



HAL
open science

La diversité génétique à l'ère de la génomique : évolution de la consanguinité et ses conséquences dans trois races bovines laitières françaises

Anna-Charlotte Doublet

► To cite this version:

Anna-Charlotte Doublet. La diversité génétique à l'ère de la génomique : évolution de la consanguinité et ses conséquences dans trois races bovines laitières françaises. Sciences agricoles. Université Paris-Saclay, 2020. Français. NNT : 2020UPASB016 . tel-03181150

HAL Id: tel-03181150

<https://pastel.hal.science/tel-03181150>

Submitted on 25 Mar 2021

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

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

La diversité génétique à l'ère de la
génomique : évolution de la
consanguinité et ses conséquences
dans trois races bovines laitières
françaises

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 581, Agriculture, alimentation, biologie,
environnement et santé (ABIES)
Spécialité de doctorat : Génétique animale
Unité de recherche : Université Paris-Saclay, INRAE, AgroParisTech, GABI,
78350, Jouy-en-Josas, France
Réfèrent : AgroParisTech

Thèse présentée et soutenue à Paris-Saclay, le 10/12/2020, par

Anna-Charlotte DOUBLET

Composition du Jury

Xavier ROGNON Professeur, AgroParisTech, Paris	Président
Frédéric AUSTERLITZ Directeur de recherche, CNRS, Paris	Rapporteur & Examineur
Pascale LE ROY Directrice de recherche, INRAE, Rennes	Rapporteur & Examinatrice
Laurence MOREAU Directrice de recherche, INRAE, Gif-sur-Yvette	Examinatrice
Laurent SCHIBLER Directeur R&D, Allice, Paris	Examineur
Denis LALOE Ingénieur de recherche, HDR, INRAE, Jouy-en-Josas	Directeur de thèse

Ce travail de thèse a été financé par APIS-GENE et Alice, dans le cadre d'une convention Cifre (Agence Nationale de la Recherche et de la Technologie, convention n° 2017/0538).



Il a été réalisé au sein de l'UMR 1313 GABI de l'INRAE de Jouy-en-Josas, dans l'équipe GiBBS et l'UMT eBIS, issue de la collaboration de l'équipe G2B, d'Alice et de l'Institut de l'Élevage.



Il a été dirigé par Denis Laloë (équipe GiBBS) et co-encadré par Gwendal Restoux (équipe GiBBS) et Pascal Croiseau (équipe G2B) côté laboratoire, et suivi par Sébastien Fritz et Chris Hozé côté entreprise, ainsi qu'Alexis Michenet durant les premiers mois.

Table des matières

Liste des figures	7
Liste des tableaux.....	8
Liste des abréviations	9
Introduction générale	10
Chapitre 1 La diversité génétique et son lien avec la sélection dans les races bovines laitières françaises.....	11
I - La diversité génétique dans les espèces domestiques sélectionnées.....	11
A - Les grands principes de la sélection génétique.....	11
1 - La genèse de la sélection génétique.....	11
2 - Le progrès génétique, indissociable de la variabilité génétique.....	12
a - La génétique quantitative classique et les évaluations polygéniques	12
b - Les évaluations génomiques.....	14
3 - Les conséquences de la sélection sur la variabilité génétique et le compromis progrès/variabilité	15
B - La diversité génétique : importance et lien avec la sélection.....	17
1 - Diversité et dépression de consanguinité.....	17
2 - Diversité et potentiel adaptatif.....	17
3 - La conception des schémas de sélection : un pilier central dans la gestion de la diversité génétique.....	19
C - La dépression de consanguinité.....	20
1 - Origine	20
2 - Les conséquences économiques de la dépression de consanguinité.....	21
3 - La détection de la dépression de consanguinité	23
a - A l'échelle individuelle.....	23
i - Les différents types de régression entre performance et consanguinité	23
ii - Quelles estimations de la consanguinité utiliser pour détecter la dépression de consanguinité ?	24
b - A l'échelle génomique.....	25

4 - Bilan.....	29
D - Bilan sur la diversité génétique dans les espèces domestiques sélectionnées.....	30
II - Sélection et diversité dans les races bovines laitières françaises	31
A - Pourquoi s'intéresser à la diversité génétique intra-race chez les races bovines laitières françaises ?.....	31
B - L'élevage et la sélection des vaches laitières en France	32
1 - Présentation générale	32
2 - Les races bovines laitières françaises de grand effectif	37
C - Le transfert embryonnaire : les principales méthodes de production d'embryons	41
1 - La production <i>in vivo</i> : Multiple Ovulation and Embryo Transfer (MOET).....	41
2 - La production <i>in vitro</i> : ovum pick-up et fécondation <i>in vitro</i> (OPU-FIV)	42
Objectifs de la thèse	44
Chapitre 2 L'impact de la sélection génomique sur la diversité et le progrès génétiques dans les races bovines laitières françaises de grand effectif	45
I - Résumé de l'étude.....	45
II - Article publié : <i>The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds</i>	46
III - Discussion.....	73
Chapitre 3 L'impact des technologies de reproduction et de la baisse du nombre de taureaux sur la diversité et le progrès génétiques dans les races bovines laitières françaises de grand effectif	74
I - Résumé de l'étude.....	74
II - Article soumis : <i>Intensifying the use of reproductive technologies and reducing the dimensions of breeding schemes put genetic diversity at risk in dairy cattle breeds</i>	75
III - Additional files.....	95
IV - Discussion	98
Chapitre 4 Dépression de consanguinité et détection du fardeau génétique – Application au cas d'une race bovine laitière française de grand effectif : la Montbéliarde.....	100
I - Introduction.....	100
II - Matériel et méthodes	102

A - Données	102
B - Calcul des taux de consanguinité et du progrès génétique	102
C - Détection de la dépression de consanguinité globale par caractère	103
D - Détection de la dépression de consanguinité le long du génome par caractère	104
E - Corrélations entre ROH et fréquences alléliques.....	105
F - Scores SIFT.....	106
III - Résultats	106
A - Taux de consanguinité et progrès génétique	106
B - Dépression de consanguinité globale par caractère	107
C - Dépression de consanguinité le long du génome par caractère.....	109
1 - Distribution des effets de la consanguinité le long du génome.....	109
2 - Liens entre MAF et ROH	114
3 - Scores SIFT.....	115
D - Une région du génome sous sélection forte : le cas de <i>DGAT1</i>	115
IV - Discussion	116
V - Conclusion	122
Chapitre 5 Discussion générale et perspectives.....	124
I - Diversité génétique et schémas de sélection dans les races locales et/ou de petit effectif	124
II - Les leviers d'action de la gestion de la diversité à l'ère de la génomique : efficacité et faisabilité sur le terrain	125
A - Les plans d'accouplement.....	126
B - Les méthodes d'évaluation des animaux.....	128
C - Le croisement et le matériel génétique issu de cryobanques.....	129
D - Quelles combinaisons de leviers actionner pour gérer au mieux la diversité génétique ?	130
III - Comment gérer la dépression de consanguinité ?	132
Conclusions.....	135
Chapitre Annexe 1 - Les mécanismes génétiques impliqués dans la dépression de consanguinité	137

Chapitre Annexe 2 - Mesurer la diversité génétique au sein d'une race.....	139
I - Les différentes métriques permettant d'évaluer la diversité génétique.....	139
II - Estimer la diversité génétique à partir des pedigrees	141
A - Coefficients de consanguinité et de parenté.....	141
B - Taille efficace	145
C - Limites de ces méthodes.....	147
III - Estimer la diversité génétique à partir des génotypes.....	148
A - Hétérozygotie, consanguinité, apparentement, taille efficace : mesures moléculaires de la diversité.....	148
B - Limites des mesures fondées sur les données moléculaires	151
C - Mesurer la consanguinité à partir des Runs of Homozygosity (ROH)	152
D - Méthodes de détection et limites des mesures fondées sur les ROH.....	155
IV - Bilan.....	156
Liste des publications.....	157
Communications orales en congrès internationaux.....	157
Articles scientifiques.....	157
Activités autour de la recherche	158
Formations suivies pendant la thèse	160
Bibliographie.....	162
Annexes.....	175

Liste des figures

Figure 1 La dépression de consanguinité : baisse de la valeur des performances et de la fitness avec l'augmentation de la consanguinité	21
Figure 2 Comment interpréter le signe du coefficient de régression (= pente) dans l'étude de la dépression de consanguinité le long du génome ?	28
Figure 3 Schéma bilan : la diversité dans les espèces domestiques sélectionnées.....	30
Figure 4 La sélection au sein d'un élevage bovin laitier en France	33
Figure 5 Les différences entre les schémas de sélection génétique et de sélection génomique dans le processus de sélection des taureaux laitiers en France	34
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).....	36
Figure 7 Vaches de race Montbéliarde (crédit photo : Clémence Gaudebout)	37
Figure 8 Vache de race Normande (crédit photo : Fanny Mollandin).....	37
Figure 9 Vaches de race Prim'Holstein (crédit photo : Clémence Gaudebout)	38
Figure 10 Effets de la consanguinité sur l'indice polygénique « cellules » le long du génome chez les vaches Montbéliardes.....	110
Figure 11 Effets de la consanguinité sur l'indice polygénique « lait » le long du génome chez les vaches Montbéliardes	110
Figure 12 Effets de la consanguinité sur l'indice polygénique « MG » le long du génome chez les vaches Montbéliardes	111
Figure 13 Effets de la consanguinité sur l'indice polygénique « MP » le long du génome chez les vaches Montbéliardes	111
Figure 14 Effets de la consanguinité sur l'indice polygénique « TB » le long du génome chez les vaches Montbéliardes	112
Figure 15 Effets de la consanguinité sur l'indice polygénique « TP » le long du génome chez les vaches Montbéliardes	112
Figure 16 Répartition des coefficients de régression linéaire correspondant aux effets de la consanguinité sur six caractères (sous forme d'indices polygéniques) chez les vaches Montbéliardes.....	113
Figure 17 Consanguinité et apparentement - définitions	139
Figure 18 Coefficients de consanguinité et de parenté – définitions dans le cas de la présence d'un seul ancêtre commun	140
Figure 19 Exemple de pedigree simple pour le calcul de coefficients de consanguinité et de parenté selon la méthode des chemins.....	142

Figure 20 Exemple de pedigree simple pour le calcul de coefficients de consanguinité et de parenté selon la méthode tabulaire.....	144
Figure 21 Arbre des probabilités d'origine des allèles tirés au sort à un même locus pour les individus A et B.....	144
Figure 22 Origine des Runs Of Homozygosity (ROH).....	142
Figure 23 Exemple de calcul du coefficient de consanguinité d'un individu I à partir des Runs of Homozygosity (ROH).....	153
Figure 24 Exemple du calcul d'un coefficient de consanguinité local à partir des Runs Of Homozygosity (ROH) pour l'individu I.....	154

Liste des tableaux

Tableau 1 Présentation des trois races bovines laitières de grand effectif étudiées dans cette thèse : effectifs, diversité et production.....	38
Tableau 2 Composition actuelle de l'Index de Synthèse Unique (ISU) dans les races Montbéliarde, Normande et Prim'Holstein en France.....	39
Tableau 3 Utilisation de l'insémination artificielle dans trois races bovines laitières françaises de grand effectif.....	40
Tableau 4 Description du jeu de données utilisé pour l'étude de la dépression de consanguinité en race Montbéliarde.....	106
Tableau 5 Progrès génétique et taux de consanguinité annuels pour les vaches génotypées de race Montbéliarde nées entre 2012 et 2017, incluses dans le jeu de données d'étude de la dépression de consanguinité.....	107
Tableau 6 Coefficients de régression linéaire pour la détection de la dépression de consanguinité chez la Montbéliarde pour six caractères.....	107
Tableau 7 Dépression de consanguinité annuelle chez la Montbéliarde pour six caractères.....	108
Tableau 8 Perte annuelle de progrès génétique due à la dépression de consanguinité pour six caractères en race Montbéliarde.....	108
Tableau 9 Statistiques descriptives des pentes correspondant à l'effet de la consanguinité le long du génome pour six caractères dans la race Montbéliarde.....	114
Tableau 10 Corrélations entre les coefficients de régression linéaire (= pentes) correspondant à l'effet de la consanguinité le long du génome sur les performances, la proportion de ROH dans la population et les fréquences alléliques pour six caractères chez les vaches Montbéliardes.....	114
Tableau 11 Caractéristiques des locus présentant un score SIFT égal à 0 ainsi qu'un coefficient de régression linéaire (= pente) de dépression de consanguinité locale fort correspondant à un pic pour au moins deux caractères différents.....	115

Tableau 12 Proportion de ROH dans la population et fréquences alléliques observées sur les SNPs inclus dans le gène *DGAT1* (chromosome 14)..... 116

Tableau 13 Coefficients de régression linéaire (= pentes) correspondant à l'effet de la consanguinité le long du génome sur les performances pour six caractères chez les vaches Montbéliarde pour les SNPs inclus dans le gène *DGAT1* (chromosome 14) 116

Liste des abréviations

CD	Coefficient de détermination
<i>DGAT1</i>	Gène acyl CoA diacylglycérol acyltransférase 1
IBD	Identity by descent, ou identité par descendance
MAF	Minor allele frequency
MOET	Multiple Ovulation and Embryo Transfer
OC	Optimal contributions
OPU-FIV	Ovum pick-up et fécondation <i>in vitro</i>
ROH	Run of homozygosity
SIFT	Sorting Intolerant from Tolerant
SNP	Single nucleotide polymorphism
TE	Transfert embryonnaire
QTL	Quantitative Trait Locus

Les termes comme « éleveur » ou autres noms de métiers sont généralement notés dans ce manuscrit au masculin mais doivent être compris comme l'ensemble de toutes les personnes exerçant le métier indiqué, quel que soit leur sexe ou leur genre.

Introduction générale

La sélection génomique, mise en place depuis 2009 en France, a drastiquement modifié les schémas de sélection dans les races bovines laitières françaises de grand effectif – la Montbéliarde, la Normande et la Prim’Holstein – en augmentant le nombre de veaux mâles candidats à la sélection et en réduisant les intervalles de génération dans ces races. L’essor de l’utilisation du transfert embryonnaire par les entreprises de sélection a également remanié ces schémas de sélection, intensifiant la sélection des mères à taureaux et réduisant une fois de plus les intervalles de génération.

Ces technologies – sélection génomique et transfert embryonnaire – constituent des opportunités attrayantes, grâce aux augmentations considérables du progrès génétique théoriquement permises par leur utilisation. Cependant, les modifications des schémas de sélection résultant de leur utilisation, si elles ne sont pas maîtrisées, mettent en danger la diversité génétique de ces races.

Cette perte de diversité potentielle, caractérisée par une augmentation de la consanguinité, peut s’accompagner d’une baisse générale de la *fitness* ou valeur sélective, causée par une augmentation de l’expression du fardeau génétique, c’est-à-dire l’ensemble des variants génétiques délétères portés par la population. C’est le phénomène de la dépression de consanguinité. Le maintien de la diversité génétique est indispensable pour limiter non seulement la dépression de consanguinité, mais aussi la baisse de l’efficacité de la sélection et la perte de potentiel adaptatif.

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

Chapitre 1 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 élités 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'introgession.

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.

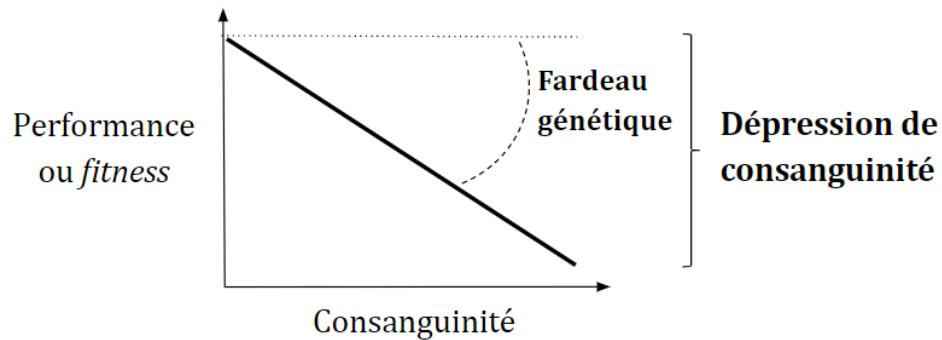


Figure 1 La dépression de consanguinité : baisse de la valeur des performances et de la *fitness* avec l'augmentation de la consanguinité

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).

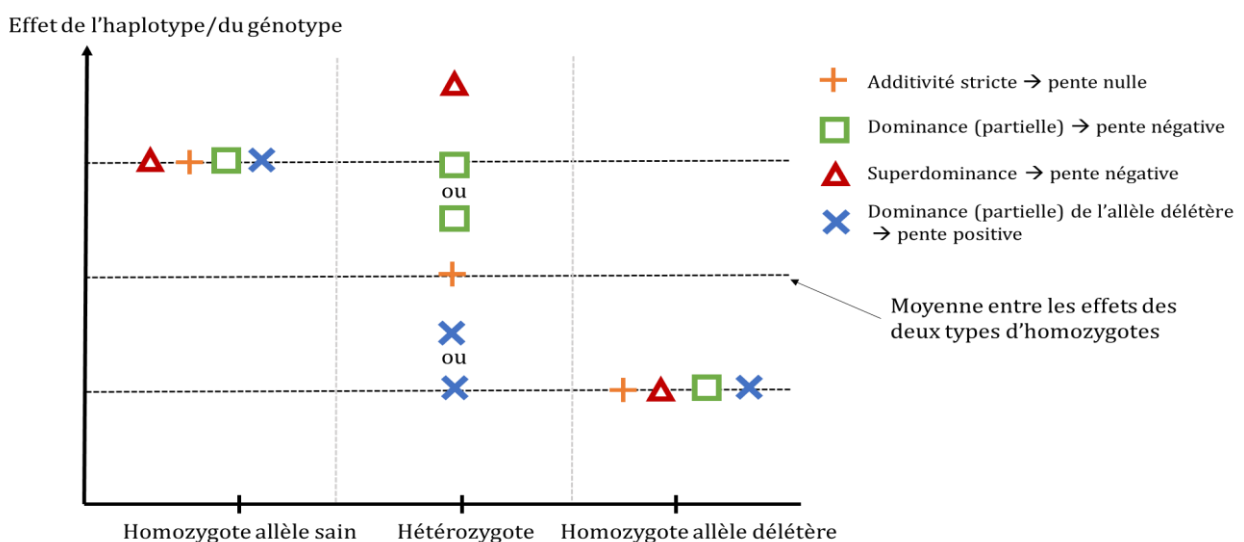


Figure 2 Comment interpréter le signe du coefficient de régression (= pente) dans l'étude de la dépression de consanguinité le long du génome ?

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 différemment 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.

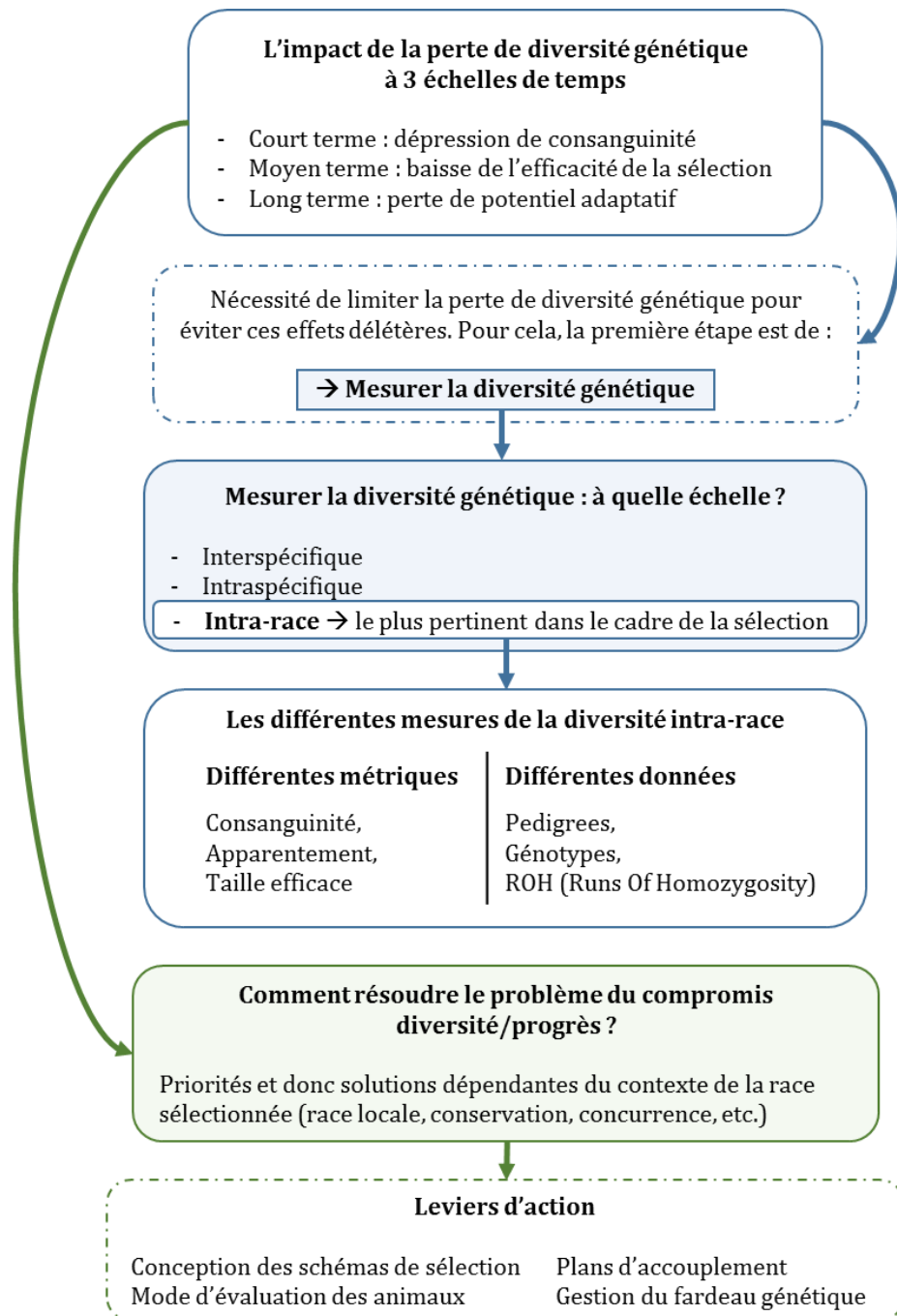


Figure 3 Schéma bilan : la diversité dans les espèces domestiques sélectionnées

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 2^{ème} 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 8^{ème} 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 décroît 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'Élevage, 2020a). En France, 75 % des veaux en élevage laitier sont nés de l'insémination (Alicce, 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.

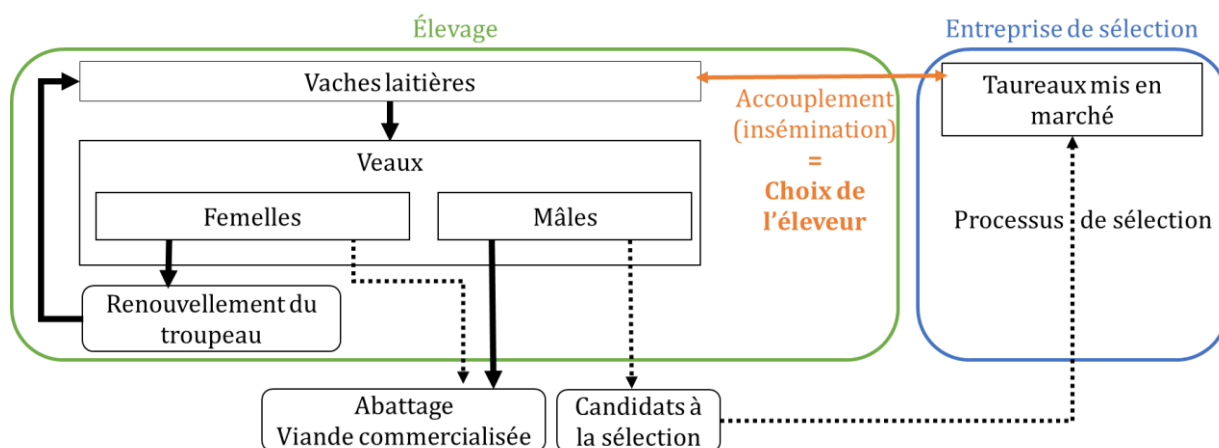


Figure 4 La sélection au sein d'un élevage bovin laitier en France

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.

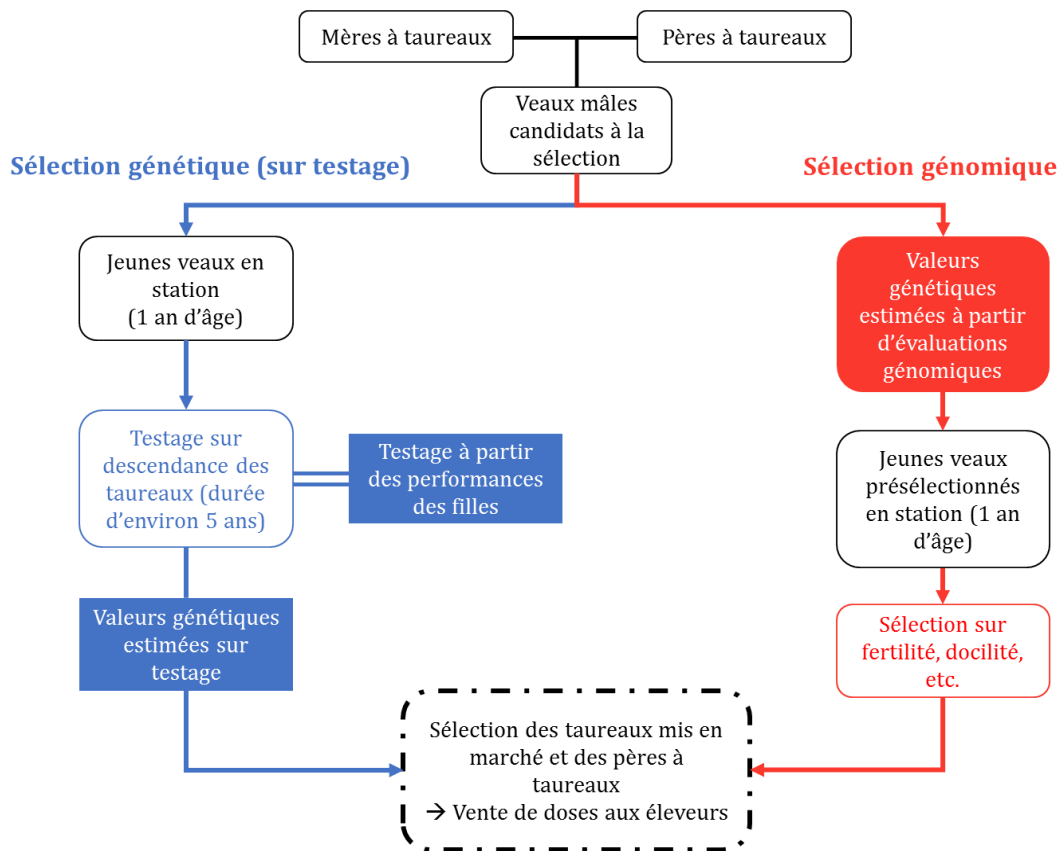


Figure 5 Les différences entre les schémas de sélection génétique et de sélection génomique dans le processus de sélection des taureaux laitiers en France

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.

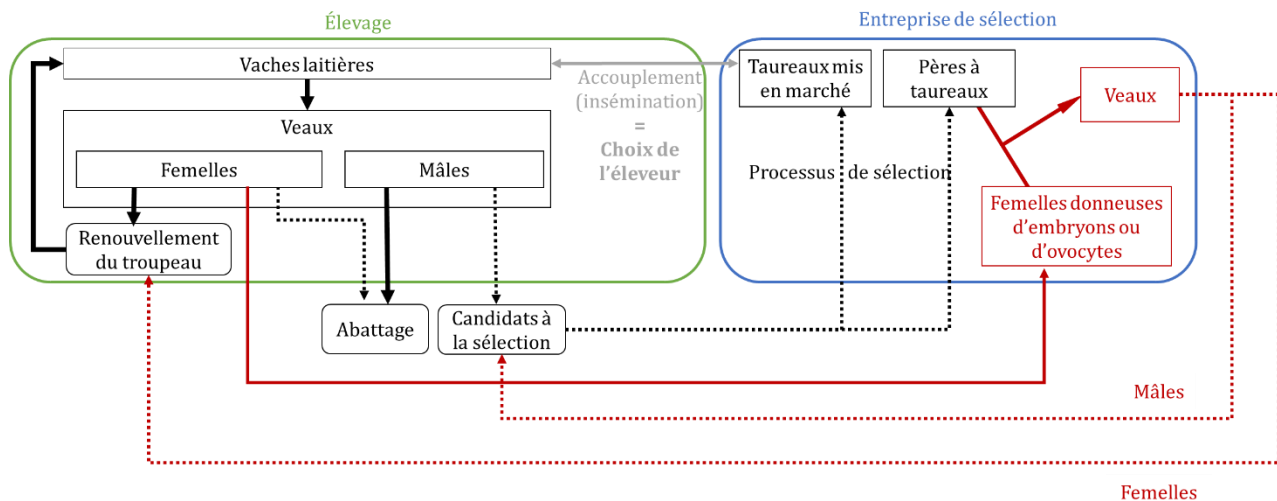


Figure 6 Les changements dans les schémas de sélection provoqués par la mise en place du transfert embryonnaire (à partir de don d'embryons ou d'ovocytes)

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 ; Thomassen 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)

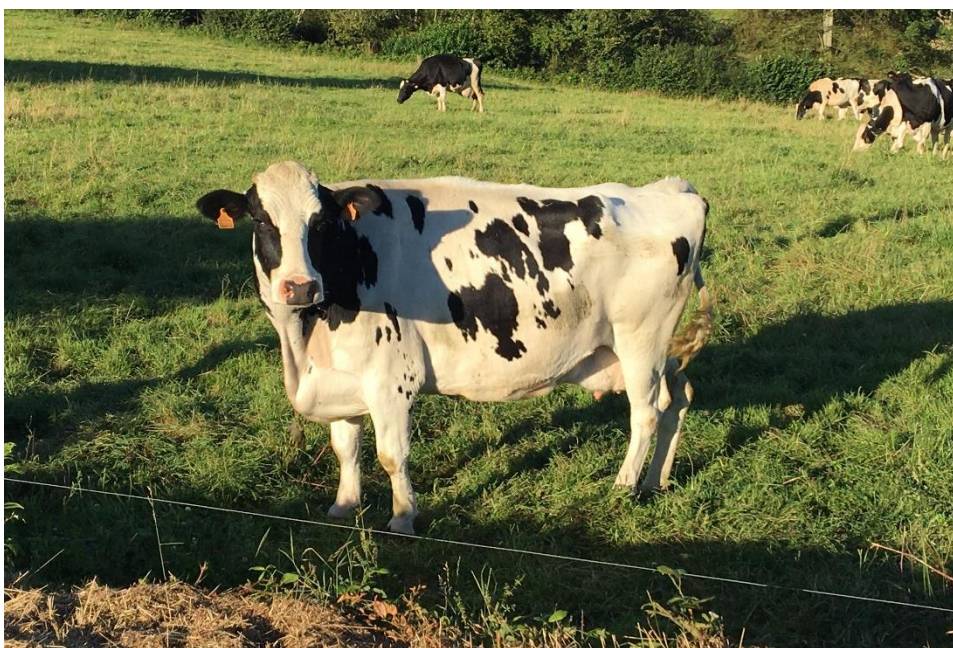


Figure 9 Vaches de race Prim'Holstein (crédit photo : Clémence Gaudebout)

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

Caractère	Race		
	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, apparemment 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'Élevage, 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'Élevage, 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 sédaturée. L'administration d'une anesthésie péridurale facilite la manipulation transrectale des ovaires. Les ovocytes sont collectés par ponction ovocytaire transvaginale échoguidée (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.

Chapitre 2 L'impact de la sélection génomique sur la diversité et le progrès génétiques dans les races bovines laitières françaises de grand effectif

I - Résumé de l'étude

En France, la mise en place des évaluations génomiques dans les races bovines laitières de grand effectif a commencé en 2009 (Boichard et al., 2012), et a modifié radicalement les schémas de sélection. Il est donc indispensable d'en évaluer les conséquences réelles sur la diversité génétique. Cela permettra aux entreprises de sélection d'adapter éventuellement ces schémas de sélection dans le futur afin d'optimiser à la fois le progrès et la diversité génétiques.

Dans ce contexte, l'objectif de cette étude est de comprendre l'impact de la sélection génomique sur la diversité génétique dans trois races bovines laitières françaises : la Montbéliarde et la Normande, deux races nationales, et la Prim'Holstein (appelée Holstein dans l'article), une race internationale. Nous avons mesuré les changements relatifs du progrès génétique et de trois estimations différentes de la diversité génétique : (i) les taux de consanguinité et d'apparentement à partir du pedigree, (ii) le taux de consanguinité à partir des ROH, et (iii) l'évolution de la longueur moyenne des ROH ; avant et après la mise en place de la sélection génomique dans ces trois races. Nous avons comparé les estimations de la consanguinité à partir des pedigrees et des ROH et exploré les changements dans les schémas de sélection qui pourraient expliquer les différences observées entre les races nationales et internationales.

