La diversité génétique et son lien avec la sélection dans les races bovines laitières françaises

I - La diversité génétique dans les espèces domestiques sélectionnées

Avant d'aborder les particularités des races bovines laitières françaises, il est primordial de nous intéresser aux grands principes sur lesquels repose la sélection génétique dans les espèces domestiques, ainsi qu'à ses liens avec la diversité génétique. En effet, la diversité génétique joue un rôle central dans la capacité à sélectionner les animaux et nous verrons en quoi l'équilibre progrès génétique/diversité génétique est difficile à satisfaire. Nous verrons ensuite quelles sont les conséquences possibles d'une perte de diversité génétique dans les populations domestiques sélectionnées et quelles sont les méthodes de gestion de la diversité génétique.

A - Les grands principes de la sélection génétique

1 - La genèse de la sélection génétique

La sélection des espèces domestiques est un processus initié avec la domestication par les populations humaines. La domestication des espèces domestiques animales a commencé par celle du chien il y a plus de 15 000 ans (Frantz et al., 2016), antérieure à la sédentarisation et à l'agriculture, et s'est poursuivie par celle de nombreuses espèces de rente du néolithique à nos jours. La domestication s'est effectuée en sélectionnant les animaux de manière empirique, à partir de caractères favorisant leur utilisation par l'Homme, par exemple la docilité. Ce processus de sélection a donc privilégié certains phénotypes et entraîné la modification des fréquences des génotypes correspondants au sein des populations domestiquées par rapport aux populations sauvages.

Une fois la domestication effectuée, la sélection s'est focalisée sur l'amélioration d'autres caractères (de production, de morphologie ou comportementaux). En France, dès la fin du XIXe siècle, les éleveurs ont initié la création des livres généalogiques ou *herd-books*, par le biais d'associations ayant fortement contribué à l'institutionnalisation des races. L'objectif de chaque association était de fixer les qualités d'une race donnée en s'assurant de la conformité des animaux reproducteurs à un standard morphologique, en passant par l'identification des animaux et l'enregistrement de leur généalogie. Au début des années 1900, des procédures de contrôle des performances individuelles des animaux sont apparues comme le contrôle laitier chez la vache ou de la ponte chez les poules. Ces procédures se sont étendues aujourd'hui à de nouveaux caractères en lien avec la santé, le bien-être ou encore l'impact environnemental des animaux. Cette préoccupation pour l'amélioration des animaux d'élevage s'est donc ancrée en France dès le XIXe

siècle dans le paysage de l'élevage, de l'enseignement et de la recherche agronomique (Gouyon, 2002).

En parallèle de ce travail des éleveurs-sélectionneurs, la génétique fait son apparition dans le monde académique avec la redécouverte des lois de Mendel en 1900 (Laloë, 2011). Cet événement a soulevé de nombreux débats sur l'antagonisme entre la nature discontinue des gènes et la variation continue des caractères mesurables, ainsi que la question de la part de responsabilité de l'environnement et de l'hérédité dans les variations observées. Mendel avait proposé une solution au paradoxe entre la discontinuité des gènes et la continuité des caractères, en suggérant qu'un caractère continu était sous dépendance polygénique. Cette hypothèse a été formalisée en 1918 par Ronald Fisher (Fisher, 1918). Selon le modèle polygénique, le phénotype d'un individu est influencé à la fois par son environnement et par sa valeur génétique, cette dernière étant la somme des effets de très nombreux gènes à effet faible. Ce modèle a jeté les bases de la génétique quantitative.

Dans ce contexte, le travail d'enregistrement des généalogies et des performances effectué par les éleveurs-sélectionneurs, l'organisation de ces derniers en associations de races, l'implication précoce des pouvoirs publics français dans l'amélioration de certaines espèces (Loi sur l'élevage, 1966), ainsi que les progrès de l'informatique, ont fourni le terreau nécessaire à la mise en place d'évaluations génétiques et de nouveaux programmes d'amélioration des animaux d'élevage.

2 - Le progrès génétique, indissociable de la variabilité génétique

a - La génétique quantitative classique et les évaluations polygéniques

L'essor de la génétique quantitative et du modèle polygénique ont offert la possibilité de prédire la valeur génétique des animaux pour chaque caractère, c'est-à-dire la partie héritable et transmissible à la descendance. L'objectif est généralement d'augmenter la valeur génétique de la population d'une génération à la suivante (sélection directionnelle). La sélection des animaux reproducteurs s'effectue alors sur la base de la valeur génétique estimée pour chaque caractère ou une combinaison de différents caractères, et non plus à partir des performances ou phénotypes bruts dont une part est due à l'environnement. Les évaluations génétiques utilisent les phénotypes, ou performances, ainsi que les pedigrees, pour prédire le plus précisément possible la valeur génétique des animaux.

Pour obtenir la valeur génétique d'un animal, il est nécessaire de corriger son phénotype par les effets fixes identifiables, comme l'âge ou l'élevage d'origine. La part de variance phénotypique expliquée par la composante génétique, appelée héritabilité, est propre à chaque caractère et à chaque population. La notion d'héritabilité résulte essentiellement des travaux menés par Fisher

(1918), Wright (1921) et Malécot (1948). On note l'héritabilité au sens strict h^2 et on la calcule comme suit :

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$$

Avec σ_p^2 la variance phénotypique observée dans la population et σ_a^2 la variance génétique additive, c'est-à-dire la part de variabilité phénotypique observée due aux gènes et transmissible génétiquement à la descendance. La variance génétique correspond à la variabilité génétique du caractère d'intérêt. Plus l'héritabilité d'un caractère est élevée, plus le caractère sera facile à sélectionner car une part importante des variations entre individus s'expliquera de manière significative par la génétique. L'efficacité de la sélection pour un caractère dépend donc de la variabilité génétique de ce caractère.

Pour sélectionner efficacement les animaux, il est nécessaire que l'estimation de leur valeur génétique soit la plus précise possible. On mesure la précision des évaluations génétiques grâce au coefficient de détermination (CD), compris entre 0 et 1, défini comme le rapport entre la variance des valeurs génétiques estimées et la variance des valeurs génétiques (Laloë, 1993). Plus le CD sera proche de 1 et plus les évaluations seront précises. La valeur du CD dépend du nombre de performances enregistrées pour chaque animal et ses apparentés, ainsi que de l'héritabilité du caractère : plus l'héritabilité est élevée, plus la précision des évaluations sera élevée. Or, plus la variabilité génétique d'un caractère est élevée, plus son héritabilité sera élevée. La précision des évaluations dépend donc fortement de la variabilité génétique du ou des caractères évalués.

En sélection animale, et surtout dans les filières où l'insémination est la plus développée comme chez les bovins laitiers, le progrès génétique est généralement diffusé par la voie mâle, principalement car les mâles peuvent avoir une descendance plus nombreuse que les femelles et qu'il est plus simple technologiquement et d'un point de vue sanitaire de manipuler et diffuser du sperme que des ovocytes. De fait, de nombreux caractères de production animale ne sont exprimés que chez les femelles (production laitière, ponte, etc.). Dans ce cas, on estimera la valeur génétique d'un mâle en utilisant la production des femelles qui lui sont apparentées, selon le principe que la valeur génétique de celles-ci est égale en espérance à la moitié de celle de leur père. La précision de l'évaluation dépend alors de l'héritabilité du caractère mais également du nombre de descendantes avec performance.

Pour évaluer l'efficacité de la sélection, on calcule le progrès génétique annuel attendu ΔG (Dickerson et Hazel, 1944) grâce à la relation suivante :

$$\Delta G = \frac{i\sqrt{CD}\sigma_a}{T}$$

avec ΔG l'augmentation du niveau génétique réalisée en une année, i l'intensité de sélection égale au différentiel de sélection standardisé entre les animaux sélectionnés et les candidats à la sélection, σ_a l'écart type génétique, qui est une mesure de la variabilité génétique du caractère d'intérêt et T l'intervalle de génération, égal à l'âge moyen des parents à la naissance de leurs descendants. Une nouvelle fois, il apparaît qu'une baisse de la variabilité génétique d'un caractère peut être préjudiciable à l'efficacité de la sélection, en réduisant le progrès génétique attendu ΔG.

Dans le cadre de la génétique quantitative et des évaluations polygéniques, l'héritabilité des caractères, la précision des évaluations et le progrès génétique annuel attendu pour un caractère sont dégradés lorsque la variabilité génétique de ce caractère diminue. La variabilité génétique de chaque caractère présente donc une importance indéniable et centrale pour une sélection efficace, car c'est la base de la sélection.

b - Les évaluations génomiques

Progressivement, les méthodes d'estimation de la valeur génétique des animaux ont évolué en relation étroite avec les nouvelles connaissances sur le génome. Dès le début des années 2000, les approches génomiques font leur apparition et leurs applications en sélection animale se dessinent. Le génome des principales espèces d'élevage est séquencé. Ces séquençages révèlent une grande variabilité de la séquence d'ADN au sein d'une même espèce, due à des changements ponctuels de nucléotides dans la séquence. Ces millions de polymorphismes localisés sont appelés *Single Nucleotide Polymorphisms* (SNPs). De nombreux animaux ont alors été génotypés grâce à des puces à ADN de densité variable allant de quelques centaines à plusieurs millions de SNPs.

Les technologies de génotypage devenant de plus en plus efficaces et abordables, ceci a conduit au développement d'une nouvelle méthodologie d'estimation de la valeur génétique : l'évaluation génomique (Meuwissen et al., 2001).

Les SNPs sont rarement causaux (i.e. responsables d'une mutation ayant un effet phénotypique) mais dès lors que leur densité sur le génome est suffisante, ils peuvent être en déséquilibre de liaison avec les gènes impliqués dans les caractères d'intérêt. On peut donc estimer l'effet de chacun des allèles des SNPs sur les caractères d'intérêt. L'estimation des effets des SNPs est réalisée à partir d'une population de référence constituée d'animaux à la fois phénotypés et génotypés. Une fois que l'effet de chacun des allèles des SNPs est estimé, la valeur génétique estimée d'un animal génotypé peut être calculée, même s'il n'est pas phénotypé (en général, un jeune candidat à la sélection). La précision de l'estimation des effets des SNPs dépend de la

précision des enregistrements des performances ainsi que de la taille de la population de référence.

L'évaluation génomique permet ainsi d'estimer la valeur génétique des animaux avec précision et dès le plus jeune âge. Ainsi, les animaux évalués dès la naissance pourront être choisis comme reproducteurs élites et mis à la reproduction dès leur maturité sexuelle, au lieu de devoir attendre que l'on enregistre leurs performances voire celles de leurs descendantes pour estimer leur valeur génétique. Les évaluations génomiques permettent ainsi une diminution de l'intervalle de génération par rapport aux évaluations génétiques classiques.

Si l'on revient au calcul du progrès génétique attendu présenté ci-dessus, une augmentation de la précision des évaluations et une diminution de l'intervalle de génération permettent une augmentation du progrès génétique annuel attendu. De plus, avec la baisse des coûts du génotypage, il est possible d'évaluer de nombreux animaux et donc d'augmenter l'intensité de la sélection et l'efficacité de la sélection, mais aussi de sélectionner des animaux parmi un grand choix de candidats plus représentatifs de la population (Bouquet et Juga, 2013) et donc de limiter la perte de variabilité génétique (Colleau et al., 2015).

