétiques sont connus, cette formule peut être utilisée pour prédire la variance de la réponse à un cycle de sélection liée à l'échantillonnage dans une population d'index de sélection indépendants et de même espérance (Razungles 1977).

#### c) Efficacité de la sélection à long terme

Les recherches sur l'optimisation des schémas de sélection doivent aussi prendre en considération la nécessité de gérer la variabilité génétique conjointement au progrès génétique, afin d'assurer l'efficacité à long terme des programmes de sélection. Les effets décrits comme des effets à court terme de la sélection continuent à avoir un impact sur l'espérance et la variance de l'efficacité de la sélection à long terme.

En l'absence de gènes majeurs,  $\Delta G$  et  $\sigma_a^2$  sont approximativement constants sur les 5 premières générations de sélection. En l'absence de nouvelle source de variation (mutation ou migration),  $\sigma_a^2$  diminue *in fine*. Un plateau de sélection est atteint, reflétant la fixation de tous les allèles favorables. Si à la fois des gènes à effets importants et faibles

influencent l'expression de la performance, une réponse initiale rapide est obtenue par changement des fréquences aux loci majeurs, suivi par une longue période de réponse plus faible due aux modifications des fréquences alléliques aux loci ayant des effets faibles sur le caractère (Walsh et Lynch 2009). Robertson (1961) montre qu'en l'absence de mutation et migration il existe une intensité de sélection optimale telle que, à nombre fixe d'individus contrôlés, toute augmentation d'intensité de sélection se fait au détriment de la taille efficace (ou effectif génétique) de la population, ce qui réduit la réponse à long terme.

L'effectif génétique  $N_e$  correspond au nombre efficace de reproducteurs d'une population idéale qui présenterait le même taux d'accroissement de consanguinité ( $\Delta F$ ) par génération que celui observé dans la population réelle. L'effectif génétique  $N_e$  d'une population se calcule directement à partir du taux asymptotique d'accroissement de consanguinité par génération (Falconer et Mackay 1996) selon la formule :

$$N_e = \frac{1}{2\Delta F}$$

Dans les populations de petite taille ( $N_e < 100$ ), la dérive génétique a un rôle considérable sur l'évolution de la réponse à la sélection. A moyen et long terme, la variance génétique diminue dans les populations de taille finie du fait de l'accumulation de la consanguinité (Verrier *et al* 1991).

Wray *et al* (1990) proposent de calculer le taux de consanguinité asymptotique ( $\Delta F = F_t - F_{t-1} / 1 - F_{t-1}$ ) dans les populations sélectionnées en l'exprimant à partir des contributions de chacun des ancêtres au pool génétique de la population sélectionnée : ces contributions varient au démarrage de la sélection, mais finissent par se stabiliser. Wray *et al* (1994) développent cette idée pour une sélection sur index BLUP et expriment le taux de consanguinité en fonction de paramètres du plan de sélection tels que les intensités de sélection et les effectifs de reproducteurs. Ces auteurs vérifient par simulation aléatoire la validité de leur modèle déterministe (erreur inférieure à 1%).

La prédiction des réponses à la sélection ignore généralement les mutations car elles sont rares. Dans de nombreuses expériences où des plateaux de sélection ont été atteints, les essais pour augmenter la réponse à la sélection en induisant des mutations par rayons X n'ont généralement pas abouti (Hill 1985). Toutefois, aucun plateau de

sélection n'a été atteint dans d'autres expériences de très longue durée (Walsh et Lynch 2009). Cela indique la validité d'un modèle avec un grand nombre de gènes à petits effets par rapport à un modèle avec un faible nombre de gènes à effets importants. Les modèles avec mutation ont été développés par Lande (1976) et Hill (1982). Hill (1985) a montré que la mutation pouvait fournir une augmentation significative de la réponse à la sélection sur le long terme (40 générations) pour des populations de taille importante.

## 4 / L'optimisation des programmes de sélection

### 4.1 / Optimisation mathématique

Au plan mathématique, l'optimisation d'un programme de sélection correspond à la maximisation d'une fonction objectif sous contraintes. Le but est de déterminer les valeurs des paramètres (nombres d'animaux contrôlés, sélectionnés...) qui correspondent à une valeur maximale de la fonction objectif, étant donné que certaines contraintes de nature biologique, démographique ou économique sont à respecter (cf. § 3.2). Quelle que soit la fonction objectif retenue, le problème d'optimisation est *in fine* un problème de maximisation d'une fonction mathématique non linéaire sous contraintes qui peut se résoudre selon les diverses méthodes développées en mathématiques appliquées.