La sélection génomique a induit une augmentation du progrès génétique annuel de 50, 71 et 33% pour les taureaux de race Montbéliarde, Normande et Holstein, respectivement, et les intervalles de génération ont été réduits d'un facteur de 1,7, 1,9 et 2, respectivement. Nous n'avons constaté aucun changement significatif du taux de consanguinité annuel pour les deux races nationales, Montbéliarde et Normande, mais une augmentation significative du taux de consanguinité annuel pour la race internationale Holstein, qui atteint désormais 0,55 % par an à partir des ROH et 0,49 % par an à partir des données généalogiques (ce qui équivaut à un taux de 1,36 et 1,39 % par génération, respectivement). Les ROH étaient en moyenne plus longs pour les taureaux de la race Holstein que pour ceux des deux autres races.

Après la mise en place de la sélection génomique, le progrès génétique annuel estimé à partir de l'ISU a augmenté pour les taureaux de trois grandes races bovines laitières françaises : la Montbéliarde, la Normande et la Prim'Holstein. Cette augmentation était attendue suite aux résultats de précédentes études prospectives sur données simulées ou rétrospectives sur données réelles. Dans le même temps, le taux de consanguinité par an et par génération a augmenté pour

les taureaux de la race internationale Prim'Holstein, comme cela a été observé dans d'autres régions du monde comme les Pays-Bas ou l'Amérique de Nord. L'augmentation du taux de consanguinité chez la Prim'Holstein était en partie due à une augmentation de la consanguinité récente (moins de 5 générations). Aucune augmentation du taux de consanguinité n'a été observée pour les taureaux de races nationales Montbéliarde et Normande. Cette différence reflète des modifications différentes des schémas de sélection dans les trois races, suggérant que l'augmentation du ratio nombre efficace/nombre réel de taureaux dans la race Prim'Holstein n'a pas suffi à compenser l'introduction substantielle de consanguinité récente dans cette race. Cette augmentation de la consanguinité récente en Prim'Holstein peut s'expliquer par l'utilisation intensive de quelques individus seulement comme grands-pères des taureaux, principalement originaires des États-Unis.


II - Article publié : The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds

RESEARCH ARTICLE

Open Access



The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds

Anna-Charlotte Doublet^{1,2*} , Pascal Croiseau¹, Sébastien Fritz^{1,2}, Alexis Michenet^{1,2}, Chris Hozé^{1,2}, Coralie Danchin-Burge³, Denis Laloë¹ and Gwendal Restoux¹

Abstract

Background: In France, implementation of genomic evaluations in dairy cattle breeds started in 2009 and this has modified the breeding schemes drastically. In this context, the goal of our study was to understand the impact of genomic selection on the genetic diversity of bulls from three French dairy cattle breeds born between 2005 and 2015 (Montbéliarde, Normande and Holstein) and the factors that are involved.

Methods: We compared annual genetic gains, inbreeding rates based on runs of homozygosity (ROH) and pedigree data, and mean ROH length within breeds, before and after the implementation of genomic selection.

Results: Genomic selection induced an increase in mean annual genetic gains of 50, 71 and 33% for Montbéliarde, Normande and Holstein bulls, respectively, and in parallel, the generation intervals were reduced by a factor of 1.7, 1.9 and 2, respectively. We found no significant change in inbreeding rate for the two national breeds, Montbéliarde and Normande, and a significant increase in inbreeding rate for the Holstein international breed, which is now as high as 0.55% per year based on ROH and 0.49% per year based on pedigree data (equivalent to a rate of 1.36 and 1.39% per generation, respectively). The mean ROH length was longer for bulls from the Holstein breed than for those from the other two breeds.

Conclusions: With the implementation of genomic selection, the annual genetic gain increased for bulls from the three major French dairy cattle breeds. At the same time, the annual loss of genetic diversity increased for Holstein bulls, possibly because of the massive use of a few elite bulls in this breed, but not for Montbéliarde and Normande bulls. The increase in mean ROH length in Holstein may reflect the occurrence of recent inbreeding. New strategies in breeding schemes, such as female donor stations and embryo transfer, and recent implementation of genomic evaluations in small regional breeds should be studied carefully in order to ensure the sustainability of breeding schemes in the future.

Background

In France, genomic evaluations started to be implemented in dairy cattle breeds in 2009. After a brief transition period for the breeding companies to adopt this technology, these new evaluation methods have drastically changed dairy cattle breeding schemes since 2011. Genomic evaluations have allowed breeding companies

to increase the number of male candidates for selection and the use of new bulls [1, 2]. At the early selection stage, genomic selection (GS) allows exploiting a major part of the genetic variance compared to the previously used models based on parents' average values [1, 3]. In fact, Mendelian sampling variation, which is defined as the difference between the genetic value of an individual and the average of the parental genetic values, can be captured from molecular markers without any information on the parents' phenotype or the offspring's phenotype. This makes it possible to obtain breeding values

*Correspondence: anna-charlotte.doublet@inra.fr

¹ GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France

Full list of author information is available at the end of the article



with higher reliability sooner in the life of an individual compared with those based on pedigree data [4]. Indeed, the estimation of pedigree-based breeding values at birth depends mostly on information about related individuals, resulting in a low reliability and a high correlation between estimated breeding values of relatives. These high correlations decrease only later in the life of an animal when information based on its own performance or performances of the progeny becomes available. Consequently, GS reduces the number of related individuals chosen as candidates for selection [5]. Moreover, collecting genotype data is less expensive than collecting phenotype data, which means that more male candidates can be evaluated with high reliability. Furthermore, methods to balance inbreeding and genetic gain in breeding schemes, such as optimal contribution selection [6], allow for better conservation of genetic diversity when based on genomic information rather than on pedigree data [7, 8]. Thus, GS was expected to reduce the rate of inbreeding and to increase genetic gain per generation at the same time [2].

However, since animals with high genomic estimated breeding values (GEBV) can be selected as parents at a very young age, generation intervals are significantly shortened [9]. As a result, the annual inbreeding rate could then increase and this has already been observed in Dutch-Flemish Holstein–Friesian cattle [10] and in North American Holstein cattle [11].

Such an increase in inbreeding rate results in a loss of genetic diversity, with lower additive genetic variance leading to a lower response to selection [12]. Moreover, a lower overall genetic variability leads also to a loss of adaptive potential for selecting new breeding goals in the context of climate change [13], and more severe inbreeding depression [14–16].

Measuring genetic diversity is necessary in order to manage it. Typically, genetic diversity is estimated and managed with the help of inbreeding and kinship coefficients based on pedigree data (e.g. [17, 18]). These coefficients reflect the proportion of genome expected to be identical-by-descent (IBD) in an individual and between two individuals, respectively. The reliability of these estimates depends on the quality of the pedigree records and their depth [7, 19, 20]. In particular, pedigree-based inbreeding and kinship may be underestimated due to incomplete pedigree data and the assumption of arbitrary non-inbred and unrelated founder individuals, thus leading to an overestimation of genetic diversity.

Falling costs in the field of high-throughput genotyping and sequencing make it possible to obtain more and more genotyping data, and thus to estimate genetic diversity more precisely. These data allow the computation of inbreeding coefficients and genomic relationship

matrices (GRM) [21, 22]. The calculation of inbreeding coefficients from GRM computed using all the allele frequencies set to 0.5 globally consists in computing observed and expected proportions of homozygous single nucleotide polymorphisms (SNPs) considered as identical-by-state (IBS).

Another method to estimate genetic diversity consists in detecting IBD segments, called runs of homozygosity (ROH), by considering that segments of consecutive homozygous SNPs originate from common ancestors [23, 24]. Thus, it is possible to compute ROH-based inbreeding coefficients defined as the proportion of the autosomal genome that is actually included in ROH above a given length threshold [24], corresponding to the realized proportion of genome that is IBD. This approach to estimate inbreeding coefficients seems to reflect IBD better than genomic inbreeding coefficients calculated from molecular data based on the GRM [11, 25–27]. The length of such ROH segments follows an inverse exponential distribution with expectation $1/2G$ Morgan, with G being the number of generations to the common ancestor from which the segment arises (e.g. [28]). Due to recombination, ROH length tends to decrease over generations [29, 30], which allows to estimate the age of inbreeding events along the genome and to trace back population history [29, 31, 32]. Shorter ROH tend to result from older inbreeding (older common ancestors) whereas longer ROH tend to result from more recent inbreeding (more recent common ancestors) [29, 31, 32]. However, different criteria have been proposed to define ROH-based inbreeding estimates and there is no consensus about which one should be preferred (e.g. the number of heterozygous loci allowed in one ROH, the minimal length, etc.). Consequently, ROH-based inbreeding estimates might differ, depending on the criteria used [28].

Pedigree data are useful to evaluate genetic diversity for individuals that are not genotyped, whereas measures based on molecular data (such as ROH) allow the study of the realized proportion of genome that is autozygous and the pattern of inbreeding along the genome. In spite of their advantage over pedigree-based inbreeding estimates, genomic-based estimates began to be used to manage genetic diversity in selection programs only recently [33–35].

In this context, the goal of our study was to understand the impact of GS on genetic diversity in three French dairy cattle breeds: Montbéliarde and Normande, two national breeds, and Holstein, an international breed. We measured the relative changes in rates of genetic merit and in three different estimates of genetic diversity: (i) pedigree-based inbreeding and kinship, (ii) ROH-based inbreeding coefficients, and (iii) mean length of ROH; and these estimates were taken before and after the

implementation of GS in these three breeds in France. Based on our findings, we compare pedigree and ROH-based inbreeding rates within breeds and explore the changes in breeding schemes that could explain the differences observed between national and international breeds.

Methods

Animals, pedigree and genotyping data

We used three datasets that consisted of all progeny-tested bulls born between 2005 and 2010 and all marketed bulls born between 2012 and 2015 from three French dairy cattle breeds: 1307 Montbéliarde bulls and 1016 Normande bulls (national breeds), and 4694 Holstein bulls (international breed) (see Table 1 and Fig. 1). We chose these categories of bulls in order to obtain cohorts of comparable sizes and to compare bulls that are selected for artificial insemination by breeding companies (either to be marketed or to be tested). Key data about these breeds in France are available in Additional file 1: Table S1.

Genealogies of all bulls in the three datasets were traced back as far as possible. The mean numbers of generations traced back for bulls in the datasets were

15.4 ± 1.7, 15.8 ± 1.6 and 15.4 ± 1.9 for Montbéliarde, Normande and Holstein, respectively (longest paths computed with R [36] package pedigree [37]).

All the bulls in the datasets were genotyped for 50 K SNPs (Illumina Infinium® BovineSNP50 BeadChip). Quality control filtering was performed by removing SNPs that were genotyped in less than 95% individuals and SNPs with a minor allele frequency (MAF) lower than 1%. No individual had less than 95% genotyped SNPs. After quality control filtering, 39,992, 40,135 and 41,377 SNPs remained in the datasets for Montbéliarde, Normande and Holstein, with on average one SNP every 62.5 ± 66.5 kb, 62.2 ± 66.4 kb and 60.4 ± 63.8 kb, respectively.

The total genetic merit index ISU (standing for *Index Synthèse Unique*) is an individual index that combines production, functional and type traits, which are weighted depending on the breed and subsequent breeding goals. The formula to obtain total genetic merit index ISU is built in order to obtain a standard error of the index of 20 points within each breed [38]. Within each breed, all genetic merit estimates were calculated relative to the same arbitrary base.

Table 1 Number of animals in each dataset

Breed	Birth period	Number of individuals born each year				Total number of individuals
		Min	Mean (± sd)	Max	Median	
Montbéliarde	2005–2010 Progeny-tested bulls	82	138 (± 34)	176	141	826
	2012–2015 Marketed bulls	81	120 (± 29)	149	125	480
Normande	2005–2010 Progeny-tested bulls	82	111 (± 21)	135	112	666
	2012–2015 Marketed bulls	73	87 (± 11)	100	87	347
Holstein	2005–2010 Progeny-tested bulls	368	573 (± 157)	697	606	3440
	2012–2015 Marketed bulls	294	312 (± 12)	322	316	1248

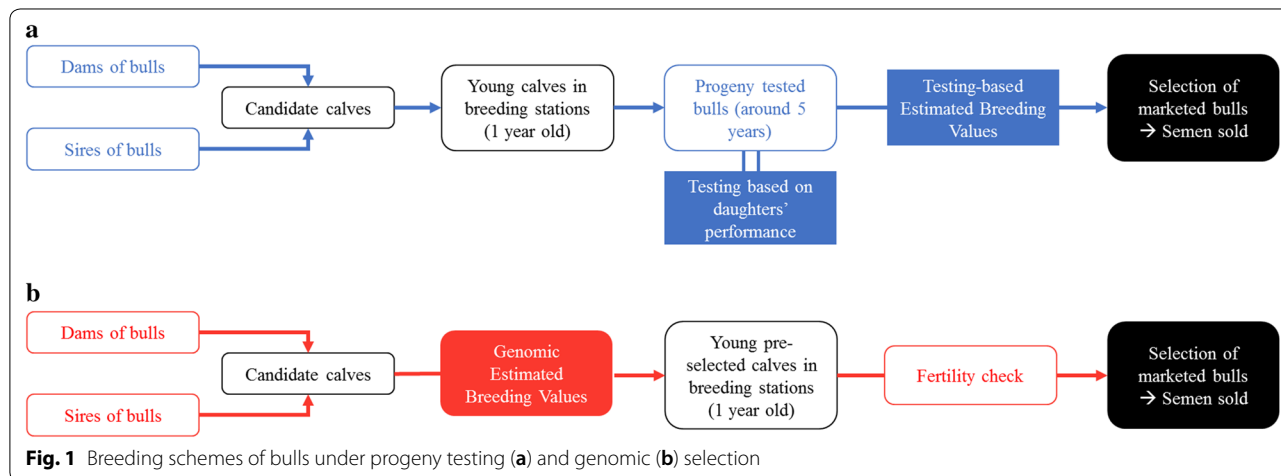


Fig. 1 Breeding schemes of bulls under progeny testing (a) and genomic (b) selection

Measures of genetic diversity

Pedigree-based measures of genetic diversity

Pedigree-based inbreeding coefficients F_{ped} and kinship coefficients f_{ped} were computed with the Pedig software [39] for all the animals in the three datasets. The mean of all kinship coefficients f_{ped} for bulls born between 2005 and 2010 and the mean of all kinship coefficients f_{ped} for bulls born between 2012 and 2015 were compared with a two-sample t test. We also computed pedigree-based coefficients by taking only the last five generations for each bull into account, with the Pedig software [39] for all animals in the three datasets.

Measures of genetic diversity based on runs of homozygosity

ROH reflect autozygous segments of the genome. A ROH was defined as a homozygous segment of at least 15 SNPs or 1000 kb long, with at least one SNP per 75 kb. Two consecutive SNPs could not be included in the same ROH if they were more than 150 kb apart. ROH were detected using the PLINK 1.9 “homozyg” function [40, 41] (command line: `plink --cow --bfile genotyping_data_filename --homozyg --homozyg-kb 1000 --homozyg-snp 15 --homozyg-window-snp 15 --homozyg-density 75 --homozyg-gap 150 --out output_filename`).

ROH-based inbreeding estimates, $F_{ROH,i}$, were computed as the proportion of genome included in ROH [10, 24] as follows:

$$F_{ROH,i} = \frac{\Sigma L_{ROH,i}}{L_{auto}}$$

where $\Sigma L_{ROH,i}$ is the total length of ROH for individual i , and L_{auto} the length of the autosomal genome covered by SNPs after withholding gaps longer than 150 kb between two SNPs, corresponding to the length of the autosomal genome on which ROH can be detected. These parameters allowed the detection of ROH on 77.0, 77.0 and 78.6% of the autosomal genome of Montbéliarde, Normande and Holstein bulls, respectively (see Additional file 2: Table S2).

For each individual, the mean ROH length, i.e. $L_{ROH,mean,i}$, was computed as:

$$L_{ROH,mean,i} = \frac{\Sigma L_{ROH,i}}{N_{ROH,i}}$$

where $\Sigma L_{ROH,i}$ is the total length of ROH for individual i in kb, and $N_{ROH,i}$ the total number of ROH for individual i .

Pearson’s product-moment correlations between pedigree- and ROH-based inbreeding coefficients were computed for each breed in order to check if it was relevant to compare both inbreeding estimates.

To compare the variability of pedigree and ROH-based inbreeding rates, we calculated the coefficient of variation

of these two parameters within each breed as the ratio between the overall standard deviation of inbreeding and its overall mean for each dataset.

We obtained approximate inbreeding rates per generation for a selection type by multiplying the mean inbreeding rate by the mean parental generation interval for this period (in years) (see Additional file 3: Table S3).

Demographic parameters

Generation interval I_1 was computed with an adapted version of Pedig software [39]. I_1 was defined as the mean difference between the birthdate of an individual and its parents, in months.

For each breed, the number of bulls born per year was divided by the number of bulls born in 2005 in the three datasets, giving S_i :

$$S_i = \frac{N_i}{N_{2005}}$$

where N_i is the number of bulls born in year i .

The effective number of bulls, $Ne_{breed, selection}$, for a given breed and a given type of selection was computed as:

$$Ne_{breed, selection} = \frac{1}{\sum_{i=1}^{n_{breed, selection}} \left(\frac{o_i}{\sum_{j=1}^{n_{breed, selection}} o_j} \right)^2}$$

where $n_{breed, selection}$ is the number of bulls for a given breed and a given type of selection and o_i the number of offspring of bull i [42, 43]. We also computed the effective/census number of bulls ratio, $Ne_{breed, selection} / n_{breed, selection}$. The two types of selection were progeny testing selection (PTS) (individuals born between 2005 and 2010), and genomic selection (GS) (individuals born between 2012 and 2014).

95% confidence intervals for $Ne_{breed, selection}$ and $Ne_{breed, selection} / n_{breed, selection}$ were generated by random resampling with replacement of the datasets (bootstrap with 1000 iterations). For a given breed, $Ne_{breed, PTS}$ and $Ne_{breed, GS}$, or $Ne_{breed, PTS} / n_{breed, PTS}$ and $Ne_{breed, GS} / n_{breed, GS}$ were statistically different if confidence intervals did not overlap.

Impact of genomic selection on genetic diversity and genetic gain

In order to assess the impact of GS on genetic diversity and genetic gain, we used the following linear model, using the R function `lm` [36]:

$$Y_i = \begin{cases} a_1 + b_1 \cdot x_i + \varepsilon_i, & 2005 \leq x_i \leq 2010 \\ a_2 + (b_1 + \delta) \cdot x_i + \varepsilon_i, & 2012 \leq x_i \leq 2015 \end{cases}$$

where Y_i is the variable of interest for bull i (F_{ped} , f_{ped} , $F_{ROH,i}$, $L_{ROH,mean,i}$ for genetic diversity and ISU for genetic gain), x_i is the birth year of bull i , and b_1 the associated coefficient of regression if bull i was born between 2005 and 2010 (PTS) or $(b_1 + \delta)$ if born between 2012 and 2015 (GS). The impact of GS was measured with the δ coefficient. Its significance was tested with an analysis of variance.

The quality of linear regressions was checked by looking at the square root of the standardized residuals of each linear regression (see Additional file 4: Figures S1–S7).

The relative change, RC , of the slopes of regression before and after GS was computed as:

$$RC = \frac{\delta}{|b_1|}$$

For a better interpretation of relative change, please note that the sign of RC quantifies the direction of the change of the slopes (if $RC > 0$, the new slope will be higher than the previous slope and if $RC < 0$, it will be lower) and its absolute value quantifies the magnitude of the change of the slopes.

Results

Genetic merit

The total genetic merit index ISU increased at a significantly higher rate in Montbéliarde, Normande and

Holstein under GS (between 2012 and 2015) than under PTS (between 2005 and 2010), with respective relative changes of 0.50, 0.71 and 0.33 (see Table 2 and Fig. 2a). The slopes of ISU during GS are equivalent to increases of ~4.4, 5.0 and 7.4% per year relative to the mean genetic merit ISU in 2012 of 111, 118 and 139 points for Montbéliarde, Normande and Holstein, respectively.

Genetic diversity

Pedigree-based measures of genetic diversity

Pedigree-based inbreeding rates increased significantly in Normande and Holstein under GS schemes (between 2012 and 2015) compared to PTS (between 2005 and 2010), i.e. from 0.059 to 0.088% per year to 0.19 and 0.49% per year for Normande and Holstein, respectively (from 0.34 to 0.46% per generation to 0.70 and 1.39% per generation, respectively) with respective relative changes of 2.20 and 4.50 (see Table 2). This was not the case in Montbéliarde, for which the inbreeding rate per year between 2005 and 2010 was 0.090% (0.52% per generation) (see Table 2). In 2005, the overall pedigree-based inbreeding level in Holstein did not differ significantly from that in Montbéliarde and Normande (see Fig. 2b). However, in 2015, the pedigree-based inbreeding level was significantly higher in Holstein than in Montbéliarde (p -value $< 1e-10$) and Normande (p -value $< 1e-10$) (see Fig. 2b). The slopes

Table 2 Trends of the different inbreeding criteria estimating genetic gain and genetic diversity per year

Parameter	Breed	b_1 (\pm SE)	b_2	δ (\pm SE)	p -value of δ	RC
ISU	MON	3.25 (\pm 0.28)	4.89	1.64 (\pm 0.72)	0.023	0.50
	NOR	3.42 (\pm 0.31)	5.85	2.43 (\pm 0.79)	2.1e-03	0.71
	HOL	7.77 (\pm 0.16)	10.31	2.54 (\pm 0.41)	6.9e-10	0.33
F_{ped} (in %) ^a	MON	0.090 (\pm 0.029)	0.11	0.019 (\pm 0.076)	0.81	0.21
	NOR	0.059 (\pm 0.024)	0.19	0.13 (\pm 0.061)	0.035	2.20
	HOL	0.088 (\pm 0.013)	0.49	0.40 (\pm 0.033)	< 1e-10	4.50
f_{ped} (in %) ^a	MON	-0.14 (\pm 0.006)	-0.013	0.13 (\pm 0.018)	< 1e-10	0.91
	NOR	-0.11 (\pm 0.008)	0.072	0.19 (\pm 0.022)	< 1e-10	1.64
	HOL	0.20 (\pm 0.002)	0.22	0.019 (\pm 0.005)	1.5e-04	0.092
F_{ROH} (in %) ^a	MON	0.17 (\pm 0.033)	0.076	-0.095 (\pm 0.085)	0.26	-0.56
	NOR	0.12 (\pm 0.033)	0.14	0.019 (\pm 0.082)	0.82	0.16
	HOL	0.080 (\pm 0.018)	0.55	0.47 (\pm 0.045)	< 1e-10	5.93
Mean ROH ^b length	MON	4.75 (\pm 1.64)	0.09	-4.66 (\pm 4.22)	0.27	-0.98
	NOR	1.32 (\pm 1.74)	1.2	-0.149 (\pm 4.36)	0.97	-0.11
	HOL	1.53 (\pm 0.965)	13	11.9 (\pm 2.43)	1.1e-06	7.77

b_1 is the slope of regression of each parameter depending on birth year for progeny-tested bulls born between 2005 and 2010 (progeny testing selection), and δ the difference between the slopes of regression of each parameter depending on birth year for progeny testing selection and for marketed bulls born between 2012 and 2015 (genomic selection), for Montbéliarde (MON), Normande (NOR) and Holstein (HOL). The relative change (RC) is equal to $\frac{\delta}{|b_1|}$. b_2 is the slope of the parameter for marketed bulls born between 2012 and 2015 (genomic selection), equal to $b_1 + \delta$. The p -value of δ corresponds to the significance of the non-nullity of δ

ISU, total merit index; F_{ped} , pedigree-based inbreeding; f_{ped} , pedigree-based kinship; F_{ROH} , inbreeding based on Runs of Homozygosity (ROH)

^a Slopes and standard errors are displayed in %. p -values and relative changes are not

^b Runs of homozygosity

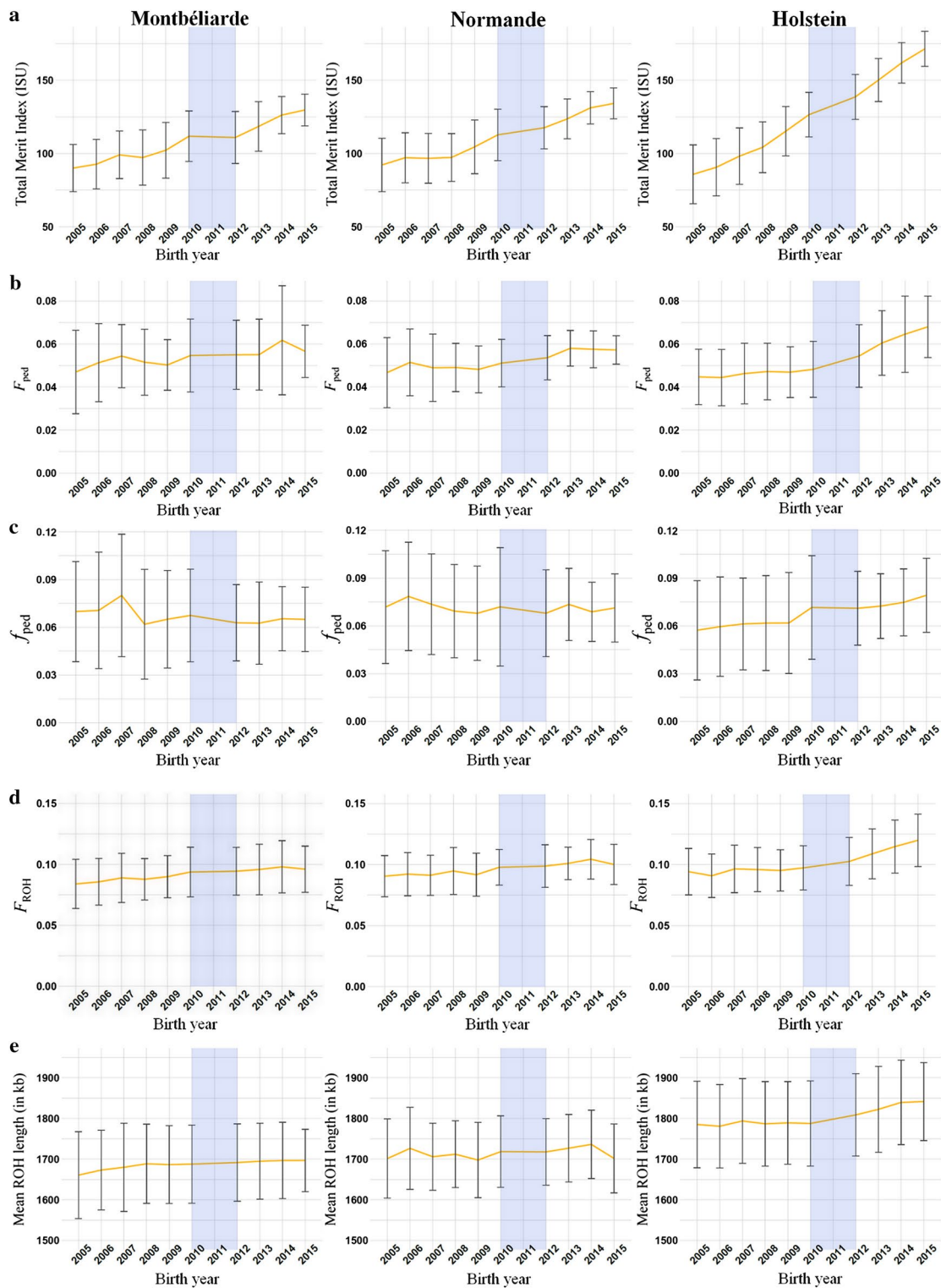


Fig. 2 Trends of the different parameters estimating genetic gain and genetic diversity. The yellow line represents the mean for each birth year and the error bars the standard deviation. Total merit indices (ISU) (a), inbreeding (b) and kinship (c) based on pedigree data, inbreeding based on runs of homozygosity (ROH) (d) and mean ROH length (in kb) (e) were retrieved or computed for progeny-tested bulls born between 2005 and 2010 and for marketed bulls born between 2012 and 2015, for Montbéliarde, Normande and Holstein

of pedigree-based inbreeding coefficients during GS are equivalent to increases of ~2.0, 3.5 and 8.9% per year of the mean pedigree-based inbreeding coefficients in 2012 for Montbéliarde, Normande and Holstein, respectively.

For inbreeding coefficients based on pedigree and calculated over the last five generations only, the slope of regression was negative during PTS for the three breeds. Under GS, it decreased significantly for Montbéliarde and Normande, reaching -0.32% and -0.23% per year, respectively (-1.22% and -0.86% per generation, respectively), whereas it increased significantly for Holstein, reaching 0.27% per year (0.78% per generation) (Fig. 3) (see Additional file 5: Tables S4, S5).

Pedigree-based kinship increased at a significantly higher rate in Montbéliarde, Normande and Holstein under GS (between 2012 and 2015) than under PTS (between 2005 and 2010), with respective relative changes of 0.91, 1.64 and 0.092 (see Table 2). The slopes of pedigree-based kinship coefficients during GS are equivalent to increases of about -0.20 , 1.1 and 3.1% per year of the mean pedigree-based kinship coefficients in 2012 for Montbéliarde, Normande and Holstein, respectively (see Fig. 2c). In Montbéliarde and Normande, pedigree-based kinship coefficients seemed to plateau during both PTS and GS, although we observed a drop in the average pedigree-based kinship in Montbéliarde in 2008 (see Fig. 2c). The mean value of pedigree-based kinship coefficients was significantly lower for Montbéliarde born between 2012 and 2015 than for Montbéliarde bulls born between 2005 and 2010 (p -value $< 1e-10$), as well as for Normande bulls (p -value $< 1e-10$) (see Fig. 2c). By contrast, this value was significantly higher for Holstein bulls born between 2012 and 2015 than for those born between 2005 and 2010 (p -value $< 1e-10$) (see Fig. 2c).

Measures of genetic diversity based on runs of homozygosity (ROH)

The rate of ROH-based inbreeding increased significantly in Holstein under GS (between 2012 and 2015) compared to under PTS (between 2005 and 2010), i.e. from 0.080% per year to 0.55% per year for Holstein (from 0.39% to 1.36% per generation) with a relative change of 5.93 (see Table 2). This was not the case for Montbéliarde and Normande, for which inbreeding rates of 0.17 and 0.12% per year were found between 2005 and 2010, respectively (0.96 and 0.71% per generation) (see Table 2). In 2005, the overall ROH-based inbreeding level in Montbéliarde was significantly lower than that in Normande and Holstein (p -values = 0.0031 and $2.1e-08$, respectively) (see Fig. 2d). However, in 2015, the level of ROH-based inbreeding was significantly higher in Holstein than in

Montbéliarde and Normande (p -values $< 1e-10$, in both cases) (see Fig. 2d). The slopes of ROH-based inbreeding coefficients during GS are equivalent to increases of ~0.8, 1.4 and 5.4% per year of the mean ROH-based inbreeding coefficients in 2012 for Montbéliarde, Normande and Holstein, respectively.

Mean ROH length increased at a significantly higher rate in Holstein under GS (between 2012 and 2015) than under PTS (between 2005 and 2010), with a relative change of 7.77 (see Table 2). This was not the case in Montbéliarde and in Normande, for which the mean ROH length plateaued under GS (see Fig. 2e). From 2005 to 2015, the overall mean ROH length for Montbéliarde was significantly shorter than that for Normande, which in turn was significantly shorter than that for Holstein (p -values $< 1e-10$ in all cases, with average mean ROH lengths of 1.68, 1.71 and 1.80 Mb, respectively) (see Fig. 2c). The slopes of mean ROH lengths during GS showed increases of ~0.005, 0.07 and 0.7% per year of the mean ROH lengths in 2012 for Montbéliarde, Normande and Holstein, respectively.

Correlations between and variability of pedigree- and ROH-based inbreeding coefficients

Pearson's product-moment correlation between pedigree- and ROH-based inbreeding coefficients were equal to 59% (95% confidence interval: [55%, 62%]), 50% [45%, 55%] and 59% [57%, 61%] for Montbéliarde, Normande and Holstein, respectively.

Coefficients of variation of pedigree-based inbreeding were equal to 0.33, 0.25 and 0.31 for Montbéliarde, Normande and Holstein, respectively. Coefficients of variation of ROH-based inbreeding were equal to 0.22, 0.18 and 0.21 for Montbéliarde, Normande and Holstein, respectively.

Demographic parameters

Generation intervals

The mean difference I_1 between the birthdate of an individual and its parents reached a plateau at ~70 months in Montbéliarde and in Normande, and 59 months in Holstein between 2005 and 2009. I_1 was reduced by a factor 1.7 in Montbéliarde, by 1.9 in Normande and by a factor 2 in Holstein between animals born between 2005 and 2009 (under PTS) and animals born in 2015 (under GS). To date, this interval has not reached a new plateau for the three breeds (see Fig. 4).