3 - Les conséquences de la sélection sur la variabilité génétique et le compromis progrès/variabilité

Pour sélectionner efficacement, il est nécessaire que la variabilité génétique dans la population soit suffisante. Or, la sélection est une force évolutive qui modifie les fréquences génotypiques d'une génération à la suivante. Lorsque que la sélection est directionnelle, comme c'est souvent le cas en sélection dans les espèces de rente, alors la sélection modifie les fréquences alléliques en augmentant la fréquence des allèles favorables pour le caractère sélectionné, qui vont se propager au sein de la population. La fréquence des allèles défavorables va baisser voire devenir nulle (fixation). Ainsi, la sélection va faire baisser la variabilité génétique du caractère sélectionné et la diversité allélique des gènes impliqués dans le caractère et des régions du génome en déséquilibre de liaison avec ces derniers, c'est le phénomène de balayage sélectif dû au mécanisme d'auto-stop génétique (Smith et Haigh, 1974).

Plus l'intensité de sélection est forte, plus la perte de variabilité génétique est rapide (Bulmer, 1971). Cela affecte l'héritabilité et la précision des évaluations, qui vont baisser également. Cela risque de provoquer à moyen et long terme une baisse du progrès génétique annuel attendu, bien qu'à court terme le procédé soit efficace pour augmenter le niveau génétique de la population. De bonnes combinaisons alléliques sont fixées rapidement lorsque l'intensité de sélection est très

forte mais cela empêche en même temps de composer de nouvelles combinaisons qui pourraient être avantageuses. Le progrès génétique est alors ralenti.

Comme la sélection est généralement effectuée sur un grand nombre de caractères, parfois antagonistes, la recherche de compromis dans la sélection de ces différents caractères atténue en partie l'effet du balayage sélectif en faisant baisser l'intensité de sélection sur le génome.

La sélection n'est pas la seule force évolutive appliquée aux populations d'animaux domestiques. En effet, la sélection d'animaux reproducteurs d'élite mène souvent à une contribution très déséquilibrée des reproducteurs d'une génération à la suivante, ce qui va réduire la taille efficace des populations (voir la partie dédiée à ce concept dans le chapitre annexe page 145) et donc accentuer l'effet de la dérive génétique (Masel, 2011). Prioriser le maintien de la variabilité génétique, en limitant le déséquilibre de contribution des reproducteurs, ne permet pas d'atteindre un progrès génétique annuel aussi élevé qu'en absence de contrainte sur la variabilité génétique.

Il apparaît alors nécessaire de trouver un équilibre entre augmentation du progrès génétique et maintien de la variabilité génétique pour atteindre et maintenir sur le long terme une sélection efficace. Le compromis entre les deux dépend fortement du contexte et des enjeux de la population sélectionnée. Dans certaines races, les objectifs économiques de progrès génétique seront prioritaires dans un contexte de forte concurrence. Dans d'autres races, locales, de petit effectif ou à valeur patrimoniale forte, c'est la conservation qui primera et donc plus d'attention sera portée au maintien de la variabilité génétique au détriment du progrès génétique.

Quand l'intensité de sélection augmente, le progrès génétique augmente à court terme et la perte de variabilité génétique du caractère sélectionné et donc de la diversité allélique de la population s'accélère. La baisse de variabilité génétique va quant à elle faire baisser le progrès génétique annuel attendu à moyen et long terme. Cependant, réduire l'intensité de la sélection pour maintenir la variabilité génétique est délétère à court terme pour le progrès génétique. Ce paradoxe est à l'origine de la difficulté à trouver un compromis acceptable entre progrès génétique et diversité génétique dans les espèces animales sélectionnées. Il n'existe pas de réponse unique à ce problème car la priorisation de l'un ou de l'autre dépend de l'espèce, de la race et du contexte socio-économique.

B - La diversité génétique : importance et lien avec la sélection

Si la diversité génétique est indispensable au progrès, la perte de diversité peut avoir d'autres conséquences : l'apparition à court terme de dépression de consanguinité, et la perte de potentiel adaptatif à long terme.

1 - Diversité et dépression de consanguinité

La consanguinité est définie comme le résultat de la reproduction sexuée entre deux individus apparentés et deux individus sont apparentés s'ils ont un ou plusieurs ancêtres communs (voir chapitre annexe page 141). La consanguinité fait augmenter la proportion de génome autozygote (ou IBD, *identical by descent*), c'est-à-dire la proportion de génome pour laquelle les deux allèles d'un même locus sont identiques chez un individu parce qu'ils proviennent d'un ancêtre commun à ses parents. La consanguinité est négativement corrélée à la survie et à la valeur sélective générale de la population (O'Grady et al., 2006; Reed et Frankham, 2003).

L'augmentation de la consanguinité peut avoir pour conséquence la diminution de la valeur des caractères d'intérêt liés à la production (production laitière, nombre d'œufs pondus, etc.), à la reproduction (fertilité, fécondité, viabilité des juvéniles) ou à la survie (santé, résistance aux maladies) des animaux à l'échelle de la population. On appelle ce phénomène la dépression de consanguinité (Leroy, 2014 ; Pryce et al., 2014 ; Wright, 1968 ; Darwin, 1876, 1868).

Pour éviter l'apparition de dépression de consanguinité, il est donc nécessaire de maintenir la diversité génétique dans les populations animales sélectionnées et en particulier de maîtriser l'augmentation de la consanguinité.

2 - Diversité et potentiel adaptatif

La perte de diversité génétique a également des effets délétères sur les capacités d'adaptation des populations animales sélectionnées.

Maintenir les capacités d'adaptation d'une population animale sélectionnée est nécessaire en cas de modifications des objectifs de sélection qui ne manqueront pas de survenir à plus ou moins long terme, comme d'autres sont déjà survenues. En effet, si les objectifs de sélection à la sortie de la seconde guerre mondiale étaient orientés vers une productivité accrue, visant à l'autonomie alimentaire, ils ont ensuite évolué pour prendre en compte des critères de qualité des produits (composition du lait par exemple) et de santé des animaux.

Aujourd'hui, de nombreux travaux sont menés pour évaluer l'impact environnemental des différentes espèces animales sélectionnées et comment cet impact pourrait diminuer, en s'intéressant par exemple à l'efficacité alimentaire des animaux ou à leurs émissions de gaz à effet

de serre (méthane chez les bovins par exemple), faisant ainsi apparaître de potentiels nouveaux objectifs de sélection (Ramayo-Caldas et al., 2020 ; Martin et al., 2019 ; Løvendahl et al., 2018 ; Renand et al., 2019 ; Herrero et al., 2013 ; Willems et al., 2013 ; Ferket et al., 2002).

Les modifications des objectifs de sélection ont le plus souvent des causes économiques. A l'avenir, les objectifs de sélection pourraient être amenés à changer en fonction :

- Des attentes sociétales : agriculture biologique, production locale, bien-être animal, qualités nutritionnelles des produits, etc. ;
- De l'environnement des animaux : conduite d'élevage, pâturage, alimentation disponible, pression parasitaire, virale et bactérienne, maladies émergentes, changements climatiques, etc.

Ces changements peuvent être imprévisibles ou soudains et nécessiter une adaptation rapide des espèces animales sélectionnées, à partir d'une diversité génétique préexistante. On retrouve ces notions dans les réflexions sur la transition agro-écologique et la création de programmes de sélection animale pour des systèmes d'élevage agro-écologiques. L'agro-écologie fait référence à une agriculture performante à la fois sur les plans économiques et environnementaux (Guillou et al., 2013), et repose sur cinq principes (Thomas et al., 2014 ; Dumont et al., 2013) :

(i) Une gestion intégrée de la santé animale, prenant en compte le bien-être ;
(ii) La réduction de l'utilisation des intrants ;
(iii) La réduction de la pollution et l'amélioration de l'efficacité métabolique des systèmes ;
(iv) L'amélioration de la diversité pour améliorer la résilience des systèmes ;
(v) La préservation de la diversité biologique.

L'objectif de la transition agro-écologique est de redonner de l'importance à l'utilisation de ressources locales pour assurer la production. Si des améliorations dans l'évaluation des animaux et la prise en compte des interactions génotype-milieu dans ces évaluations sont nécessaires pour sélectionner des animaux adaptés aux particularités locales de leur élevage, la transition agro-écologique repose sur l'utilisation de la diversité génétique au sein des espèces et races sélectionnées afin de maximiser les capacités d'adaptation des animaux (Phocas et al., 2017). En effet, l'agro-écologie ne consiste pas en la création d'un animal agro-écologique type mais en la possibilité d'obtenir les animaux les plus complémentaires et adaptés possible aux particularités locales, et donc des animaux aux profils variés entre localités et au sein des troupeaux (Phocas et al., 2017).

Maintenir la diversité génétique au sein des populations animales sélectionnées est nécessaire non seulement au maintien de l'efficacité de la sélection et du potentiel adaptatif de ces populations à long terme dans un contexte de transition agro-écologique et climatique, mais aussi pour limiter l'effet de la dépression de consanguinité sur la santé et les performances des animaux à court terme. Il apparaît donc nécessaire d'estimer la diversité génétique pour être capable de la prendre en compte dans la construction de programmes de sélection animale durables.

Un chapitre annexe intitulé « Mesurer la diversité génétique au sein d'une race », disponible à la page 139 de ce manuscrit, détaille les différentes méthodes permettant d'estimer la diversité génétique.

3 - La conception des schémas de sélection : un pilier central dans la gestion de la diversité génétique

Il existe plusieurs leviers d'action pour limiter la perte de diversité génétique au sein des races. Ils se situent à trois échelons différents : la **conception des schémas de sélection** en termes de dimension, le choix des animaux sélectionnés et les plans d'accouplement. Les schémas de sélection sont conçus et mis en œuvre par les entreprises de sélection et permettent le choix ou la procréation des futurs reproducteurs. Un plan d'accouplement consiste en une liste des accouplements à réaliser dans la population ou au sein du schéma de sélection, généralement avec l'objectif d'obtenir les meilleurs produits possibles à la génération suivante, en termes de progrès et/ou de diversité génétiques. Il est également possible de conserver du matériel génétique d'animaux originaux ou représentatifs de la race ou de la population à un instant donné dans des cryobanques, autorisant leur réutilisation ultérieure comme reproducteurs, mais aussi de réinjecter de la diversité génétique dans une race, grâce au croisement ou à l'introgression.

La conception des schémas de sélection est centrale dans la gestion de la diversité génétique. Elle consiste en grande partie à décider du nombre d'animaux candidats à la sélection et du nombre d'animaux sélectionnés pour devenir reproducteurs. Assez logiquement, à contribution équilibrée des reproducteurs, on s'attend à ce que moins il y ait d'animaux sélectionnés, plus la taille efficace de la population diminue et donc plus la diversité génétique de la race baisse rapidement. Pour le visualiser, prenons une formule approchée de la taille efficace N_e en fonction du nombre de mâles et de femelles reproducteurs :

$$N_e = \frac{4N_f N_m}{N_f + N_m}$$

Avec N_f le nombre de femelles reproductrices et N_m le nombre de mâles reproducteurs.