### 4.2 / Choix de la fonction objectif

La fonction objectif peut être un critère soit d'efficacité technique (progrès génétique annuel ou cumulé sur un laps de temps donné, taux de consanguinité, variabilité génétique) soit d'efficacité économique (associant coûts et recettes occasionnées pour la mise en œuvre d'un programme de sélection) dont l'intérêt dépend de l'échéance considérée (court, moyen ou long terme) ainsi que de la taille et la structure de la population soumise au plan de la sélection.

La mise en œuvre d'un programme de sélection ne peut se faire qu'après en avoir évalué l'efficacité économique. En effet, la sélection est une activité coûteuse et donc toute prise de décision demande au préalable une analyse des coûts et des recettes espérés. Les pionniers dans ce domaine ont été Poutous et Vissac (1962). L'évaluation économique nécessite d'actualiser les coûts et les recettes de la sélection. En effet, les coûts interviennent dès la mise en place

d'un schéma (construction de stations de contrôle, mesure des performances, acquisition et entretien des candidats, calcul des indices...) alors que les recettes sont différées dans le temps : elles ne sont perçues qu'au fur et à mesure que les générations expriment et accumulent le progrès génétique (supplément de valeur des animaux améliorés, nombre d'expressions du caractère par animal amélioré). Le principe de l'actualisation est qu'un gain réalisé aujourd'hui a plus de valeur que le même gain réalisé demain, ce qui se traduit par l'inclusion d'un taux d'actualisation dans le calcul des coûts et des recettes cumulés sur l'horizon de temps considéré.

Pour le décideur, l'intérêt d'un schéma croît avec la réponse attendue à la sélection, sous réserve d'une prise de risques limitée. Il faudrait donc intégrer dans la prédiction de l'efficacité d'un schéma le risque que la réponse réalisée soit nettement plus faible que la réponse attendue. Aussi, la variance de la réponse à la sélection (ou du progrès génétique) est un critère d'efficacité technique d'un schéma de sélection. En raison du comportement des acteurs de la sélection face au risque, l'intérêt économique et sociologique d'un programme de sélection n'est pas forcément une fonction linéaire de la réponse à la sélection attendue. Un critère d'efficacité d'un plan de sélection intégrant le risque que la réponse attendue ne soit pas obtenue peut être établi à partir d'une fonction d'utilité dont l'espérance fait intervenir l'espérance et la variance de la réponse à la sélection (Schneeberger *et al* 1981, Wyler *et al* 1994).

Un critère d'efficacité qui peut être important à moyen et long terme (au-delà de 5 générations) et/ou dans les populations de petite taille est le maintien d'une variance génétique suffisante afin de préserver de futures possibilités de sélection. Dans ce contexte, certains auteurs comparent l'efficacité relative de schémas à niveau de consanguinité fixe (Quinton *et al* 1992, Smith et Quinton 1993) ou à variance de la réponse fixée (Meuwissen et Woolliams 1994). L'un ou l'autre de ces paramètres peut indifféremment être regardé en raison de l'étroite relation entre taux de consanguinité et variance du progrès génétique.

### 4.3 / Quelques résultats importants en optimisation d'un programme de sélection

Les décisions de sélection optimale dépendent de l'horizon considéré. Des années 60 à la fin des années 80, l'optimisation des plans de sélection a concerné la maximisation de la réponse à la sélection à court terme, sans envisa-

ger de restriction sur la variabilité de cette réponse. Mais la mise en œuvre de méthodes très précises d'évaluation des reproducteurs (BLUP-MA puis SAM et SG) et des techniques modernes de reproduction (superovulation, transfert embryonnaire et sexage) rend très efficaces des schémas de sélection où la variance de la réponse et l'accumulation de la consanguinité peuvent être très élevées. Dans ce cadre, ainsi que dans un but de maximisation de la réponse à moyen et long terme, des optimisations de l'espérance de la réponse à la sélection avec restriction sur la variance de la réponse ou sur le taux de consanguinité ont été entreprises depuis les années 90.

#### a) Sélection intra ou entre-populations

Une stratégie de sélection souvent étudiée est l'intérêt d'une sélection en population (ou lignée) unique par rapport à une gestion et une sélection de plusieurs sous-populations (sous-lignées), issues de la subdivision d'une population. La mise en commun des ressources de plusieurs populations permet d'augmenter l'effectif disponible de candidats à la sélection et donc le taux de sélection. Cette stratégie peut être avantageuse pour le choix des animaux fondateurs d'un nouveau programme de sélection ou pour sélectionner les meilleurs mâles utilisés en IA en jouant sur les taux de reproduction élevés possibles en IA (Smith et Banos 1991). A taux de sélection global constant et pour un objectif de sélection commun à toutes les lignées, la gestion simultanée de plusieurs sous-lignées peut être une assurance contre les risques (anomalies génétiques, dérive génique défavorable). Toutefois, sous les hypothèses du modèle génétique infinitésimal, Smith et Quinton (1993) montrent que, à même niveau de consanguinité, la sélection en lignée unique engendre toujours une réponse à la sélection supérieure à celle obtenue en sélection sur plusieurs sous-lignées.