Number and use of bulls

A histogram of the proportion of the total number of offspring per bull for each breed and selection type is available in Additional file 6: Figure S8 and Additional file 7:

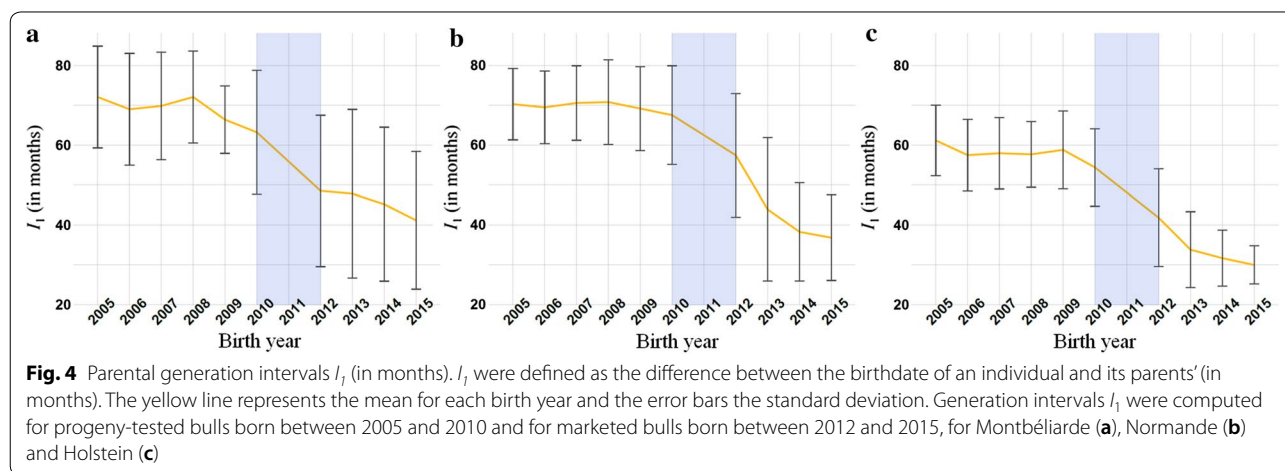
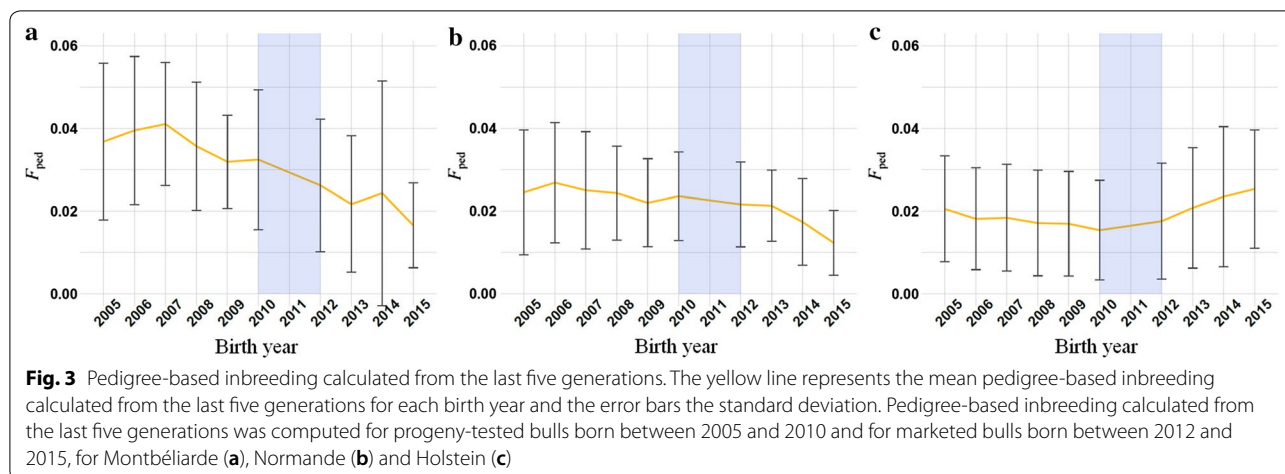


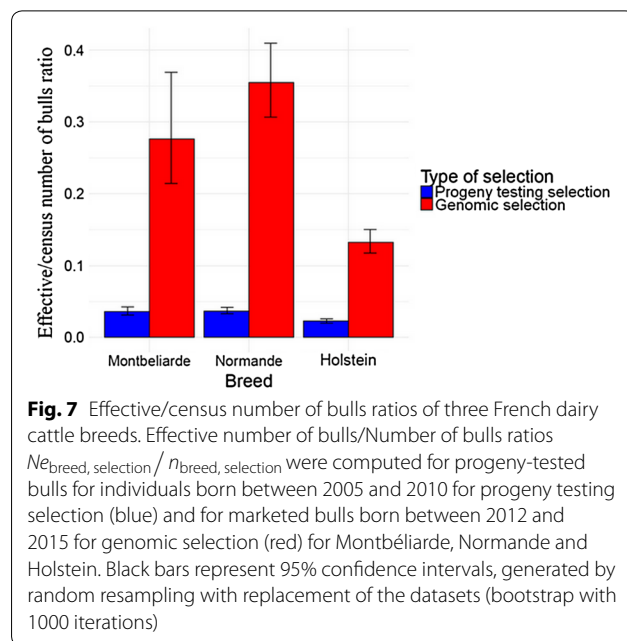
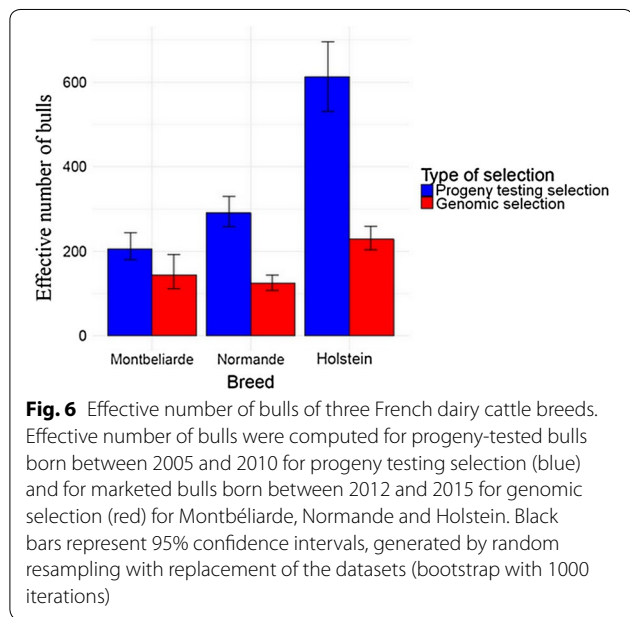
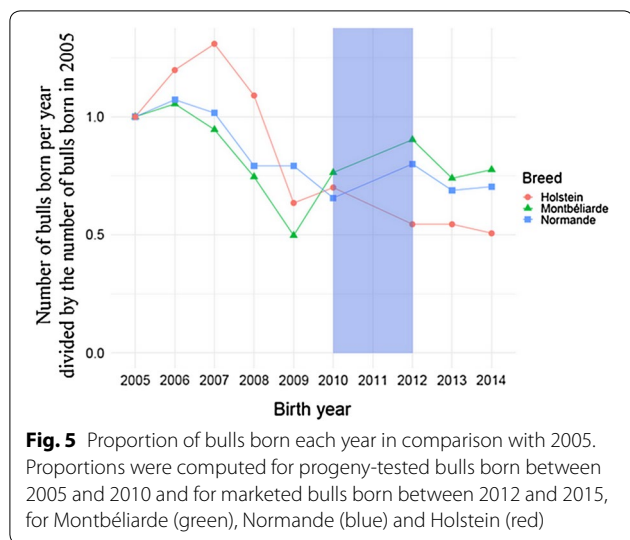
Table S6. The ratio between number of marketed bulls born in 2014 and number of progeny-tested bulls born in 2005 was equal to 77.6% (128 compared to 165) in Montbéliarde, 70.4% (88 compared to 125) in Normande, and 50.7% (294 compared to 580) in Holstein (see Fig. 5).

The effective number of bulls used was significantly larger between 2005 and 2010 (under PTS) than between 2012 and 2014 (under GS) in Normande and Holstein, but not in Montbéliarde (see Fig. 6). The ratio of the effective and the actual number of bulls ratio under PTS, $\frac{N_{e,breed,PTS}}{n_{breed,PTS}}$, was significantly lower than the same ratio under GS, $\frac{N_{e,breed,GS}}{n_{breed,GS}}$, in Montbéliarde, Normande and Holstein, with $\frac{N_{e,breed,GS}}{n_{breed,GS}}$ of 0.009, 0.021 and 0.003%, respectively (see Fig. 7).

Discussion

In dairy cattle breeds, GS was expected to increase the annual genetic gain by up to 70 to 108% depending on the scenarios used for the breeding schemes (with different

proportions of young genotypes bulls and proven bulls used) [1, 2, 44] but only a few studies have examined its impact on genetic diversity. These studies were preliminary and mainly predictive since they were conducted before GS had really started. Genomic evaluations were implemented in 2009 in French dairy cattle breeds [4] and they have drastically changed dairy cattle breeding schemes for the three major French dairy cattle breeds. Thus, after a decade of GS, we can evaluate the impact of new breeding schemes on both the genetic gain and diversity in these three breeds in France. The Holstein, Montbéliarde and Normande breeds in France vary in population size with 628,000, 319,000, 2,402,000 cows, respectively, as of January 2018 [45]. These breeds differ also in terms of breeding program: the Holstein breed is widespread and genetic material is exchanged globally (*i.e.* it is an international breed), whereas Montbéliarde and Normande are French typical breeds managed at the country level (*i.e.* national breeds). Thus, the results drawn from our study can compare these two situations.



To conduct this study, we focused on bulls of these three breeds born between 2005 and 2010 for PTS, and bulls born between 2012 and 2015 for GS.

An increase in annual genetic gain with GS has been observed in most dairy cattle populations. For instance, in the US Holstein cattle breed, the annual genetic gain for milk, fat and protein yield in cows born before or after 2010 has increased by 71, 111 and 81%, respectively, between 2008 and 2014 [46, 47]. This can be explained by a drastic decrease in the generation interval by ~50% on average for Holstein bulls [10, 46]. This effective decrease in generation interval is consistent with previous predictions [9] and is mainly due to the withdrawal of progeny testing prior to marketing of bulls.

We obtained similar results for Montbéliarde, Normande and Holstein bulls. Their mean annual genetic gains (ΔG) increased by 50, 71 and 33%, respectively, after implementation of GS. At the same time, the bulls' parental generation intervals were reduced by a factor 1.7, 1.9 and 2 for the Montbéliarde, Normande and Holstein breeds, respectively, suggesting that, similar to the US Holstein situation, most of the extra genetic gain is due to reduced generation intervals. However, there are differences between breeds in terms of increase in annual genetic gain and decrease in generation interval, with the Holstein breed having less increase in gain relative to the reduction in generation interval. This may be due to different formulas used for total merit indices and different breeding goals.

Given this global increase in genetic gain for all the breeds studied here along with the decrease in generation interval, one might be concerned about the trend in genetic diversity. Indeed, on the one hand, an accelerated decrease in genetic diversity could be expected due to the decrease in generation interval. On the other hand, genomic evaluations can lead to a more diverse offer in terms of candidates for selection, and thus a moderation of inbreeding rate over time [5]. To study the possible change of inbreeding rate, we first looked at various estimates of inbreeding and diversity to recognize whether and why they differed. ROH-based inbreeding estimates aligned well with pedigree-based inbreeding estimates when averaged per breed. However, pedigree-based inbreeding estimates had a higher coefficient of

variation between bulls within a breed (0.33, 0.25 and 0.31 for Montbéliarde, Normande and Holstein, respectively) than ROH-based inbreeding estimates (0.22, 0.18 and 0.21 for Montbéliarde, Normande and Holstein, respectively), for the three breeds. Pedigree depths were substantial and similar in the three breeds and there was limited variation among individual bulls; the mean number of generations traced for Montbéliarde, Normande and Holstein bulls was 15.4 ± 1.7 , 15.8 ± 1.6 and 15.4 ± 1.9 , respectively. Therefore, pedigree completeness was not a major cause of the higher coefficient of variation of pedigree-based inbreeding estimates in comparison with ROH-based inbreeding estimates. Pedigree-based estimations of individual inbreeding coefficients are based on expected IBD probabilities and because of Mendelian sampling these values might differ from ROH-based estimates, which reflects more the realized or effective inbreeding [48]. The correlations between pedigree- and ROH-based inbreeding estimates ranged from 50 to 59%, depending on the breed. These correlations were in the range of those previously reported for Jersey and Danish red cattle (52 and 54%, respectively) [49] but lower than those reported for Holstein (around 80%) [49] and simulated cattle (around 70%) [11]. Errors and depth of pedigree are very plausible reasons for these differences between reported and presented correlations. Moreover, such differences in these correlations are expected because pedigree-based inbreeding corresponds to the expected value of inbreeding while ROH-based inbreeding corresponds to its realized value. In groups of highly selected bulls, linkage disequilibrium and homozygosity can be strong, which leads to large discrepancies between expected (based on pedigree) and observed (based on ROH) inbreeding and subsequently lower correlations.

In all three breeds, the ratio between the effective and actual number of bulls (N_e/N) increased after implementation of GS. This reflects changes in breeding schemes in all three breeds, which should allow for a decrease in inbreeding rate and more generally for a slowdown in the annual loss of genetic diversity. However, we distinguished two contrasting types of behavior of inbreeding based on ROH. First, for the two national breeds, Montbéliarde and Normande, we did not detect any significant change in the annual inbreeding rate (ΔF). Second, for the Holstein international breed there was a significant increase in ΔF , which has increased from 0.080% to 0.55% per year based on ROH, and from 0.088% to 0.49% per year based on pedigree. This result does not align with the fact that all three breeds experienced a significant increase in annual genetic gain, with the lowest increase observed in Holstein while the inbreeding rate

in Holstein was highest. This result suggests that changes in breeding schemes reflected by an increase of N_e/N of bulls in Holstein did not allow for a slowdown in the annual loss of genetic diversity.

For Holstein, ΔF per generation also increased from 0.39% to 1.36% based on ROH, and from around 0.46% to 1.39% based on pedigree. Thus, the decrease in generation interval did not explain all the increase in ΔF per year for Holstein, since ΔF per generation increased too. The observed inbreeding rate in Holstein is higher than the 1% per generation acceptable rate of inbreeding according to FAO guidelines [50]. Since the accumulation of inbreeding through ROH is not linear, and the observed level and rate of inbreeding could be underestimated when based on ROH, the inbreeding problem is likely even more severe than suggested by these figures.

We used a linear regression to estimate the rate of ROH-based inbreeding because it gave better results than a logarithmic regression in terms of residuals of the model (see Additional file 4: Figures S1–S7). When estimated from pedigree data, the slope of inbreeding increased significantly for both Normande and Holstein bulls, reaching 0.19% and 0.49% per year, respectively. The inbreeding rate reached 0.70% and 1.39% per generation, respectively, so the decrease in generation interval did not explain all the increase in inbreeding rate per year. Thus, the Normande breed remained under the recommended FAO threshold whereas Holstein exceeded this when estimating inbreeding on pedigree data as well as on ROH data. Why there are such differences between these two breeds should be further investigated, and in particular, the differences in breeding schemes, such as the use of foreign bulls or the setting-up of quotas for the number of inseminations allowed per bull.

From 2005 to 2015, the overall mean ROH length for Montbéliarde (1.68 Mb) was significantly shorter than that for Normande (1.71 Mb), which in turn was significantly shorter than that for Holstein (1.80 Mb). The length of ROH depends on two main criteria, for given detection parameters and genotyping density: (i) the effective population size and the level of linkage disequilibrium [51]; and (ii) the number of generations that separates an individual from the ancestor from which the ROH originated, with longer ROH reflecting more recent inbreeding [28, 52]. We investigate these two possibilities to explain the differences in mean ROH lengths between breeds.

Estimated effective population sizes (N_e) of French dairy cattle breeds were equal to 79 for Montbéliarde, 87 for Normande and 96 for Holstein in 2017 [53]. Based on these results, effective/census population sizes ratios

(N_e/N) of 0.009, 0.021 and 0.003% for these breeds, respectively, were calculated [53]. The effective number of bulls born under GS for the period between 2012 and 2015 was estimated to be 144 for Montbéliarde, 124 for Normande and 229 for Holstein, and effective/census number of bull ratios were 28, 35 and 13%, respectively. In both cases, the ratios between effective and census effective population sizes were lowest for Holstein and higher for Normande and Montbéliarde. We can hypothesize that these differences are probably linked to the small numbers of bulls that have strong contributions and the low ratios between effective and census number of bulls, particularly in Holstein [54]. Moreover, whereas the distribution of the number of offspring per bull became less skewed for the three breeds under GS than under PTS, this change seemed to be less pronounced in Holstein (see Additional file 6: Figure S8, Additional file 7: Table S6). This was consistent with a relatively low ratio between effective and census number of bulls in Holstein under GS in comparison with Montbéliarde and Normande, and with the hypothesis of a remaining effect of a few bulls with strong contributions. There can be two effects explaining the greater observed ROH length for Holstein. The first is the small effective population size relative to census population size resulting in a highly structured genome in this breed with long haploblocks [55] where recombination is unlikely to have occurred leading to long ROH. The second is recent inbreeding, likely to have occurred because of the massive use of few elite bulls. This was suggested by kinship estimated from pedigree being higher for the period between 2012 and 2015 than for the period between 2005 and 2010 for Holstein and not for national breeds, which can be explained by the intensive use of a few individuals as grandfathers of bulls, mostly coming from the USA [56]. We chose not to compute kinship from GRM because of their high sensitivity to the estimation of allelic frequencies (data not shown) [21]. The hypothesis that a few individuals were used as grandfathers of bulls in Holstein was consistent with the increase by a factor 8.77 (with a relative change of 7.77) of the annual rate of mean ROH length in Holstein breed after the beginning of GS, whereas it remained unchanged in the Montbéliarde and Normande breeds. This increase might be counterintuitive. Indeed, GS may lead to relatively shorter ROH than PTS [10]. Therefore, the increase in mean ROH length in Holstein might be underestimated and was corroborated by the increase in the number of long ROH per individual in Holstein with the beginning of GS, whereas it was not the case in Montbéliarde and Normande (see Additional file 8:

Table S7, Additional file 9: Figures S9–S11)). Therefore, the increase in the number of long ROH per individual in Holstein might be responsible for the increased mean ROH length within the breed. Shorter ROH tend to result from older inbreeding (older common ancestors) whereas longer ROH tend to result from more recent inbreeding (more recent common ancestors) [29, 31, 32]. Thus, the increase in the mean ROH length and in the number of long ROH in Holstein suggests an increase in the number of ROH segments arising from recent inbreeding. This introduction of recent inbreeding was confirmed by the observed increase in pedigree-based inbreeding calculated from the last five generations that started when GS was implemented, with an inbreeding rate ranging from -0.083% per year for PTS to 0.27% per year for GS (-0.40% and 0.78% per generation, respectively). In conclusion, in Holstein, this increase in the mean ROH length might reflect the substantial introduction of recent inbreeding, but also a reduction in the number of selected bulls.

Conclusions

After the introduction of genomic selection, the annual genetic gain increased based on estimated breeding values for bulls of three large French dairy cattle breeds: Montbéliarde, Normande and Holstein. At the same time, the inbreeding rate per year and per generation increased in Holstein bulls, an international breed, but not in Montbéliarde and Normande bulls, two national breeds. Small effective population sizes resulted in highly structured genomes with long haploblocks, particularly in Holstein, as a consequence of the low effective/census number of bulls ratios. A joint increase in genetic gain and conservation of genetic diversity is possible by (i) applying optimal contribution selection; (ii) carrying out on-farm mating plans; and (iii) ensuring that these tools are flexible enough to be used effectively in the field. Our study highlights the fact that changes in breeding schemes can impact genetic diversity, and thus, it is important to continue monitoring genetic diversity and to study the future impact of different breeding schemes strategies on genetic diversity. The recent implementation of genomic evaluations in small regional breeds, for which effective population size can be limited, should be studied carefully in order to ensure sustainability of breeding schemes in the future, in terms of both genetic gain and diversity. The spread of new technologies in breeding schemes, such as female reproductive technologies and embryo transfer that might increase global selection intensity through the dams of bulls, represents a risk in terms of genetic diversity that should be anticipated in the light of our conclusions.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12711-019-0495-1>.

Additional file 1: Table S1. Description of the three studied dairy cattle breeds: Montbéliarde, Normande and Holstein.

Additional file 2: Table S2. Length of the autosomal genome in three French dairy cattle breeds. Length of autosomal genome on which ROH can be detected corresponds to the length of the autosomal genome covered by SNPs after withholding gaps longer than 150 kb between two SNPs. The proportion of autosomal genome on which ROH can be detected is the length of autosomal genome on which ROH can be detected divided by the total length of the autosomal genome (from *Bos taurus* assembly 3.1.1).

Additional file 3: Table S3. Parental generation intervals. Mean parental generation intervals for bulls born between 2005 and for bulls born between 2010 and 2015 in months and in years, for Montbéliarde, Normande and Holstein.

Additional file 4: Figure S1. Square root of standardized residuals for the linear regression of pedigree-based inbreeding depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S2.** Square root of standardized residuals for the linear regression of pedigree-based inbreeding for the last five generations depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S3.** Square root of standardized residuals for the linear regression of pedigree-based kinship depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S4.** Square root of standardized residuals for the linear regression of F_{ROH} inbreeding based on ROH depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S5.** Square root of standardized residuals for the logarithmic regression of F_{ROH} inbreeding based on ROH depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right) ($F_{ROH, transformed} = \log_{10}(1 - F_{ROH})$). **Figure S6.** Square root of standardized residuals for the linear regression of the mean length of ROH depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S7.** Square root of standardized residuals for the linear regression of total merit index (ISU) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right).

Additional file 5: Table S4. Pedigree-based inbreeding calculated from the last five generations in percent per year. **Table S5.** Estimation of the slopes of pedigree-based inbreeding calculated from the last five generations in percent per generation.

Additional file 6: Figure S8. Proportion of the total number of offspring per bull for each breed and selection type. Progeny testing selection corresponds to bulls born between 2005 and 2010 and genomic selection to bulls born between 2012 and 2014.

Additional file 7: Table S6. Number of bulls and offspring for each birth period in the dataset.

Additional file 8: Table S7. ROH categories lengths thresholds.

Additional file 9: Figure S9. Confidence intervals of the number of ROH per ROH length category per selection type for Montbéliarde bulls. **Figure S10.** Confidence intervals of the number of ROH per ROH length category per selection type for Normande bulls. **Figure S11.** Confidence intervals of the number of ROH per ROH length category per selection type for Holstein bulls.

Acknowledgements

The authors would like to acknowledge Aurélie Baur, for her help in retrieving the genealogies of bulls, and Didier Boichard, for his helpful comments about this study. The authors would also like to acknowledge Valogene for providing genotyping data. This study was partially funded by the INRA SELGEN funding

metaprogram (Project GDivSelGen). The authors thank two anonymous reviewers as well as the editors for their valuable and useful comments about this study. This study was funded by ALLICE and APIS-GENE.

Authors' contributions

PC and GR conceived and designed the research. CH provided data. ACD performed the statistical analysis and drafted the manuscript. PC, DL and GR supervised the statistical analyses. ACD, PC, DL, AM, CH, SF and GR contributed to the interpretation of the results. ACD, PC, DL, AM, CH, CDB, SF and GR contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France. ² ALLICE, Paris, France. ³ Institut de l'Elevage, Paris, France.

Received: 14 February 2019 Accepted: 11 September 2019

Published online: 23 September 2019

References

- Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: genomic selection in dairy cattle: progress and challenges. *J Dairy Sci.* 2009;92:433–43.
- de Roos APW, Schrooten C, Veerkamp RF, van Arendonk JAM. Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. *J Dairy Sci.* 2011;94:1559–67.
- Wang J. Pedigrees or markers: which are better in estimating relatedness and inbreeding coefficient? *Theor Popul Biol.* 2016;107:4–13.
- Boichard D, Guillaume F, Baur A, Croiseau P, Rossignol MN, Boscher MY, et al. Genomic selection in French dairy cattle. *Anim Prod Sci.* 2012;52:115–20.
- Daetwyler HD, Villanueva B, Bijma P, Woolliams JA. Inbreeding in genome-wide selection. *J Anim Breed Genet.* 2007;124:369–76.
- Meuwissen THE. Maximizing the response of selection with a predefined rate of inbreeding. *J Anim Sci.* 1997;75:934–40.
- Engelsma KA, Veerkamp RF, Calus MPL, Windig JJ. Consequences for diversity when prioritizing animals for conservation with pedigree or genomic information: prioritization with pedigree or genomic data. *J Anim Breed Genet.* 2011;128:473–81.
- Eynard SE, Windig JJ, Hiemstra SJ, Calus MPL. Whole-genome sequence data uncover loss of genetic diversity due to selection. *Genet Sel Evol.* 2016;48:33.
- Schaeffer LR. Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet.* 2006;123:218–23.
- Doekes HP, Veerkamp RF, Bijma P, Hiemstra SJ, Windig JJ. Trends in genome-wide and region-specific genetic diversity in the Dutch–Flemish Holstein–Friesian breeding program from 1986 to 2015. *Genet Sel Evol.* 2018;50:15.
- Forutan M, Ansari Mahyari S, Baes C, Melzer N, Schenkel FS, Sarogolzaei M. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genomics.* 2018;19:98.
- Dickerson GE, Hazel LN. Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J Agric Res.* 1944;69:459–76.
- Barrett RDH, Schluter D. Adaptation from standing genetic variation. *Trends Ecol Evol.* 2008;23:38–44.
- Leroy G. Inbreeding depression in livestock species: review and meta-analysis. *Anim Genet.* 2014;45:618–28.
- Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet Sel Evol.* 2014;46:71.
- Reed DH, Frankham R. Correlation between fitness and genetic diversity. *Conserv Biol.* 2003;17:230–7.

17. Meuwissen THE, Luo Z. Computing inbreeding coefficients in large populations. *Genet Sel Evol.* 1992;24:305–13.
18. Danchin-Burge C, François L, Laloë D, Leroy G, Verrier E. Using the information collected for genetic evaluation to assess the French ruminant and equine within-breeds genetic variability. *Interbull Bull.* 2013;47:111–9.
19. Boichard D, Maignel L, Verrier E. The value of using probabilities of gene origin to measure genetic variability in a population. *Genet Sel Evol.* 1997;29:5–23.
20. Oliehoek PA, Bijma P. Effects of pedigree errors on the efficiency of conservation decisions. *Genet Sel Evol.* 2009;41:9.
21. Eynard SE, Windig JJ, Leroy G, van Binsbergen R, Calus MP. The effect of rare alleles on estimated genetic relationships from whole genome sequence data. *BMC Genet.* 2015;16:24.
22. Eynard SE, Windig JJ, Leroy G, Verrier E, Hiemstra SJ, Van Binsbergen R, et al. The use of whole genome sequence data to estimate genetic relationships including rare alleles information. In: Proceedings of the 10th world congress of genetics applied to livestock production: 17–22 August 2014; Vancouver. 2014.
23. de Cara MÁR, Villanueva B, Toro MÁ, Fernández J. Using genomic tools to maintain diversity and fitness in conservation programmes. *Mol Ecol.* 2013;22:6091–9.
24. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. Runs of homozygosity in European populations. *Am J Hum Genet.* 2008;83:359–72.
25. Curik I, Ferenčaković M, Sölkner J. Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livest Sci.* 2014;166:26–34.
26. Ferenčaković M, Hamzić E, Gredler B, Solberg TR, Klemetsdal G, Curik I, et al. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J Anim Breed Genet.* 2013;130:286–93.
27. Keller MC, Visscher PM, Goddard ME. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics.* 2011;189:237–49.
28. Howrigan DP, Simonson MA, Keller MC. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics.* 2011;12:460.
29. Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. *BMC Genet.* 2012;13:70.
30. Ferenčaković M, Sölkner J, Curik I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genet Sel Evol.* 2013;45:42.
31. Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF. Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet.* 2018;19:220–34.
32. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic runs of homozygosity record population history and consanguinity. *PLoS One.* 2010;5:e13996.
33. Bosse M, Megens HJ, Madsen O, Crooijmans RPMA, Ryder OA, Austerlitz F, et al. Using genome-wide measures of coancestry to maintain diversity and fitness in endangered and domestic pig populations. *Genome Res.* 2015;25:970–81.
34. Gómez-Romano F, Villanueva B, Fernández J, Woolliams JA, Pong-Wong R. The use of genomic coancestry matrices in the optimisation of contributions to maintain genetic diversity at specific regions of the genome. *Genet Sel Evol.* 2016;48:2.
35. Sonesson AK, Woolliams JA, Meuwissen TH. Genomic selection requires genomic control of inbreeding. *Genet Sel Evol.* 2012;44:27.
36. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018. <https://www.R-project.org/>.
37. Coster A. Pedigree: Pedigree functions. 2012. <https://CRAN.R-project.org/package=pedigree>.
38. Launay A, Rocland M. Genetic evaluation—methods and definitions. http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/genetic-evaluation-methods-and-definitions.html. Accessed 17 Dec 2018.
39. Boichard D. Pedig: a fortran package for pedigree analysis suited for large populations. 2002. <https://prodinra.inra.fr/?locale=fr#ConsultNotice:7255>. Accessed 3 Oct 2018.
40. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7.
41. Purcell S, Chang C. PLINK 1.9. <http://www.cog-genomics.org/plink/1.9/>. Accessed 3 Oct 2018.
42. Tarpy DR, Nielsen DI. Sampling error, effective paternity, and estimating the genetic structure of Honey bee colonies (Hymenoptera: Apidae). *Ann Entomol Soc Am.* 2002;95:513–28.
43. Starr CK. Sperm competition, kinship, and sociality in the Aculeate Hymenoptera. In: Smith RL, editor. Sperm competition and the evolution of animal mating systems. New York: Academic Press; 1984. p. 427–64.
44. Colleau JJ, Fritz S, Guillaume F, Baur A, Dupassieux D, Boscher MY, et al. Simulation des potentialités de la sélection génomique chez les bovins laitiers. *INRA Prod Anim.* 2015;28:251–8.
45. Groupe Economie du Bétail. Chiffres-clés Bovins 2018—Productions Bovines lait et viande. <http://idele.fr/filieres/publication/idelesolr/recommends/chiffres-cles-bovins-2018.html>. Accessed 22 Nov 2018.
46. García-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, Van Tassell CP. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc Natl Acad Sci USA.* 2016;113:E3995–4004.
47. Howard JT, Pryce JE, Baes C, Maltecca C. Invited review: inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability. *J Dairy Sci.* 2017;100:6009–24.
48. Hill WG, Weir BS. Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res (Camb).* 2011;93:47–64.
49. Zhang Q, Calus MP, Guldbandsen B, Lund MS, Sahana G. Estimation of inbreeding using pedigree, 50 k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genet.* 2015;16:88.
50. FAO. Secondary guidelines for development of national farm animal genetic resources management plans: Management of small populations at risk. Rome: FAO and UNEP; 2000. p. 219.
51. Waples RS, England PR. Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics.* 2011;189:633–44.
52. Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. *Hum Mol Genet.* 2006;15:789–95.
53. Danchin-Burge C. Indicateurs de variabilité génétique—races bovines—Edition 2018. http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/indicateurs-de-variabilite-genetique-races-bovines-edition-2018-1.html. Accessed 22 Nov 2018.
54. Boichard D, Ducrocq V, Fritz S, Colleau JJ. Where is dairy cattle breeding going? A vision of the future. *Bull Int Bull Eval Serv.* 2010;41:63–8.
55. Jónás D, Ducrocq V, Croiseau P. Short communication: the combined use of linkage disequilibrium-based haploblocks and allele frequency-based haplotype selection methods enhances genomic evaluation accuracy in dairy cattle. *J Dairy Sci.* 2017;100:2905–8.
56. Le Mézec P, Danchin-Burge C. Davantage de diversité avec la génomique ? plutôt non... <http://idele.fr/contact/publication/idelesolr/recommends/les-programmes-de-selection-et-de-diffusion-de-taureaux-dia-ler-de-la-genomique-et-leurs-effets.html>. Accessed 21 Jan 2019.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Additional file 1 Table S1: Description of the three studied dairy cattle breeds: Montbéliarde, Normande and Holstein

Breed	Montbéliarde	Normande	Holstein	Reference
Number of cows in France (on 2018, January 1)	628	319	2,402,000	Groupe Economie du Bétail, 2018
Number of milk-recorded cows in France (in 2017)	430,186	194,524	1,572,614	Thomas et al., 2018
Total number of AI performed by pure bred bulls in France (in 2017)	723,782	460,32	3,581,695	Moureaux and Le Mezec, 2018
Number of first AI performed by pure bred bulls in France (in 2017)	442,235	262,594	1,982,418	Moureaux and Le Mezec, 2018
Total number of pure bred bulls (in 2017)	1,061	725	4,38	Moureaux and Le Mezec, 2018
Mean milk yield per cow per lactation in France in 2017 (in kg)	6,933	6,59	9,241	Thomas et al., 2018
Fat/protein content in milk in 2017 (in g/kg)	38.9/33.1	42.4/34.7	39.7/32.0	Thomas et al., 2018
Estimated number of effective ancestors for the females born between 2014 and 2017	18	21	20	Danchin-Burge, 2018
Mean inbreeding for the females born in France between 2014 and 2017 (in %)	4.8	5.0	4.4	Danchin-Burge, 2018

Danchin-Burge, C., 2018. Indicateurs de variabilité génétique - races bovines - Edition 2018.

Groupe Economie du Bétail, 2018. Chiffres-clés Bovins 2018 - Productions lait et viande.

Moureaux, S., Le Mezec, P., 2018. BGIAP 2017 Bilan Génétique de l'Insémination en races bovines laitières.

Thomas, G., Bourrigan, X., Astruc, J.-M., 2018. Résultats de Contrôle Laitier - France 2017.

Additional file 2 Table S2: Length of the autosomal genome in three French dairy cattle breeds

Length of autosomal genome on which ROH can be detected corresponds to the length of the autosomal genome covered by SNPs after withholding gaps longer than 150kb between two SNPs. The proportion of autosomal genome on which ROH can be detected is the length of autosomal genome on which ROH can be detected divided by the total length of the autosomal genome (from *Bos taurus* assembly 3.1.1).