Nous voyons que la taille efficace de la population sera d'autant plus faible que le déséquilibre entre le nombre de mâles et de femelles se reproduisant est important. Prenons l'exemple d'une population comportant 1000 femelles reproductrices ($N_f = 1000$) et 100 mâles reproducteurs ($N_m = 100$) : la taille efficace de la population selon cette équation sera d'environ 363 individus. Si le nombre de mâles reproducteurs est divisé par 10 ($N_m = 10$) et bien que le nombre total d'individus reproducteurs passera de 1100 à 1010 (soit une baisse d'environ 8 % du nombre total de reproducteurs), la taille efficace de la population sera alors d'environ 39 individus seulement, soit une taille efficace divisée par plus de 9. Pour simplifier, **on peut considérer que si $N_m \ll N_f$, alors $N_e \approx 4\ N_m$.**

Cet exemple ne prend pas en compte toute la complexité des populations réelles. **Il permet néanmoins d'illustrer simplement l'importance de la conception des schémas de sélection** et en particulier, du nombre de mâles reproducteurs.

De nombreux travaux de simulation ont rendu compte de l'effet d'augmenter ou de réduire le nombre de taureaux mis en marché par les entreprises de sélection sur la diversité génétique des races bovines laitières (Thomasen et al., 2016 ; Bouquet et al., 2015; Colleau et al., 2015). Ainsi, le nombre d'individus mis à la reproduction doit être maximal à chaque génération et les reproducteurs doivent être renouvelés rapidement dans le but de diffuser plus largement les gènes fondateurs des mères. En effet, l'utilisation majoritaire d'un seul mâle pour une génération, par exemple, crée un goulot d'étranglement et fait baisser la taille efficace de la population et donc sa diversité génétique.

C - La dépression de consanguinité

1 - Origine

Le phénomène de dépression de consanguinité, illustré sur la Figure 1 ci-dessous, résulte majoritairement de l'homozygotie d'allèles délétères et récessifs. Les mutations, parmi lesquelles les mutations délétères (les plus fréquentes), apparaissent au hasard le long du génome et il y a très peu de chances de rencontrer la même mutation au même locus chez deux individus différents uniquement par hasard. En effet, dans un modèle panmictique, si la fréquence p d'un allèle est petite, alors la fréquence du génotype homozygote pour ce même allèle p^2 est négligeable. Dans ce contexte, la seule façon pour une mutation récessive d'être rencontrée à l'état homozygote chez un individu, et donc d'être exprimée, est d'être autozygote et donc de résulter d'un accouplement consanguin. Ainsi, la fréquence des mutations récessives et délétères reste faible car les hétérozygotes ne sont pas contre-sélectionnés tandis que les homozygotes le sont.

La dépression de consanguinité est due au fardeau génétique, qui correspond à l'ensemble de ces allèles délétères, le plus souvent récessifs, portés par le génome des animaux à l'échelle de la population. Plus le fardeau génétique est grand, plus la dépression de consanguinité sera forte.

Les mécanismes génétiques impliqués dans la dépression de consanguinité sont présentés dans un chapitre annexe page 137.

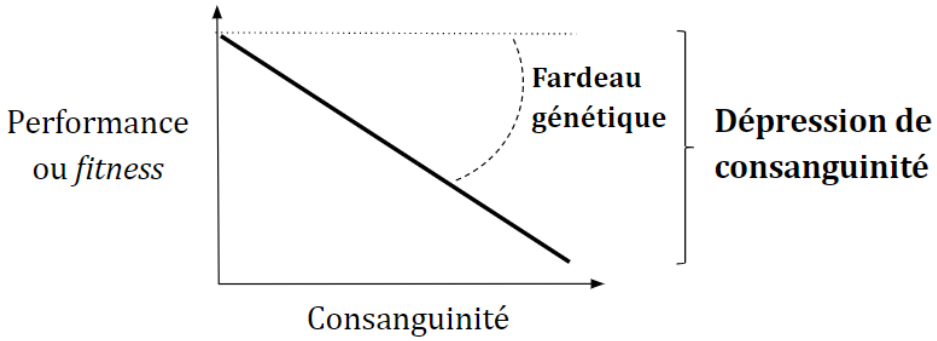

Figure 1 La dépression de consanguinité : baisse de la valeur des performances et de la *fitness* avec l'augmentation de la consanguinité

2 - Les conséquences économiques de la dépression de consanguinité

La dépression de consanguinité n'affecte pas seulement les caractères liés à la valeur sélective mais aussi les caractères de production de nombreuses espèces de rente, comme la production laitière, la croissance, la conformation ou la production d'œufs (Doekes et al., 2019 ; Leroy, 2014 ; Pryce et al., 2014 ; Croquet et al., 2006 ; Ercanbrack et Knight, 1991). L'ampleur des conséquences économiques liées à la dépression de consanguinité dépend donc de l'évolution de la consanguinité dans la population (Gama et Smith, 1993).

Ainsi, la dépression de consanguinité aura des conséquences plus délétères dans des systèmes en race pure, comme on en trouve chez les ruminants, que chez les volailles ou les porcs, chez lesquels les systèmes de production prennent majoritairement appui sur le croisement terminal et les populations en exploitations sont découplées du schéma de sélection. La dépression de consanguinité peut globalement être négligée dans ce type d'espèces, alors que chez les ruminants, en particulier les races laitières, la dépression de consanguinité peut être plus préoccupante.

Dans une race bovine telle que la Holstein, il a été estimé qu'au taux de consanguinité actuel, la dépression de consanguinité fait perdre 4,4 % du progrès génétique en production laitière par an, soit une perte d'environ 1,4 kg de lait par an par vache (Leroy, 2014). D'autres études ont montré

qu'une augmentation de 1 % de la consanguinité avait pour effet une baisse de la production laitière allant de 20 à 63 kg par an (Pryce et al., 2014 ; Bjelland et al., 2013). Un impact délétère significatif de la dépression de consanguinité a aussi été observé sur la fertilité avec, pour une augmentation de 1 % de la consanguinité, une augmentation de l'intervalle vêlage-vêlage de 0,12 à 1,7 jours (Pryce et al., 2014 ; Bjelland et al., 2013), une augmentation des difficultés au vêlage (Pryce et al., 2014 ; Bjelland et al., 2013), et une baisse de 1 % de la moyenne phénotypique du nombre de spermatozoïdes par éjaculat (Ferenčaković et al., 2017). Dans l'espèce ovine, une augmentation de 1 % de la consanguinité affecte la fertilité avec une baisse de 0,014 % (Lamberson et Thomas, 1984) à 0,301 % (Ercanbrack et Knight, 1991), le poids de naissance, qui diminue de 4 à 8 g selon les races (van Wyk et al., 1993, 2009 ; Ercanbrack et Knight, 1991), le taux de survie des agneaux, qui baisse de 2,8 % (Lamberson et Thomas, 1984) et en particulier de 0,31 % à 15 jours et 0,36 % à un mois (Mandal et al., 2004). La production de laine est également impactée, avec un effet sur le poids de la toison (baisse de 9 à 17 g de laine pour une augmentation de 1 % de la consanguinité) et sur la qualité de la laine (Ercanbrack et Knight, 1991).

Pour chaque caractère pris séparément, l'impact de la dépression de consanguinité peut être considéré comme relativement limité. Cependant, la rentabilité économique dépend de la combinaison des tous les caractères d'intérêt (Kristensen et Sørensen, 2005). Plusieurs études ont ainsi examiné l'impact économique de la dépression de consanguinité en prenant en compte soit un seul caractère, soit une combinaison de plusieurs caractères. Prenons l'exemple de deux études ayant évalué l'impact économique de la dépression de consanguinité sur toute la durée de vie d'une vache Holstein. L'une a estimé que la perte de revenu net à vie pour une vache était d'environ 22,7 € (Croquet et al., 2006) quand une autre a estimé cette perte autour de 60 € (Smith et al., 1998). Cette différence peut s'expliquer par l'absence de prise en compte des caractères liés à la fertilité, la longévité et la santé dans la première étude. Chez les ovins, la perte économique moyenne par brebis due à la dépression de consanguinité a quant à elle été estimée à environ 15 € (Ercanbrack et Knight, 1991).

Chez les espèces animales de rente, tout type de caractère sélectionné peut être affecté par la dépression de consanguinité, et un caractère est d'autant plus affecté par la dépression de consanguinité qu'il est sélectionné de manière intensive (Leroy, 2014). L'intensité de la dépression de consanguinité dépend également de la structure génétique des caractères et donc des gènes impliqués et de leur mode d'action (Curik et al., 2001).

La dépression de consanguinité peut donc réduire le revenu des éleveurs en altérant la valeur des caractères d'intérêt économique. Il est alors nécessaire de contrôler le taux de consanguinité dans les programmes de sélection. Il ne s'agit cependant pas de la seule manière de limiter l'impact de la dépression de consanguinité. Il est également pertinent de mieux comprendre les mécanismes

génétiques impliqués dans la dépression de consanguinité, dans le but de limiter ses conséquences délétères, non seulement à l'échelle des individus, par la maîtrise du taux de consanguinité, mais aussi à l'échelle du génome, en commençant par identifier les régions génomiques impliquées dans ce phénomène.

3 - La détection de la dépression de consanguinité

a - A l'échelle individuelle

De manière générale, la dépression de consanguinité et son amplitude (dépendant du fardeau génétique) sont évaluées en mesurant la vitesse à laquelle la valeur du caractère d'intérêt diminue quand le coefficient de consanguinité augmente (Charlesworth et Willis, 2009) (voir Figure 1). Dans les espèces d'animaux de rente, cette évaluation consiste à réaliser la régression de la performance individuelle sur le coefficient de consanguinité individuel. Ainsi, rien qu'en partant de ce principe, il existe de nombreuses façons de détecter la dépression de consanguinité :

(i) En fonction du **type de régression** réalisée (linéaire ou non),

(ii) En fonction du **type de coefficient de consanguinité** utilisé (à partir des pedigrees ou des données de génotypes, ancienneté de la consanguinité).

i - Les différents types de régression entre performance et consanguinité

Selon la prédominance de l'une des hypothèses présentées en Chapitre Annexe 1 pour expliquer les mécanismes génétiques impliqués dans la dépression de consanguinité, cette dernière est supposée être une fonction linéaire (hypothèses de dominance partielle ou de superdominance) ou non linéaire (hypothèse d'épistasie) de la consanguinité (Croquet et al., 2007; Kristensen et Sørensen, 2005).

En effet, lorsque les effets des allèles aux différents locus se combinent de manière additive, indépendamment du fait que la dominance soit partielle ou complète, la relation entre consanguinité et valeur génétique est censée être linéaire. Ce n'est cependant pas le cas dans le cadre d'interactions épistatiques car la valeur génétique peut par exemple diminuer proportionnellement au carré de la consanguinité. Des régressions logarithmiques peuvent également être utilisées dans certains cas.

Des écarts plus ou moins importants à la linéarité ont été observés dans les espèces ovines et bovines selon les caractères étudiés (Silva et al., 2019; Wiener et al., 1992). Cette déviation à la linéarité peut être due à des artefacts statistiques tout comme à de l'épistasie (Lynch, 1988).

Enfin, la consanguinité affecte la variance des caractères quantitatifs (Abney et al., 2000). Cela peut introduire un biais dans l'analyse de régression, qui suppose l'homogénéité des variances

(Burdon et Russell, 1998). De plus, une variance faible de la consanguinité dans la population peut conduire à une puissance statistique faible de l'analyse de régression (Restoux et al., 2012 ; Keller et al., 2011). En pratique, différentes études sur les espèces d'animaux de rente prenant en compte à la fois des modèles linéaires et non linéaires (Carrillo et Siewerdt, 2010; Croquet et al., 2007) ont conclu que, pour des valeurs de consanguinité faibles à modérées, les modèles linéaires étaient adaptés pour détecter la dépression de consanguinité.

ii - Quelles estimations de la consanguinité utiliser pour détecter la dépression de consanguinité ?