#### b) Taille et ouverture des noyaux de sélection

En générations discrètes, Shepherd et Kinghorn (1992) étudient la taille optimale d'un noyau ouvert pour une sélection sur index BLUP, avec prise en compte de l'effet Bulmer. Ils montrent que la taille optimale du noyau et le gain d'efficacité lié à l'ouverture sont indépendants de la précision de la sélection. Pour des schémas bovins ou ovins, la taille optimale du noyau serait entre 2 et 11% de la population totale, d'autant plus importante que l'intensité de sélection des mâles est forte. En autorisant à la fois les échanges de femelles mais aussi de mâles entre noyau et base et en

optimisant ces échanges chaque année, Shepherd et Kinghorn (1992) trouvent des gains de progrès génétiques asymptotiques variant entre 8 et 13% en faveur d'un noyau de sélection ouvert sur la base.

L'ouverture des noyaux de sélection permet aussi de réduire les taux de consanguinité (James 1977, 1978). Pour la structure optimale d'un noyau ouvert définie par James (1977), le taux de consanguinité est réduit d'un facteur 2 par rapport au taux obtenu dans un noyau fermé de même taille. James (1978) formalise l'expression de la taille efficace dans un noyau ouvert.

#### c) Maintien de la variabilité génétique

L'augmentation du progrès génétique attendu dans les schémas modernes de sélection se fait souvent aux dépens du risque que le progrès réalisé soit plus faible que dans les schémas plus classiques. Autrement dit, une plus grande espérance de la réponse à la sélection est généralement liée à une plus grande variance de cette même réponse.

Ainsi, l'efficacité de la sélection assistée par marqueurs a été comparée à celle d'une sélection BLUP en modèle polygénique à plus ou moins long terme. De nombreuses études ont montré qu'à court terme la SAM permet d'augmenter le progrès génétique en raison de la fixation plus rapide des allèles favorables. Cet avantage est important pour les caractères peu hérissables ou exprimés dans un seul sexe ou tardivement dans la vie de l'animal (Lande et Thompson 1990). Toutefois, sur le long terme la SAM est moins performante que la sélection polygénique (Ruane et Colleau 1995, Verrier 2001), à moins d'appliquer une SAM dynamique, c'est-à-dire où l'importance accordée aux marqueurs diminue avec la fréquence de l'allèle défavorable (Dekkers et van Arendonk 1998, Manfredi *et al* 1998).

Par rapport à d'autres méthodes moins précises comme la sélection sur le phénotype, la sélection sur index BLUP entraîne une augmentation de la probabilité de cosélection d'individus apparentés issus des meilleures familles parce qu'elle prend en compte toute l'information familiale disponible (Quinton *et al* 1992). La sélection de reproducteurs plus fortement apparentés augmente le niveau moyen de consanguinité dans la population et réduit la variance génétique de la population. Dans une grande population (~150 reproducteurs efficaces), il a été montré par simulation (Verrier *et al* 1993) que la sélection sur index BLUP est toujours

plus performante qu'une sélection phénotypique à court terme (30 générations). Cependant, dans des populations de petite taille en termes d'effectif génétique (~15 reproducteurs efficaces) dans lesquelles une intensité de sélection élevée est appliquée, la réduction de la variance génétique liée à la sélection sur index BLUP peut être telle que la sélection phénotypique donne de meilleurs résultats en considérant un horizon d'une dizaine de générations (Verrier *et al* 1993).

En évitant l'utilisation de l'information généalogique, la sélection génomique réduit la probabilité de co-sélection d'individus apparentés par rapport à une sélection classique sur BLUP polygénique. Toutefois, le progrès génétique étant accru par la SG en particulier quand l'intervalle de génération est drastiquement réduit par l'abandon du contrôle sur descendance, la variabilité génétique va rapidement décroître si aucune stratégie de maintien de limitation de la consanguinité n'est mise en œuvre (Nielsen *et al* 2011).

Dans des noyaux de sélection fermés et de petite taille, des stratégies optimisées pour le choix et l'accouplement des reproducteurs doivent donc être mises en place pour maintenir progrès génétique et variabilité génétique (Toro et Perez-Enciso 1990). Les stratégies d'accouplement limitant l'accumulation de la consanguinité au cours des générations consistent à maximiser l'espérance du progrès génétique en posant une contrainte sur le taux de consanguinité (Meuwissen 1997), ou vice-versa à minimiser le taux de consanguinité à progrès génétique désiré (Colleau *et al* 2004, 2009). Les principaux travaux en la matière ont utilisé des algorithmes d'optimisation par recuit simulé. Seuls Pong-Wong et Woolliams (2007) et Kinghorn (2011) proposent des approches par programmation semi-définie et algorithme génétique, respectivement.