Breed	Total length of the autosomal genome (in Mb)	Length of autosomal genome on which ROH can be detected (in Mb)	Proportion of autosomal genome on which ROH can be detected
Montbéliarde	2,493.12	1,919.31	77.0%
Normande	2,493.12	1,919.65	77.0%
Holstein	2,493.12	1,959.21	78.6%

Additional file 3 Table S3: Parental generation intervals

Breed	Mean generation interval between 2005 and 2010		Mean generation interval between 2012 and 2015	
	In months	In years	In months	In years
Montbéliarde	69.1	5.76	46.2	3.85
Normande	69.8	5.81	44.9	3.74
Holstein	58.0	4.84	34.3	2.86

Additional file 4: Residuals of regressions of genetic diversity and genetic gain parameters depending on birth year and selection type.

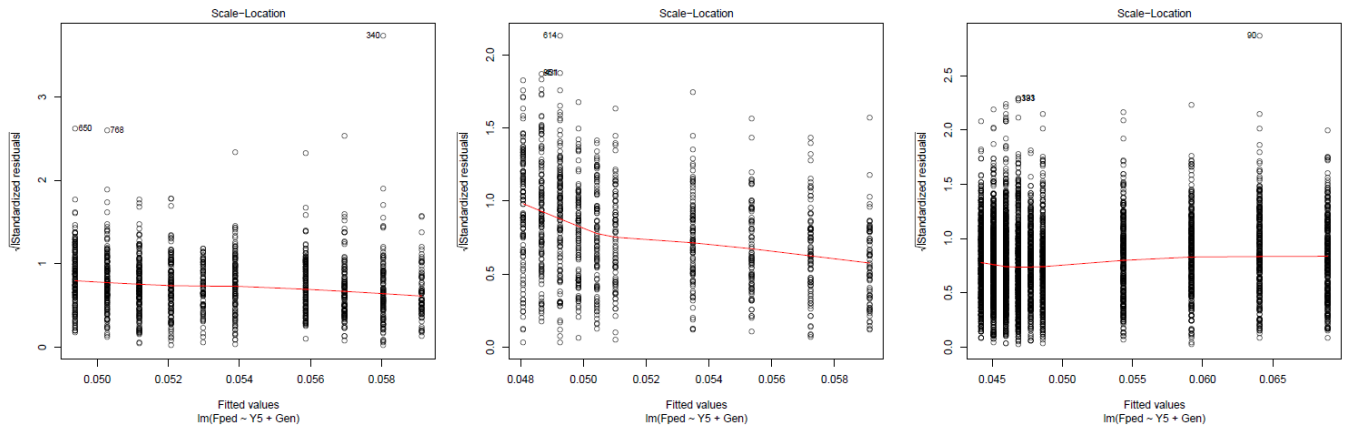


Figure S1: Square root of standardized residuals for the linear regression of pedigree-based inbreeding depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

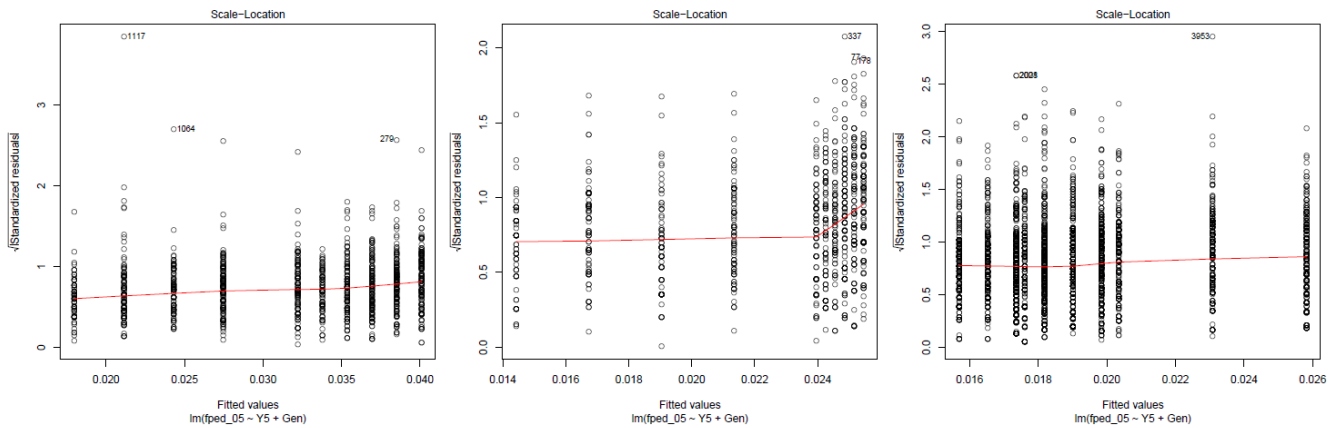


Figure S2: Square root of standardized residuals for the linear regression of pedigree-based inbreeding for the last five generations depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

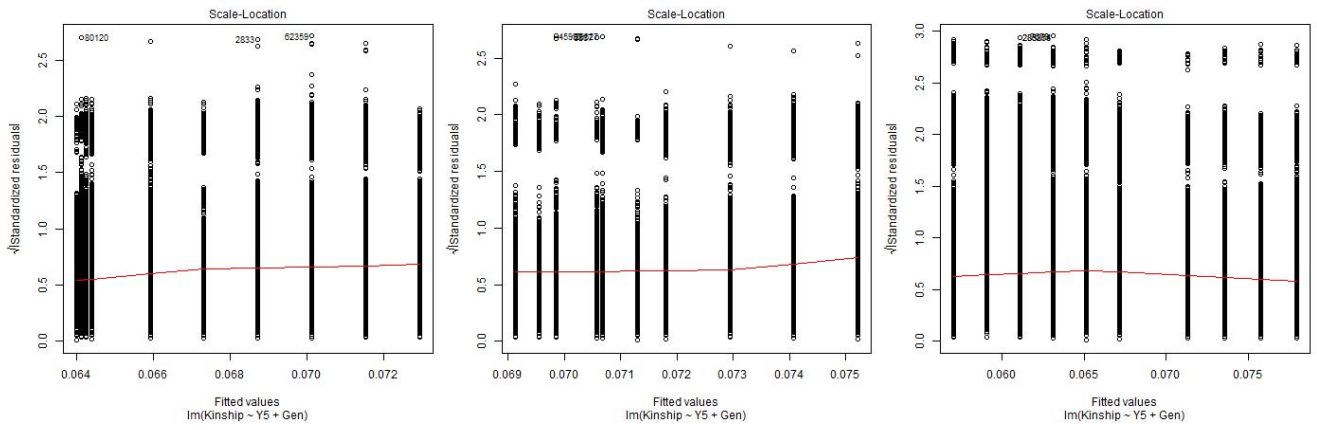


Figure S3: Square root of standardized residuals for the linear regression of pedigree-based kinship depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

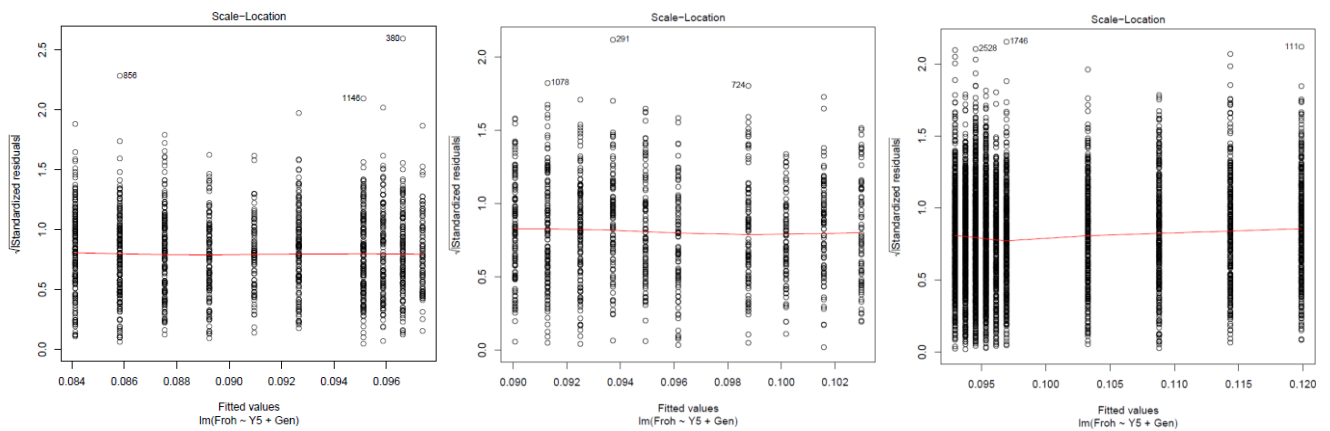


Figure S4: Square root of standardized residuals for the linear regression of F_{ROH} , inbreeding based on Runs of Homozygosity (ROH) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

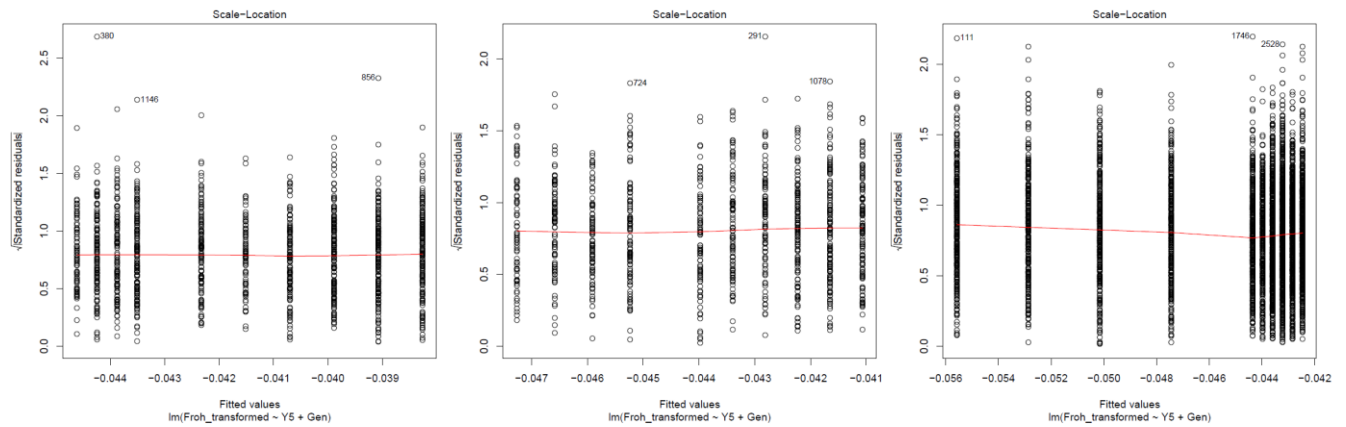


Figure S5: Square root of standardized residuals for the logarithmic regression of F_{ROH} , inbreeding based on Runs of Homozygosity (ROH) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

$$F_{ROH, \text{transformed}} = \log_{10}(1 - F_{ROH})$$

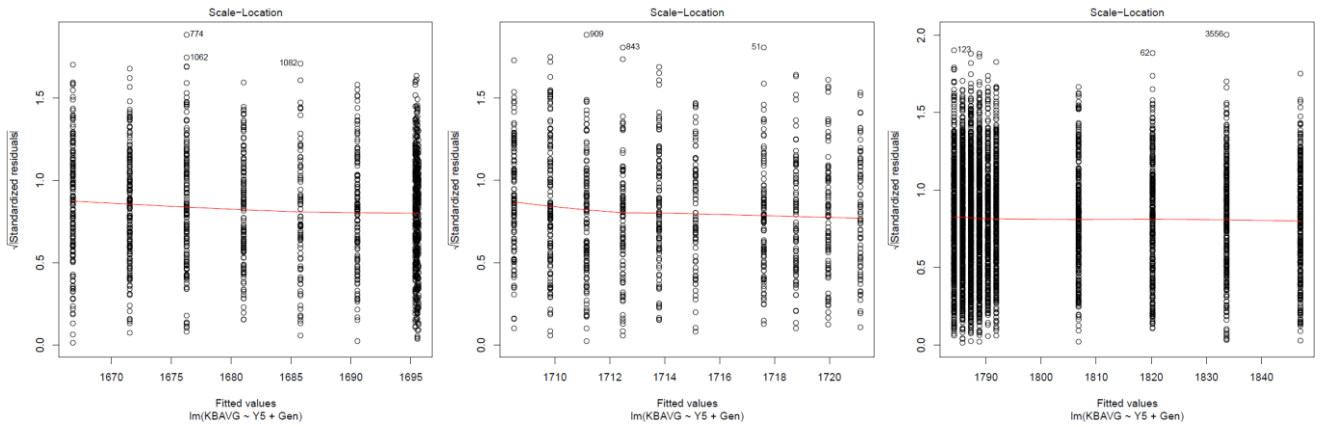


Figure S6: Square root of standardized residuals for the linear regression of the mean length of Runs of Homozygosity (ROH) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

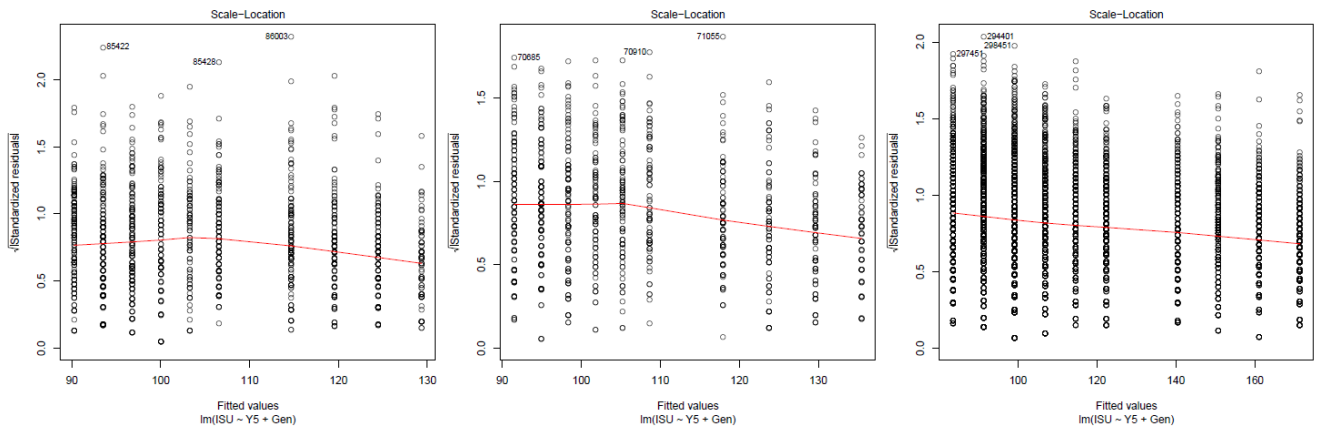


Figure S7: Square root of standardized residuals for the linear regression of Total Merit Index (ISU) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

Additional file 5: Pedigree-based inbreeding calculated from the last five generations per year and per generation.

Table S4: Pedigree-based inbreeding calculated from the last five generations in percent per year.

Breed	b_1 (\pm standard error)	p -value of b_1	δ (\pm standard error)	p -value of δ	b_2	Relative change
Montbéliarde	-0.16 (\pm 0.030)	< 2e-16	-0.16 (\pm 0.077)	0.039	-0.32	-1.00
Normande	-0.030 (\pm 0.024)	< 2e-16	-0.20 (\pm 0.059)	7.4e-04	-0.23	-6.65
Holstein	-0.083 (\pm 0.013)	9.6e-11	0.36 (\pm 0.032)	< 2.2e-16	0.27	4.30

b_1 is the slope of regression of pedigree-based inbreeding calculated from the last five generations depending on birth year for progeny tested bulls born between 2005 and 2010 (progeny testing selection), and δ the difference between the slopes of regression of pedigree-based inbreeding calculated from the last five generations depending on birth year for progeny testing selection and for marketed bulls born between 2012 and 2015 (genomic selection), for Montbéliarde, Normande and Holstein. The relative change is equal to $\frac{\delta}{|b_1|}$. b_2 is the slope of the pedigree-based inbreeding for the last five generations for marketed bulls born between 2012 and 2015 (genomic selection), equal to $b_1 + \delta$. The p -value of δ corresponds to the significance of the non-nullity of δ .

Table S5: Estimation of the slopes of pedigree-based inbreeding calculated from the last five generations in percent per generation.

Breed	b_1 per generation	δ per generation	b_2 per generation	Relative change per generation
Montbéliarde	-0.91%	-3.1e-03	-1.22%	-0.34
Normande	-0.18%	-6.9e-03	-0.86%	-3.94
Holstein	-0.40%	1.2e-02	0.78%	2.95

Additional file 6: Proportion of the total number of offspring per bull for each breed and selection type.

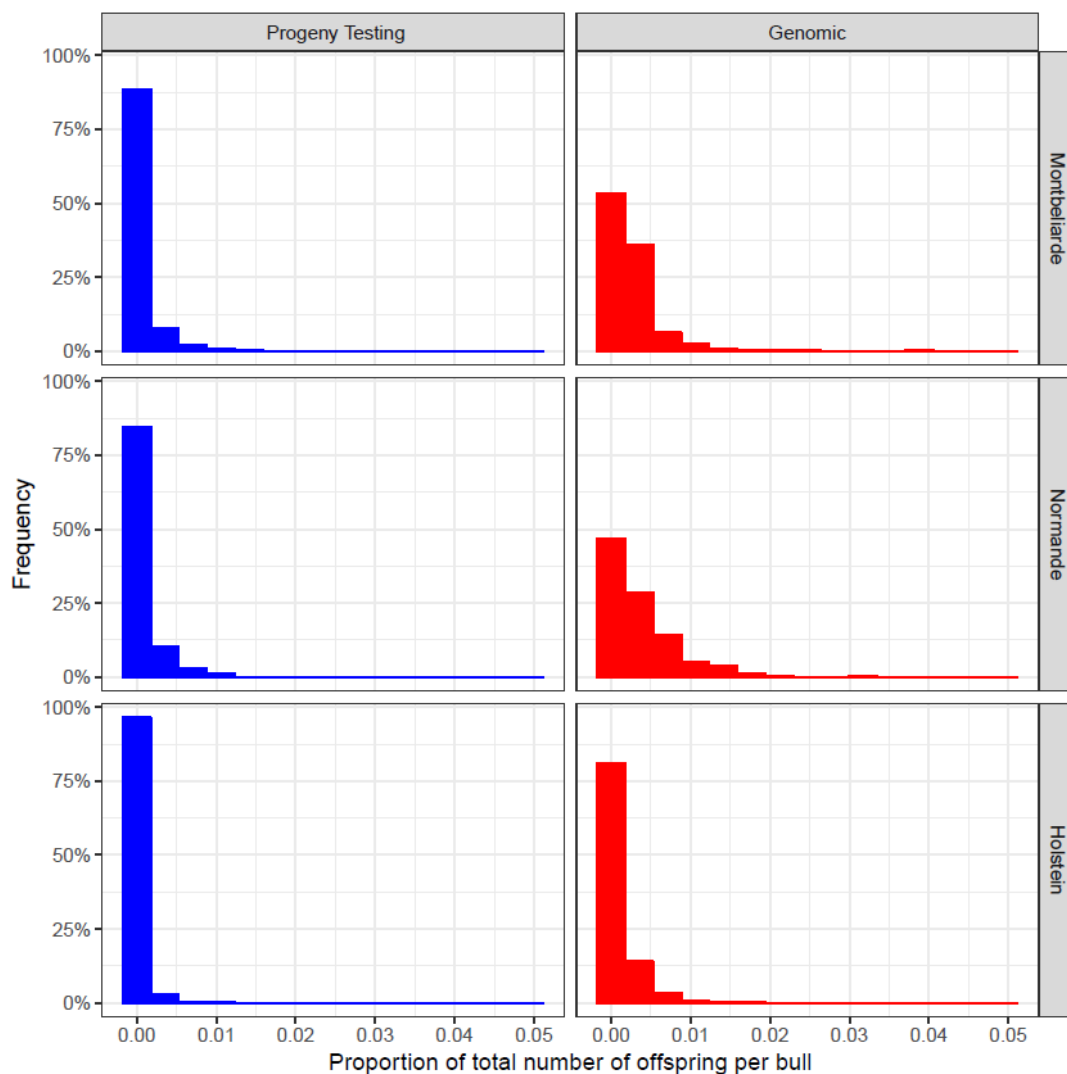


Figure S8: Proportion of the total number of offspring per bull for each breed and selection type

Progeny testing selection corresponds to bulls born between 2005 and 2010 and genomic selection to bulls born between 2012 and 2014.

For example, for Montbéliarde, around 50% of the bulls born under genomic selection (y-axis) are the fathers of less than $0.003 = 0.3\%$ of all calves (x-axis) and around 30% of the bulls are the fathers of 0.3% to 0.6% of all calves.

Additional file 7: Number of bulls and offspring for each birth period in the dataset.

Table S6: Number of bulls and offspring for each birth period in the dataset.

Breed	Birth period	Total number of individuals	Total number of offspring
Montbéliarde	2005 – 2010 Progeny-tested bulls	826	1,102,266
	2012 – 2014 Marketed bulls	399	562,178
	2005 – 2010 Progeny-tested bulls	666	542,939
Normande	2012 – 2014 Marketed bulls	347	437,558
	2005 – 2010 Progeny-tested bulls	3440	4,573,893
	2012 – 2014 Marketed bulls	1248	2,390,778

Additional file 8: ROH categories lengths thresholds

For each breed, we selected bulls born in 2005, obtaining three cohorts: Montbéliarde bulls born in 2005, Normande bulls born in 2005 and Holstein bulls born in 2005. For each of these three cohorts, five ROH length categories were defined with R function kmeans (R Core Team, 2018) by applying it to the list of all ROH for all individuals of each cohort, for each breed separately. We obtained length thresholds for each category.

Table S7: ROH categories lengths thresholds

Category	Montbéliarde		Normande		Holstein	
	Minimum ROH (kb)	Maximum ROH length (kb)	Minimum ROH (kb)	Maximum ROH length (kb)	Minimum ROH (kb)	Maximum ROH length (kb)
1	1000.079	1426.286	1000.051	1458.437	1000.079	1465.802
2	1426.639	2080.933	1459.128	2121.894	1466.58	2163.376
3	2084.642	3046.285	2127.988	3087.975	2164.072	3247.136
4	3055.704	4723.705	3091.578	4730.121	3249.867	5031.288
5	4749.088	8718.715	4761.585	8217.436	5040.034	9126.138

REFERENCES

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Additional file 9: Number of ROH in different length categories

For each breed, we selected bulls born in 2005, obtaining three cohorts: Montbéliarde bulls born in 2005, Normande bulls born in 2005 and Holstein bulls born in 2005. For each of these three cohorts, five ROH length categories were defined with R function kmeans (R Core Team, 2018) by applying it to the list of all ROH for all individuals of each cohort, for each breed separately. We obtained length thresholds for each category [see Additional file 8 Table S7]. Then, we counted how many ROH were in each category for each individual. This allowed computing the mean number of ROH per individual per category for all individuals born during progeny testing selection (between 2005 and 2010) or during genomic selection (between 2012 and 2015) for each breed separately. We then performed an anova to compare the number of ROH in each category between progeny testing and genomic selection and estimated confidence intervals of these numbers with R function emmeans (Russell, 2018).

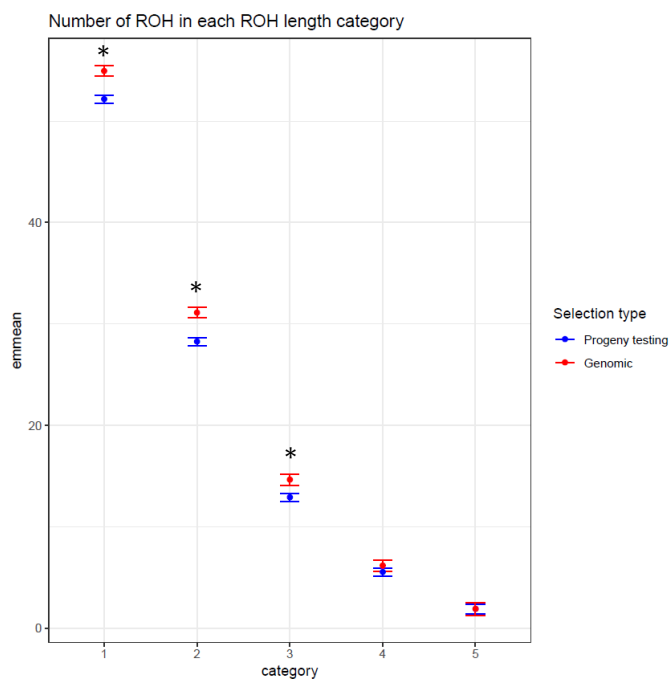


Figure S9: Confidence intervals of the number of ROH per ROH length category per selection type for Montbéliarde bulls. *: mean ROH numbers were significantly different between progeny testing and genomic selection for this category.

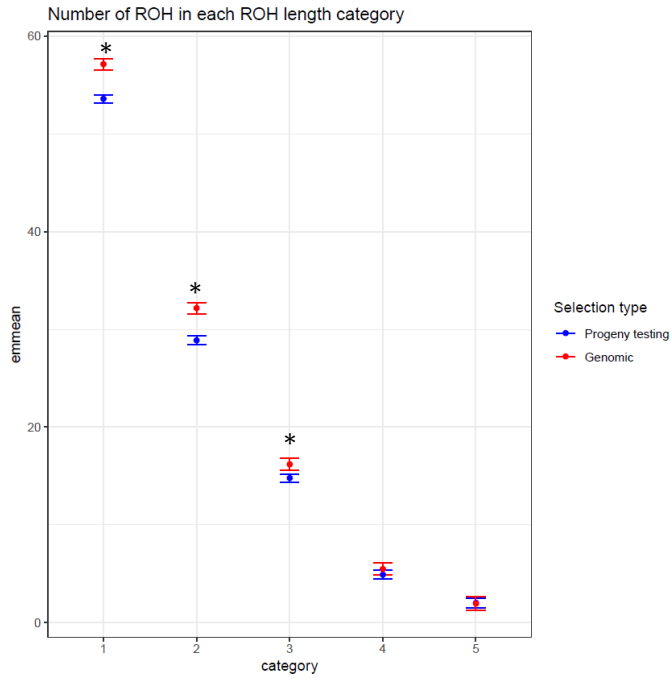


Figure S10: Confidence intervals of the number of ROH per ROH length category per selection type for Normande bulls. *: mean ROH numbers were significantly different between progeny testing and genomic selection for this category.

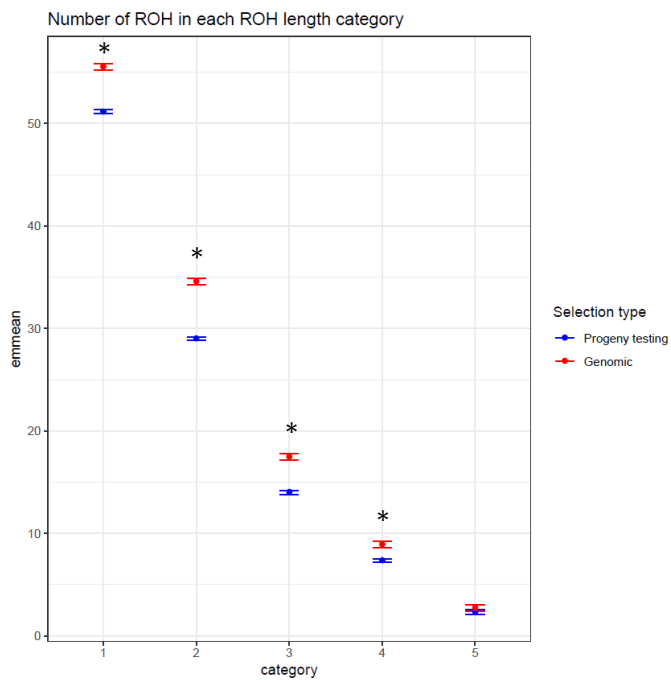


Figure S11: Confidence intervals of the number of ROH per ROH length category per selection type for Holstein bulls. *: mean ROH numbers were significantly different between progeny testing and genomic selection for this category.

In conclusion, small (categories 1 and 2) and medium length (category 3) ROH were significantly more numerous during genomic selection than during progeny testing selection for all three breeds. However, long ROH (category 4, between 3.25 and 5.03 Mb) were more significantly numerous during genomic selection than during progeny testing selection only for Holstein.

REFERENCES

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Russell Lenth (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.0. <https://CRAN.R-project.org/package=emmeans>

III - Discussion

Les évaluations génomiques permettent d'augmenter le nombre de candidats à la sélection voire le nombre de taureaux diffusés, avec un renouvellement plus rapide du catalogue des taureaux diffusés. Cela devrait permettre de tirer parti des avantages des évaluations génomiques sur le plan du progrès génétique tout en ménageant la diversité génétique. Or, si le nombre de taureaux ou de candidats à la sélection n'augmente pas voire diminue, ou si les taureaux sélectionnés sont plus apparentés qu'auparavant, l'effet d'accélération de la perte de diversité génétique dû au raccourcissement important des intervalles de génération n'est plus compensé et la perte de diversité génétique s'accélère, comme nous l'avons observé dans la race Prim'Holstein.

La petite taille efficace des populations a entraîné une forte structuration de leurs génomes avec de l'apparition de longs blocs haplotypiques, en particulier chez la Prim'Holstein. C'est le résultat d'une intensité de sélection élevée (faible nombre de taureaux), de la consanguinité qui en résulte ainsi que de la dérive génétique, d'autant plus forte que la taille efficace de la population est petite. Une augmentation conjointe du progrès génétique et de la conservation de la diversité génétique est possible (i) en appliquant des méthodes de contribution optimale (OC) ; (ii) en réalisant des plans d'accouplement en ferme pertinents ; et (iii) en veillant à ce que ces outils soient suffisamment souples pour être utilisés efficacement sur le terrain.

Cette étude souligne le fait que les conséquences de la sélection génomique sur la diversité génétique dépendent de la manière d'utiliser les évaluations génomiques, et que la conception et les modifications des schémas de sélection, tout comme la manière de diffuser les taureaux sélectionnés, jouent un rôle majeur dans la gestion de la diversité génétique. La sélection génomique est donc un outil permettant l'augmentation du progrès génétique sans nécessairement mettre en danger la diversité génétique des races bovines laitières françaises.

Les entreprises de sélection sont responsables du maintien de la diversité génétique mais aussi du progrès génétique dans les races qu'elles sélectionnent. Ces résultats leur ont permis de savoir quel a été l'impact de leurs choix au moment de la mise en place de la sélection génomique, et si les modifications apportées aux schémas de sélection leur ont permis d'atteindre leurs objectifs en termes de progrès et de diversité génétiques.

Chapitre 3 L'impact des technologies de reproduction et de la baisse du nombre de taureaux sur la diversité et le progrès génétiques dans les races bovines laitières françaises de grand effectif

I - Résumé de l'étude

Les schémas de sélection sont des facteurs clés dans la gestion de la diversité et du progrès génétiques. Dans les races bovines laitières, les technologies de transfert embryonnaire (MOET et OPU-FIV) sont de plus en plus utilisées depuis 2015. En outre, les entreprises de sélection ont tendance à réduire à la fois le nombre de taureaux dans les schémas de sélection (pères à taureaux) et de taureaux mis en marché, et ce pour des raisons économiques. Les conséquences de ces changements dans la gestion des schémas de sélection, tant en termes de progrès génétique que de diversité génétique, ne sont pas simples à prévoir.

Nous avons simulé 15 schémas de sélection génomique similaires à ceux admis dans les grandes races laitières françaises. Les scénarios différaient en termes de dimensions (c'est-à-dire en nombre de pères à taureaux et de taureaux mis en marché, et en nombre de femelles donneuses d'embryons ou d'ovocytes), d'intensité d'utilisation des technologies de reproduction (c'est-à-dire le nombre de veaux par femelle donneuse) et de type de technologie employée (MOET ou OPU-FIV). Par rapport à la MOET, l'OPU-FIV permet l'obtention d'embryons provenant de plus de pères différents, pour un même nombre d'embryons issus d'une même mère. Ils ont ensuite été évalués en termes de progrès génétique et de taux de consanguinité estimée à partir des ROH.

D'après nos simulations, l'utilisation intensive du transfert embryonnaire entraîne une amélioration du progrès génétique et une détérioration de la diversité génétique. La réduction du nombre de taureaux a renforcé l'impact négatif du transfert embryonnaire sur la diversité génétique. La supériorité de l'OPU-FIV sur la MOET en termes de progrès génétique était principalement due à la réduction de l'intervalle de génération consécutive permise par l'OPU-FIV. En effet, la récolte d'embryons par OPU-FIV ne nécessite pas que les donneuses aient atteint la puberté, contrairement à la MOET, et peut être pratiquée dès deux mois d'âge.

La réduction de l'intervalle de génération a deux conséquences qui se combinent pour augmenter le progrès génétique et le taux de consanguinité : (i) un effet multiplicateur sur le taux de consanguinité et le progrès génétique annuels et (ii) une augmentation de la proportion de donneuses nées elles-mêmes de donneuses, ce qui a pour effet de créer des « lignées star » de mères à taureaux qui seront surutilisées. Cela entraîne une forte augmentation du taux de

consanguinité, ce qui laisse supposer que cette pratique pourrait avoir des effets délétères importants sur la diversité génétique des races bovines laitières française.

II - Article soumis : Intensifying the use of reproductive technologies and reducing the dimensions of breeding schemes put genetic diversity at risk in dairy cattle breeds

Anna-Charlotte Doublet^{1,2}, Gwendal Restoux², Sébastien Fritz^{1,2}, Laura Balberini³, Guillaume Fayolle⁴, Chris Hozé^{1,2}, Denis Laloë² and Pascal Croiseau²

¹ ALLICE, Paris, France

² Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350, Jouy-en-Josas, France.