La procédure standard pour détecter et évaluer la dépression de consanguinité consiste à réaliser la régression de la performance individuelle sur le coefficient de consanguinité individuel estimé à partir des pedigrees (Curik et al., 2001). Toutefois, les estimations de la consanguinité à partir des pedigrees peuvent être imprécises, à cause de l'inexactitude et de la non-exhaustivité des pedigrees (Leroy et al., 2012 ; Engelsma et al., 2011 ; Leroy, 2011 ; Boichard et al., 1997). Ainsi, l'utilisation des pedigrees pour estimer la consanguinité peut mener à une réduction de la précision de l'estimation de la dépression de consanguinité (Cassell et al., 2003). Certaines corrections des coefficients de consanguinité estimés à partir des pedigrees permettent a priori de surmonter la sous-estimation de la dépression de consanguinité due à ces défauts (Santana et al., 2012 ; Gómez et al., 2009 ; Gutiérrez et al., 2009 ; González-Recio et al., 2007).

La consanguinité estimée à partir des pedigrees suppose l'identité par descendance (IBD) pour un nombre infini de locus non sélectionnés et donc l'homogénéité de la distribution de la consanguinité le long du génome. Elle mesure la consanguinité attendue des individus. Elle ne réagit donc pas à la sélection ou au déséquilibre de liaison comme les mesures réalisées à partir de l'hétérozygotie ou de l'autozygotie réalisée (estimée à partir des *runs of homozygosity* ou ROH, voir la partie consacrée aux ROH dans le chapitre « Mesurer la diversité génétique au sein d'une race », page 152) (Curik et al., 2001).

L'utilisation de coefficients de consanguinité estimés à partir des ROH permet de passer outre ces limites. De plus, ils autorisent l'estimation de la consanguinité le long du génome sur des segments chromosomiques de la taille souhaitée, et leur longueur est liée à l'âge de la consanguinité de laquelle ils sont issus. Bien que leur détection ne soit pas complètement standardisée, ces propriétés en font des candidats de choix pour estimer la consanguinité dans le cadre de l'étude de la dépression de consanguinité (Keller et al., 2011).

En plus d'estimer la consanguinité à partir de différents types de données, il est possible d'estimer l'impact de la consanguinité ancienne et de la consanguinité récente séparément sur la dépression de consanguinité. Une étude récente sur la race bovine Holstein a estimé la dépression de

consanguinité due à la composante ancestrale de la consanguinité à partir des pedigrees et des ROH courts, et celle due à la composante récente de la consanguinité à partir des pedigrees et des ROH longs (Doekes et al., 2019). Pour des caractères de production laitière, elle a montré que la dépression de consanguinité due à une consanguinité récente estimée à partir des pedigrees était beaucoup plus forte que celle due à une consanguinité ancestrale estimée à partir des pedigrees, avec une baisse de 2,42 kg contre une baisse de 0,03 kg de matière grasse lors d'une lactation pour une augmentation de 1 % de consanguinité. Aucune différence significative n'a été mise en évidence entre l'effet des ROH longs (consanguinité récente) et des ROH courts (consanguinité ancienne) dans cette étude. Dans d'autres études, les ROH longs étaient associés avec une baisse de production laitière chez les vaches mais pas les ROH courts (Pryce et al., 2014).

Toutes les méthodes d'analyse de la dépression de consanguinité comme une régression de la valeur de caractères d'intérêt sur la consanguinité globale, qu'elle soit calculée à partir des pedigrees ou des ROH, supposent que l'autozygotie est répartie uniformément le long du génome. Or, des preuves expérimentales ont été apportées pour montrer qu'il s'agit d'une hypothèse irréaliste, en particulier dans les populations sélectionnées, et que les niveaux d'autozygotie varient fortement le long du génome (Curik et al., 2017; Pemberton et al., 2012). L'autozygotie à l'échelle d'un SNP, d'une région du génome ou d'un chromosome est ainsi faiblement corrélée avec les coefficients de consanguinité, quelle que soit la méthode d'estimation. En estimant la consanguinité à l'échelle globale du génome, il n'est pas possible de prendre en compte cette hétérogénéité. En particulier, deux individus peuvent avoir exactement le même coefficient de consanguinité tout en ayant une autozygotie répartie très différemment le long du génome. Ils n'exprimeront donc pas les mêmes baisses individuelles de valeur des caractères d'intérêt par rapport à des individus moins consanguins.

b - A l'échelle génomique

Les développements récents de la génomique ont permis d'envisager d'autres approches de la dépression de consanguinité et du fardeau génétique. Plusieurs types de méthodes ont ainsi été développés pour détecter les régions du génome participant au fardeau génétique.

- Détecter le fardeau génétique par des **études d'association pangénomique (GWAS)**

Certaines méthodes de détection du fardeau génétique à partir de GWAS se basent sur des modèles de dominance (Doekes, 2020). D'autres méthodes prennent appui sur la détection des ROH des animaux de la population étudiée. Ensuite, à chaque position, c'est-à-dire à chaque SNP, une régression linéaire est effectuée pour expliquer les performances des individus à partir des effets fixes classiquement pris en compte (troupeau, année de naissance, etc.), du génotype à ce SNP ainsi que du statut ROH du SNP, c'est-à-dire : le SNP est-il inclus dans un ROH ou non pour

cet individu, autrement dit, l'animal est-il consanguin à ce SNP (Ferenčaković et al., 2017; Howard et al., 2015; Pryce et al., 2014) ? Il est possible de prendre en compte plus de nuances en passant de deux statuts ROH (présence/absence de ROH) à des catégories plus précises, par exemple : absence de ROH, présence d'un ROH d'une catégorie de longueur donnée, avec autant de catégories de longueur qu'il est pertinent d'en inclure.

Plusieurs études ont ainsi montré une association entre la présence de ROH à certains SNPs et la baisse de la valeur de certains caractères d'intérêt chez des bovins laitiers (production laitière, fertilité des mâles et femelles) (Ferenčaković et al., 2017; Howard et al., 2015; Pryce et al., 2014). Ces SNPs participent probablement au fardeau génétique. Il s'est avéré que la covariance entre la présence de ROH à certains SNPs et la valeur des caractères d'intérêt n'était pas forcément négative, son signe étant hétérogène le long du génome (Howard et al., 2015). Être consanguin dans certaines régions du génome présenterait alors un avantage. Cela peut être le signe de signatures de sélection (Kim et al., 2013). Les SNPs participant au fardeau ayant été détecté ne sont pas les mêmes selon les races, indiquant une hétérogénéité du fardeau génétique entre les races (Howard et al., 2015; Pryce et al., 2014).

Il est intéressant de noter que les animaux ayant fait l'objet de ces études sont des animaux possédant des performances enregistrées sur plusieurs années. Les positions du génome ainsi détectées comme faisant partie du fardeau génétique ont donc un impact délétère sur les performances des animaux mais n'entraînent pas leur mort et n'ont pas un impact individuel très fort sur les animaux (Howard et al., 2015). En effet, si des allèles récessifs létaux ou à effet majeur se trouvent à l'état homozygote, leur expression ne permettra pas à l'animal les portant d'avoir une carrière normale au sein d'un troupeau (mort embryonnaire, fœtale ou prématurée, santé très fragile, infertilité forte).

- Détecter le fardeau génétique par l'**étude des haplotypes**

Les allèles récessifs létaux sont l'une des causes génétiques de l'inefficacité de la reproduction, car ils entraînent la mort lorsqu'un embryon est homozygote. Ils peuvent également réduire la viabilité des embryons et des animaux vivants en cas de pénétrance incomplète. L'expression de ces allèles se traduit par des taux de conception réduits, des intervalles entre les vêlages allongés, ou une survie plus faible pour les animaux nés vivants. Ces allèles sont généralement rares, mais si leur fréquence augmente, le nombre d'embryons homozygotes affectés augmente, tout comme la perte économique associée. La fréquence des allèles récessifs létaux peut augmenter dans une population en raison de la dérive génétique, du lien avec des allèles favorables (déséquilibre de liaison) et de l'avantage des hétérozygotes (hypothèse de superdominance). Cela conduirait à la sélection de porteurs hétérozygotes comme parents de la prochaine génération, propagerait

l'allèle létal récessif dans toute la population et, à son tour, affecterait considérablement la reproduction.

Il est possible de détecter les régions portant ces allèles récessifs létaux ou sub-létaux (causant une viabilité faible des jeunes animaux) en s'intéressant aux régions du génome présentant un déficit ou une absence d'homozygotes. En effet, si tous les animaux homozygotes pour un allèle ou presque meurent avant même leur naissance, ou très jeunes, on observera un déficit d'animaux homozygotes pour cet allèle dans la population par rapport à l'attendu.

Cette approche est généralement élargie à l'étude d'haplotypes et il existe plusieurs façons d'identifier les haplotypes pour lesquels il existe un déficit en homozygotes (Jenko et al., 2019 ; Fritz et al., 2013 ; VanRaden et al., 2011). Identifier ces haplotypes à effet létal peut être une première étape dans l'identification d'anomalies génétiques, et permet aussi d'identifier les porteurs de ces haplotypes parmi les candidats à la reproduction pour éviter l'augmentation de la fréquence de l'haplotype létal dans la population, ou au moins la maîtriser (par exemple, si on veut garder l'avantage des hétérozygotes dans les cas de superdominance).

On ne s'attend donc pas nécessairement à trouver des régions communes entre ces méthodes de recherche des déficits en homozygotes et celles utilisant des GWAS.

Certaines approches de détection du fardeau génétique cherchent le lien entre les haplotypes et l'impact de la présence dans un ROH pour chacun de ces haplotypes sur les performances (Howard et al., 2017). L'objectif n'est pas cette fois de découvrir des haplotypes récessifs létaux mais de trouver des haplotypes récessifs délétères, donc avec un effet plus faible, n'empêchant pas les animaux d'être productifs. Un exemple de telles méthodes (Howard et al., 2017) consiste à identifier les haplotypes défavorables, en scannant le génome sur la base de fenêtres de taille décroissante (en nombre de SNPs) et en identifiant les haplotypes pour lesquels la présence de ROH est associée à un phénotype défavorable, jusqu'à des haplotypes de la plus petite longueur possible.

Ces méthodes, bien que performantes, et permettant de détecter non seulement les régions portant le fardeau génétique mais également les haplotypes responsables de la dépression de consanguinité, sont très gourmandes en temps de calcul et en mémoire.

- Détecter le fardeau génétique par l'**étude de l'autozygotie par régions du génome**

Si l'objectif de la détection du fardeau génétique est de savoir dans quelles régions du génome se situent les allèles délétères constituant le fardeau, sans forcément s'intéresser aux haplotypes responsables, il n'est pas nécessaire d'utiliser de telles approches. En revanche, il est possible de

garder la notion de régression par fenêtre du génome entre la consanguinité ROH et les performances.