#### 4.4 / Perspectives en optimisation des programmes de sélection

Beaucoup de travaux sont à mener pour permettre une optimisation des programmes de sélection incluant les outils de la génomique. Ces outils constituent une révolution dans la gestion des populations que cela soit pour accroître l'efficacité de la sélection intra-population ainsi qu'entre-popula-

tions ou permettre l'introgression de gènes d'une population vers une autre (Odegaard *et al* 2008).

Si l'on sait à l'heure actuelle que l'efficacité de la sélection génomique dépend de la taille de la population de référence, il est indispensable d'assurer le renouvellement des animaux génotypés avec performances si l'on veut garantir la pérennité du système (Muir 2007, Goddard 2009). En effet, l'efficacité de cette sélection dépend du degré d'association existant entre marqueurs et QTL, c'est-à-dire du déséquilibre de liaison entre loci des marqueurs et des QTL. Cette association diminue au fur et à mesure des générations sous l'effet des recombinaisons méiotiques. Calus *et al* (2008) ont ainsi montré par simulation que pour des caractères de faible héritabilité, la précision (CD) d'un index génomique peut être doublée lorsque le déséquilibre de liaison entre marqueurs contigus est doublé. La structure de la population de référence et ses modalités de renouvellement restent à préciser en fonction des contraintes biologiques de l'espèce, des caractères à évaluer et du niveau de précision souhaité pour les évaluations génomiques (Sonesson et Meuwissen 2009).

Les programmes de sélection doivent être profondément repensés afin de ne pas accroître l'évolution de la consanguinité dans les populations soumises à sélection génomique. La poursuite du raisonnement classique du sélectionneur, c'est-à-dire la forte diffusion d'un petit nombre de reproducteurs élites serait extrêmement dangereuse en termes de maintien d'une diversité génétique raisonnable au sein des populations. Il s'agira donc de favoriser un changement culturel profond dans les pratiques des sélectionneurs en augmentant en particulier le nombre et l'importance des caractères autres que les caractères de production dans les objectifs de sélection animale.

## Conclusion

Cette synthèse fait un état des lieux nécessairement simplifié et incomplet des connaissances en organisation et optimisation des programmes de sélection sous l'hypothèse classique d'un modèle génétique infinitésimal. Ce survol montre qu'il existe une grande diversité possible de programmes de sélection, mais qu'il y a toutefois des

points clés à optimiser pour obtenir à la fois une bonne efficacité à court terme et un maintien de variabilité génétique raisonnable dans les populations sélectionnées. A l'heure actuelle, de grandes questions demeurent tant autour de l'objectivation des critères d'efficacité associant l'espérance de la réponse à la variance de la réponse ou au maintien de la variabilité génétique, que de la prise en compte des informations génomiques dans les programmes de sélection à plus ou moins brève échéance selon les espèces. En effet, l'arrivée de la sélection génomique, dans les méthodes de sélection devrait induire une accélération sensible du progrès génétique dans les populations tant pour les caractères de production que pour les caractères d'adaptation qui ont souvent une héritabilité faible et où la sélection classique est alors peu efficace. En conséquence, la poursuite du raisonnement classique du sélectionneur, c'est-à-dire la forte diffusion d'un petit nombre de reproducteurs élites est extrêmement dangereuse en termes de maintien d'une diversité génétique raisonnable au sein des populations animales, notamment si l'on raccourcit dans le même temps l'intervalle entre générations en supprimant l'étape de contrôle sur descendance chez les bovins. Les objectifs de sélection doivent être adaptés, pour intégrer de nouveaux caractères, au fur et à mesure qu'ils seront évalués. Ce point est d'autant plus critique que des évolutions défavorables ont été constatées sur divers caractères fonctionnels et que la génétique devra s'intégrer dans une plus grande diversité de systèmes de production pour répondre aux attentes sociétales. Par ailleurs, en raison de la complexité croissante des enjeux socio-économiques, il devient impossible aux seuls maîtres d'œuvre des programmes de sélection de juger de l'importance relative à accorder aux caractères à améliorer génétiquement. Pour actualiser les objectifs de sélection, il est alors indispensable de développer des outils et démarches communs aux divers acteurs des filières, mais aussi d'établir un dialogue avec les responsables des politiques d'aménagement des territoires ainsi qu'avec le secteur associatif pour considérer les attentes sociales fortes. En effet, aucun changement opérationnel ne peut être effectué sans l'implication totale de toutes les parties concernées et une représentation adéquate de leurs points de vue.