³ AURIVA-Elevage, 69126 Brindas, France

⁴ Groupe UMOTEST, 01250, Ceyzériat, France

Article

Intensified Use of Reproductive Technologies and Reduced Dimensions of Breeding Schemes Put Genetic Diversity at Risk in Dairy Cattle Breeds

Anna-Charlotte Doublet ^{1,2}, Gwendal Restoux ², Sébastien Fritz ^{1,2}, Laura Balberini ³, Guillaume Fayolle ⁴, Chris Hozé ^{1,2}, Denis Laloë ²  and Pascal Croiseau ^{2,*} 

¹ ALLICE, 149 rue de Bercy, 75012 Paris, France; anna-charlotte.doublet@inrae.fr (A.-C.D.); sebastien.fritz@allice.fr (S.F.); chris.hoze@allice.fr (C.H.)

² Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France; gwendal.restoux@inrae.fr (G.R.); denis.laloe@inrae.fr (D.L.)

³ AURIVA-Elevage, 61 Chemin des Hoteaux, 69126 Brindas, France; laura.balberini@auriva-elevage.fr

⁴ Groupe UMOTEST, Les Soudanières, 01250 Ceyzériat, France; gfayolle@umotest.com

* Correspondence: pascal.croiseau@inrae.fr

Received: 23 September 2020; Accepted: 15 October 2020; Published: 17 October 2020



Simple Summary: Maintaining genetic diversity in dairy cattle breeds is essential to support good performance and avoid inbreeding depression. This diversity could be threatened, however, by the recent increase in the use of reproductive technologies and the limited number of bulls in dairy cattle breeds. This study aimed to investigate the effects of these practices by simulating 15 breeding schemes similar to those carried out in the main dairy cattle breeds in France. We found that intensive use of reproductive technologies resulted in improved genetic gain, but diminished genetic diversity. However, this detrimental effect could be mitigated by maintaining a sufficient number of bulls in the breeding scheme.

Abstract: In the management of dairy cattle breeds, two recent trends have arisen that pose potential threats to genetic diversity: the use of reproductive technologies (RT) and a reduction in the number of bulls in breeding schemes. The expected outcome of these changes, in terms of both genetic gain and genetic diversity, is not trivial to predict. Here, we simulated 15 breeding schemes similar to those carried out in large French dairy cattle breeds; breeding schemes differed with respect to their dimensions, the intensity of RT use, and the type of RT involved. We found that intensive use of RT resulted in improved genetic gain, but deteriorated genetic diversity. Specifically, a reduction in the interval between generations through the use of ovum pick-up and in vitro fertilization (OPU-IVF) resulted in a large increase in the inbreeding rate both per year and per generation, suggesting that OPU-IVF could have severe adverse effects on genetic diversity. To achieve a given level of genetic gain, the scenarios that best maintained genetic diversity were those with a higher number of sires/bulls and a medium intensity of RT use or those with a higher number of female donors to compensate for the increased intensity of RT.

Keywords: inbreeding; MOET; OPU-IVF; simulation study; genomic selection

1. Introduction

The degree of inbreeding and, more generally, the loss of genetic diversity in domestic animal populations have been the subject of particular attention since the first half of the last century [1]. Low genetic diversity leads to lower expected genetic gain [2,3], altered adaptive potential, and increased inbreeding depression, which all have a direct cost for breeders and breeding companies [4,5].

Human-designed breeding schemes may play a key role as strong selection in domestic populations may result in low genetic diversity [6]. This is especially true in dairy cattle breeds, for which the management of genetic diversity has become an important challenge.

One of the main approaches to managing genetic diversity in a dairy cattle breed is through breeding schemes, particularly with respect to the number of animals considered at each of the selection steps [7,8]. One important factor in the design of breeding schemes is the number of sires of bulls and marketed bulls included in the scheme. In addition, breeders must determine if they want to use reproductive technologies (RT hereafter), and if so, they must make decisions regarding the type and intensity of RT and the number of females to be used as embryo or oocyte donors [8–10].

In dairy cattle breeds, breeding schemes are designed and implemented by breeding companies with the goal of producing elite bulls whose high-genetic-quality semen can be commercialized. The use of artificial insemination is widespread, with the result that genetic diversity in a breed is highly correlated with genetic diversity in breeding schemes [11]. Currently, genetic diversity in French breeding programs is mostly managed by controlling the number of half or full siblings chosen as male candidates for selection, but kinship ties further back in the pedigree tend to be largely ignored [12].

Genetic diversity in French dairy cattle breeds is typically monitored using pedigree-based inbreeding estimates (*VARIabilité génétique des RUMinants et des Equidés*, genetic variability in ruminants and equines, VARUME) [13]. However, more accurate and comprehensive evaluations can be obtained using genomic data [6,14,15]. In particular, inbreeding estimates based on genotypes and runs of homozygosity (ROH), in particular, have been shown to be as efficient, if not more, in evaluating and managing genetic diversity [16]. Although pedigree information is still very useful, genomic or ROH-based inbreeding estimates can provide more detailed data for better monitoring of genetic diversity.

The aim of a breeding company is to increase the genetic gain. This is generally achieved by an increase in selection intensity using RT, which is accompanied by a decrease in the number of bulls considered and the interval between generations. The use of RT drastically modifies breeding schemes and can have significant impacts on genetic diversity. The same can also be said of other technologies such as genomic selection (GS). On one hand, it is possible to use GS to preserve genetic diversity (e.g., by increasing the number of candidates for selection and incorporating a wider range of bulls). On the other, however, the use of GS can also pose a threat to genetic diversity (due to shorter generation intervals, for example). Some studies have shown that the implementation of GS programs resulted in an accelerated loss of genetic diversity, with severe increases in inbreeding rates (both per year and per generation) in breeds such as Holstein [7,17,18]. However, this loss was not observed in other breeds because of differences in the breeding schemes and the ways bulls are chosen. For instance, in France, Montbéliarde and Normande breeds were able to maintain their inbreeding rates while increasing their genetic gain [7]. Thus, the beneficial or detrimental effects of such technologies depend on the breeding scheme(s) and the population in question. The management of breeding schemes therefore represents a key factor in the management of genetic diversity and genetic gain in dairy cattle breeds.

In large dairy cattle breeds, RTs such as multiple ovulation and embryo transfer (MOET) and ovum pick-up and in vitro fertilization (OPU-IVF) are commonly used. In combination with genomic selection, RT appears to be very beneficial for genetic gain in dairy breeds [19–23]. The increased genetic gain associated with RT is due to an intensification in the use of the best females, with higher selection intensity [24], and to a reduction in the intervals between generations. With OPU-IVF, generation intervals can be shortened even more, as OPU can be performed on immature young heifers as soon as two months of age [25]. Taken together, these effects tend to increase the loss of genetic diversity, but this can be at least partially counterbalanced with the use of genomic selection to more-accurately evaluate candidates for selection and thus reduce within-family selection [15]. Moreover, the true effects of the different types of RT on genetic gain and genetic diversity depend on the number of female donors in the breeding scheme and the intensity of their use (number of calves born from RT) [10]. This is also true for the selection of bulls. The cost of breeding one bull has decreased dramatically

since the implementation of genomic selection, as this process has nearly put an end to expensive progeny testing (€45,000 on average per bull in France, although this value varies depending on the breed). Reducing the number of bulls also reduces the total cost of breeding bulls, allowing breeding companies to invest more money in RT, for example. The complexity of the system, and the number of factors involved, can make it difficult to determine the outcome of these types of choices with respect to their effects on genetic gain and genetic diversity.

The objective of this study was, in the context of genomic selection of dairy cattle breeds, to quantify the genetic gain and the loss of genetic diversity associated with different breeding schemes. Our aim was to provide useful recommendations to breeding companies to help with the overall management of breeding programs and, specifically, with the implementation of RT (MOET or OPU-IVF). We simulated 15 genomic selection breeding schemes that differed in dimensions (numbers of sires of bulls, marketed bulls, and female donors), the intensity of RT use, and the type of RT involved. These scenarios were then evaluated in terms of genetic gain and ROH-based inbreeding rate using stochastic simulations.

2. Materials and Methods

2.1. Simulated Population and Scenarios

Multiple scenarios were designed that differed in the extent of the use of embryo transfer and that relied on different numbers of sires of bulls and marketed bulls. These were tested by stochastic simulations of a breeding program that was similar in dimensions to those carried out in large French dairy cattle breeds.

Stochastic simulations were conducted using the MoBPS R package, version 1.0.2 (Göttigen, Germany) [26,27]. We simulated the same number of genetic markers as are present on the publicly available medium-density SNP chip (Illumina Infinium® BovineSNP50 BeadChip, San Diego, CA, USA), which is widely used in genomic selection programs around the world. We only considered the 29 bovine autosomes, on which we analyzed 41,377 SNP markers that were distributed regularly with respect to recombination rates (i.e., the distances in centimorgans between neighboring SNPs were approximately equal) [28]. On average, one SNP was simulated every 60.4 kb. We simulated 500 additive quantitative trait loci (QTLs), randomly distributed along the genome, whose effects were drawn from a gamma distribution (shape = 0.4 and scale = 5) [29] (see Supplementary Table S1).

Animals were selected according to the breeding goal, which can mimic either a single trait or a multi-trait synthetic index (e.g., [30]); the latter consists of an estimated breeding value (EBV) based on the true breeding values generated by the simulation of genotypes. The gross true breeding value for individual i ($gTBV_i$) was calculated as the sum of all additive effects of QTLs of individual i (see Table S1). We mimicked genomic evaluations with a constant coefficient of determination (CD) of 0.7.

Assuming a constant CD for each individual, the joint distribution of EBV_i and $gTBV_i$ for an individual i is [31]:

$$\begin{pmatrix} gTBV_i \\ EBV_i \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & CD\sigma_a^2 \\ CD\sigma_a^2 & CD\sigma_a^2 \end{pmatrix} \right],$$
 where σ_a^2 is the additive genetic variance (or the variance of $gTBV$ in the founding population).

Knowing $gTBV_i$, the conditional distribution of EBV_i is then [32]:

$$EBV_i | gTBV_i \sim N (CD \times gTBV_i, (1 - CD)CD\sigma_a^2) \quad (1)$$

For each individual i , EBV_i was drawn once from this normal distribution. Genotypes were used for the evaluation of ROH-based inbreeding.

Before we applied the different scenarios, we instituted a burn-in process: the objective was to obtain 10 simulated populations (one for each of the 10 replicates for each scenario) with ROH-based inbreeding around 10–11%, similar to what is observed in real French dairy cattle breeds [7]. For this, we simulated 10 founding populations of 100 cows and 10 bulls and performed selection on EBV until the size (20,000 elite cows and heifers spread over four consecutive birth years) and inbreeding of the

simulated populations reached our goal values. This generated linkage disequilibrium and genomic structure in the simulated populations.

Next, we simulated 10 replicates for each of the 15 different scenarios of genomic selection schemes based on embryo transfer, conducted for 20 years. Selection was performed using genomic evaluation for both young bulls and potential dams of bulls (see Figure 1 and Table 1). The 15 scenarios differed in their dimensions, intensity of RT use, and the type of RT applied.

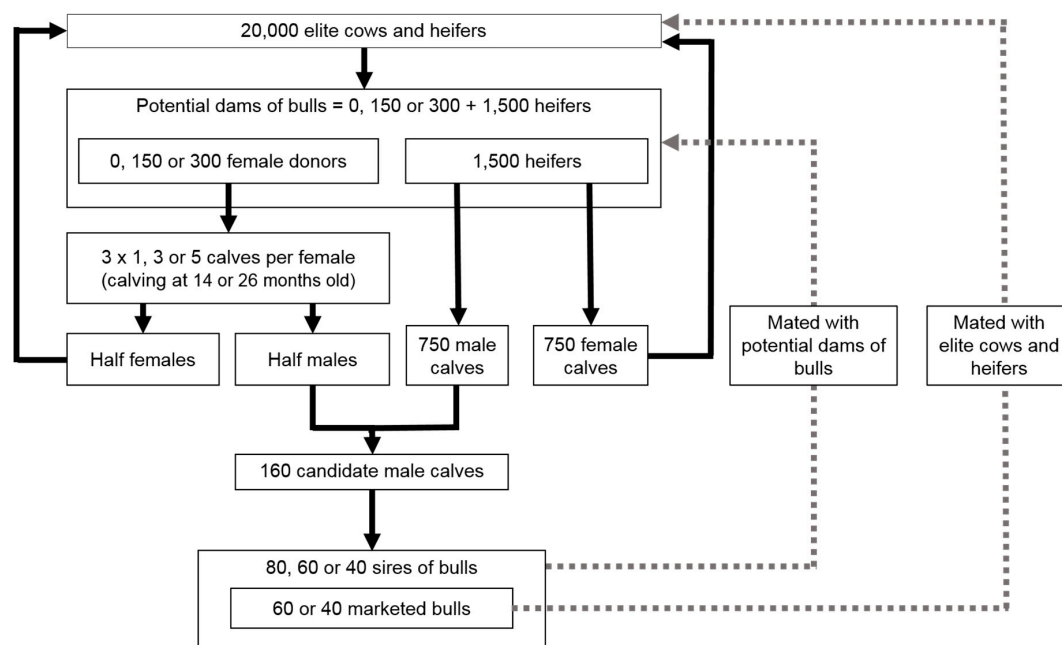


Figure 1. Selection steps in the simulated dairy cattle breeding scheme using embryo transfer.

Each simulated population consisted of 20,000 elite cows and heifers spread over four consecutive birth years, obtained at the end of the burn-in process described above. Only heifers could be selected as potential dams of bulls. In the reference scenario (REF), no RT was used. In all other scenarios, the 150 or 300 heifers with the highest estimated breeding values (EBV) were chosen to be embryo donors. We compared different intensities in the use of embryo transfer: low-intensity scenarios (three calves per female donor), medium-intensity scenarios (nine calves per female donor), and high-intensity scenarios (15 calves per female donor). In the context of high-intensity RT use, two types of RT were compared: MOET-like scenarios (five bulls used to obtain 15 calves) and OPU-IVF-like scenarios (15 bulls used to obtain 15 calves), with either medium or short generation intervals between female donors (calving at either 26 or 14 months, respectively). In most cases, female donors were mated randomly with 80 sires of bulls. We also simulated high-intensity scenarios with a medium or low number of sires of bulls (60 or 40 sires, respectively) to estimate the effect of a reduction in the number of sires. The resulting 1500 heifers with the highest EBV were randomly mated with the same number of sires of bulls by conventional insemination.

In all but three scenarios (H, I, N; see Table 1), all heifers and cows not selected as potential dams of bulls were randomly mated with 60 marketed bulls, chosen from the sires of bulls with the highest EBV. In high-intensity scenarios, we also estimated the effect of decreasing the number of marketed bulls by simulating scenarios with a low number of marketed bulls (40 bulls) (scenarios H, N, and I; see Table 1). Each cow reproduced three years in a row and obtained one calf per reproductive year (except when they were chosen as female donors). The cow population was renewed by keeping heifers with the highest EBV for replacement.

Table 1. Parameters used in different scenarios of the simulated dairy cattle breeding program based on the use of embryo transfer.

Scenario	Total Number of Calves Born per Female Donor	Age of Female Donors at Birth of Their Embryo-Transfer Calves	Number of Distinct Sires of Bulls Mated with Each Female Donor	Number of Female Donors (Embryos or Oocytes)	Number of Sires of Bulls	Number of Marketed Bulls	
REF	No use of RT	/	/	/	0	80	60
A	Low intensity	3 (1 flushing)	26 months	1 (1 per flushing)	150	80	60
B	Low intensity 300	3 (1 flushing)	26 months	1 (1 per flushing)	300	80	60
C	Medium intensity	9 (3 flushings)	26 months	3 (1 per flushing)	150	80	60
D	Medium intensity 300	9 (3 flushings)	26 months	3 (1 per flushing)	300	80	60
E	High intensity MOET-like	15 (5 flushings)	26 months	5 (1 per flushing)	150	80	60
F	High intensity 300 MOET-like	15 (5 flushings)	26 months	5 (1 per flushing)	300	80	60
G	High intensity MOET-like, medium sires	15 (5 flushings)	26 months	5 (1 per flushing)	150	60	60
H	High intensity MOET-like, medium sires, low bulls	15 (5 flushings)	26 months	5 (1 per flushing)	150	60	40
I	High intensity MOET-like, low sires, low bulls	15 (5 flushings)	26 months	5 (1 per flushing)	150	40	40
J	High intensity OPU-IVF-like	15	26 months	15 (1 per calf)	150	80	60
K	High intensity 300 OPU-IVF-like	15	26 months	15 (1 per calf)	300	80	60
L	High intensity MOET-like, short interval	15 (5 flushings)	14 months	5 (1 per flushing)	150	80	60
M	High intensity OPU-IVF-like, short interval	15	14 months	15 (1 per calf)	150	80	60
N	High intensity OPU-IVF-like, short interval, low sires, low bulls	15	14 months	15 (1 per calf)	150	40	40

RT: Reproductive technologies. MOET: Multiple ovulation and embryo transfer. OPU-IVF: Ovum pick-up and in vitro fertilization.

Sires of bulls were selected randomly from the 160 male calves with highest EBV born from the potential dams of bulls. The sires of bulls with the highest EBV became marketed bulls. These young sires and bulls were two years old when their calves were born and used only one year.

A bull (or cow) could not have more than five daughters (three daughters, respectively) chosen as female donors and five sons (three sons, respectively) chosen as male candidates for selection. All sires had at least one female offspring chosen as a female donor and one male offspring chosen as a candidate for selection. These constraints were based on current practices in large French dairy cattle breeding schemes.

2.2. Evaluation of Genetic Gain and Genetic Diversity Outcomes

Since different genetic gains could be expected between the sexes, we evaluated the outcomes of each scenario in terms of genetic gain and genetic diversity in two distinct subpopulations: (i) cows, defined as all females with at least one calf; and (ii) sires and bulls, defined as all males that became sires of bulls and/or marketed bulls.

Genetic gain was evaluated based on the evolution of the true breeding value TBV_i , expressed as:

$$TBV_i = \frac{gTBV_i - \text{mean}(gTBV_{\text{year } 1})}{\sigma(gTBV_{\text{year } 1})} \quad (2)$$

with $\text{mean}(gTBV_{\text{year } 1})$ the mean and $\sigma(gTBV_{\text{year } 1})$ the standard derivation of $gTBV$ for individuals born in the first year after implementation of the scenario. As we relied on TBV , the Bulmer effect was taken into account in this estimation of genetic gain.

Genetic diversity was evaluated based on inbreeding estimated from runs of homozygosity (ROH-based inbreeding) [33]. ROHs correspond to autozygous portions of the genome [33,34]. A ROH was defined as a homozygous segment with a minimum length of at least 15 SNPs or 1000 kb, with a minimum density of one SNP per 1000 kb. Two consecutive SNPs could not be included in the same ROH if they were over 1000 kb apart. We detected ROHs using the “homozyg” function of PLINK 1.9 [35,36] (command line: `plink -cow -bfile genotyping_data_filename -homozyg -homozyg-kb 1000 -homozyg-snp 15 -homozyg-window-snp 15 -homozyg-density 1000 -out output_filename`).

ROH-based inbreeding estimates, $F_{\text{ROH},i}$, were computed as follows [33]:

$$F_{\text{ROH},i} = \frac{\Sigma L_{\text{ROH},i}}{L_{\text{auto}}} \quad (3)$$

with $\Sigma L_{\text{ROH},i}$ the total length of ROHs for animal i , and L_{auto} the length of the autosomal genome covered by SNPs, after removing gaps of more than 1000 kb between two SNPs. ROH-based inbreeding estimates were expressed in percentages.

Both annual genetic gain (based on TBV) and annual ROH-based inbreeding rate were modeled using the following linear model:

$$Y_{ijk} = \beta_j^{[1]} + \beta_{klj}^{[2]} + \alpha_j \times \text{Year}_i + \epsilon_{ijk} \quad (4)$$

with Y_{ijk} the variable of interest (TBV or ROH-based inbreeding) for individual i in the k th replicate of scenario j and born in the year Year_i , $\beta_j^{[1]}$ the intercept of the model for scenario j , and $\beta_{klj}^{[2]}$ the intercept of the model for the k th replicate of scenario j . α is the regression coefficient of the model, corresponding to the slope of the model; it represents the annual trend of either genetic gain or genomic inbreeding. All computations were performed using the “lm” function of R [37]. Slopes were compared using the “emtrends” function from the emmeans package [38] and the “cld” function of the multcomp package [39].

To facilitate comparison between scenarios, these slopes were transformed with respect to the reference scenario, REF (see Table 1), in other words, we calculated the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario (no use of embryo transfer).

3. Results

3.1. General Observations

The annual ROH-based inbreeding rate (ΔF_{ROH}) for the reference scenario (no use of RT) was 0.095% in cows (see Table 2) and 0.047% in sires and bulls (see Table 3) (0.235% and 0.094% per generation, respectively; see Supplementary Table S2). The annual genetic gain in TBV (ΔG_{TBV}) for the reference scenario was 0.298 in cows and 0.280 in sires and bulls (see Tables 2 and 3). The ΔF_{ROH} for cows was thus higher by a factor of 2.02 than that of sires and bulls, while ΔG_{TBV} values for cows and for sires and bulls were more similar (ΔG_{TBV} of cows was slightly higher, by a factor of 1.06). The smallest values of ΔF_{ROH} and ΔG_{TBV} were always observed when no RT was used (scenario REF), for cows as well as for sires and bulls.

The highest values of ΔF_{ROH} , for both cows and sires/bulls, were observed in the high-intensity OPU-IVF-like scenario with short generation intervals and low numbers of sires and bulls, with $\Delta F_{ROH} = 0.219\%$ and 0.220% , respectively (scenario N; see Tables 2 and 3) (0.501% and 0.356% per generation, respectively; see Table S2). The highest values of ΔG_{TBV} for both cows and sires/bulls were observed in the high-intensity OPU-IVF-like, short interval scenario, with $\Delta G_{TBV} = 0.415$ and 0.405 , respectively (scenario M, see Tables 2 and 3).

The comparison of ΔF_{ROH} from the different scenarios with that of scenario REF ($\Delta F_{ROH}/REF$) ranged between 1.05 and 2.31 for cows, and between 1.05 and 4.67 for sires and bulls (see Table 3). The comparison of ΔG_{TBV} from the different scenarios with that of scenario REF ($\Delta G_{TBV}/REF$) ranged between 1.08 and 1.48 for cows, and between 1.08 and 1.45 for sires and bulls (see Table 3).

In Figures 2 and 3, for a given genetic gain (ΔG_{TBV}), the scenarios further to the left were those with the lowest inbreeding rates (ΔF_{ROH}); these are therefore the scenarios that offer the best compromise between genetic gain and genetic diversity. For instance, if we consider three clusters with different values of $\Delta G_{TBV}/REF$ (between 1.2 and 1.25; between 1.25 and 1.3; between 1.35 and 1.4), the scenarios that best balance genetic gain and inbreeding rate within each cluster were C (medium intensity), D (medium intensity 300), and F (high intensity 300 MOET-like), respectively (see Figures 2 and 3).

Table 2. Annual runs of homozygosity (ROH)-based inbreeding rate versus annual genetic gain among cows.

	Scenario	Annual ROH-Based Inbreeding Rate in % [95% Confidence Interval]	Annual ROH-Based Inbreeding Rate Compared with Scenario REF	Annual Genetic Gain in TBV [95% Confidence Interval]	Annual Genetic Gain in TBV Compared with Scenario REF
REF	No use of RT	0.095 ^a [0.094;0.096]	1.00 ^a	0.298 ^a [0.297;0.298]	1.00 ^a
A	Low intensity	0.100 ^b [0.099;0.101]	1.05 ^b	0.323 ^b [0.322;0.323]	1.08 ^b
B	Low intensity 300	0.103 ^c [0.102;0.104]	1.08 ^c	0.332 ^c [0.332;0.333]	1.11 ^c
C	Medium intensity	0.127 ^d [0.126;0.128]	1.34 ^d	0.364 ^d [0.363;0.364]	1.22 ^d
D	Medium intensity 300	0.124 ^e [0.123;0.125]	1.31 ^e	0.381 ^e [0.381;0.382]	1.28 ^e
E	High intensity MOET-like	0.140 ^f [0.139;0.141]	1.47 ^f	0.384 ^f [0.384;0.384]	1.29 ^f
F	High intensity 300 MOET-like	0.134 ^g [0.133;0.135]	1.41 ^g	0.403 ^g [0.403;0.403]	1.35 ^g
G	High intensity MOET-like, medium sires	0.146 ^h [0.145;0.147]	1.54 ^h	0.373 ^h [0.372;0.373]	1.25 ^h
H	High intensity MOET-like, medium sires, low bulls	0.155 ⁱ [0.154;0.156]	1.63 ⁱ	0.389 ⁱ [0.389;0.389]	1.31 ⁱ
I	High intensity MOET-like, low sires, low bulls	0.152 ^j [0.151;0.153]	1.60 ^j	0.365 ^j [0.365;0.366]	1.22 ^j
J	High intensity OPU-IVF-like	0.135 ^g [0.134;0.136]	1.42 ^g	0.386 ^k [0.386;0.386]	1.30 ^k
K	High intensity 300 OPU-IVF-like	0.146 ^h [0.145;0.147]	1.54 ^h	0.407 ^l [0.407;0.407]	1.37 ^l
L	High intensity MOET-like, short interval	0.196 ^k [0.195;0.197]	2.06 ^k	0.438 ^m [0.438;0.439]	1.47 ^m
M	High intensity OPU-IVF-like, short interval	0.196 ^k [0.195;0.197]	2.06 ^k	0.440 ⁿ [0.440;0.440]	1.48 ⁿ
N	High intensity OPU-IVF-like, short interval, low sires, low bulls	0.219 ^l [0.218;0.220]	2.31 ^l	0.415 ^o [0.415;0.416]	1.39 ^o

Comparisons with REF are calculated as the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario REF (no use of embryo transfer). RT: Reproductive technologies. MOET: Multiple ovulation and embryo transfer. OPU-IVF: Ovum pick-up and in vitro fertilization. ROH: Runs of Homozygosity. TBV: True Breeding Value. ^{a-o} Within each population (cows or sires and bulls), values within a column with different superscripts are significantly different (p -value < 0.05).

Table 3. Annual ROH-based inbreeding rate versus annual genetic gain among sires and bulls.

Scenario	Annual ROH-based Inbreeding Rate in % [95% Confidence Interval]	Annual ROH-based Inbreeding Rate Compared with Scenario REF	Annual Genetic Gain in TBV [95% Confidence Interval]	Annual Genetic Gain in TBV Compared with Scenario REF	
REF	No use of RT	0.04 ^a [0.038;0.056]	1.00 ^a	0.280 ^a [0.277;0.282]	1.00 ^a
A	Low intensity	0.050 ^a [0.041;0.058]	1.05 ^a	0.303 ^b [0.301;0.305]	1.08 ^b
B	Low intensity 300	0.057 ^{ab} [0.048;0.066]	1.21 ^{ab}	0.314 ^c [0.312;0.316]	1.12 ^c
C	Medium intensity	0.076 ^b [0.067;0.084]	1.60 ^b	0.340 ^d [0.338;0.342]	1.21 ^d
D	Medium intensity 300	0.078 ^b [0.069;0.086]	1.65 ^b	0.357 ^e [0.355;0.359]	1.28 ^e
E	High intensity MOET-like	0.109 ^c [0.100;0.117]	2.31 ^c	0.356 ^e [0.354;0.359]	1.27 ^e
F	High intensity 300 MOET-like	0.109 ^c [0.100;0.117]	2.31 ^c	0.374 ^f [0.372;0.376]	1.34 ^f
G	High intensity MOET-like, medium sires	0.108 ^c [0.098;0.118]	2.29 ^c	0.344 ^d [0.342;0.347]	1.23 ^d
H	High intensity MOET-like, medium sires, low bulls	0.117 ^c [0.107;0.127]	2.48 ^c	0.361 ^e [0.359;0.364]	1.29 ^e
I	High intensity MOET-like, low sires, low bulls	0.130 ^c [0.118;0.142]	2.76 ^c	0.340 ^d [0.337;0.343]	1.21 ^d
J	High intensity OPU-IVF-like	0.112 ^c [0.103;0.120]	2.38 ^c	0.359 ^e [0.357;0.361]	1.28 ^e
K	High intensity 300 OPU-IVF-like	0.116 ^c [0.107;0.125]	2.46 ^c	0.380 ^g [0.378;0.382]	1.36 ^g
L	High intensity MOET-like, short interval	0.200 ^d [0.191;0.209]	4.25 ^d	0.404 ^h [0.402;0.406]	1.44 ^h
M	High intensity OPU-IVF-like, short interval	0.204 ^d [0.195;0.212]	4.33 ^d	0.405 ^h [0.403;0.407]	1.45 ^h
N	High intensity OPU-IVF-like, short interval, low sires, low bulls	0.220 ^d [0.208;0.232]	4.67 ^d	0.380 ^{fg} [0.377;0.383]	1.36 ^{fg}

Comparisons with REF are calculated as the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference. scenario REF (no use of embryo transfer). RT: Reproductive technologies. MOET: Multiple ovulation and embryo transfer. OPU-IVF: Ovum pick-up and in vitro fertilization. ROH: Runs of Homozygosity. TBV: True Breeding Value. ^{a-h} Within each population (cows or sires and bulls), values within a column with different superscripts are significantly different (p -value < 0.05).

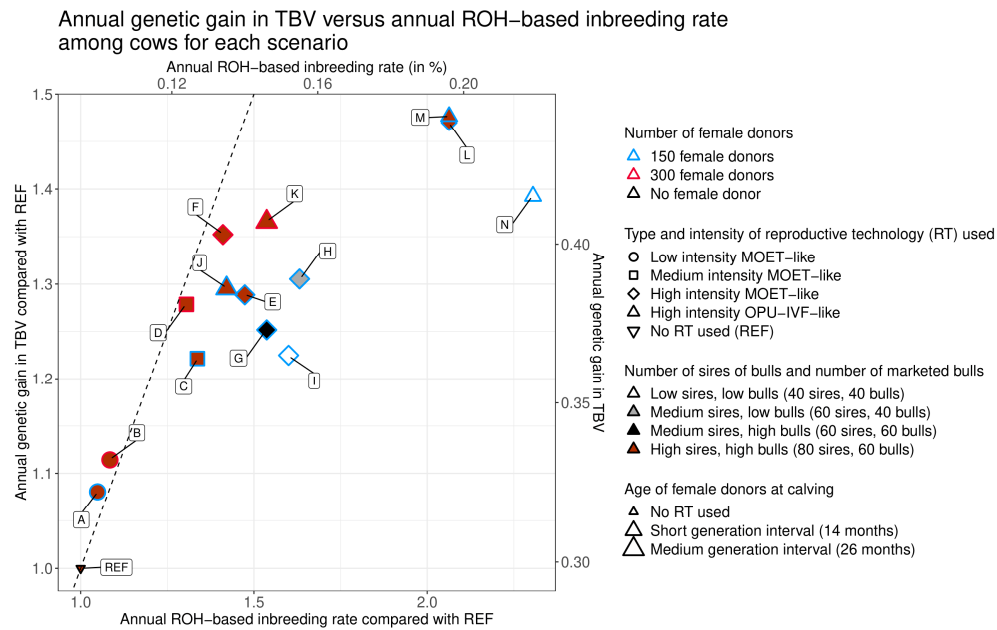


Figure 2. Annual genetic gain versus annual ROH-based inbreeding rate among cows. Letters indicate the names of the tested scenarios. Comparisons with REF were calculated as the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario (no use of embryo transfer). The dotted line is the line of identity between the annual ROH-based inbreeding rate compared with REF and the estimated annual genetic gain in true breeding value compared with REF.

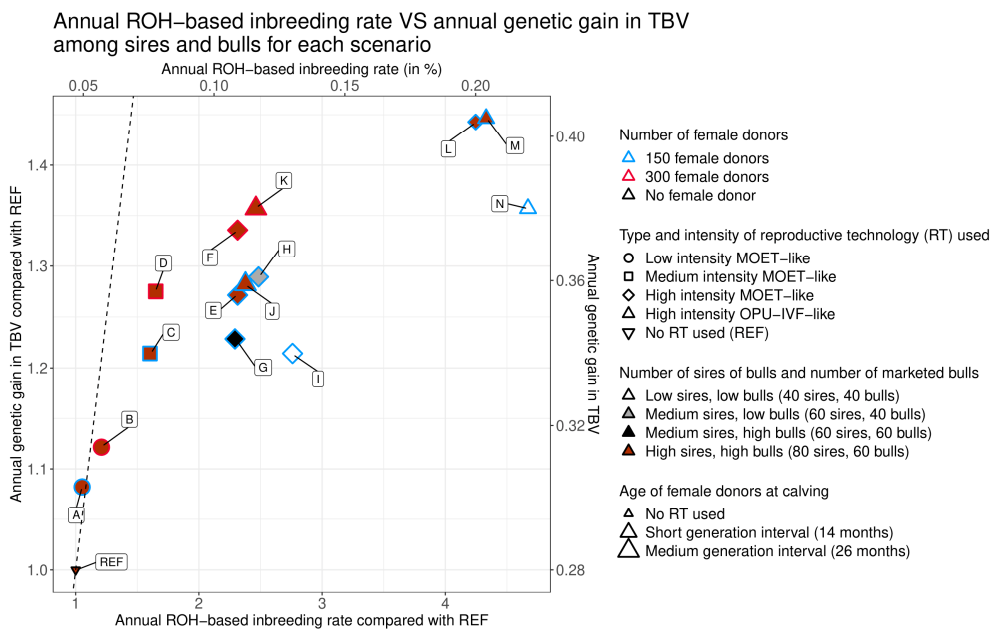


Figure 3. Annual genetic gain versus annual ROH-based inbreeding rate among sires and bulls. Letters indicate the names of the tested scenarios. Comparisons with REF were calculated as the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario (no use of embryo transfer). The dotted line is the line of identity between the annual ROH-based inbreeding rate compared with REF and the estimated annual genetic gain in true breeding value compared with REF.