Ce type de méthode consiste à faire coulisser une fenêtre d'une longueur donnée le long du génome et pour chaque fenêtre, effectuer la régression des performances des animaux par des effets fixes classiques et par la consanguinité locale sur cette fenêtre à partir des ROH (voir Figure 24 page 154). On obtient alors l'effet de la consanguinité sur les performances pour chaque région du génome, correspondant au coefficient de la régression linéaire, c'est-à-dire à la pente du modèle. Ce type d'approche a déjà été utilisé dans l'étude de maladies humaines comme la schizophrénie (Keller et al., 2012). Cela permet d'identifier les régions impliquées dans la dépression de consanguinité et de savoir quelle est l'ampleur de l'effet de la consanguinité dans ces régions. Ce type de méthode est simple à mettre en place et demande des temps de calcul comparables aux GWAS (avec les mêmes corrections et les mêmes données).

Tout comme les approches de type GWAS, la présence de ROH semble favorable au phénotype d'intérêt dans certaines régions du génome (Keller et al., 2012). De plus, le nombre de régions impliquées dans l'expression de la dépression de consanguinité est grand (Keller et al., 2012 ; Ayroles et al., 2009), avec des effets variables de la consanguinité le long du génome.

La valeur des performances des hétérozygotes pour un haplotype délétère et un haplotype sain est supérieure à la moyenne des valeurs des deux types d'homozygotes, dans les cas de dominance (partielle ou complète) ou de superdominance de l'haplotype sain. Dans ces deux cas, la pente de régression sera a priori négative (voir Figure 2).

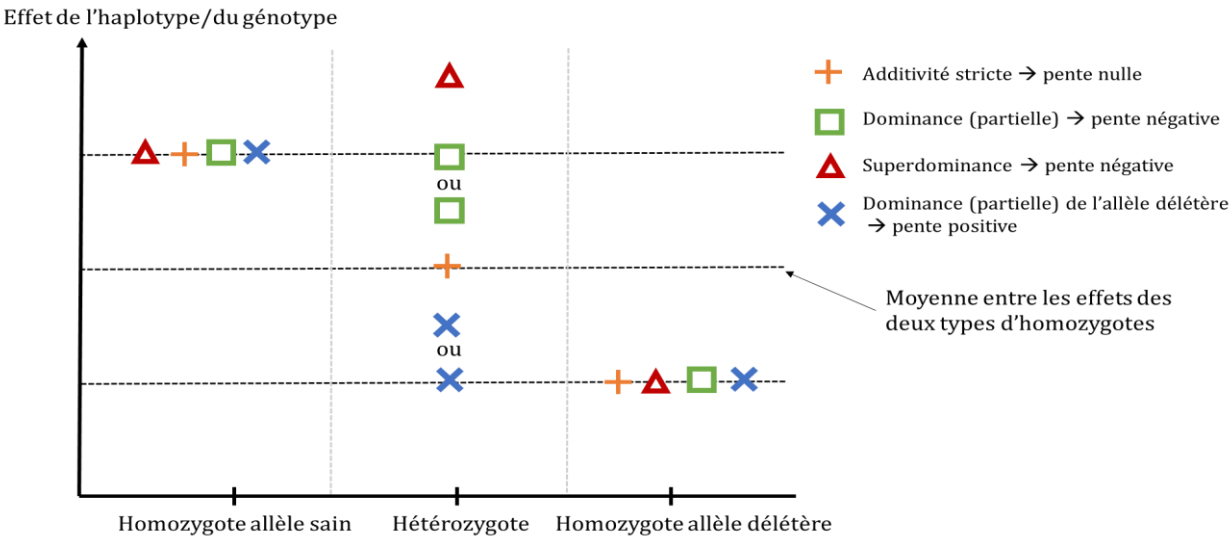

Figure 2 Comment interpréter le signe du coefficient de régression (= pente) dans l'étude de la dépression de consanguinité le long du génome ?

La détection des régions du génome portant le fardeau génétique ou tout du moins, celles au sein desquelles la consanguinité a un effet délétère, permettrait à terme de prendre en compte ces régions différentiellement dans les évaluations génomiques ou dans les plans d'accouplement pour limiter l'expression du fardeau génétique et donc la dépression de consanguinité.

4 - Bilan

La **dépression de consanguinité** correspond à la réduction de la valeur phénotypique moyenne des caractères liés à la *fitness* (valeur sélective) quand la consanguinité augmente dans une population. La dépression de consanguinité affecte également les caractères de production (production laitière, croissance, etc.). La dépression de consanguinité est due au **fardeau génétique**, constitué par l'ensemble des allèles délétères, le plus souvent récessifs, portés par le génome des animaux à l'échelle de la population. Plus le fardeau génétique est grand, plus la dépression de consanguinité est forte. La dépression de consanguinité peut réduire le revenu des éleveurs en altérant la valeur des caractères d'intérêt économique.

Il est donc pertinent de mieux comprendre les mécanismes génétiques impliqués dans la dépression de consanguinité, dans le but de limiter ses conséquences délétères, non seulement à l'échelle des individus, par la **maîtrise du taux de consanguinité**, mais aussi à l'échelle du génome, en commençant par **identifier les régions génomiques impliquées** dans ce phénomène.

Dans les espèces d'animaux de rente, l'évaluation de la dépression de consanguinité consiste à réaliser la régression de la performance individuelle sur le coefficient de consanguinité individuel. Il existe de nombreuses façons de détecter la dépression de consanguinité en fonction du type de régression réalisée (linéaire ou non), mais aussi du type de coefficient de consanguinité utilisé (à partir des pedigrees ou des données de génotypes, ancienneté de la consanguinité).

Les développements récents de la génomique ont permis d'envisager d'autres approches de la dépression de consanguinité et du fardeau génétique. Plusieurs types de méthodes ont ainsi été développés pour détecter les régions du génome participant au fardeau génétique, la plupart utilisant les ROH comme indicateurs de l'autozygotie le long du génome.

D - Bilan sur la diversité génétique dans les espèces domestiques sélectionnées

Les notions abordées dans cette première partie à propos de la diversité génétique dans les espèces domestiques sélectionnées sont résumées dans la Figure 3. Cette partie nous a permis de mettre en évidence que la sélection dans les espèces animales de rente met à mal la diversité génétique. Il est donc nécessaire d'étudier la diversité génétique dans ces espèces et de la gérer, afin de maintenir le progrès génétique et la santé des populations à plus ou moins long terme.

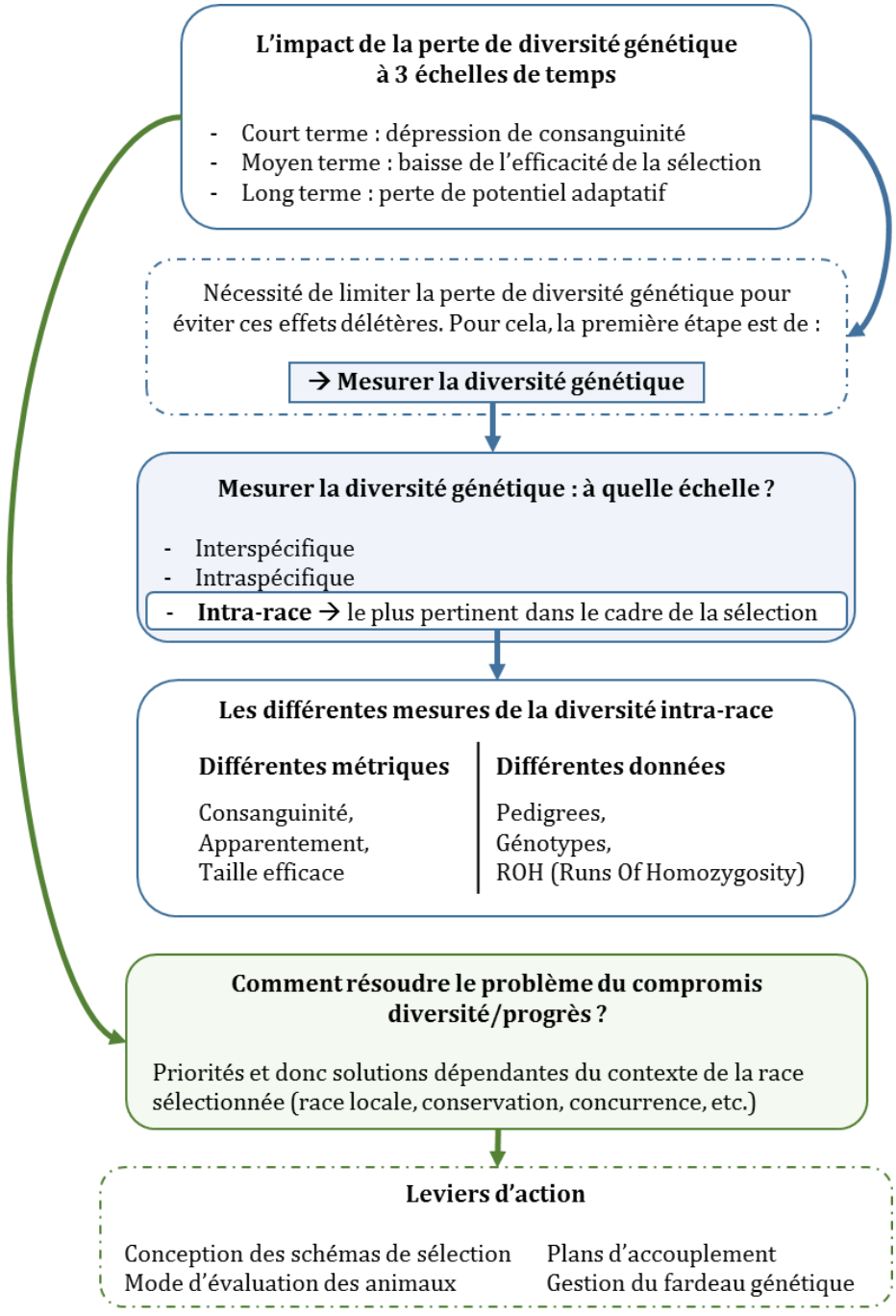

Figure 3 Schéma bilan : la diversité dans les espèces domestiques sélectionnées

II - Sélection et diversité dans les races bovines laitières françaises

Dans le cadre de ma thèse, je me suis intéressée à la diversité génétique au sein des races bovines laitières de grand effectif. Ces populations de grande taille présentent une diversité génétique faible (Danchin-Burge, 2020) et une importance économique élevée. La disponibilité d'une quantité importante de données ainsi que la mise en place précoce de la sélection génomique par rapport aux autres espèces et l'essor récent de l'utilisation du transfert embryonnaire dans ces races en font des objets d'étude pertinents pour analyser le lien entre la diversité génétique, les schémas de sélection et leurs changements.

A - Pourquoi s'intéresser à la diversité génétique intra-race chez les races bovines laitières françaises ?

Il est possible d'estimer la diversité des populations animales à différentes échelles, plus ou moins pertinentes pour évaluer la variabilité génétique dans le cadre de la sélection animale.

La diversité peut être évaluée en s'intéressant directement à la variabilité génétique de l'espèce considérée et, dans le cas des espèces sélectionnées, à la variété de races constituant l'espèce. A l'échelle de l'espèce, il existe peu d'espèces de rente avec une variabilité génétique assez faible pour poser un problème, en dehors de l'âne par exemple, espèce chez laquelle la diversité des races est minimale en comparaison des autres espèces de rente (FAO, 2007).