## Références

- Avery P.J., Hill W.G., 1977. Variation in genetic parameters in small populations. *Genet. Res.*, 29, 198-213.
- Bichard M., 1971. Dissemination of genetic improvement through a livestock industry. *Anim. Prod.*, 13, 401-411.
- Bichard M., Pease A.H.R., Swales P.H., Ozkütük K., 1973. Selection in a population with overlapping generations. *Anim. Prod.*, 17, 215-227.
- 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. 9<sup>th</sup> World Congr. Genet. Applied Livest. Prod.*, Leipzig, Germany August 1-6, Comm. 716.
- Bulmer M.G., 1971. The effect of selection on genetic variability. *Am. Nat.*, 105, 201-211.
- 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.
- Colleau J.J., Regaldo D., 2001. Définition de l'objectif de sélection dans les races bovines laitières. *Renc. Rech. Rum.*, 8, 329-332.
- Colleau J.J., Heyman Y., Renard J.P., 1998. Les biotechnologies de la reproduction chez les bovins et leurs applications réelles ou potentielles en sélection. *INRA Prod. Anim.*, 11, 41-56.
- Colleau J.J., Moureaux S., Briend M., Béchu J., 2004. A method for the dynamic management of genetic variability in dairy cattle. *Genet. Sel. Evol.*, 36, 373-394.
- Colleau J.J., Tual K., de Preumont H., Regaldo D., 2009. A mating method accounting for inbreeding and multi-trait selection in dairy cattle population. *Genet. Sel. Evol.*, 41:7 doi:10.1186/1297-9686-41-7
- Dekkers J.C.M., 1992. Asymptotic response to selection on best linear unbiased predictors of breeding values. *Anim. Prod.*, 54, 351-360.
- Dekkers J., van Arendonk J., 1998. Optimizing selection for quantitative traits with information on an identified locus in outbred populations. *Genet. Res.*, 71, 257-275.
- De Roos A.P.W., Hayes B.J., Goddard M.E., 2009. Reliability of genomic predictions across multiple populations. *Genetics*, 183, 1545-1553.
- Dickerson G.E., Hazel L.N., 1944. Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J. Agric. Res.*, 69, 459-476.
- Dockès AC, Magdelaine P, Daridan D, Guillaume A., Rémondet M., Selmi A., Gilbert H., Mignon-Grasteau S., Phocas F., 2011. Attentes en matière d'élevage des acteurs de la sélection animale, des filières de l'agroalimentaire et des associations. In : Numéro spécial, Amélioration génétique. Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M. (Eds). *INRA Prod. Anim.*, 24, 285-296.
- Ducos A., 1995. Evolutions of selection objective for pig breeding in France. Consequences on expected genetic changes. *Rev. Méd. Vét.*, 146, 715-722.
- Ducrocq V., 1994. Multiple trait selection: principles and problems. *Proc. 5<sup>th</sup> World Congr. Genet. Applied Livest. Prod.*, August 7-12, University of Guelph, Ontario, Canada, 18, 455-466.
- Elsen J.M., 1992. De l'optimisation au progrès génétique réalisé dans les schémas de sélection. In : Numéro Hors-série, *Eléments de Génétique Quantitative et applications aux populations animales*. INRA Prod. Anim., 237-242.
- Elsen J.M., 1993. Prediction of annual genetic gain and improvement lag between populations. *Genet. Sel. Evol.*, 25, 75-82.
- Elsen J.M., Mocquot J.C., 1974. Recherches pour une rationalisation technique et économique des schémas de sélection des bovins et ovins. *Bull. Techn. Départ. Génétique animale*, 17.
- Falconer D.S., Mackay T.F.C., 1996. *Introduction to quantitative genetics*. Longman (Ed), England, UK, 464p.
- Finney D.J., 1956. The consequences of a selection for a variate subject to errors of measurement. *Inst. Int. Stat.*, 24, 1-10.
- Foulley J.L., Hanocq E., Boichard D., 1992. A criterion for measuring the degree of connectedness in linear models of genetic evaluation. *Genet. Sel. Evol.*, 24, 315-330.
- 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 trois principales races bovines laitières françaises. *Renc. Rech. Rum.*, 10, 53-56.