3.2. Intensity of Reproductive Technologies (RT) Use and Number of Female Donors

For cows, all scenarios with female donors and embryo transfer had values of ΔF_{ROH} and ΔG_{TBV} that were significantly higher than those of the scenario with no RT (scenario REF; see Tables 2 and 3 and Figures 2 and 3). For sires and bulls, the difference was significant only for ΔG_{TBV} . Even at a low intensity of RT use and with 150 female donors, the inbreeding rate increased by a factor of 1.05 and the genetic gain for both cows and sires/bulls increased by a factor of 1.08 (scenario A; see Tables 2 and 3).

Regardless of whether 150 or 300 female donors were used, the more intensive the use of RT (from scenarios A to F), the higher ΔF_{ROH} and ΔG_{TBV} were, with all scenarios being significantly different from one another (see Tables 2 and 3 and Figures 2 and 3). However, with an increase in the intensity of RT, values of $\Delta F_{ROH}/REF$ increased more than $\Delta G_{TBV}/REF$. In the high-intensity MOET-like scenario with 150 female donors (scenario E), these values reached, respectively, 1.47 and 1.29 for cows and 2.31 and 1.27 for sires and bulls, while in the same RT scheme with 300 donors (scenario F), $\Delta F_{ROH}/REF$ and $\Delta G_{TBV}/REF$ were 1.41 and 1.35 for cows and 2.31 and 1.34 for sires and bulls, respectively (see Tables 2 and 3).

All other parameters being equal, the scenarios with 300 female donors (scenarios B, D, F, and K; see Tables 2 and 3) always achieved significantly higher values of ΔG_{TBV} than scenarios with 150 female donors (scenarios A, C, E, and J; see Tables 2 and 3). However, this was not true for ΔF_{ROH} , which, depending on the comparison, was higher, lower, or not significantly different between scenarios with 300 female donors (scenarios B, D, F, and K; see Tables 2 and 3) and scenarios with 150 female donors (scenarios A, C, E, and J; see Tables 2 and 3).

3.3. Generation Interval between Female Donors and Type of RT Used

The differences between MOET and OPU-IVF are that (i) OPU-IVF can be performed earlier in the life of the heifer than MOET, leading to shorter generation intervals between female donors (14 months instead of 26), and (ii) OPU-IVF allows each oocyte to be fertilized by a different sire, whereas MOET only permits fertilization by one sire per flushing. To compare the use of MOET and OPU-IVF, we simulated high-intensity MOET-like and OPU-IVF-like scenarios.

3.3.1. Generation Interval between Female Donors

We first estimated the impact of generation intervals between female donors by comparing high-intensity scenarios that had either medium (calving at 26 months) or short generation intervals (calving at 14 months). For high-intensity MOET-like scenarios, ΔF_{ROH} and ΔG_{TBV} were both significantly higher in the scenario with the short interval (scenario L) than in the scenario with the medium interval (scenario E), both for cows (see Table 2 and Figure 2) and for sires and bulls (see Table 3 and Figure 3). For cows, $\Delta F_{ROH}/REF$ reached 2.06 and 1.47 for the short- and medium-interval scenarios, respectively, while for sires and bulls, this value reached 4.25 and 2.31, respectively. Values of $\Delta G_{TBV}/REF$ were slightly lower: for cows, 1.47 and 1.29 for the short- and medium-interval scenarios, respectively (see Table 2), and for sires and bulls, 1.44 and 1.27, respectively (see Table 3). We observed similar results when we performed the same comparison using high-intensity OPU-IVF-like scenarios (medium generation intervals: scenario J, short generation intervals: scenario M; see Tables 2 and 3). As with MOET, simulations of OPU-IVF had higher inbreeding rates and increased genetic gain per generation in scenarios with short generation intervals compared to those with medium generation intervals (scenarios E, J, L, and M). The proportion of female donors that were born from female donors was higher for scenarios with short generation intervals than for those with medium generation intervals (see Supplementary Figure S1).

3.3.2. Type of RT Used

We then evaluated high-intensity scenarios to determine the impact of the type of RT used: MOET versus OPU-IVF. Specifically, we compared the effect on ΔF_{ROH} and ΔG_{TBV} of differences in the number

of distinct sires used to generate 15 calves (five for MOET versus 15 for OPU-IVF). For high-intensity scenarios with a medium generation interval, ΔF_{ROH} and ΔG_{TBV} were not significantly different between the MOET-like (scenario E) and the OPU-IVF-like (scenario J) scenarios for sires and bulls (see Table 3 and Figure 3). For cows, though, ΔF_{ROH} was significantly higher and ΔG_{TBV} significantly lower for the MOET-like scenario ($\Delta F_{ROH}/REF = 1.47$, $\Delta G_{TBV}/REF = 1.29$) than for the OPU-IVF-like scenario ($\Delta F_{ROH}/REF = 1.42$, $\Delta G_{TBV}/REF = 1.30$; see Table 2 and Figure 2). However, the difference between the two scenarios was small for both values, with an increase of $\Delta F_{ROH}/REF$ by a factor of 1.04 ($=1.47/1.42$) for the OPU-IVF-like scenario with respect to the MOET-like scenario, and an increase of $\Delta G_{TBV}/REF$ by a factor of 1.01 ($=1.30/1.29$) for the MOET-like scenario compared to the OPU-IVF-like scenario. For high-intensity scenarios with a short generation interval, we observed only one small but significant difference in ΔG_{TBV} for cows: $\Delta G_{TBV}/REF$ was 1.47 in the MOET-like scenario (scenario L) and 1.48 in the OPU-IVF-like scenario (scenario M; see Table 2 and Figure 2).

3.4. Number of Sires of Bulls and of Marketed Bulls

We compared the impact of the number of sires of bulls and of marketed bulls on ΔF_{ROH} and ΔG_{TBV} in the context of high-intensity MOET-like scenarios with 150 female donors and medium generation intervals.

Two types of comparisons were performed. First, we compared the impact of a reduction in the number of sires of bulls while keeping constant the number of marketed bulls. Second, we compared the impact of a reduction in the number of marketed bulls while keeping constant the number of sires of bulls.

3.4.1. Number of Sires of Bulls

With the number of marketed bulls held constant, a reduction in the number of sires of bulls resulted in significantly lower values of ΔG_{TBV} , both for cows and for sires and bulls. A reduction from 80 (scenario E) to 60 (scenario G) sires of bulls, with 60 marketed bulls, resulted in a decrease in $\Delta G_{TBV}/REF$ from 1.29 to 1.25 for cows and from 1.27 to 1.23 for sires and bulls. A reduction from 60 (scenario H) to 40 (scenario I) sires of bulls, with 40 marketed bulls, resulted in a decrease of $\Delta G_{TBV}/REF$ from 1.31 to 1.22 for cows (see Table 2 and Figure 2) and from 1.29 to 1.21 for sires and bulls (see Table 3 and Figure 3). No clear trend was observed for ΔF_{ROH} : differences between scenarios/populations were significant or not, and changed in direction, depending on the number of marketed bulls and on the population under consideration.

3.4.2. Number of Marketed Bulls

With the number of sires of bulls held constant, a reduction in the number of marketed bulls resulted in significantly higher values of ΔG_{TBV} , both for cows and for sires and bulls, and significantly higher values of ΔF_{ROH} for cows only. A reduction from 60 (scenario G) to 40 (scenario H) marketed bulls, with 60 sires of bulls, resulted in an increase in $\Delta G_{TBV}/REF$ from 1.25 to 1.31 for cows (see Table 2 and Figure 2) and from 1.23 to 1.29 for sires and bulls (see Table 3 and Figure 3). It also resulted in an increase in ΔF_{ROH} from 1.54 to 1.63 for cows (see Table 2 and Figure 2), and from 2.29 to 2.48 for sires and bulls, but this latter change was not significant (see Table 3 and Figure 3).

3.4.3. Effect of RT Type on Reductions in the Number of Sires of Bulls and Marketed Bulls

To further examine the high-intensity scenarios, we next evaluated the differences between the two types of RT (MOET with medium generation interval (E and I) and OPU-IVF with short generation interval (M and N)) in two extreme cases: scenarios with a high number of sires of bulls and marketed bulls (E and M) versus those with a low number of sires and bulls (I and N). Two types of comparisons were performed. First, we compared the impact of the type of RT with the numbers of sires of bulls and marketed bulls held constant; then, we compared the impact of a reduction in the numbers of sires of bulls and marketed bulls within each type of RT.

With the numbers of sires of bulls and marketed bulls held constant, the OPU-IVF-like scenario had significantly higher values of both ΔF_{ROH} and ΔG_{TBV} than the MOET-like scenario for both cows and sires/bulls (see Table 4).

Table 4. Annual ROH-based inbreeding rate versus annual genetic gain among cows and sires/bulls for four different scenarios.

Number of Sires and Bulls	Annual ROH-Based Inbreeding Rate Compared with Scenario REF		Annual Genetic Gain in TBV Compared with Scenario REF	
	MOET, Medium Interval	OPU-IVF, Short Interval	MOET, Medium Interval	OPU-IVF, Short Interval
80 sires of bulls and 60 marketed bulls	Scenario E 1.47 for cows 2.31 for sires/bulls	Scenario M 2.06 for cows 4.33 for sires/bulls	Scenario E 1.29 for cows 1.27 for sires/bulls	Scenario M 1.48 for cows 1.45 for sires/bulls
	Scenario I 1.60 for cows 2.76 for sires/bulls	Scenario N 2.31 for cows 4.67 for sires/bulls	Scenario I 1.22 for cows 1.21 for sires/bulls	Scenario N 1.39 for cows 1.36 for sires/bulls

Comparisons with REF were calculated as the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario REF (no use of embryo transfer). RT: Reproductive technologies. MOET: Multiple ovulation and embryo transfer. OPU-IVF: Ovum pick-up and in vitro fertilization. ROH: Runs of Homozygosity. TBV: True Breeding Value.

The impact of RT type was stronger on $\Delta F_{ROH}/REF$ than on $\Delta G_{TBV}/REF$; the use of OPU-IVF instead of MOET increased $\Delta F_{ROH}/REF$ in cows by a factor of 1.40 ($= 2.06/1.47$) in the scenario with 80 sires of bulls and 60 marketed bulls and a factor of 1.44 ($= 2.31/1.60$) in the scenario with 40 sires of bulls and 40 marketed bulls. In sires and bulls in those same two scenarios, the use of OPU-IVF instead of MOET increased $\Delta F_{ROH}/REF$ by a factor of 1.87 ($= 4.33/2.31$) and 1.69 ($= 4.67/2.76$), respectively. In comparison, $\Delta G_{TBV}/REF$ increased for cows by a factor of 1.15 ($= 1.48/1.29$) in the scenario with 80 sires of bulls and 60 marketed bulls and a factor of 1.14 ($= 1.39/1.22$) in the scenario with 40 sires of bulls and 40 marketed bulls. For sires and bulls in the same two scenarios, the use of OPU-IVF increased $\Delta G_{TBV}/REF$ by a factor of 1.14 ($= 1.45/1.27$) and 1.12 ($= 1.36/1.21$), respectively.

For both MOET and OPU-IVF, drastic reductions in the number of sires of bulls and of marketed bulls resulted in a significant increase in ΔF_{ROH} for cows only (see Table 4), and significant decreases in ΔG_{TBV} for both cows and sires/bulls (see Table 4).

4. Discussion

The objective of this study was to quantify the genetic gain and loss of genetic diversity resulting from changes in different parameters of breeding schemes (use of RT and dimensions of breeding schemes), in the context of genomic selection in dairy cattle breeds. We simulated 15 breeding schemes for genomic selection that differed in their dimensions, and the type and intensity of RT used, and we evaluated these scenarios in terms of annual genetic gain and annual ROH-based inbreeding rate. Although the simulations were designed based on breeding schemes used in French dairy cattle breeds, they reflect the breeding schemes currently used in large dairy cattle breeds using RT.

When no RT was used, the annual ROH-based inbreeding rate (ΔF_{ROH}) was 0.095% in cows and 0.047% in sires and bulls (0.235% and 0.094% per generation, respectively). From the tested scenarios, the highest values of ΔF_{ROH} for both cows and sires/bulls were observed in the high-intensity OPU-IVF-like scenario with a short generation interval and low numbers of both sires and bulls, with $\Delta F_{ROH} = 0.219\%$ and 0.220% , respectively (0.501% and 0.356% per generation, respectively). These values were close to the real ones observed in France for the Montbéliarde and Normande breeds, for which the annual ROH-based inbreeding rates in sires and bulls born between 2012 and 2015 were 0.076% and 0.14% , respectively [7]. This similarity supports the relevance of our simulations in mimicking the loss of genetic diversity in dairy cattle breeding schemes. In all studied scenarios, the per-generation ROH-based inbreeding rate stayed under the maximum acceptable threshold of 1%

(set by FAO (Food and Agriculture Organization) guidelines [40]), probably thanks to the constraints that we implemented in the simulated breeding schemes (i.e., maximum number of calves per bull or per cow that could become candidates for selection). These constraints were based on current practices in large French dairy cattle breeding schemes, specifically those for Montbéliarde and Normande.

The smallest values of ΔF_{ROH} and ΔG_{TBV} were always observed when no RT was used. As the intensity of RT increased, both ΔF_{ROH} and ΔG_{TBV} also increased, regardless of the number of female donors used (150 or 300). In addition, increasing the number of female donors from 150 to 300 always resulted in significantly higher ΔG_{TBV} regardless of the intensity of RT use. These results were all consistent with the findings of previous studies [9,10,19,20]. Instead, changes in the number of female donors did not seem to have any consistent effect on inbreeding rate, ΔF_{ROH} ; this was also in agreement with a previous report that increasing the intensity of RT use (i.e., the number of calves per female donor) had a greater impact on ΔF_{ROH} than increasing the number of female donors [9].

These results can probably be explained by the fact that the intensity of selection was less affected by the number of female donors than by the number of calves per donor. Regardless of the total number of donors (150 or 300), the offspring of the best females were selected first. This means that it was largely the same 150 donors that experienced the most intense selection in each scenario, and thus, total selection intensity might not have changed very much among different scenarios. Instead, an increase in the number of calves born per female intensified the use of the best female donors based on EBV. Therefore, the higher the intensity of RT use, the higher the selection intensity for the top female donors. Here, constraints on the numbers of sons and daughters that could be selected per donor made it possible to limit the selection intensity. Without these constraints, ΔF_{ROH} and ΔG_{TBV} would probably have been even higher in the scenarios with high intensity RT, with the risk of severe reductions in genetic diversity and, ultimately, too little genetic variability to be able to increase genetic gain in the medium- or long-term [2].

We then explored the differences between MOET and OPU-IVF. OPU-IVF can be performed sooner in the life of the heifer than MOET as it does not require the donor to have matured enough to develop ovulation cycles [25]. The breeding schemes that we simulated included up to five flushings (corresponding to five sessions) of MOET or 15 sessions of OPU-IVF per year. In France, as of 2018, an average of 5.3 viable embryos was retrieved per flushing for MOET, with one bull used per flushing, while an average of 1.95 viable embryos were retrieved per session for OPU-IVF, with one bull per session [41]. In our simulations, success rates of gestation and birth were both set to 40% following both techniques. This meant that, over the course of a year, it would be possible to obtain 15 calves from five different bulls (three calves times five flushings) with MOET, or 15 calves from 15 different sires (one calf times 15 sessions) using OPU-IVF. Over the same amount of time, then, it is possible to obtain calves from more bulls using OPU-IVF than using MOET. We investigated the use of these techniques by simulating high-intensity MOET-like and OPU-IVF-like scenarios using different parameters.

We first estimated the impact of the interval between generations of female donors by comparing high-intensity scenarios that had either medium- or short-generation intervals (calving at 26 months and calving at 14 months, respectively). Unsurprisingly, we observed that a reduction in the generation interval led to significantly higher values of ΔF_{ROH} and ΔG_{TBV} for both types of RT, as was reported in a previous study [10]. For a given generation interval, we did not observe clear differences between the types of RT in terms of either ΔF_{ROH} or ΔG_{TBV} . Therefore, it seems that the differences between MOET and OPU-IVF were probably mostly due to the reduction in generation interval that the latter method permits. Moreover, reducing the generation interval had a stronger effect on ΔF_{ROH} than on ΔG_{TBV} , which indicates that drastic reductions in generation intervals between female donors could have a strong detrimental effect on genetic diversity for relatively little genetic gain.

When all other factors were held constant, the differences in the numbers of distinct sires of bulls mated with each female donor (higher for OPU-IVF than for MOET) had a low impact on both genetic gain and genetic diversity. In all comparisons, we held constant the total numbers of sires of bulls, female donors, and calves born per female donor, which were all dependent on the size of the breeding

scheme. Therefore, the only difference between MOET-like and OPU-IVF-like scenarios, for a given generation interval, was how sires of bulls and female donors were mated. For both methods, all sires of bulls and female donors were used, and the constraints we applied ensured that all sires of bulls and female donors had male calves chosen as candidates for selection and female calves chosen as female donors. Moreover, matings between the female donors and sires of bulls were performed randomly, which probably homogenized the genetic differences between their calves to a certain extent. It is therefore possible that our simulations did not generate a sufficient degree of difference between the MOET-like and OPU-IVF-like scenarios in the genetic quality of calves chosen as candidates for selection or as female donors.

Instead, a reduction in the generation interval between female donors had a much higher impact on the annual change in genetic diversity and genetic gain. Indeed, given a constant per-generation genetic gain or inbreeding rate, a shortening of the generation interval will increase the subsequent annual rates. However, our simulations showed that a reduction in the generation interval (from medium to short) also increased inbreeding rates and genetic gain per generation. Therefore, the impact of the reduced generation interval between female donors was not only due to the subsequent increase in annual inbreeding rates and genetic gain. In short-interval scenarios, the proportion of female donors that were born from female donors was higher than in medium-interval scenarios. Shortening the generation interval thus has two consequences that combine to increase the genetic gain and the inbreeding rate: (i) a multiplier effect on the annual inbreeding rate and genetic gain and (ii) an increase in the proportion of female donors born from female donors and the overuse of a subset of female donors.

Taken together, our results suggest that the use of OPU-IVF instead of MOET to reduce the generation interval between female donors might lead to detrimental effects on genetic diversity for little genetic gain. OPU-IVF is more invasive than MOET, as it first requires epidural anesthesia and then oocyte aspiration by needle [42,43]. This raises questions about the impact of ovum pick-up on the welfare of heifers, especially after repeated samplings [44]. Consequently, considering its adverse effects on both genetic diversity and animal welfare, OPU-IVF is probably not the optimal RT for dairy cattle breeding schemes.

For both types of RT, a reduction in the number of sires of bulls and marketed bulls resulted in a significant decrease in ΔG_{TBV} and a significant increase in ΔF_{ROH} . Regardless of the numbers of sires of bulls and marketed bulls used in the breeding scheme, the impact of using OPU-IVF rather than MOET was more deleterious to ΔF_{ROH} than it was beneficial to ΔG_{TBV} , even though both values increased significantly. It had previously been predicted that increasing the numbers of sires and bulls in a breeding scheme could alleviate the impact of RT on genetic diversity without seriously compromising genetic gain [10]. However, this approach would increase the costs of breeding schemes, which breeding companies might find unacceptable. In this context, maintaining the number of sires and bulls at least at current levels would keep costs constant, for a higher genetic gain than would be achieved with a lower (as shown in this study) or higher [10] number of sires and bulls. This would not decrease the loss of genetic diversity, but would at least avoid its acceleration.

The costs of the different scenarios were not taken into account in our simulations as the objective of this study did not include an evaluation of the financial aspects of RT. However, for breeding companies, the more intensive the use of RT and the higher the numbers of female donors, the more expensive the breeding of bulls will be. These increased costs could require a corresponding increase in the price of semen doses. This increase in price would be acceptable only if the use of RT yielded bulls of significantly higher genetic quality. The balance between the costs and benefits of RT will vary among species and breeds as well as based on the breeding goals and the monetary value of the trait under selection. In addition, the overall organization of selection and reproduction processes (e.g., use of artificial insemination, veterinary support, etc.) could also play a role. For example, a study by Granleese et al. (2019) [19] of Australian sheep showed that for two different breeding objectives, the genetic gain enabled by the use of RT had different costs and benefits for different kinds

of breeders. Therefore, the profitability of RT use for breeding companies and breeders depends on many parameters. Decisions on the use of RT must be made on a case-by-case basis, after careful consideration of the costs and benefits generated according to the breeding objective and breed.

In our simulations, intensification of RT use combined with a reduction in the number of sires and bulls had a detrimental impact on genetic diversity by increasing inbreeding rates. For a given genetic gain, the scenarios that best maintained genetic diversity were those with a medium intensity of RT use or those with a higher number of female donors to compensate for the increase in RT intensity.

Our results were obtained from simulations of a breeding scheme that used random mating and that imposed constraints on the number of calves per bull and female donor that could become bulls or donors themselves. With this design, we might have overestimated genetic gains and inbreeding rates in comparison with breeding schemes that use avoidance mating [45] or optimal contributions (OC) [19,46,47]. The recommendations provided by OC algorithms can be quite useful, as this technique aims to produce a list of individuals for breeding in which global relatedness is minimized. However, any substitutions in the breeding population can completely change the expected overall relatedness, which can limit the practical application of this approach. Although changes in the mating strategy could certainly affect the values obtained by our simulations, such modifications would affect all of the studied scenarios equally, and we would thus not expect them to change the outcomes of any of the comparisons performed here.

Inbreeding depression depends on both the inbreeding rate and the inbreeding load (i.e., the set of deleterious variants carried by the population). In future studies, it would be interesting to identify the genomic regions carrying these variants and to incorporate this information into genomic evaluations. With this knowledge, genomic selection programs would be able to focus more on maintaining genetic diversity in the regions where it is truly important. Such a strategy could limit the impact of the loss of genetic diversity on the health and performance of dairy cattle breeds even with an intensification in the use of RT.

5. Conclusions

Our simulations predicted that intensive use of reproductive technologies could lead to improved genetic gain, but this would be accompanied by a deterioration in genetic diversity. The larger genetic gain that we found in scenarios based on OPU-IVF compared to MOET appeared to be, in large part, the result of the significant reduction in the generation interval enabled by OPU-IVF. However, this shortened generation interval led to significant increases in the inbreeding rate, suggesting that using RT to drastically reduce the generation intervals between female donors could have severe detrimental effects on genetic diversity in dairy cattle breeds. In addition, reducing the number of sires of bulls and marketed bulls in the breeding scheme had a detrimental impact on both genetic gain and genetic diversity.

These results led us to the conclusion that two of the major trends in dairy cattle breeding—the intensified use of RT and the cost-conscious reduction in the number of sires of bulls and marketed bulls used by breeding companies—pose serious risks to the genetic diversity of these breeds. In the context of genomic selection and according to the level of genetic gain breeding companies want to reach, avoiding OPU-IVF in favor of MOET, limiting the intensity of use of MOET while maintaining the number of sires and bulls are good practices to maintain genetic diversity. It is also possible to compensate (to a certain point) for an increase in the intensity of use of MOET by increasing the number of female donors and putting constraints on the number of bulls having the same mother, while still maintaining the number of sires and bulls.

It is necessary to find solutions to the loss of genetic diversity, or, at a minimum, approaches that can mitigate the consequences of this loss (i.e., inbreeding depression). One approach that might be effective would involve managing genetic diversity at the genome level by locating the genetic load of deleterious mutations and focusing on the diversity of these regions in particular.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-2615/10/10/1903/s1>, Table S1: Positions and effects of additive quantitative trait loci (QTLs) along the simulated genome used as the basis for simulations of an embryo transfer breeding scheme and variations thereof in French dairy cattle breeds, Table S2: ROH-based inbreeding rates per generation and genetic gains in true breeding value per generation for each scenario of the simulated embryo transfer breeding program in French dairy cattle, Figure S1: Proportion of female donors and of sires and bulls born from embryo transfer over all birth years and replicates for each scenario of the simulated embryo transfer breeding scheme in French dairy cattle breeds.

Author Contributions: Conceptualization: P.C., G.R., and A.-C.D.; Design of simulations: A.-C.D., G.F., and L.B.; Software: A.-C.D.; Statistical analysis: A.-C.D.; Writing—original draft preparation: A.-C.D.; Writing—review and editing: A.-C.D., G.R., S.F., G.F., L.B., C.H., D.L., and P.C.; Supervision: G.R., S.F., G.F., L.B., C.H., D.L., and P.C. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially funded by the INRA SELGEN funding metaprogram (Project GDivSelGen). This study was funded by ALLICÉ and APIS-GENE.

Acknowledgments: The authors would like to acknowledge Torsten Pook for his invaluable help with the use of the MoBPS R package.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stachowicz, K.; Sargolzaei, M.; Miglior, F.; Schenkel, F.S. Rates of inbreeding and genetic diversity in Canadian Holstein and Jersey cattle. *J. Dairy Sci.* **2011**, *94*, 5160–5175. [[CrossRef](#)]
2. Dickerson, G.E.; Hazel, L.N. Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J. Agric. Res.* **1944**, *69*, 459–476.
3. Notter, D.R. The importance of genetic diversity in livestock populations of the future. *J. Anim. Sci.* **1999**, *77*, 61. [[CrossRef](#)]
4. Leroy, G. Inbreeding depression in livestock species: Review and meta-analysis. *Anim. Genet.* **2014**, *45*, 618–628. [[CrossRef](#)]
5. Pryce, J.E.; Haile-Mariam, M.; Goddard, M.E.; Hayes, B.J. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet. Sel. Evol.* **2014**, *46*, 71. [[CrossRef](#)]
6. Eynard, S.E.; Windig, J.J.; Hiemstra, S.J.; Calus, M.P.L. Whole-genome sequence data uncover loss of genetic diversity due to selection. *Genet. Sel. Evol.* **2016**, *48*. [[CrossRef](#)]
7. Doublet, A.-C.; Croiseau, P.; Fritz, S.; Michenet, A.; Hozé, C.; Danchin-Burge, C.; Laloë, D.; Restoux, G. The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genet. Sel. Evol.* **2019**, *51*, 52. [[CrossRef](#)]
8. Weigel, K.A. Controlling Inbreeding in Modern Breeding Programs. *J. Dairy Sci.* **2001**, *84*, E177–E184. [[CrossRef](#)]
9. Bouquet, A.; Sørensen, A.C.; Juga, J. Genomic selection strategies to optimize the use of multiple ovulation and embryo transfer schemes in dairy cattle breeding programs. *Livest. Sci.* **2015**, *174*, 18–25. [[CrossRef](#)]
10. Thomasen, J.R.; Willam, A.; Egger-Danner, C.; Sørensen, A.C. Reproductive technologies combine well with genomic selection in dairy breeding programs. *J. Dairy Sci.* **2016**, *99*, 1331–1340. [[CrossRef](#)]
11. Maignel, L.; Boichard, D.; Verrier, E. Genetic variability of French dairy breeds estimated from pedigree information. *Interbull Bull.* **1996**, *49*.
12. Le Mézec, P.; Danchin-Burge, C.; Moureaux, S. Davan tage de diversité avec la génomique? plutôt non... Available online: <http://idele.fr/contact/publication/idelesolr/recommends/les-programmes-de-selection-et-de-diffusion-de-taureaux-dia-a-lere-de-la-genomique-et-leurs-effets.html> (accessed on 21 January 2019).
13. Danchin-Burge, C.; Danvy, S.; Laloë, D.; Verrier, E. Création d'un observatoire de la VARiabilité génétique des RUMinants et des Equidés (VARUME). *Innov. Agron.* **2017**, *55*, 235–245.
14. Eynard, S.E.; Windig, J.J.; Leroy, G.; van Binsbergen, R.; Calus, M. The effect of rare alleles on estimated genomic relationships from whole genome sequence data. *BMC Genet.* **2015**, *16*, 24. [[CrossRef](#)] [[PubMed](#)]
15. Thomasen, J.R.; Liu, H.; Sørensen, A.C. Genotyping more cows increases genetic gain and reduces rate of true inbreeding in a dairy cattle breeding scheme using female reproductive technologies. *J. Dairy Sci.* **2020**, *103*, 597–606. [[CrossRef](#)] [[PubMed](#)]
16. Kardos, M.; Luikart, G.; Allendorf, F.W. Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. *Heredity* **2015**, *115*, 63–72. [[CrossRef](#)]

17. Doekes, H.P.; Veerkamp, R.F.; Bijma, P.; Hiemstra, S.J.; Windig, J.J. Trends in genome-wide and region-specific genetic diversity in the Dutch-Flemish Holstein–Friesian breeding program from 1986 to 2015. *Genet. Sel. Evol.* **2018**, *50*, 15. [[CrossRef](#)]
18. Forutan, M.; Ansari Mahyari, S.; Baes, C.; Melzer, N.; Schenkel, F.S.; Sargolzaei, M. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genom.* **2018**, *19*, 98. [[CrossRef](#)]
19. Granleese, T.; Clark, S.A.; Kinghorn, B.P.; Werf, J.H.J. vander Optimizing female allocation to reproductive technologies considering merit, inbreeding and cost in nucleus breeding programmes with genomic selection. *J. Anim. Breed. Genet.* **2019**, *136*, 79–90. [[CrossRef](#)]
20. Pedersen, L.D.; Kargo, M.; Berg, P.; Voergaard, J.; Buch, L.H.; Sørensen, A.C. Genomic selection strategies in dairy cattle breeding programmes: Sexed semen cannot replace multiple ovulation and embryo transfer as superior reproductive technology. *J. Anim. Breed. Genet.* **2012**, *129*, 152–163. [[CrossRef](#)]
21. Pryce, J.E.; Goddard, M.E.; Raadsma, H.W.; Hayes, B.J. Deterministic models of breeding scheme designs that incorporate genomic selection. *J. Dairy Sci.* **2010**, *93*, 5455–5466. [[CrossRef](#)]
22. Sørensen, A.C.; Sørensen, M.K. Inbreeding Rates in Breeding Programs with Different Strategies for Using Genomic Selection. In Proceedings of the 2009 Interbull Meeting, Barcelona, Spain, 21–24 August 2009; p. 4.
23. Thomasen, J.R.; Egger-Danner, C.; Willam, A.; Guldbrandtsen, B.; Lund, M.S.; Sørensen, A.C. Genomic selection strategies in a small dairy cattle population evaluated for genetic gain and profit. *J. Dairy Sci.* **2014**, *97*, 458–470. [[CrossRef](#)] [[PubMed](#)]
24. Land, R.B.; Hill, W.G. The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Anim. Sci.* **1975**, *21*, 1–12. [[CrossRef](#)]
25. Galli, C.; Crotti, G.; Notari, C.; Turini, P.; Duchi, R.; Lazzari, G. Embryo production by ovum pick up from live donors. *Theriogenology* **2001**, *55*, 1341–1357. [[CrossRef](#)]
26. Pook, T. *MoBPS: Simulation of Breeding Programs*, version 1.0.2; R Package: Göttingen, Germany, 2018.
27. Pook, T.; Schlather, M.; Simianer, H. MoBPS-Modular Breeding Program Simulator. *G3 Genes Genomes Genet.* **2020**. [[CrossRef](#)]
28. Arias, J.A.; Keehan, M.; Fisher, P.; Coppieters, W.; Spelman, R. A high density linkage map of the bovine genome. *BMC Genet.* **2009**, *10*, 18. [[CrossRef](#)]
29. Hayes, B.; Goddard, M. The distribution of the effects of genes affecting quantitative traits in livestock. *Genet. Sel. Evol.* **2001**, *33*, 209–229. [[CrossRef](#)]
30. Eynard, S.E.; Croiseau, P.; Laloë, D.; Fritz, S.; Calus, M.P.L.; Restoux, G. Which Individuals to Choose to Update the Reference Population? Minimizing the Loss of Genetic Diversity in Animal Genomic Selection Programs. *G3 Bethesda Md* **2018**, *8*, 113–121. [[CrossRef](#)]
31. Laloë, D.; Phocas, F. A proposal of criteria of robustness analysis in genetic evaluation. *Livest. Prod. Sci.* **2003**, *80*, 241–256. [[CrossRef](#)]
32. Mardia, K.V.; Kent, J.T.; Bibby, J.M. *Multivariate Analysis*; Probability and Mathematical Statistics; Academic Press: London, UK, 1979; Volume 15.
33. McQuillan, R.; Leutenegger, A.-L.; Abdel-Rahman, R.; Franklin, C.S.; Pericic, M.; Barac-Lauc, L.; Smolej-Narancic, N.; Janicijevic, B.; Polasek, O.; Tenesa, A.; et al. Runs of Homozygosity in European Populations. *Am. J. Hum. Genet.* **2008**, *83*, 359–372. [[CrossRef](#)]
34. de Cara, M.Á.R.; Villanueva, B.; Toro, M.Á.; Fernández, J. Using genomic tools to maintain diversity and fitness in conservation programmes. *Mol. Ecol.* **2013**, *22*, 6091–6099. [[CrossRef](#)]
35. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* **2015**, *4*, 7. [[CrossRef](#)]
36. Purcell, S.; Chang, C. *PLINK 1.9*. Available online: www.cog-genomics.org/plink/1.9/ (accessed on 17 October 2020).
37. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
38. Lenth, R.; Singmann, H.; Love, J.; Buerkner, P.; Herve, M. *emmeans: Estimated Marginal Means, aka Least-Squares Means*, version 1.5.1; R package, 2020; Available online: <https://CRAN.R-project.org/package=emmeans> (accessed on 16 October 2020).