Or, chez les bovins laitiers en France, la sélection s'effectue exclusivement à l'échelle de la race, pas à celle de l'espèce. En effet, l'évaluation des animaux reproducteurs élite y est effectuée uniquement en intra-race depuis le milieu du XXe siècle et il n'existe pour le moment pas d'évaluation des animaux croisés ou des animaux non contemporains dont la semence est conservée en cryobanque.

De plus, si le nombre de races bovines laitières est élevé (12 races en programme officiel de sélection et cinq en programme de conservation) (France Génétique Elevage), trois races seulement constituent une très grande part de l'effectif total de bovins laitiers en France : la Prim'Holstein, la Normande et la Montbéliarde, représentant environ 3,316 millions des 3,714 millions de vaches laitières en France, soit plus de 89 % (Groupe Economie du Bétail, 2019). Une diversité intra-race suffisante est une condition sine qua non à sa conservation et donc à la conservation de la diversité inter-race.

Il est donc nécessaire de limiter la perte de diversité à l'échelle intra-race au sein des races bovines laitières en France, pour lesquelles la sélection elle-même s'effectue principalement en intra-race.

B - L'élevage et la sélection des vaches laitières en France

1 - Présentation générale

La filière bovine laitière française présente une importance économique élevée. En effet, la France est le 2ème producteur européen de lait après l'Allemagne, avec 23,8 milliards de litres de lait de vache collectés par an (FranceAgriMer, 2020). Elle est également le 8ème producteur de lait dans le monde (CNIEL, 2020). Elle est aussi le premier consommateur de beurre et de fromage dans le monde (IDF, 2019). La production de valeur du produit lait en France est de 10,5 milliards d'euros par an, soit 14 % de la production agricole totale (Insee, 2019).

La France compte plus de 53 700 exploitations laitières livrant du lait de vache (Agreste, 2019), dans lesquelles vivent 3,55 millions de vaches laitières (CNIEL, 2020). Le nombre d'exploitations livrant du lait de vache décroit depuis plusieurs décennies (plus de 120 000 exploitations en 2000 (CNIEL, 2020)) mais la taille des exploitations laitières est en expansion ces dernières années. Une ferme laitière française comportait en moyenne 66 vaches laitières en 2018 contre 35,9 en 2000 (CNIEL, 2020) et livre en moyenne 443 500 litres de lait par an, soit 6 800 litres de lait en moyenne par vache laitière (Agreste, 2019).

Une exploitation bovine laitière est généralement constituée de vaches de race pure. Dans les races bovines laitières en France, l'éleveur est également sélectionneur. C'est en effet l'éleveur qui, à l'échelle de son élevage, choisit les génisses qui vont participer au renouvellement de son troupeau, ou encore les accouplements qu'il souhaite effectuer entre ses vaches ou génisses et des taureaux (présents dans l'exploitation et/ou via l'utilisation de l'insémination artificielle (IA) pour des taureaux extérieurs à l'exploitation). En France, en 2019, 6 859 624 IA ont été réalisées au total sur des génisses et vaches laitières, dont 3 719 701 IA premières (Institut de l'Elevage, 2020a). En France, 75 % des veaux en élevage laitier sont nés de l'insémination (Allice, 2020).

Le processus de sélection au sein d'un élevage bovin laitier français de race pure utilisant majoritairement l'IA est décrit dans la Figure 4 ci-dessous. L'éleveur choisit avec quels taureaux mis en marché par les entreprises de sélection il souhaite accoupler ses vaches via IA, en achetant des doses de semence.

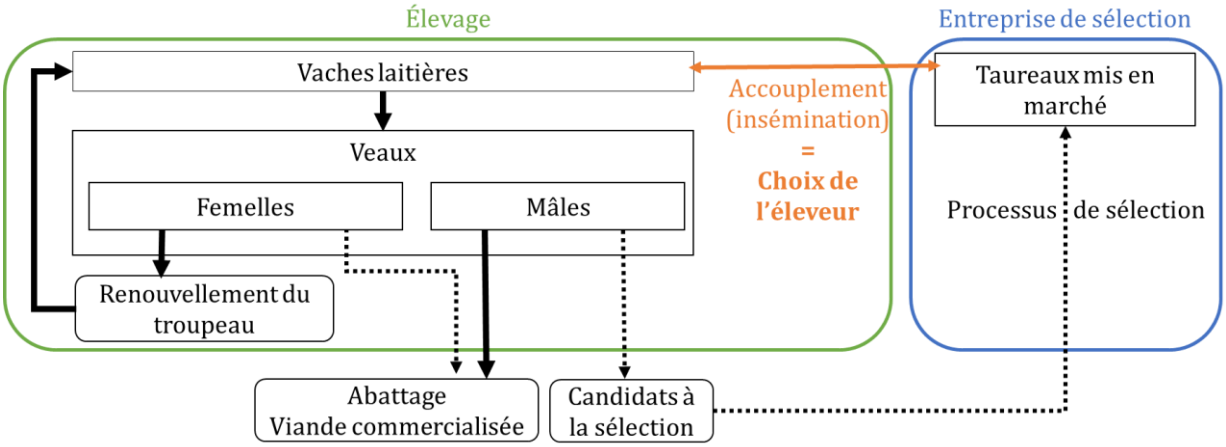

Figure 4 La sélection au sein d'un élevage bovin laitier en France

L'éleveur obtient alors un certain nombre de veaux. Les veaux femelles vont servir au renouvellement du troupeau suite à la réforme des vaches laitières en fin de carrière (baisse de production, blessures, etc.), être vendues ou destinées à l'abattoir. Les veaux mâles seront majoritairement destinés à l'abattoir (après engraissement ou non, au sein de l'élevage ou après avoir été vendus) mais certains vont devenir candidats à la sélection après repérage par les entreprises de sélection. Il arrive qu'une entreprise de sélection, par le biais d'un contrat avec l'éleveur, choisisse certains accouplements avec les vaches et génisses de l'éleveur afin d'obtenir des veaux mâles candidats de manière plus ciblée. Après un processus de sélection, une partie de ces veaux deviendront les futurs taureaux dont la semence sera mise en marché par les entreprises de sélection.

Le processus de sélection des taureaux au sein d'une entreprise de sélection est schématisé dans la Figure 5, avec à gauche en bleu le processus de sélection génétique (basée sur des évaluations polygéniques, à partir des performances des descendants des taureaux), et à droite en rouge le processus de sélection génomique (basée sur des évaluations génomiques).

Dans le cadre de la sélection génétique classique, les veaux candidats à la sélection sont sélectionnés en fonction de la moyenne des valeurs génétiques de leurs parents (mères et pères à taureaux), on parle alors d'évaluation sur ascendants, et entrent en station de testage, au sein de l'entreprise de sélection. Ils produisent alors des filles, qui produisent chacune au moins un veau avant de pouvoir enregistrer leurs performances : production laitière, morphologie, etc. Les caractères de production laitière ou de morphologie de la mamelle n'étant pas exprimés chez les mâles, la valeur génétique d'un mâle sera alors déterminée à partir des performances de ses filles, c'est l'évaluation sur descendance. C'est un processus long : deux générations, soit 5 ou 6 ans, sont nécessaires à l'obtention d'une valeur génétique d'une précision satisfaisante pour choisir les

taureaux qui seront mis en marché et/ou deviendront pères à taureaux. De par sa longueur, ce processus est très coûteux et ne permet pas de tester de nombreux taureaux, ce qui limite le nombre de candidats à la sélection.

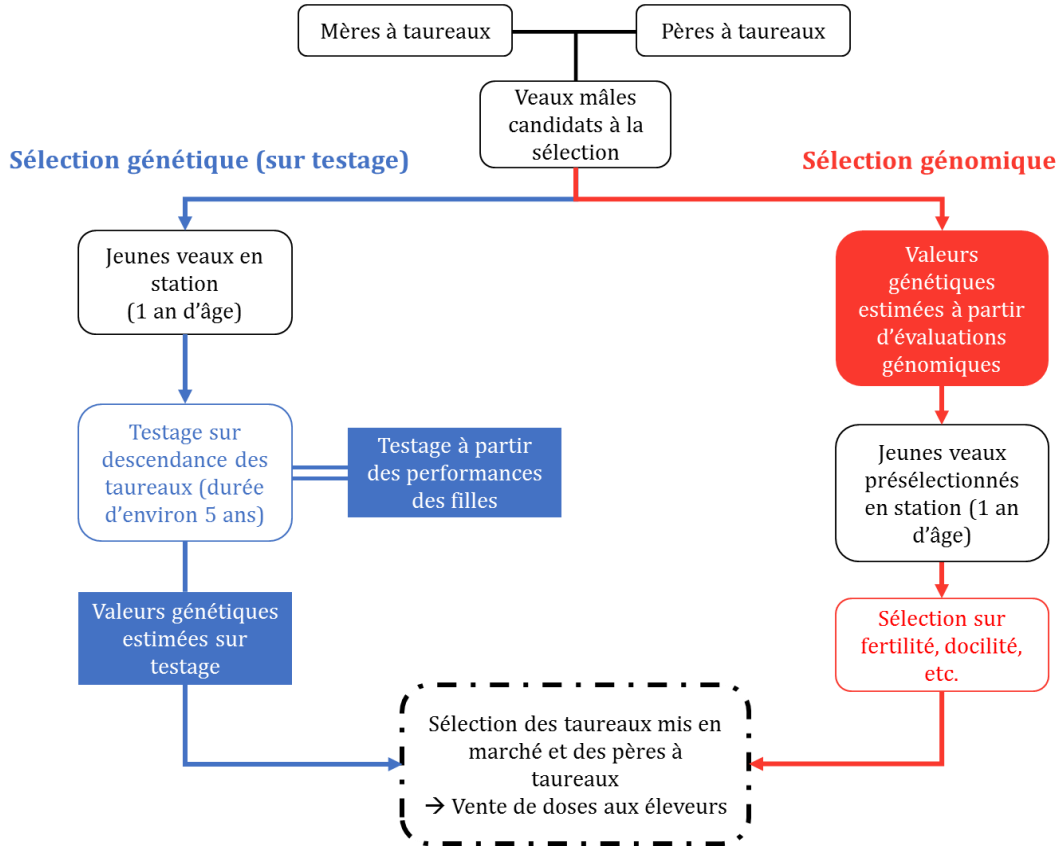

Figure 5 Les différences entre les schémas de sélection génétique et de sélection génomique dans le processus de sélection des taureaux laitiers en France

Les évaluations génomiques ont été mises en place en France en 2009 dans les races bovines laitières de grand effectif (Prim'Holstein, Montbéliarde et Normande) et en 2016 dans les races de plus petit effectif. Jusqu'en octobre 2018, c'est l'INRA (devenue INRAE en janvier 2020) qui était en charge de ces évaluations. L'évaluation des animaux était donc centralisée et public. Depuis novembre 2018, c'est à chaque organisme de sélection d'effectuer les évaluations, suite à la mise en application du nouveau règlement zootechnique de l'Union Européenne (RZUE, règlement (UE) 2016/1012 relatif aux conditions zootechniques et généalogiques applicables à l'élevage, aux échanges et à l'entrée dans l'Union de reproducteurs de race pure, de reproducteurs porcins hybrides et de leurs produits germinaux) (Ministère de l'agriculture, de l'agroalimentaire et de la forêt, 2016). Les évaluations sont donc désormais décentralisées et privées, ce qui peut faire baisser la taille des populations de référence à partir desquelles les animaux sont évalués et donc

faire baisser la précision des évaluations. En France, c'est l'entreprise GenEval qui a été chargée par les organismes de sélection d'effectuer les évaluations.