- Gibson J.P., Kennedy B.W., 1990. The use of constrained selection indexes in breeding for economic merit. *Theor. Appl. Genet.*, 80, 801-805.
- Goddard M., 1983. Selection indices for non-linear profit functions. *Theor. Appl. Genet.*, 80, 339-344.
- Goddard M., 2009. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*, 136, 245-257.
- Goffinet B., Elsen J.M., 1984. Critères optimaux de sélection : quelques résultats généraux. *Genet. Sel. Evol.*, 16, 307-318.
- Groen A.F., Steine T., Colleau J.J., Pedersen J., Pribyl J., Reinsch N., 1997. Economic values in dairy cattle breeding, with special reference to functional traits. *Livest. Prod. Sci.*, 49, 1-21.
- Harris B., Johnson D., 1998. Approximate reliability of genetic evaluations under an animal model. *J. Dairy Sci.*, 81, 2723-2728.
- Harris B., Johnson D., 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *J. Dairy Sci.*, 93, 1243-1252.
- Harris B., Newman S., 1994. Breeding for profit: synergism between genetic improvement and livestock production (a review). *J. Anim. Sci.*, 72, 2178-2200.
- Hayes B., Bowman P., Chamberlain A., Verbyla K., Goddard M., 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.*, 41, 51.
- Hazel L.N., 1943. The genetic basis for constructing selection indexes. *Genetics*, 28, 476-490.
- Henderson C.R., 1963. Selection index and expected genetic advance. In: *Statistical genetics and plant breeding*. NAS-NRC, Washington DC, Publ 982, 141-163.
- Henderson, C.R., 1977. Prediction of future records. *Proc. Int. Conf. Quantitative Genetics*, Ames, Iowa, Aug. 16-21, 1976. Pollack E., Kempthorne O., Bailey T.B. Jr. (Eds) Iowa State University Press, Ames, USA, 615-638.
- Hill W.G., 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.*, 18, 117-139.
- Hill W.G., 1982. Prediction of response to selection from new mutations. *Genet. Res.*, 40, 255-278.
- Hill W.G., 1985. Effects of population size on response to short and long term selection. *Z. Tierzüchtg. Züchtgsbiol.*, 102, 161-173.
- Houle D., 1992. Comparing evolvability and variability of quantitative traits. *Genetics*, 130, 195-204.
- Itoh Y., Yamada Y., 1987. Comparisons of selection indices achieving predetermined proportional gains. *Génét. Sél. Evol.*, 19, 69-82.
- James J.W., 1977. Open nucleus breeding systems. *Anim. Prod.*, 24, 285-305.
- James J.W., 1978. Effective population size in open nucleus breeding schemes. *Acta Agric. Scand.*, 28, 387-392.
- Kanis E., De Greef K.H., Hiemstra A., van Arendonk J.A.M., 2005. Breeding for societally important traits in pigs. *J. Anim. Sci.*, 83, 946-957.
- Kempthorne O., Nordskog A.W., 1959. Restricted selection indices. *Biometrics*, 15, 10-19.
- Kinghorn B.P., 2011. An algorithm for efficient constrained mate selection. *Genet. Sel. Evol.*, 43, 4. <http://www.gsejournal.org/content/43/1/4>
- Kizilkaya K., Fernando R., Garrick D., 2010. Genomic prediction of simulated multi-breed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. *J. Anim. Sci.*, 88, 544-551.
- Laloë D., 1993. Precision and information in linear models of genetic evaluation. *Genet. Sel. Evol.*, 25, 557-576.
- Laloë D., 2011. La genèse et le développement des concepts de l'évaluation génétique classique. In : Numéro spécial, Amélioration génétique. Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M. (Eds). *INRA Prod. Anim.*, 24, 323-330.
- Lande R., 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.*, 28, 221-235.
- Lande R., Thompson R., 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. *Genetics*, 124, 743-756.
- Manfredi E., Barbieri M., Fournet F., Elsen J.M., 1998. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genet. Sel. Evol.*, 30, 127-148.
- Meuwissen T.H.E., 1997. Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci.*, 75, 934-940.