39. Hothorn, T.; Bretz, F.; Westfall, P.; Heiberger, R.M.; Schuetzenmeister, A.; Scheibe, S. *multcomp: Simultaneous Inference in General Parametric Models*, version 1.4-14; R package, 2020; Available online: <https://CRAN.R-project.org/package=multcomp> (accessed on 16 October 2020).
40. FAO. *Secondary Guidelines for the National Farm Animal Genetic Resources Management Plans: Management of Small Populations at Risk*; FAO: Rome, Italy, 2000; p. 219.
41. AETE. *Commercial Embryo Transfer Activity in Europe 2018*. 2019. Available online: <https://www.aete.eu/publications/statistics/> (accessed on 17 October 2020).
42. Petyim, S.; Bage, R.; Forsberg, M.; Rodriguez-Martinez, H.; Larsson, B. The Effect of Repeated Follicular Puncture on Ovarian Function in Dairy Heifers. *J. Vet. Med. Ser. A* **2000**, *47*, 627–640. [[CrossRef](#)]
43. Qi, M.; Yao, Y.; Ma, H.; Wang, J.; Zhao, X.; Liu, L.; Tang, X.; Zhang, L.; Zhang, S.; Sun, F. Transvaginal Ultrasound-guided Ovum Pick-up(OPU) in Cattle. *J. Biomim. Biomater. Tissue Eng.* **2013**, *18*, 118.
44. Petyim, S.; Båge, R.; Madej, A.; Larsson, B. Ovum Pick-up in Dairy Heifers: Does it Affect Animal Well-being? *Reprod. Domest. Anim.* **2007**, *42*, 623–632. [[CrossRef](#)]
45. Bérodiér, M.; Berg, P.; Meuwissen, T.; Brochard, M.; Ducrocq, V. Improving mating plans at herd level using genomic information. In Proceedings of the Annual Meeting of the European Association for Animal Production (EAAP), Ghent, Belgium, 26–30 August 2019; p. 210.
46. Meuwissen, T.H.E. Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci.* **1997**, *75*, 934–940. [[CrossRef](#)]
47. Granleese, T.; Clark, S.A.; Swan, A.A.; van der Werf, J.H.J. Increased genetic gains in sheep, beef and dairy breeding programs from using female reproductive technologies combined with optimal contribution selection and genomic breeding values. *Genet. Sel. Evol.* **2015**, *47*, 70. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

III - Additional files

Additional file 1 Table S1 Positions and effects of additive quantitative trait loci (QTLs) along the simulated genome for the simulated embryo transfer French dairy cattle breeding scheme and its different variations.

Voir Annexe 1 du manuscrit de thèse.

Additional file 2 Table S2 ROH-based inbreeding rates per generation and genetic gains in true breeding value per generation for each scenario of the simulated embryo transfer French dairy cattle breeding program.

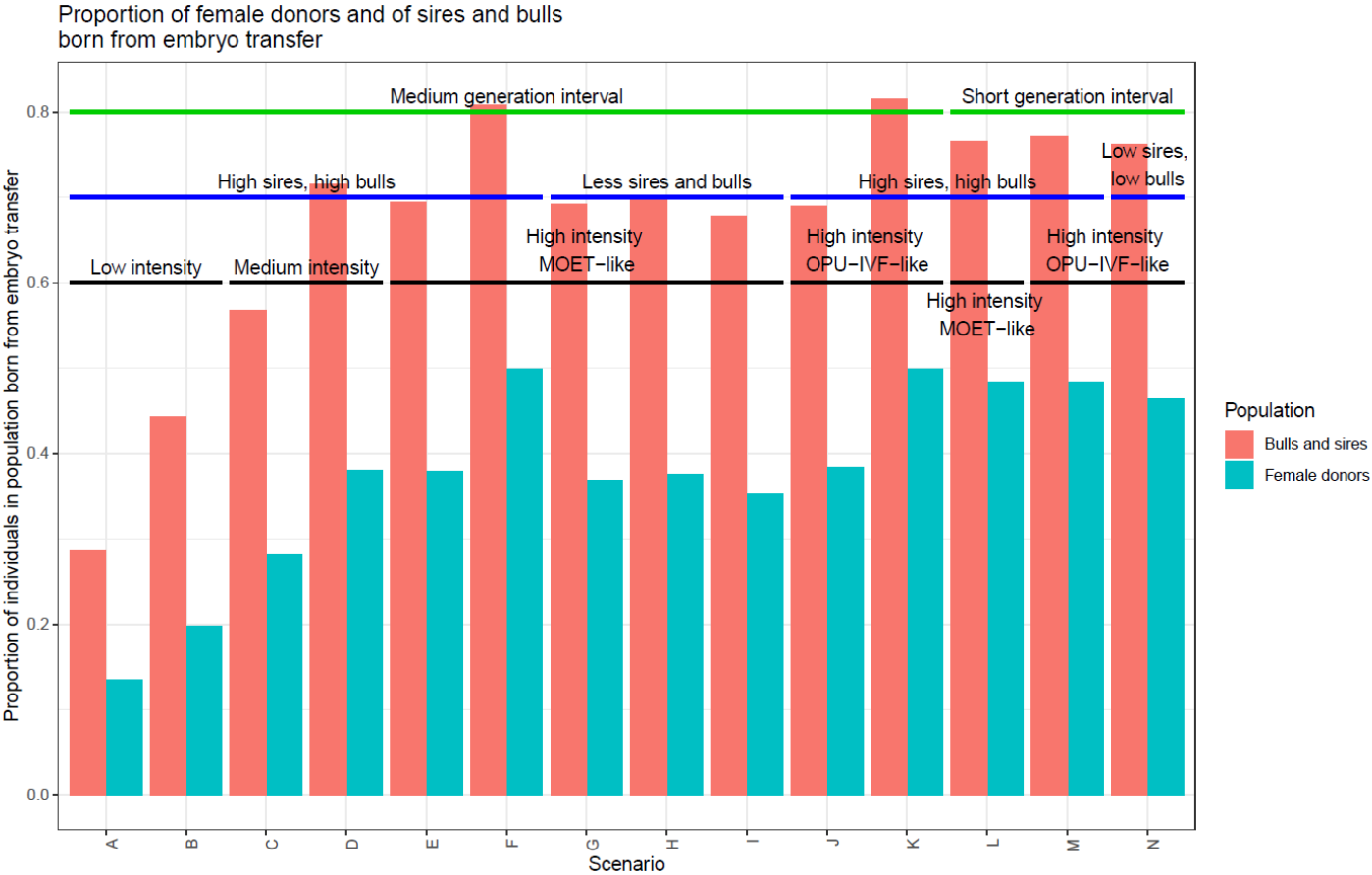
Comparisons with REF are equal to the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario REF (no use of embryo transfer).

Population	Scenario	ROH-based inbreeding rate per generation in %	ROH-based inbreeding rate per generation compared with scenario REF	Genetic gain in TBV per generation	Genetic gain in TBV per generation compared with scenario REF
Cows	REF	0.235	1.00	0.738	1.00
	A	0.245	1.04	0.793	1.07
	B	0.253	1.07	0.814	1.10
	C	0.308	1.31	0.882	1.20
	D	0.295	1.25	0.907	1.23
	E	0.334	1.42	0.917	1.24
	F	0.312	1.32	0.937	1.27
	G	0.349	1.48	0.89	1.21
	H	0.370	1.57	0.929	1.26
	I	0.363	1.54	0.872	1.18
	J	0.322	1.37	0.922	1.25
	K	0.340	1.44	0.947	1.28
	L	0.448	1.90	1.001	1.36
	M	0.448	1.90	1.005	1.36
N	0.501	2.13	0.949	1.28	
Sires and bulls	REF	0.094	1.00	0.560	1.00
	A	0.099	1.05	0.606	1.08
	B	0.114	1.21	0.628	1.12
	C	0.151	1.60	0.680	1.21
	D	0.155	1.65	0.714	1.28
	E	0.218	2.31	0.712	1.27
	F	0.218	2.31	0.748	1.34
	G	0.216	2.29	0.688	1.23
	H	0.234	2.48	0.722	1.29
	I	0.260	2.76	0.680	1.21
	J	0.224	2.38	0.718	1.28
	K	0.232	2.46	0.760	1.36
	L	0.323	3.43	0.653	1.17
	M	0.329	3.50	0.654	1.17
N	0.356	3.78	0.615	1.10	

Comparisons with REF are equal to the ratio between the estimated value of the parameter for a given scenario and the estimated value of the parameter for the reference scenario REF (no use of embryo transfer).

ROH Runs of Homozygosity; TBV True Breeding Value

Additional file 3 Figure S1 Proportion of female donors and of sires and bulls born from embryo transfer over all birth years and runs for each scenario of the simulated embryo transfer French dairy cattle breeding scheme and its different variations.



IV - Discussion

Si les études rétrospectives permettent de connaître a posteriori les conséquences de modifications des schémas, des études prospectives, via simulations, permettent aux entreprises de sélection de modifier leurs schémas de sélection au mieux en fonction de leurs objectifs, grâce aux tendances dégagées par les résultats des simulations. Ces simulations, réalisées dans le cadre d'un partenariat fort avec les entreprises de sélection, ont permis de quantifier l'évolution relative du progrès et de la diversité génétique pour 15 scénarios différents et donc de produire des conseils aux entreprises de sélection dont le schéma se rapprochait du schéma de base simulé dans le cadre de ma thèse.

Les contraintes, réalistes et réellement imposées dans les schémas de sélection, sur le nombre de fils et de filles sélectionnés par femelle donneuse ont permis de limiter l'augmentation du taux de consanguinité et du progrès génétique dans les scénarios avec une forte intensité d'utilisation du transfert embryonnaire. Ces contraintes sont une bonne pratique des entreprises de sélection pour limiter la perte de diversité génétique dans les schémas de sélection.

La réduction du nombre de pères à taureaux et de taureaux mis en marché a entraîné une diminution significative du progrès génétique et une augmentation significative du taux de consanguinité. Toutefois, pour réduire les coûts des schémas de sélection, il est peu probable que le nombre de taureaux augmente. Dans ce contexte, le maintien du nombre de taureaux au moins au niveau actuel permettrait de maintenir des coûts constants tout en atteignant un progrès génétique plus élevé que si le nombre de taureaux était inférieur (comme le montre cette étude) ou supérieur (Thomassen et al., 2016). Cela ne diminuerait pas la perte de diversité génétique par rapport à la situation actuelle mais éviterait au moins son accélération. Le maintien du nombre de taureaux ne suffit toutefois pas à garantir le maintien de la diversité génétique. En effet, comme nous l'avons vu dans le chapitre précédent avec le cas de la Prim'Holstein, il est nécessaire de maîtriser également l'équilibre de contribution des taureaux et donc la manière de les diffuser pour éviter les effets délétères du « star system », c'est-à-dire l'utilisation massive de quelques taureaux élite.

Dans nos simulations, l'intensification de l'utilisation du transfert embryonnaire et la réduction du nombre de taureaux ont eu un impact négatif sur la diversité génétique, en augmentant le taux de consanguinité (par an et par génération). Différentes stratégies peuvent être utilisées pour limiter cet impact tout en permettant un progrès génétique accru dans les races bovines laitières en général, telles qu'une meilleure conception des plans d'accouplement (Bérodier et al., 2019) ou les méthodes d'OC (Granleese et al., 2019; Meuwissen, 1997). Toutefois, ces stratégies ne prennent en compte que la diversité génétique globale des animaux, alors que la dépression de la

consanguinité dépend à la fois de la consanguinité et du fardeau génétique, c'est-à-dire de l'ensemble des mutations délétères portées par la population. Il serait intéressant d'identifier les régions génomiques porteuses de ces mutations et de prendre en compte ces informations dans les évaluations génomiques. Cela devrait permettre d'être plus strict sur la perte de diversité génétique dans ces régions du génome tout en étant plus souple dans les autres régions. Ainsi, l'impact de la perte de diversité génétique sur la santé et les performances des races bovines laitières pourrait être limité dans le cadre d'une intensification de l'utilisation du transfert embryonnaire.

Une limite de cette étude de simulation des schémas de sélection est le nombre de réplicats pour chaque scénario. Nous avons simulé 10 réplicats pour chacun des 15 scénarios étudiés. Nous aurions aimé en simuler davantage. Or, cela est compliqué car garder en mémoire les informations de pedigree et de génotypage des 150 populations simulées ainsi que les informations sur les positions des ROH et les valeurs génétiques des animaux utilise actuellement près de 700 Go de mémoire sous forme compressée.

La mémoire a également été une limite pendant les simulations. En effet, chaque réplicat de chaque scénario utilisait près de 70 Go de mémoire vive pendant plusieurs jours. Ce temps peut augmenter jusqu'à plusieurs semaines si le nombre d'animaux simulés augmente ou qu'on calcule des valeurs génétiques prenant en compte l'originalité des animaux les uns par rapport aux autres (un individu très apparenté au reste de la population est peu original tandis qu'un individu très peu apparenté au reste de la population est très original) comme nous l'avons testé. Ces durées longues ainsi que les grandes quantités de mémoire nécessaires ont rendu difficile la cohabitation avec les autres utilisateurs des infrastructures informatiques (clusters de calculs), la gestion des bugs ou la paramétrisation d'options mal documentées.

Dans le cadre d'autres études de simulation, servant à évaluer l'efficacité de nouvelles méthodes d'évaluation ou de gestion de la diversité génétique comme des plans d'accouplements par exemple, il serait pertinent de simuler des populations de petite taille et de comparer les résultats à des populations simulées de grande taille pour évaluer la nécessité de travailler avec des populations simulées d'une taille comparable aux populations réelles ou si une diminution des effectifs à l'échelle suffirait pour obtenir de bons résultats. Cela permettrait de réduire les temps de calcul et la mémoire nécessaire, et d'optimiser également le temps passé à résoudre les bugs qui ne manqueraient pas de survenir.

Chapitre 4 Dépression de consanguinité et détection du fardeau génétique – Application au cas d’une race bovine laitière française de grand effectif : la Montbéliarde

I - Introduction

Une des conséquences de la perte de diversité dans les populations animales est la dépression de consanguinité, définie comme la réduction de la valeur phénotypique moyenne des caractères liés à la *fitness* ou valeur sélective, c’est-à-dire à la capacité de reproduction ou à l’efficacité physiologique, due à la consanguinité (Wright, 1968 ; Darwin, 1876, 1868).

Tous les caractères peuvent être affectés par la dépression de consanguinité, d’autant plus s’ils sont sélectionnés (Leroy, 2014). Or, dans les races bovines laitières, les caractères sélectionnés, constituant les objectifs de sélection, sont généralement les plus importants économiquement. C’est le cas en particulier des caractères de production laitière (quantité de lait produit, concentration en matière grasse ou en matière protéique) ou de caractères de santé associés par exemple à la santé de la mamelle, comme le taux de cellules somatiques dans le lait.

Un nombre élevé de cellules somatiques dans le lait correspond généralement à une mammite subclinique, c’est-à-dire sans autre symptôme apparent (douleur, etc.), et déprécie le prix du lait payé par les laiteries aux éleveurs (des pénalités sont appliquées par tranche de concentration de cellules somatiques dans le lait). De plus, si une vache présentant une mammite clinique est traitée avec des antibiotiques, son lait n’est pas commercialisable durant le traitement et pendant une période d’attente réglementaire suivant la fin du traitement, d’une durée variable selon le traitement appliqué. Pour la production laitière, moins une vache produit de lait ou moins celui-ci est concentré en matières grasse et protéique, moins l’éleveur va gagner d’argent, car il vendra moins de lait. De plus, le prix du lait est déprécié si sa composition n’est pas jugée satisfaisante par les laiteries. La dépression de consanguinité peut donc détériorer le progrès génétique annuel dans ces races et altérer le revenu des éleveurs.

Ainsi, il est important pour les entreprises de sélection et les éleveurs de savoir si la dépression de consanguinité est détectable au sein des races qu’ils sélectionnent et élèvent, et de connaître la part de progrès génétique perdue à cause de la dépression de consanguinité, afin de pouvoir si nécessaire modifier leurs schémas de sélection pour limiter le taux de consanguinité dans la population.

La détection de la dépression de consanguinité à l’échelle globale s’effectue généralement comme une régression de la valeur de caractères d’intérêt sur la consanguinité globale, qu’elle soit calculée à partir des pedigrees ou des données moléculaires (Doekes et al., 2019; Pryce et al.,

2014 ; Charlesworth et Willis, 2009 ; Curik et al., 2001). Cette méthode suppose que l'autozygotie est répartie uniformément le long du génome. Or, il s'agit d'une hypothèse irréaliste, en particulier dans les populations sélectionnées, où les niveaux d'autozygotie varient fortement le long du génome (Curik et al., 2017; Pemberton et al., 2012) en fonction des taux de recombinaison variables et des blocs haplotypiques présents dans les régions du génome sous sélection. En estimant la consanguinité à l'échelle globale, il n'est pas possible de prendre en compte cette hétérogénéité.

La disponibilité croissante de données de génotypage dans les races bovines laitières françaises permet d'envisager l'étude de cette hétérogénéité. En effet, il est désormais possible d'estimer la dépression de consanguinité non seulement à l'échelle globale mais aussi à l'échelle du génome, en détectant les régions du génome impliquées dans la dépression de consanguinité, c'est-à-dire portant le fardeau génétique, grâce aux ROH (Jenko et al., 2019 ; Ferenčaković et al., 2017 ; Pryce et al., 2014 ; Howard et al., 2015 ; Fritz et al., 2013 ; Keller et al., 2012). Cela permettrait à terme de prendre en compte ces régions différenciellement dans les évaluations génomiques ou encore dans les plans d'accouplement conçus à partir d'informations génomiques (Gómez-Romano et al., 2016).

Chez la Montbéliarde, race bovine laitière française de grand effectif, le taux de consanguinité est bien maîtrisé (Doublet et al., 2019). On peut donc se demander si dans cette race, il existe de la dépression de consanguinité à l'échelle globale pour des caractères de production et de santé, et si oui, quels sont les caractères les plus affectés. Si certains caractères sont touchés par la dépression de consanguinité et que ces effets sont importants (alors que la consanguinité est déjà bien gérée chez la Montbéliarde), cela signifie qu'il sera difficile de maîtriser la dépression de consanguinité en réduisant le taux de consanguinité. Dans ce cas, on peut se demander quelles régions du génome sont impliquées dans le fardeau génétique, dans l'optique à terme d'agir sur le fardeau si limiter l'accroissement de la consanguinité ne suffit pas à éviter la dépression de consanguinité. De plus, il est intéressant de savoir comment se comporte la dépression de consanguinité à proximité des régions du génome fortement sélectionnées car ayant un fort effet sur les caractères de production, comme le QTL de production laitière situé au niveau du gène *DGAT1* (Gautier et al., 2007 ; Kühn et al., 2004 ; Grisart et al., 2002 ; Cases et al., 1998).

Pour répondre à ces interrogations, nous avons estimé le niveau de dépression de consanguinité globale en race Montbéliarde pour six caractères : un caractère de santé et cinq caractères de production. Nous avons ensuite estimé l'effet de la consanguinité le long du génome sur ces mêmes caractères pour identifier les régions du génome de la Montbéliarde impliquées dans la dépression de consanguinité. Nous nous sommes enfin intéressés au comportement de la dépression de consanguinité dans les régions du génome fortement sélectionnées comme *DGAT1*.

II - Matériel et méthodes

A - Données

Nous avons étudié la dépression de consanguinité pour six caractères en race Montbéliarde. Le premier caractère est un caractère de santé, correspondant à la quantité de cellules somatiques présentes dans le lait. Il sera par la suite noté « cellules ». Les cinq autres caractères sont des caractères de production (caractères laitiers) : la production laitière totale pour une lactation (caractère noté « lait »), la quantité totale de matière grasse (notée MG) et de matière protéique (notée MP) dans le lait, ainsi que la concentration du lait en matière grasse (le taux butyreux, noté TB) et en matière protéique (le taux protéique, noté TP).

Afin d'étudier ces caractères, nous avons utilisé les indices polygéniques officiels. Ils sont calculés à partir des performances corrigées des animaux, qui correspondent à la part des performances expliquée par la génétique. Les performances sont donc corrigées pour la part de la performance expliquée par l'effet d'environnement permanent, les effets de milieux identifiés (le troupeau par campagne, le rang de lactation et la durée de tarissement par région et par campagne, le mois de lactation et l'âge au vêlage par rang de lactation, région et campagne) et un résidu (Launay, 2015). Les indices polygéniques sont ensuite estimés à partir des performances corrigées après la prise en compte de l'apparentement entre les individus. Nous avons utilisé les indices polygéniques des vaches nées entre 2012 et 2017 et ayant réalisé au moins une lactation. Le jeu de données était constitué de 90 102 vaches pour l'étude du caractères cellules et de 97 504 vaches pour l'étude des caractères de production laitière.

Nous disposons du génotype imputé, à partir du génotypage basse densité 10K, de toutes les vaches du jeu de données étudié pour 53 469 SNPs basés sur le génome de référence de l'assemblage ARS-UCD1.2 (Rosen et al., 2020).

B - Calcul des taux de consanguinité et du progrès génétique

La consanguinité des vaches a été estimée à partir des ROH. Un ROH était défini comme un segment homozygote d'au moins 15 SNPs ou d'une longueur minimale de 1000 kilobases, avec un écart maximal entre deux SNPs au sein d'un ROH de 1 Mb. Les ROH ont été détectés grâce au logiciel PLINK 1.9 via la fonction "homozyg" (Chang et al., 2015; Purcell et Chang) (ligne de commande : `plink --cow --bfile genotyping_data_filename --homozyg --homozyg-kb 1000 --homozyg-snp 15 --homozyg-window-snp 15 --homozyg-density 1000 --out output_filename`).

La consanguinité estimée à partir des ROH pour un individu i , $F_{ROH,i}$ a été calculée comme la proportion du génome inclus dans les ROH (McQuillan et al., 2008) :

$$F_{ROH,i} = \frac{\sum L_{ROH,i}}{L_{auto}}$$

avec $\sum L_{ROH,i}$ la longueur totale des ROH pour l'individu i , et L_{auto} la longueur du génome autosomique couvert par les SNPs après retrait des espaces de plus de 1000 kb entre deux SNPs consécutifs.

Pour la suite de l'étude, nous avons centré et réduit les valeurs des indices pour chaque caractère dans le jeu de données. Ces indices centrés-réduits seront notés indices CR par la suite.

Note : Plus la valeur de ces indices polygéniques est élevée, meilleure est la vache pour le caractère concerné. Dans le cas des cellules, plus la concentration en cellules somatiques dans le lait est élevée, plus la valeur de l'indice correspondant est basse.

Le progrès génétique annuel observé dans l'échantillon a été calculé en réalisant une régression linéaire simple entre les indices polygéniques bruts de chaque vache et l'année de naissance, selon le modèle qui suit :

$$Y_i = \beta_1 + \alpha_1 \times Année_i + \epsilon_i$$

avec Y_i l'indice polygénique brut, $Année_i$ l'année de naissance de l'individu i , β_1 l'ordonnée à l'origine du modèle, α_1 le coefficient de régression du modèle, correspondant au progrès génétique annuel, et ϵ_i le résidu du modèle. Le progrès génétique annuel a également été calculé à partir de l'indice polygénique CR. Le taux de consanguinité annuel a été estimé de la même façon à partir de la consanguinité estimée à partir des ROH.

C - Détection de la dépression de consanguinité globale par caractère

Afin de détecter une potentielle dépression de consanguinité pour chacun des caractères étudiés en race Montbéliarde, nous avons réalisé une régression linéaire avec comme variable à expliquer l'indice polygénique CR pour chaque vache et comme variable explicative la consanguinité estimée à partir des ROH pour chaque vache, en %, selon le modèle qui suit :

$$Y_i = \beta_2 + \alpha_2 \times F_{ROH,i} + \epsilon_i$$

Avec Y_i l'indice polygénique CR de l'individu i , $F_{ROH,i}$ le coefficient de consanguinité estimé à partir des ROH (en %) de l'individu i , β_2 l'ordonnée à l'origine du modèle, α_2 le coefficient de régression du modèle, et ϵ_i le résidu du modèle.

Les coefficients de régression obtenus correspondent à la baisse de la moyenne de l'indice polygénique de chaque caractère en unités d'écart-type pour une augmentation de 1 % de la consanguinité estimée à partir des ROH.

Pour chaque caractère, trois résultats sont possibles :

- (i) Obtenir un coefficient de régression significativement négatif. Dans ce cas, on peut raisonnablement conclure que la dépression de consanguinité s'exprime dans la population étudiée pour ce caractère ;
- (ii) Obtenir un coefficient de régression non significativement différent de 0. Dans ce cas, deux conclusions sont possibles : soit il n'y a pas de dépression de consanguinité, soit son effet est trop faible pour être détecté grâce à notre jeu de données ;
- (iii) Obtenir un coefficient de régression significativement positif. Ce cas est plus difficile à interpréter. Il peut, entre autres, être le signe d'une purge du fardeau génétique.

Nous avons ensuite calculé le pourcentage de perte annuelle de progrès génétique due à la dépression de consanguinité. L'objectif était d'obtenir une valeur de dépression de consanguinité « annuelle », à partir du coefficient de régression obtenu précédemment (baisse de l'indice polygénique en unités d'écart-type pour une augmentation de 1% de la consanguinité), du taux de consanguinité annuel estimé à partir des ROH (en %) ainsi que du progrès génétique annuel (en unités d'écart-type).

Le pourcentage de perte annuelle de progrès génétique due à la dépression de consanguinité était calculé comme suit :

$$\text{Perte annuelle de progrès génétique (en \%)} = \frac{\text{Coefficient de régression} \times \text{Taux de consanguinité annuel (en \%)}}{\text{Progrès génétique annuel (en unités d'écart-type)}} \times 100.$$

D - Détection de la dépression de consanguinité le long du génome par caractère

Le deuxième objectif de cette étude était de savoir dans quelles régions du génome se situent les allèles délétères constituant le fardeau. Pour cela, nous avons réalisé une régression par fenêtre du génome entre la consanguinité estimée à partir des ROH et les indices polygéniques CR utilisés précédemment.

Cette méthode consiste à faire coulisser une fenêtre d'une longueur donnée le long du génome et pour chaque fenêtre, effectuer la régression des indices polygéniques (= performances corrigées) des animaux par la consanguinité locale sur cette fenêtre à partir des ROH. Ce type d'approche a déjà été utilisé dans l'étude de maladies humaines comme la schizophrénie (Keller et al., 2012). Cette méthode permet d'estimer l'effet de la consanguinité sur les performances dans des régions du génome de taille définie.

Pour estimer les effets de la consanguinité sur les performances le long du génome, nous avons défini 53 063 régions chevauchantes d'une longueur de 15 SNPs, soit environ 1 Mb. Cette longueur correspond à la longueur minimale des ROH détectés dans cette étude. La première région de chaque chromosome commence au premier SNP de ce chromosome, les autres régions sont ensuite définies selon une fenêtre coulissante glissant d'un SNP à la fois. La seconde région de chaque chromosome commence donc un SNP plus loin soit au deuxième SNP du chromosome et ainsi de suite. La dernière région de chaque chromosome est constituée des 15 derniers SNPs du chromosome.

Pour chacune des 53 063 régions et pour chacun des six caractères, nous avons calculé $F_{\text{ROH local}}$ la consanguinité locale estimée à partir des ROH de toutes les vaches de la population étudiée comme suit :

$$F_{\text{ROH},i,k} = \frac{\sum L_{\text{ROH},i,k}}{L_k}$$

avec $\sum L_{\text{ROH},i,k}$ la longueur totale des ROH inclus dans la région k pour l'individu i , et L_k la longueur de la région k . $F_{\text{ROH},i,k}$ était ainsi inclus entre 0 et 1.

Ensuite, pour chaque région et chaque caractère séparément, nous avons réalisé une régression linéaire avec l'indice polygénique CR de chaque individu comme variable à expliquer et $F_{\text{ROH local}}$ comme variable explicative, selon le modèle qui suit :

$$Y_i = \beta_3 + \alpha_3 \times F_{\text{ROH},i,k} + \epsilon_{i,k}$$

avec Y_i l'indice polygénique CR de l'individu i , $F_{\text{ROH},i,k}$ le coefficient de consanguinité local estimé à partir des ROH pour l'individu i dans la région génomique k , β_3 l'ordonnée à l'origine du modèle, α_3 le coefficient de régression du modèle, et $\epsilon_{i,k}$ le résidu du modèle.

Les coefficients de régression obtenus correspondent à la différence entre les indices des animaux pour lesquels la région étudiée est complètement consanguine (soit un $F_{\text{ROH local}}$ égal à 1) et ceux des animaux pour lesquels cette région n'est pas du tout consanguine (soit un $F_{\text{ROH local}}$ égal à 0). Dans la suite, ces coefficients de régression seront aussi appelés pentes.

E - Corrélations entre ROH et fréquences alléliques

Nous avons calculé la proportion de vaches possédant un ROH à chaque SNP. Cela permet d'obtenir une information sur les ROH à l'échelle de la population, alors qu'ils sont à l'origine une mesure individuelle de la consanguinité le long du génome. Nous avons également calculé les fréquences alléliques minoritaires (MAF) pour tous les SNPs dans le jeu de données avec la fonction « freq » du logiciel PLINK 1.9 (Chang et al., 2015; Purcell et Chang).

Nous avons ensuite calculé pour chaque caractère les corrélations de Pearson entre la pente estimée précédemment dans chaque région de 15 SNPs du génome, la proportion de vaches possédant un ROH pour le SNP central de chaque région (c'est-à-dire le 8^{ème} SNP de la région) et la MAF pour ce même SNP.

F - Scores SIFT

Nous avons sélectionné les régions du génome correspondant aux pics de pentes obtenus pour la détection de la dépression de consanguinité le long du génome. Pour cela, nous avons recherché les maximums (pour les pentes positives) et minimums (pour les pentes négatives) locaux des pentes (pour les pics positifs ou négatifs) et avons gardé pour chaque pic la région correspondant à l'extremum local. Les pics de pentes correspondaient généralement aux pics de $-\log_{10}(p\text{-valeur})$. Nous avons conservé pour chaque caractère les 50 régions correspondant aux pics de pentes les plus forts et les 50 régions correspondant aux pics de pentes les plus forts en négatif.

Nous avons ensuite cherché à savoir si ces régions « pics de pentes » possédaient des locus auxquels des mutations peuvent avoir des effets délétères très forts. Pour cela, nous avons récupéré les scores SIFT (*Sorting Intolerant from Tolerant*) des variants potentiels pour tous les SNPs inclus dans les régions pics de pentes grâce à l'outil Variant Effect Predictor d'Ensembl (McLaren et al., 2016). Le score SIFT permet de prédire si la substitution d'un acide aminé dans une protéine affecte sa fonction et à quel point. Ce score va de 0 (substitution délétère) à 1 (substitution tolérée).

III - Résultats

A - Taux de consanguinité et progrès génétique

La moyenne et l'écart-type des valeurs des indices pour chaque caractère et de la consanguinité estimée à partir des ROH pour les vaches du jeu de données sont présentées dans le Tableau 4 ci-dessous.

Tableau 4 Description du jeu de données utilisé pour l'étude de la dépression de consanguinité en race Montbéliarde

Caractère	Moyenne [écart-type] des indices de la population étudiée	Moyenne [écart-type] de la consanguinité estimée à partir des ROH (en %)
Cellules	0,101 [0,682]	20,4 [2,46]
Lait	167 [391]	
MG	7,42 [16,1]	
MP	6,92 [13,0]	20,4 [2,47]
TB	0,0769 [1,65]	
TP	0,167 [0,980]	

Les valeurs de progrès génétique et de taux de consanguinité annuels pour les vaches incluses dans le jeu de données sont présentées dans le Tableau 5 ci-dessous.

Tableau 5 Progrès génétique et taux de consanguinité annuels pour les vaches génotypées de race Montbéliarde nées entre 2012 et 2017, incluses dans le jeu de données d'étude de la dépression de consanguinité

Caractère	Progrès génétique annuel à partir des indices polygéniques bruts [erreur standard]	Progrès génétique annuel en unités d'écart-type de l'échantillon [erreur standard]	Taux de consanguinité annuel (en %) [erreur standard]
Cellules	0,0259 [0,00164] *	0,0380 [0,00240] *	0,193 [0,00588] *
Lait	23,8 [0,901] *	0,0608 [0,00231] *	
MG	1,05 [0,0372] *	0,0649 [0,00230] *	
MP	1,06 [0,0298] *	0,0818 [0,00230] *	0,190 [0,00568] *
TB	0,000357 [0,00382] ns	0,000216 [0,00231] ns	
TP	0,0273 [0,00226] *	0,0279 [0,00231] *	

* significativement différent de zéro au risque d'erreur de 5 %
ns non significativement différent de zéro au risque 5 %

Le TB était le seul caractère pour lequel le progrès génétique n'était pas significativement différent de zéro au risque 5% (erreur standard plus grande que la valeur estimée du progrès).

B - Dépression de consanguinité globale par caractère

Les coefficients de régression obtenus dans l'étude de la consanguinité globale par caractère sont présentés dans le Tableau 6 ci-dessous.