Dans le cadre de la sélection génomique, de nombreux veaux mâles sont génotypés et donc évalués quasiment dès leur naissance avec une très bonne précision. Les meilleurs veaux entrent alors en station. Les veaux en station sont ensuite sélectionnés sur des critères comme leur fertilité ou leur docilité et leur semence pourra alors être commercialisée à un très jeune âge (environ un ou deux ans, dès la maturité sexuelle atteinte). Ce fonctionnement rend le processus de sélection des taureaux beaucoup plus court et moins coûteux, et permet potentiellement d'augmenter fortement le nombre de veaux candidats à la sélection à coût fixé (de Roos et al., 2011 ; Hayes et al., 2009). La sélection génomique permettrait donc d'augmenter le progrès génétique dans les races bovines laitières françaises (de Roos et al., 2011).

La mise en place des évaluations génomiques a donc bouleversé l'organisation de la sélection des taureaux mis en marché et des pères à taureaux en augmentant le nombre de veaux mâles candidats et en réduisant drastiquement l'âge de mise en marché des taureaux et donc l'intervalle de génération dans les races bovines laitières françaises.

Dès 2015 dans les races bovines laitières françaises de grand effectif, l'essor de l'utilisation d'autres technologies a perturbé l'organisation de la sélection au sein des entreprises de sélection comme les technologies de transfert embryonnaire. L'utilisation de routine de ces technologies a modifié la sélection et l'utilisation des mères à taureaux.

Les modifications provoquées par l'utilisation du transfert embryonnaire sont présentées dans la Figure 6. Parmi les génisses naissant dans les élevages, les meilleures sont sélectionnées pour entrer en station de donneuses dans les entreprises de sélection. Ces génisses deviennent alors des donneuses d'embryons (ou d'ovocytes ensuite fécondés *in vitro*) de haut niveau génétique qui seront implantés dans l'utérus de vaches receveuses, généralement de niveau génétique inférieur. En 2018, 28 411 embryons produits *in vivo* ont, par exemple, été implantés sur des vaches laitières receveuses en France (AETE, 2019). Ces technologies permettent une intensification de l'utilisation des mères à taureaux car une femelle peut produire environ 30 veaux, voire plus, par an grâce au transfert embryonnaire (Merton et al., 2003), quand elle n'en produit que 3 à 4 sur toute sa carrière reproductive sans transfert embryonnaire.

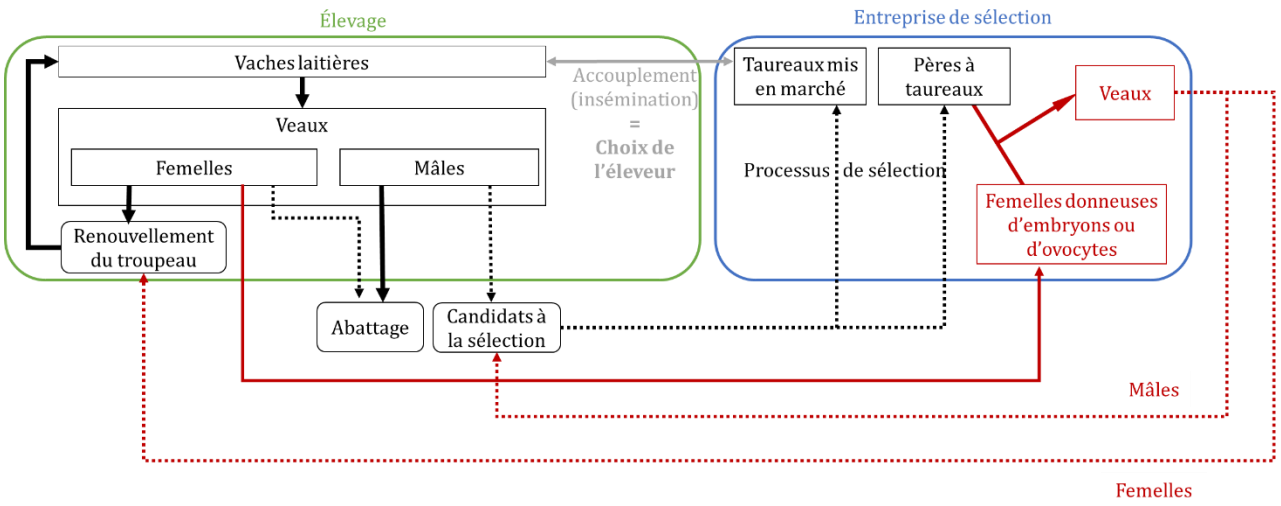

Figure 6 Les changements dans les schémas de sélection provoqués par la mise en place du transfert embryonnaire (à partir de don d'embryons ou d'ovocytes)

Le transfert embryonnaire (TE) permet donc d'augmenter les chances d'obtenir des veaux de très haut niveau génétique et donc des meilleurs taureaux, pères à taureaux et mères à taureaux. En association avec l'utilisation de la sélection génomique, les technologies permettant le transfert embryonnaire sont très bénéfiques pour le progrès génétique des races bovines laitières (Granleese et al., 2019 ; Thomasen et al., 2016 ; Pedersen et al., 2012 ; Pryce et al., 2010), grâce à une **augmentation de l'intensité de sélection sur la voie femelle** (Land et Hill, 1975) et une diminution de l'intervalle de génération. Certains des procédés de récolte d'ovocytes permettent une réduction extrême des intervalles de génération car réalisables sur des génisses immatures dès 2 mois d'âge (Galli et al., 2001).

Notons que le fonctionnement de l'élevage bovin laitier français est organisé dans un système de race pure, avec une forte implication des éleveurs dans la sélection des taureaux. Il n'y a pas de séparation réelle entre la population générale et le schéma de sélection, car l'ensemble des animaux d'une race font partie de la population de référence de cette race et participent à la production des futurs reproducteurs.

Dans les schémas de sélection des bovins laitiers en France, les points à risque pour la diversité génétique se situent principalement dans les effectifs d'animaux à chaque étape de la sélection (nombre de candidats à la sélection, nombre de veaux en station, nombre de taureaux, nombre de pères et de mères à taureaux) mais également dans l'apparentement entre les animaux sélectionnés et accouplés.

2 - Les races bovines laitières françaises de grand effectif

Les trois races bovines laitières de grand effectif que j'ai étudiées durant ma thèse sont la Montbéliarde (principalement) (voir Figure 7), la Normande (voir Figure 8) et la Prim'Holstein (voir Figure 9). Le Tableau 1 regroupe les principales caractéristiques de ces races : effectifs, production, ainsi que leur a consanguinité et leur taille efficace établies à partir des pedigrees.

Figure 7 Vaches de race Montbéliarde (crédit photo : Clémence Gaudebout)

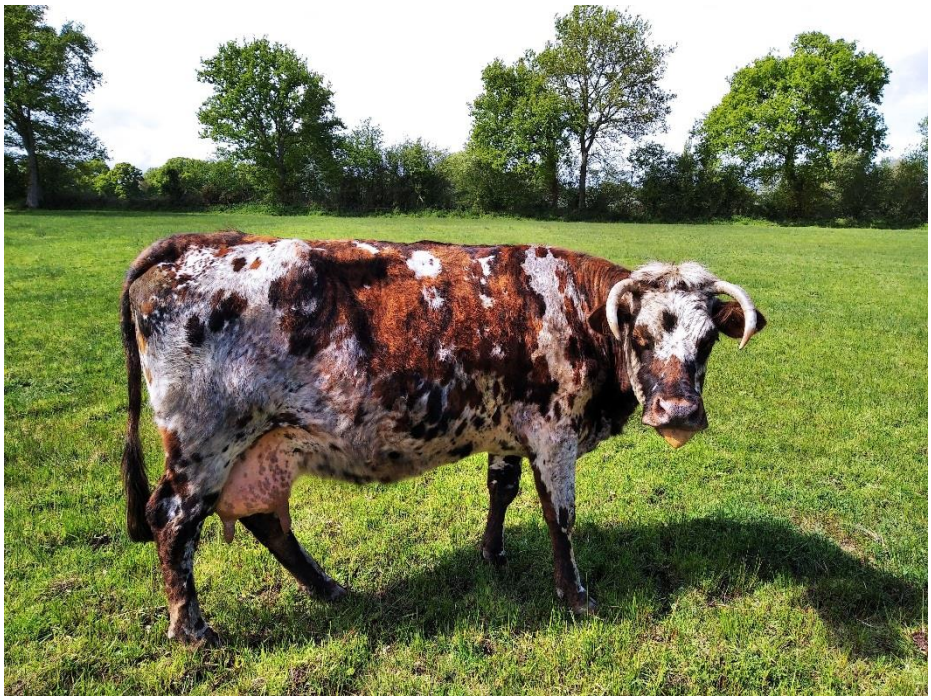

Figure 8 Vache de race Normande (crédit photo : Fanny Mollandin)

Figure 9 Vaches de race Prim'Holstein (crédit photo : Clémence Gaudebout)

Tableau 1 Présentation des trois races bovines laitières de grand effectif étudiées dans cette thèse : effectifs, diversité et production

	Montbéliarde	Normande	Prim'Holstein	Référence
Nombre total de femelles en France (génisses + vaches laitières) (au 1er janvier 2018)	628 000	319 000	2 402 000	(Groupe Economie du Bétail, 2018)
Nombre de vaches laitières en France (en 2017)	430 186	194 524	1 572 614	(Thomas et al., 2018)
Consanguinité moyenne estimée à partir des pedigrees pour les femelles nées en France entre 2016 et 2019 (en %)	5,1	5,2	4,7	(Danchin-Burge, 2020)
Taille efficace (pour les animaux nés entre 2016 et 2019)	81	89	96	(Danchin-Burge, 2020)
Production laitière moyenne par vache par lactation en France en 2017 (en kg de lait)	6 933	6 590	9 241	(Thomas et al., 2018)
Taux butyreux/protéique du lait en France en 2017 (en g/kg de lait)	38,9/33,1	42,4/34,7	39,7/32,0	(Thomas et al., 2018)

Les objectifs de sélection de ces trois races sont combinés pour chacune dans un Index de Synthèse Unique (ISU) dédié, incluant des caractères de production, de morphologie et des caractères fonctionnels (santé de la mamelle, longévité, fertilité), et pondérés différemment selon les races en fonction des objectifs de sélection et des corrélations génétiques. Ces pondérations sont présentées dans le Tableau 2.

Tableau 2 Composition actuelle de l'Index de Synthèse Unique (ISU) dans les races Montbéliarde, Normande et Prim'Holstein en France

Race / Caractère	Montbéliarde	Normande	Prim'Holstein
Production laitière	45 %	26 %	35 %
Morphologie (incluant les aptitudes bouchères)	12,50 %	33,75 %	15 %
Santé de la mamelle	14,50 %	16 %	18 %
Reproduction, fertilité	18 %	13,50 %	22 %
Longévité	5 %	6 %	5 %
Vitesse de traite	5 %	4,75 %	5 %
Référence	(Montbéliarde Association)	(Organisme de Sélection en race Normande, 2018)	(Prim'Holstein France, 2020)

La race Montbéliarde est élevée majoritairement dans l'Est de la France (Franche-Comté et Auvergne-Rhône-Alpes) (Races de France, 2014a). La Montbéliarde est sélectionnée pour ses qualités laitières, en particulier à destination de la fabrication de fromages, le plus connu étant le Comté. L'objectif de sélection principal en race Montbéliarde en France est la capacité à produire du lait de qualité en grandes quantités en conservant une bonne résistance aux mammites et un rapport taux butyreux/protéique assez faible. En effet, la part du taux protéique dans la synthèse laitière de l'ISU en Montbéliarde est 3 fois supérieure à celle du taux butyreux (Montbéliarde Association).