- Meuwissen T.H.E., 1989. A deterministic model for the optimization of dairy cattle breeding based on BLUP breeding value estimates. *Anim. Prod.*, 49, 193-202.
- 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., Sonesson A.K., 1998. Maximizing the response of selection with a predefined rate of inbreeding: overlapping generations. *J. Anim. Sci.*, 76, 2575-2583.
- Meuwissen T.H.E., Sonesson A.K., 2004. Genotype-assisted optimum contribution selection to maximize selection response over a specified time period. *Genet. Res.*, 84, 109-116.
- 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.
- Meuwissen T.H.E., Woolliams J.A., 1994. Maximizing genetic response in breeding schemes of dairy cattle with constraints on variance of response. *J. Dairy Sci.*, 77, 1905-1916.
- Muir W.M., 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *J. Anim. Breed. Genet.*, 124, 342-355.
- Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M., 2011. Amélioration génétique. Numéro spécial, *INRA Prod. Anim.*, 24, 128p.
- Neimann-Sorensen A., Robertson A., 1961. The association between blood groups and several production characteristics in three Danish cattle breeds. *Acta Agric. Scand.*, 11, 163-196.
- Nielsen H.M., Christensen L.G., Groen A.F., 2005. Derivation of sustainable breeding goals for dairy cattle using selection index theory. *J. Dairy Sci.*, 88, 1882-1890.
- Nielsen H.M., Sonesson A.K., Meuwissen T.H.E., 2011. Optimum contribution selection using traditional best linear unbiased prediction and genomic breeding values in aquaculture breeding schemes. *J. Anim. Sci.*, 89, 630-638.
- Odegard J., Yazdi M.H., Sonesson A.K., Meuwissen T.H.E., 2009. Incorporating desirable genetic characteristics from an inferior into a superior population using genomic selection. *Genetics*, 181, 737-745.
- Olesen I., Groen A.F., Gjerde B., 2000. Definition of animal breeding goals for sustainable production systems. *J. Anim. Sci.*, 78, 570-582.
- Olesen I., Alfnes F., Rora M., Navrud S., Kolstad K., 2006. Economic values of fish welfare and application of market experiments. In: Ethics and the politics of food. Kaiser M., Lien M. (Eds), Wageningen Academic Publishers, The Netherlands, 446-451.
- Ollivier L., 1974. Optimum replacement rates in animal breeding. *Anim. Prod.*, 19, 257-271.
- Phocas F., Colleau J.J., 1995. Approximating selection differentials and variances for correlated selection indices. *Genet. Sel. Evol.*, 27, 551-565.
- Phocas F., Colleau J.J., 1996. Genetic steady-state under BLUP selection for an infinite and homogeneous population with discrete generations. *Theor. Appl. Genet.*, 93, 287-294.
- Phocas F., Colleau J.J., Ménéssier F., 1995. Expected efficiency for growth in a French beef cattle breeding scheme. 2. Prediction of asymptotic genetic gain in a heterogeneous population. *Genet. Sel. Evol.*, 27, 171-188.
- Phocas F., Hanocq E., Bouix J., Renand G., Poivey J.P., Elsen J.M., Bibé B., Ménéssier F., 1997. Détermination des objectifs de sélection chez les ruminants allaitants : situation actuelle et perspectives d'évolution. *Renc. Rech. Rum.*, 171-178.
- Pong-Wong R., Woolliams J., 2007. Optimisation of contribution of candidate parents to maximise genetic gain and restricting inbreeding using semidefinite programming. *Genet. Sel. Evol.*, 39, 3-25.
- Poutous M., Vissac B., 1962. Recherche théorique des conditions de rentabilité maximum de l'épreuve de descendance des taureaux d'insémination artificielle. *Ann. Zootech.*, 11, 233-256.
- Quinton M., Smith C., Goddard M.E., 1992. Comparison of selection methods at the same level of inbreeding. *J. Anim. Sci.*, 70, 1060-1067.
- Razungles J., 1977. Variation de la différentielle phénotypique de sélection. *Ann. Génét. Sél. Anim.*, 9, 105-112.
- Rendel J.M., Robertson A., 1950. Estimation of genetic gain in milk yield by selection in a close herd of dairy cattle. *J. Genet.*, 50, 1-8.
- Robertson A., 1961. Inbreeding in artificial selection programmes. *Genet. Res.*, 2, 189-195.
- Ruane J., Colleau J., 1995. Marker-assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genet. Res.*, 66, 78-83.
- Schaeffer L.R., 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.*, 123, 218-223.
- Schneeberger M., Freeman A.E., Boehlje M.D., 1981. Estimation of a utility function from semen purchases from Holstein sires. *J. Dairy Sci.*, 64, 1713-1718.
- Schwob S., Riquet J., Bellec T., Kernaléguen L., Tributou T., Bidanel J.P., 2009. Mise en place d'un programme de sélection assistée par marqueurs dans la population sino-européenne Duochan. *Journ. Rech. Porcine Fr.*, 41, 29-30.
- Shepherd R.K., Kinghorn B.P., 1992. Optimising multi-tier open nucleus breeding schemes. *Theor. Appl. Genet.*, 85, 372-378.
- Smith C., Banos G., 1991. Selection within and across populations in livestock improvement. *J. Anim. Sci.*, 69, 2387-2394.
- Smith C., Quinton M., 1993. Effect of selection in sables and crossing on genetic response and inbreeding. *J. Anim. Sci.*, 71, 2631-2638.
- Sonesson A.K., Meuwissen T.H.E., 2009. Testing strategies for genomic selection in aquaculture breeding programs. *Genet. Sel. Evol.*, 41, 37.
- Toosi A., Fernando R.L., Dekkers J.C.M., 2010. Genomic selection in admixed and crossbred populations. *J. Anim. Sci.*, 88, 32-46.
- Toro M. A., Perez-Enciso M., 1990. Optimizing selection under restricted inbreeding. *Genet. Sel. Evol.*, 22, 93-107.
- Tribout T., Bidanel J.P., Phocas F., Schwob S., Guillaume F., Larzul C., 2011. La sélection génomique : principe et perspectives d'utilisation pour l'amélioration des populations porcines. *Journ. Rech. Porcine Fr.*, 13-25.
- Verrier E., 2001. Marked-assisted selection for the improvement of antagonistic traits under mixed inheritance. *Genet. Sel. Evol.*, 33, 17-38.
- Verrier E., Colleau J.J., Foulley J.L., 1991. Methods for predicting response to selection in small populations under additive genetic models. *Livest. Prod. Sci.*, 29, 93-114.
- Verrier E., Colleau J.J., Foulley J.L., 1993. Long-term effects of selection based on the animal BLUP model in a finite population. *Theor. Appl. Genet.*, 87, 446-454.
- Walsh B., Lynch M., 2009. Evolution and selection of quantitative traits. [http://nitro.biosci.arizona.edu/zbook/NewVolume\\_2/newvol2.html](http://nitro.biosci.arizona.edu/zbook/NewVolume_2/newvol2.html)
- Woolliams J., Bijma P., Villanueva B., 1999. Expected genetic contributions and their impact on gene flow and genetic gain. *Genetics*, 153, 1009-1020.
- Woolliams J.A., Meuwissen T.H.E., 1993. Decision rules and variance of response in breeding schemes. *Anim. Prod.*, 56, 179-186.
- Wray N.R., Hill W.G., 1989. Asymptotic rates of response from index selection. *Anim. Prod.*, 49, 217-227.
- Wray N.R., Woolliams J.A., Thompson R., 1990. Methods for predicting rates of inbreeding in selected populations. *Theor. Appl. Genet.*, 80, 503-512.
- Wray N.R., Woolliams J.A., Thompson R., 1994. Prediction of rate of inbreeding in populations undergoing index selection. *Theor. Appl. Genet.*, 87, 878-892.
- Wyler A., Hagger C., Stranzinger G., 1994. Different risk-utility functions in animal breeding. Proc. 5<sup>th</sup> World Congr. Genet. Appl. Livest. Prod., August 7-12, University of Guelph, Ontario, Canada, 18, 244-247.

---

## Résumé

---

La finalité d'un programme de sélection est de générer du progrès génétique dans les générations animales futures afin de répondre aux besoins alimentaires, nutritionnels, sociaux et environnementaux de demain. Les travaux pionniers en optimisation technique et économique des programmes de sélection des années 1970 à 1990 sont décrits dans cette synthèse. Maintenant, il s'agit d'accroître à la fois le nombre de caractères inclus dans les objectifs de sélection et la valeur prédictive des critères de sélection et des valeurs génétiques prédites pour permettre le choix le plus fiable possible dès le plus jeune âge, de reproducteurs améliorateurs parmi les candidats à la sélection. A ce titre, la sélection génomique est une véritable révolution dans la gestion des populations apparue dans la première décennie du 21<sup>ème</sup> siècle. L'organisation des programmes de sélection doit être revue en profondeur : la poursuite du raisonnement classique, c'est-à-dire la forte diffusion d'un petit nombre de reproducteurs élites est extrêmement dangereuse en termes de maintien d'une diversité génétique raisonnable au sein des populations animales, notamment si l'on raccourcit dans le même temps l'intervalle entre générations en supprimant l'étape de contrôle sur descendance chez les bovins. Les objectifs de sélection doivent être adaptés, pour intégrer les nouveaux caractères, au fur et à mesure qu'ils seront évalués. Ce point est d'autant plus critique que des évolutions défavorables ont été constatées sur divers caractères fonctionnels et que la génétique devra s'intégrer dans une plus grande diversité de systèmes de production pour répondre aux attentes sociétales.

---

## Abstract

---

### *Optimization of breeding schemes*

The aim of a breeding scheme is to generate livestock improvement in future generations in order to fulfill the future needs of the human population in terms of alimentation, environment and society. In this review, pioneer works conducted between the 1970s and the 1990s are summarized in terms of technical and economic optimization of breeding schemes. Now, it is time to increase both the number of traits considered in breeding goals and the predictive value of selection criteria and estimated breeding values to allow the best choice of reproducers among candidates for selection at the youngest age. For these purpose, genomic selection is a true revolution in population genetic management that arose in the first decade of the 21st century. The design of breeding schemes must be deeply revisited: the pursute of classical reasoning, i.e. the large diffusion of a small number of elite reproducers is extremely dangerous in terms of sensible maintenance of genetic diversity among animal populations, especially when generation intervals are reduced at the same time as that which is currently happening in cattle breeding with the abandon of progeny testing. Furthermore, breeding objectives must be adapted to account for new traits as soon as they are evaluated. This point is critical when unfavorable trends are observed for functional traits and genetics has to integrate a large diversity of production systems to answer to society's demands.

PHOCAS F., 2011. L'optimisation des programmes de sélection. In : Numéro spécial, Amélioration génétique. Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M. (Eds). INRA Prod. Anim., 24, 341-356.