La Normande est principalement élevée en Normandie ainsi que dans l'Est de la Bretagne et le Nord des Pays de la Loire (Races de France, 2014b). Le lait des vaches de race Normande est particulièrement riche en matière protéique. Cette race présente une bonne fertilité et une bonne facilité de vêlage. L'objectif de sélection principal est d'augmenter la production laitière tout en maintenant un taux protéique élevé, comme en race Montbéliarde. La Normande étant une race mixte, le poids de la production laitière dans l'ISU est toutefois plus faible qu'en Montbéliarde et en Prim'Holstein, au profit des aptitudes bouchères (voir Tableau 2).

La Prim'Holstein est élevée dans une grande partie de la France, avec des effectifs plus importants en Normandie, Bretagne, Pays de la Loire, Grand-Est et Auvergne, autrement dit les grandes régions laitières françaises (Races de France, 2014c). La Prim'Holstein présente de très grandes capacités de production, c'est d'ailleurs son objectif de sélection principal. Cette race présente une production par vache en France quasiment 50% supérieure à celle des races Montbéliarde et Normande (voir Tableau 1). C'est également une race précoce, permettant des vêlages dès 2 ans et donc des intervalles de génération très courts.

Dans ces trois races, l'utilisation de l'IA en race pure est largement majoritaire (voir Tableau 3). Le nombre de taureaux mis en marché est très inférieur à la taille de la population femelle dans ces races (entre 60 et 200 taureaux mis en marché chaque année selon les races). Les taureaux constituent donc le facteur limitant de la diversité génétique dans ces races.

Ainsi, la diversité génétique dans les races laitières françaises de grand effectif est très dépendante des choix effectués par les entreprises de sélection dans leurs schémas de sélection pour obtenir les futurs taureaux : nombre de taureaux mis en marché et critères de sélection, nombre de pères et de mères à taureaux, apparentement entre les candidats à la sélection, entre les reproducteurs et entre les taureaux sélectionnés et la population générale, etc.

Tableau 3 Utilisation de l'insémination artificielle dans trois races bovines laitières françaises de grand effectif

	Montbéliarde	Normande	Prim'Holstein	Référence
Nombre total d'inséminations effectuées par des taureaux de race pure en France (en 2019)	673 388	433 967	3 353 902	(Institut de l'Elevage, 2020a)
Nombre de premières inséminations effectuées par des taureaux de race pure en France (en 2019)	408 957	243 996	1 865 490	(Institut de l'Elevage, 2020a)

Nous voyons donc que les décisions prises par les entreprises de sélection dans la création de leurs schémas de sélection, et les changements qui peuvent subvenir dans ces derniers, sont cruciaux dans la gestion de la diversité génétique des races bovines laitières françaises. C'est la raison pour laquelle, au cours de ma thèse, j'ai beaucoup travaillé à l'échelle du schéma de sélection, dans la partie entreprise de sélection, afin de comprendre et de prévoir les conséquences de certains changements dans les schémas de sélection sur la diversité génétique des races bovines laitières de grand effectif.

Les races Montbéliarde et Normande sont quasiment exclusivement élevées en France, ce qui fait des organismes de sélection et des entreprises de sélection français responsables de ces races et de leurs schémas de sélection les garants de la diversité génétique de ces deux races dans leur ensemble. La race Prim'Holstein est quant à elle une sous-population de la Holstein, première race laitière mondiale avec plus de 26 millions de femelles dans le monde en 2018 (WHFF, 2019).

Avec moins de 10 % de ses effectifs mondiaux présents en France, la responsabilité qui pèse sur les épaules des organismes et des entreprises de sélection français pour le maintien de la diversité génétique à l'échelle de la race internationale Holstein est plus faible que pour des races nationales. De plus, la concurrence y est plus féroce, poussant à obtenir un progrès génétique au moins aussi important que celui des autres pays pour rester compétitifs.

La mise en place précoce des évaluations génomiques, dès 2009, dans ces trois races permet un recul suffisant pour étudier leur impact sur la diversité et le progrès génétique chez la Montbéliarde, la Normande et la Prim'Holstein. Le transfert embryonnaire est utilisé en routine depuis quelques années. Cela soulève la question de son impact sur la diversité génétique par rapport à l'augmentation du progrès génétique qu'il permet. Les schémas de sélection des races bovines laitières de grand effectif sont donc des objets d'étude intéressants et pertinents pour étudier le lien entre diversité génétique et schémas de sélection.

Dans les races bovines laitières françaises de grand effectif comme la Montbéliarde, la Normande et la Prim'Holstein, les changements dans les schémas de sélection construits et appliqués par les entreprises de sélection sont cruciaux dans la gestion de la diversité génétique de ces races.

C - Le transfert embryonnaire : les principales méthodes de production d'embryons

1 - La production *in vivo* : Multiple Ovulation and Embryo Transfer (MOET)

La production d'embryons *in vivo* est le mode de production des embryons le plus fréquent en France. En 2018 en France, sur 37 980 embryons produits à partir de vaches donneuses vivantes (laitières ou allaitantes), 36 769 ont été produits *in vivo* (soit 96,8 % des embryons produits) et seulement 1 211 *in vitro* (soit 3,2 % des embryons produits) (AETE, 2019).

La production d'embryons *in vivo*, aussi appelée Multiple Ovulation and Embryo Transfer (MOET), se déroule selon les étapes suivantes :

- La sélection des donneuses,
- Un traitement hormonal de superovulation des donneuses,

Différents protocoles hormonaux de superovulation existent, généralement à base d'eCG ou de FSH (Galli et al., 2003), chacun présentant des avantages et des inconvénients. Les traitements hormonaux sont généralement administrés au milieu de la phase lutéale du cycle et induisent une lutéolyse.

- L'insémination des donneuses,

L'insémination a lieu classiquement au moment de l'œstrus, après observation des chaleurs ou à l'aveugle en fonction du protocole de synchronisation utilisé. Elle peut être réalisée avec de la semence conventionnelle ou de la semence sexée.

- La collecte des embryons,

Les embryons sont collectés par lavage des cornes utérines une semaine après l'insémination (Dubois, 2016; Galli et al., 2003). Une sonde 3 voies introduite par la vulve de manière stérile est mise en place de manière à ce que son extrémité arrive au-delà du col de l'utérus et dans l'une des cornes utérines. La sonde est guidée par palpation transrectale. Du liquide est envoyé dans l'utérus par la sonde puis ré-aspiré. Il est alors filtré pour récupérer les embryons. Cette opération est ensuite réalisée sur la deuxième corne utérine. Un double lavage permettrait d'augmenter le taux de récupération d'embryons de 40 % (Neto et al., 2005). En France, 5,3 embryons étaient récoltés par collecte en moyenne chez les bovins en 2018 (AETE, 2019).

- La sélection des embryons viables,

Les embryons sont notés selon des systèmes de classification standards décrivant le stade de développement embryonnaire (du stade ovocyte non fertilisé au stade blastocyste éclos) ainsi que la qualité de l'embryon (par exemple, note allant de 1 à 4 avec une qualité décroissante) (Robertson et Nelson, 2010). Les embryons viables sont lavés pour des raisons sanitaires. Le sexage des embryons viables est possible à ce stade.

- Le transfert d'embryons frais dans l'utérus des receveuses ou la congélation des embryons pour un TE ultérieur.

Dans tous les cas, la mise en place de l'embryon dans l'utérus de la receveuse consiste à implanter l'embryon dans la corne utérine de la receveuse du côté de l'ovaire portant un corps jaune.

2 - La production *in vitro* : *ovum* pick-up et fécondation *in vitro* (OPU-FIV)

La production d'embryons *in vitro* appelée *ovum* pick-up et fécondation *in vitro* (OPU-FIV) s'effectue selon les étapes suivantes :

- La sélection des donneuses (pouvant être très jeunes, jusqu'à deux mois),

- Une stimulation hormonale, en particulier pour les donneuses pré-pubères,
- Le prélèvement d'ovocytes immatures sur les donneuses (*ovum* pick-up, OPU),

Une bonne contention de la donneuse est indispensable, au moyen d'une cage de contention. La donneuse est sédatée. L'administration d'une anesthésie péridurale facilite la manipulation transrectale des ovaires. Les ovocytes sont collectés par ponction ovocytaire transvaginale écho-guidée (Qi et al., 2013 ; Galli et al., 2003, 2001), par aspiration du liquide folliculaire, l'ovocyte s'y trouvant. Dans le cas de donneuses pré-pubères âgées de 2 à 3 mois, l'OPU peut être réalisé via laparotomie.

- La maturation des ovocytes *in vitro*, qui dure 20 à 24h à 38,5 °C,
- La fécondation *in vitro* des ovocytes matures (FIV),

Les ovocytes sont fécondés in vitro grâce à de la semence congelée, via une co-incubation des ovocytes avec le sperme d'environ 18 à 20 heures (Galli et al., 2003). Dans l'absolu, il est donc possible de féconder chaque ovocyte d'une femelle avec le sperme d'un taureau différent.

- La culture des ovocytes fécondés (zygotes),

L'ovocyte fécondé est cultivé jusqu'à atteindre le stade blastocyste (7 jours). Différents protocoles existent et peuvent être appliqués pour ce faire, *in vivo* (dans les oviductes ligaturés d'une brebis comme receveuse provisoire) ou *in vitro* (Galli et al., 2003, 2001).

- La sélection des embryons (étape similaire à la production *in vivo*, sexage possible),

En France, en 2018, sur 622 sessions d'OPU, 5 308 ovocytes ont été récoltés, donnant 1211 embryons viables, soit 1,95 embryons par session (AETE, 2019).

- Le transfert des embryons dans l'utérus des receveuses.

Objectifs de la thèse

Des changements ont eu lieu récemment et ont toujours lieu dans les schémas de sélection des races bovines laitières françaises de grand effectif : la Montbéliarde, la Normande et la Prim'Holstein. Ces changements peuvent provenir de l'essor de nouvelles technologies, comme la sélection génomique ou le transfert embryonnaire, ou de décisions économiques, comme la réduction du nombre de taureaux, et présentent des risques pour le maintien de la diversité génétique dans ces races. Or, le maintien de la diversité génétique y est indispensable, pour limiter la dépression de consanguinité, la baisse de l'efficacité de la sélection et la perte de potentiel adaptatif.

En France, les races bovines laitières de grand effectif présentent un bon cas d'étude de ces liens entre modifications des schémas de sélection et diversité génétique, grâce à la disponibilité d'une grande quantité de données, l'importante et récente utilisation de technologies d'évaluation ou de reproduction dans leurs schémas de sélection, ainsi que par leur importance économique.

L'objectif principal de ma thèse était de comprendre et de **prédire l'impact des changements** évoqués ici **sur la diversité génétique** des races bovines laitières françaises de grand effectif dans le **contexte de la sélection génomique**, afin de mieux la gérer et la maintenir.