Tableau 6 Coefficients de régression linéaire pour la détection de la dépression de consanguinité chez la Montbéliarde pour six caractères

Caractère	Coefficient de régression linéaire [Intervalle de confiance à 95%]	Dépression de consanguinité détectée ?
Cellules	-0,00839 [-0,0110; -0,00574]	Oui
Lait	-0,00380 [-0,00634; -0,00126]	Oui
MG	-0,00712 [-0,00966; -0,00458]	Oui
MP	-0,00334 [-0,00588; -0,000800]	Oui
TB	-0,003878 [-0,00642; -0,00134]	Oui
TP	0,0000584 [-0,00248 ; 0,00260]	Non car coefficient de régression non significativement différent de 0

Les coefficients de régression présentés dans le Tableau 6 ci-dessus correspondent à la baisse de la moyenne de l'indice polygénique de chaque caractère en unités d'écart-type pour une augmentation de 1 % de la consanguinité estimée à partir des ROH.

Les valeurs de la dépression de consanguinité annuelle pour les six caractères étudiés sont présentées dans le Tableau 7 ci-dessous. Elles correspondent à la baisse annuelle de l'indice polygénique pour chaque caractère en unités d'écart-type causée par la dépression de consanguinité.

Tableau 7 Dépression de consanguinité annuelle chez la Montbéliarde pour six caractères

Caractère	Coefficient de régression linéaire correspondant à la dépression de consanguinité	Taux de consanguinité annuel (en %)	Dépression de consanguinité annuelle
Cellules	-0,00839	0,193	-0,00161
Lait	-0,00380		-0,000723
MG	-0,00712		-0,00135
MP	-0,00334	0,190	-0,000636
TB	-0,00388		-0,000738
TP	/		/

Les valeurs de perte annuelle du progrès génétique due à la dépression de consanguinité sont présentées dans le Tableau 8 ci-dessous.

Tableau 8 Perte annuelle de progrès génétique due à la dépression de consanguinité pour six caractères en race Montbéliarde

Caractère	Dépression de consanguinité annuelle en unités d'écart-type	Progrès génétique annuel en unités d'écart-type	Perte annuelle de progrès génétique due à la dépression de consanguinité (en %)
Cellules	-0,00161	0,0380	4,25
Lait	-0,00072	0,0608	1,19
MG	-0,00135	0,0649	2,09
MP	-0,00064	0,0818	0,78
TB	-0,00074	/	/
TP	/	0,0279	/

On observe que la dépression de consanguinité a entraîné chez les vaches Montbéliarde un manque à gagner de 4,25 % du progrès génétique annuel pour les cellules, de 1,19 % pour le lait, de 2,09 % pour la MG et de 0,78 % pour la MP. Pour le TB et le TP, nous n'avons pas calculé de perte annuelle du progrès génétique. En effet, pour le TB, le progrès génétique annuel n'étant pas

significativement différent de 0 ; et pour le TP, nous n'avons pas détecté de dépression de consanguinité à l'échelle individuelle.

C - Dépression de consanguinité le long du génome par caractère

1 - Distribution des effets de la consanguinité le long du génome

Les coefficients de régression linéaire obtenus dans l'étude de la dépression de consanguinité le long du génome ainsi que les p-valeurs associées sont présentés de la Figure 10 à la Figure 15 pour les caractères cellules, lait, MG, MP, TB et TP. La distribution globale des pentes pour chacun de ces caractères est présentée sous forme d'histogrammes dans la Figure 16.

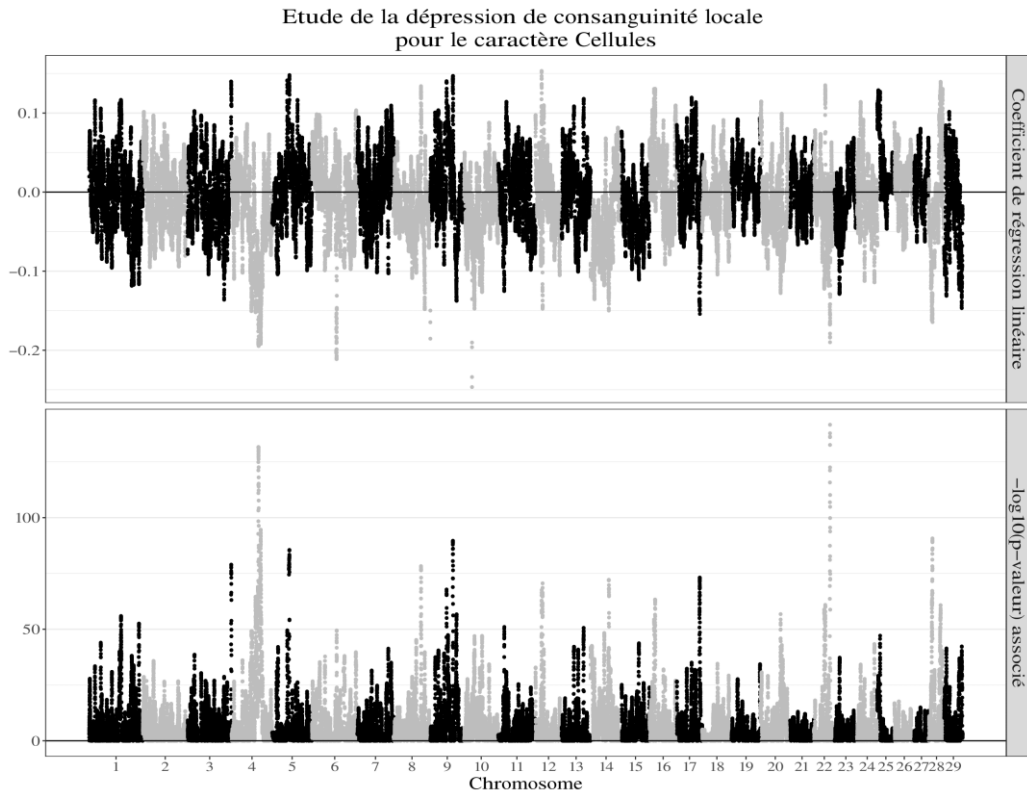


Figure 10 Effets de la consanguinité sur l'indice polygénique « cellules » le long du génome chez les vaches Montbéliardes

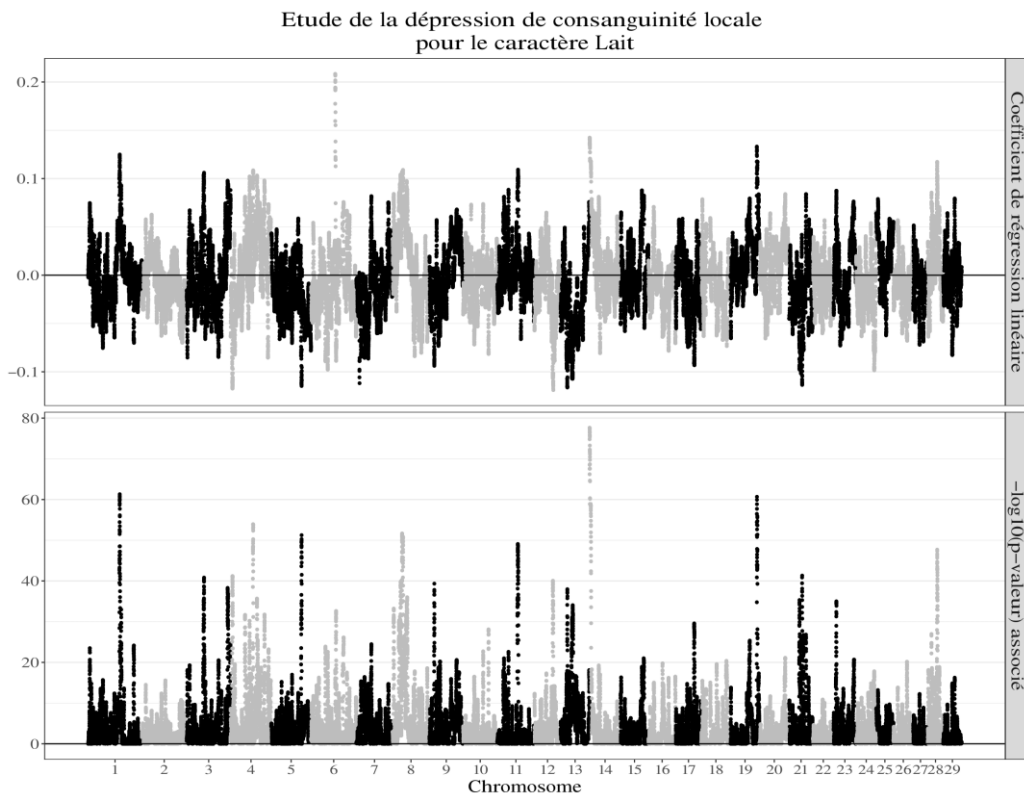


Figure 11 Effets de la consanguinité sur l'indice polygénique « lait » le long du génome chez les vaches Montbéliardes

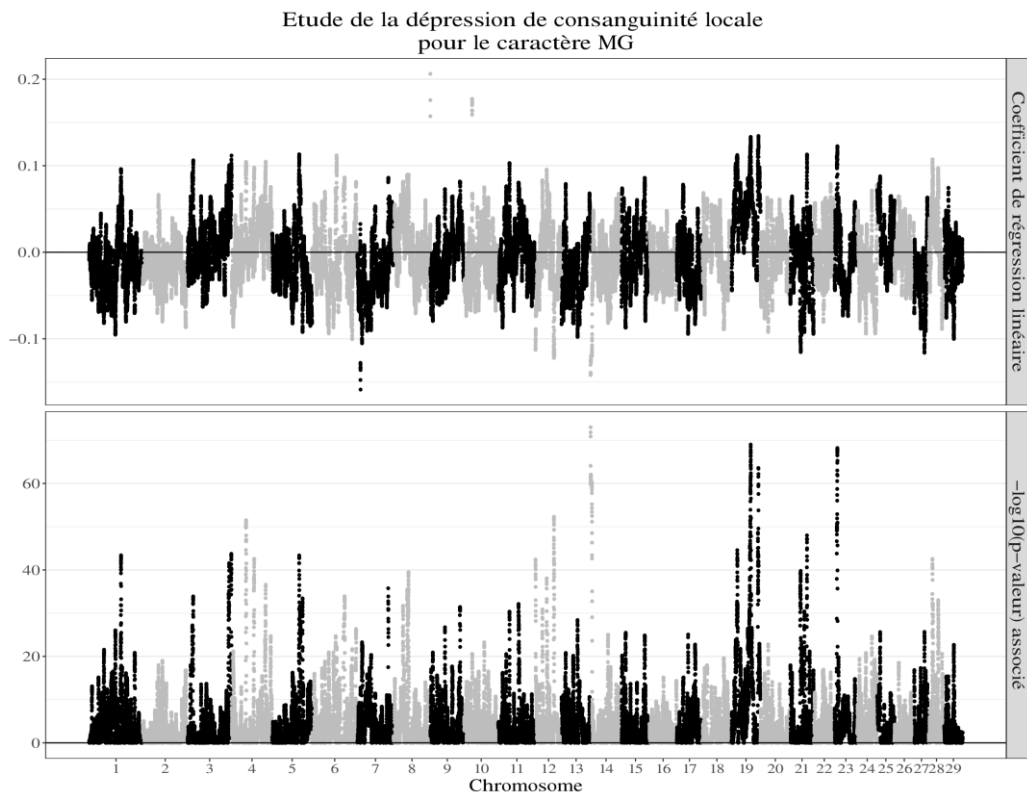


Figure 12 Effets de la consanguinité sur l'indice polygénique « MG » le long du génome chez les vaches Montbéliardes

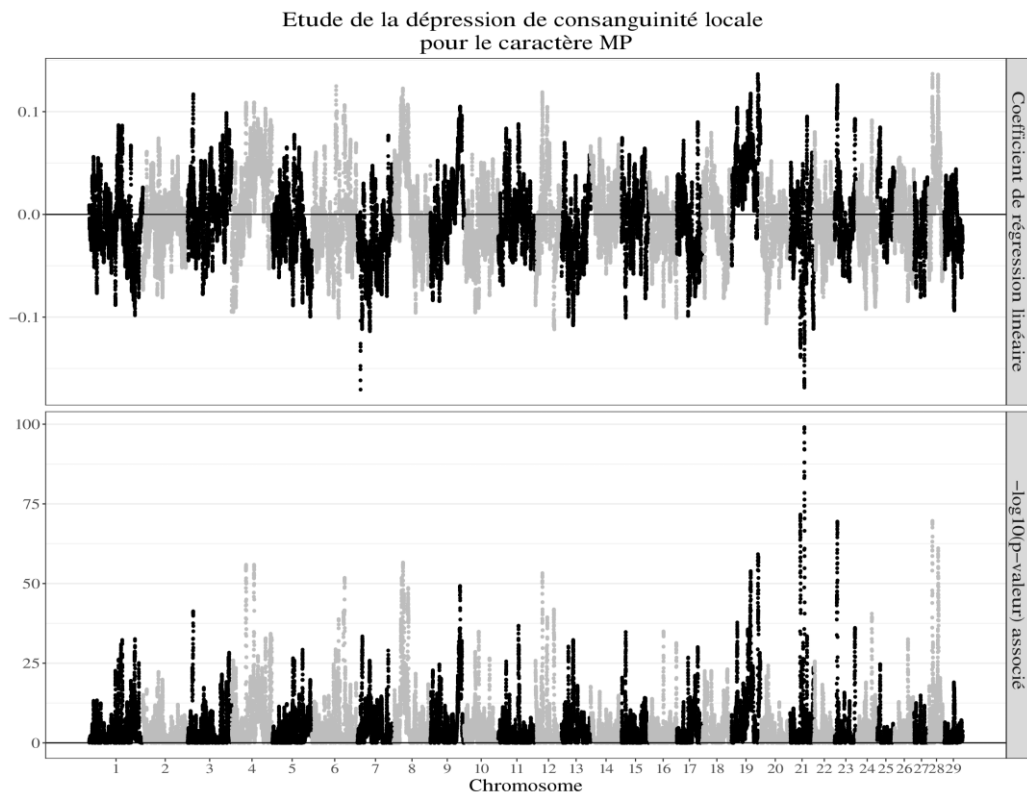


Figure 13 Effets de la consanguinité sur l'indice polygénique « MP » le long du génome chez les vaches Montbéliardes

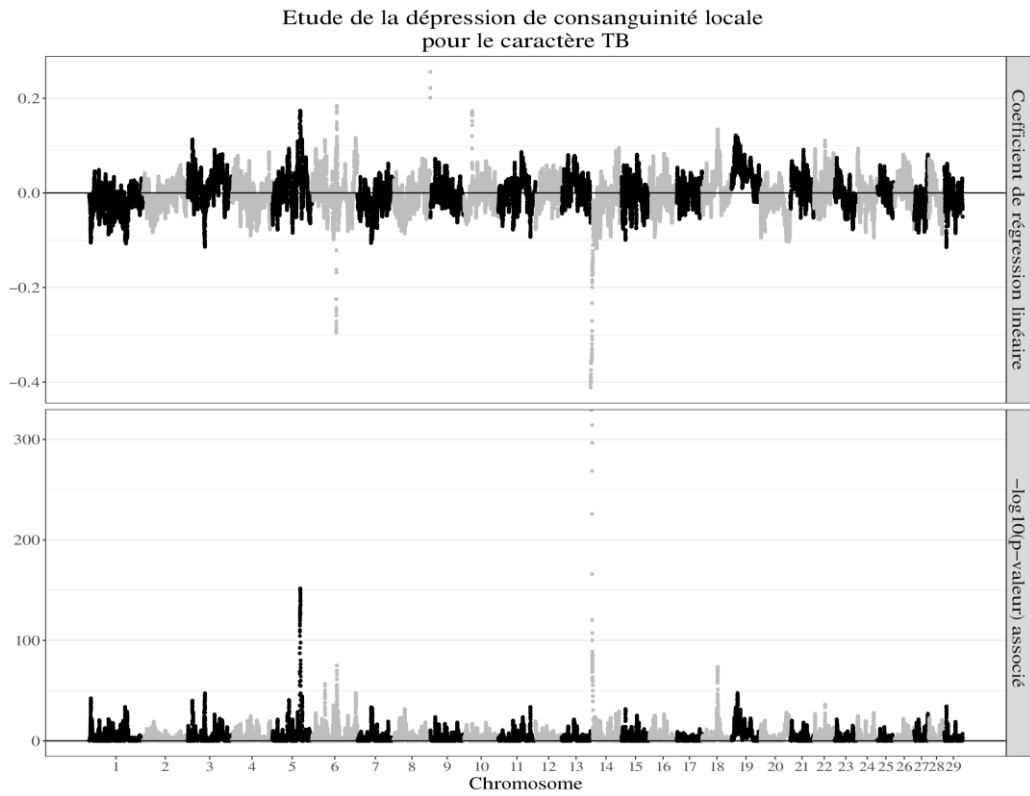


Figure 14 Effets de la consanguinité sur l'indice polygénique « TB » le long du génome chez les vaches Montbéliardes

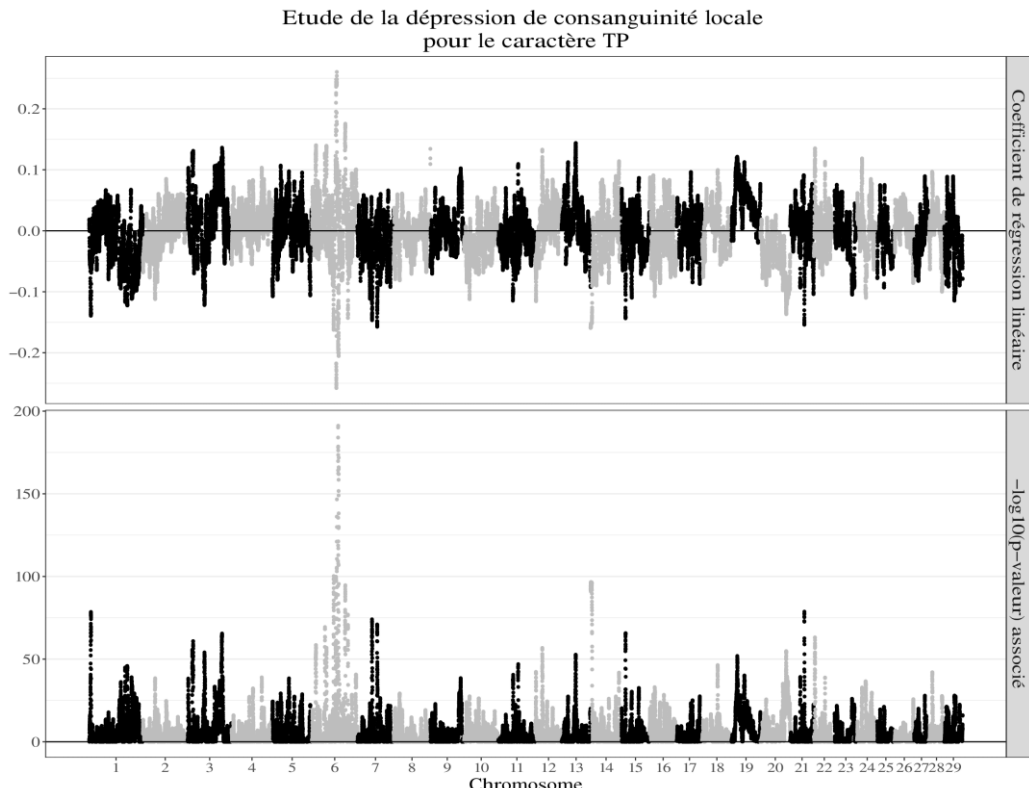


Figure 15 Effets de la consanguinité sur l'indice polygénique « TP » le long du génome chez les vaches Montbéliardes

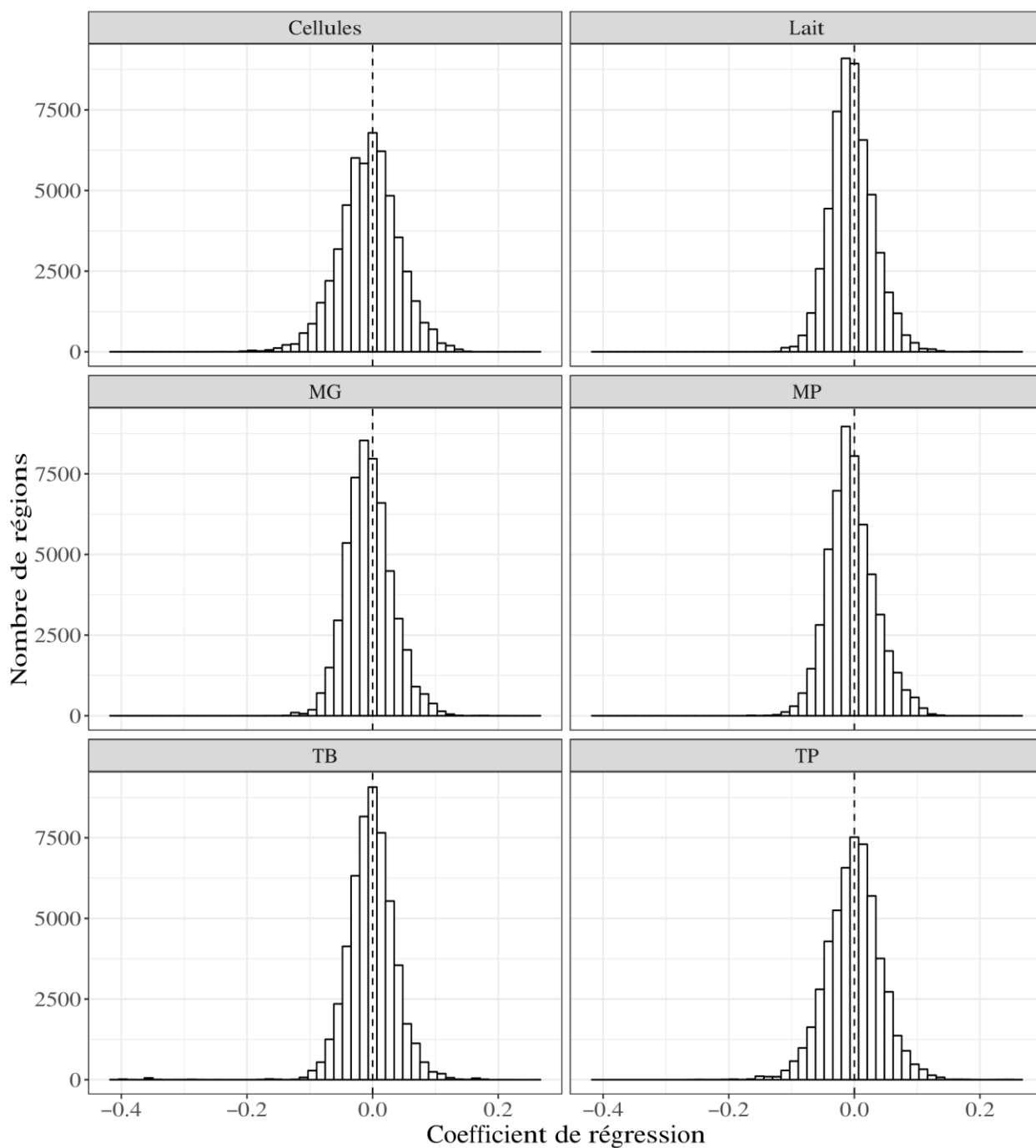


Figure 16 Répartition des coefficients de régression linéaire correspondant aux effets de la consanguinité sur six caractères (sous forme d'indices polygéniques) chez les vaches Montbéliardes

Pour tous les caractères, la répartition des pentes était très hétérogène le long du génome, avec des pentes négatives ou positives selon les régions étudiées. Quelques statistiques descriptives de ces pentes sont présentées dans le Tableau 9.

Tableau 9 Statistiques descriptives des pentes correspondant à l'effet de la consanguinité le long du génome pour six caractères dans la race Montbéliarde

Caractère	Pente minimale	Pente maximale	Moyenne des pentes	Ecart-type des pentes	Nombre de régions avec pente négative	Nombre de régions avec pente positive
Cellules	-0.247	0.153	-0.00595	0.0469	28949	24114
Lait	-0.119	0.208	-0.00335	0.0348	30462	22601
MG	-0.159	0.206	-0.00512	0.0363	30738	22325
MP	-0.171	0.137	-0.00379	0.0379	30788	22275
TB	-0.412	0.256	-0.00214	0.0390	27796	25267
TP	-0.258	0.260	-0.000941	0.0428	26457	26606

2 - Liens entre MAF et ROH

En complément de l'estimation des effets de la consanguinité le long du génome, nous avons calculé la proportion de vaches possédant un ROH à chaque SNP. C'est sur le chromosome 6, entre la 37 180 233 et la 37 776 464^{ème} paire de base, que le maximum de plus de 98 % de vaches possédant un ROH a été atteint. Nous avons également calculé les fréquences alléliques minoritaires (MAF) pour tous les SNPs dans le jeu de données.

Les corrélations de Pearson entre les pentes estimées précédemment dans chaque région, la proportion de vaches possédant un ROH pour le SNP central de chaque région et la MAF pour ce même SNP sont présentées dans le Tableau 10.

Tableau 10 Corrélations entre les coefficients de régression linéaire (= pentes) correspondant à l'effet de la consanguinité le long du génome sur les performances, la proportion de ROH dans la population et les fréquences alléliques pour six caractères chez les vaches Montbéliardes

Caractère	Pentes – MAF	Pentes – proportion de vaches possédant un ROH	Valeur absolue des pentes – MAF	Valeur absolue des pentes – proportion de vaches possédant un ROH
Cellules	-0,0089 ns	0,050 ***	-0,0220 ***	0,152 ***
Lait	-0,0481 ***	0,220 ***	-0,0154 *	0,089 ***
MG	-0,0497 ***	0,240 ***	-0,0143 *	0,062 ***
MP	-0,0559 ***	0,270 ***	-0,0161 *	0,121 ***
TB	-0,0051 ns	0,044 ***	-0,0366 ***	0,089 ***
TP	-0,0174 **	0,101 ***	-0,0293 ***	0,118 ***

*** p -valeur < 10^{-6} , ** p -valeur < 10^{-4} , * p -valeur < 10^{-3}

Nous avons observé que, de manière générale, ni la MAF ni la proportion de vaches possédant un ROH n'étaient fortement corrélées aux pentes ou à leur valeur absolue.

3 - Scores SIFT

Parmi les pics de pente, nous avons trouvé 34 SNPs pour lesquels une mutation peut résulter en un score SIFT égal à 0. Parmi ces 34, pour 14 SNPs, nous avons observé au moins un caractère pour lequel la pente de la région « pic de pente » correspondant était négative, et pour 20 SNPs, au moins un caractère ayant une pente positive. Nous avons observé 5 SNPs pour lesquels le score SIFT était égal à 0 et qui appartenaient à des régions présentant des pentes fortes (positives et/ou négatives) pour plusieurs caractères. Ces positions sont présentées dans le Tableau 11.

Tableau 11 Caractéristiques des locus présentant un score SIFT égal à 0 ainsi qu'un coefficient de régression linéaire (= pente) de dépression de consanguinité locale fort correspondant à un pic pour au moins deux caractères différents

Chromosome	Position sur le chromosome en paires de bases	Caractère	Pente	Proportion de vaches avec un ROH à chaque position	MAF
4	63320920	MP	0,0625	0,137	0,00526
4	63320920	TB	-0,0746		
5	116630358	Cellules	-0,0928	0,0977	0,453
5	116630358	MG	-0,0851		
7	70901266	Cellules	0,087	0,117	0
7	70901266	MP	-0,0779		
8	98530413	MG	0,0602	0,296	0,0169
8	98530413	TB	0,0697		
19	40982250	MP	0,101	0,216	0,00593
19	40982250	TP	0,0743		
19	40982250	Lait	0,0578		

D - Une région du génome sous sélection forte : le cas de *DGAT1*

Nous nous sommes ensuite intéressés au cas particulier d'un QTL sous forte sélection dans la région du gène *DGAT1*. Le gène acyl CoA diacylglycérol acyltransférase 1 (*DGAT1*) est situé sur le chromosome 14, entre la 603 813 et la 612 791^{ème} paire de base. Les caractéristiques de la population étudiée au locus du gène *DGAT1* sont présentées dans le Tableau 12 ci-dessous.

Les pentes mesurant l'effet de la consanguinité au locus du gène *DGAT1* sur les performances des vaches pour les six caractères étudiés sont présentées dans le Tableau 13 ci-dessous. Dans ce tableau, la position donnée correspond au SNP central de la région du chromosome prise en compte, c'est-à-dire le 8^{ème} SNP de la région.

Tableau 12 Proportion de ROH dans la population et fréquences alléliques observées sur les SNPs inclus dans le gène *DGAT1* (chromosome 14)

Position sur le chromosome en paires de bases	Caractère « cellules »		Caractères laitiers	
	Proportion de vaches avec un ROH à chaque position	MAF	Proportion de vaches avec un ROH à chaque position	MAF
603930	0,264	0,456	0,263	0,457
608230	0,264	0,00475	0,263	0,00478
609153	0,264	0,00443	0,263	0,00444
609870	0,264	0,00533	0,263	0,00537
611019	0,264	0,00573	0,263	0,00581
611020	0,264	0,00503	0,263	0,00506

Tableau 13 Coefficients de régression linéaire (= pentes) correspondant à l'effet de la consanguinité le long du génome sur les performances pour six caractères chez les vaches Montbéliarde pour les SNPs inclus dans le gène *DGAT1* (chromosome 14)

Position sur le chromosome en paires de bases	Pente cellules	Pente lait	Pente MG	Pente MP	Pente TB	Pente TP
603930						
608230						
609153	-0,103	0,118	-0,120	0,0438	-0,355	-0,152
609870						
611019						
611020						

L'effet de la consanguinité au locus du gène *DGAT1* était favorable pour les caractères de production laitière (lait) et MP, et défavorable pour les autres caractères : cellules, MG, TP et en particulier TB, avec une pente de -0,355.

Plus de 26 % des vaches du jeu de données possédaient un ROH au locus du gène *DGAT1*. La MAF des SNPs inclus dans le gène *DGAT1* était très faible, aux alentours de 0,5 %. Le premier SNP inclus dans le gène présentait en revanche une MAF très élevée (plus de 45 %).

IV - Discussion

Nous avons estimé le niveau de dépression de consanguinité globale en race Montbéliarde pour six caractères : un caractère de santé et cinq caractères de production. Pour la production laitière (caractère « lait »), la dépression de consanguinité a fait perdre 1,19 % du progrès génétique annuel en race Montbéliarde dans ce jeu de données. Nous avons vu dans le Chapitre 2 que le progrès génétique annuel des taureaux nés entre 2012 et 2015 était 2,1 fois plus grand en Prim'Holstein qu'en race Montbéliarde et le taux de consanguinité annuel estimé à partir des ROH était environ 3 fois plus grand en race Prim'Holstein qu'en race Montbéliarde (Doublet et al.,

2019). En prenant en compte ces taux plus élevés et en gardant la même valeur de coefficient de régression (impact de la dépression de consanguinité pour une augmentation de 1 % de la consanguinité), la baisse du progrès génétique annuel due à la dépression de consanguinité en Prim'Holstein serait d'environ 2,6 %, soit plus de 2 fois celle observée en race Montbéliarde. Dans une autre étude, la baisse du progrès génétique annuel de la production laitière a été estimée à 4,4 % en utilisant des valeurs de consanguinité estimées à partir des pedigrees (Leroy, 2014). La bonne gestion de la consanguinité en race Montbéliarde semble donc limiter efficacement l'impact de la dépression de consanguinité sur la production laitière.

Pour le caractère « cellules », nous avons estimé la perte de progrès génétique annuel à 4,25 % dans la race Montbéliarde. C'est le caractère pour lequel la perte relative de progrès génétique est la plus forte parmi les six caractères étudiés.

Nous n'avons pas pu calculer la perte annuelle du progrès génétique causée par la dépression de consanguinité pour le TB car le progrès génétique mesuré n'était pas significativement différent de 0 mais la dépression de consanguinité a tout de même été détectée. Sachant que le TB n'a pas un poids fort dans l'ISU en race Montbéliarde (Montbéliarde Association), il n'est pas étonnant que le progrès génétique pour ce caractère ne soit pas détectable. En ce qui concerne le TP, caractère fortement sélectionné en race Montbéliarde (Montbéliarde Association), le coefficient de régression linéaire n'était pas significativement différent de 0, nous n'avons donc pas observé de dépression de consanguinité pour ce caractère, soit parce qu'il n'en existe pas dans la population étudiée, soit parce que son effet est trop faible pour être détecté.

Cette méthode d'analyse de la dépression de consanguinité comme une régression de la valeur de caractères d'intérêt sur la consanguinité globale donne des valeurs de dépression de consanguinité intéressantes pour les éleveurs. Elle permet en effet d'avoir une bonne idée de l'impact global de la consanguinité sur les performances des animaux et si le manque à gagner est important par rapport au progrès génétique.

Si l'on s'intéresse à la localisation du fardeau génétique le long du génome, cette méthode ne suffit pas. Elle suppose que l'autozygotie est répartie uniformément le long du génome. Or, des preuves expérimentales ont montré qu'il s'agit d'une hypothèse irréaliste 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 pour la détection et l'estimation de la dépression de consanguinité, il n'est pas possible de prendre en compte cette hétérogénéité.

Nous avons vu que les valeurs de dépression de consanguinité dépendaient des caractères étudiés et que pour certains d'entre eux comme la production laitière ou la concentration de cellules somatiques dans le lait, l'effet de la consanguinité était important malgré une bonne gestion de la consanguinité en race Montbéliarde. Nous nous sommes donc demandé quelles régions du génome étaient impliquées du fardeau génétique pour les différents caractères.

Nous avons donc estimé l'effet de la consanguinité le long du génome sur ces mêmes caractères pour identifier les régions du génome de la Montbéliarde impliquées dans la dépression de consanguinité.

Pour tous les caractères, la répartition des effets de la consanguinité était très hétérogène le long du génome, avec des pentes négatives ou positives selon les régions étudiées, quel que soit l'effet de la consanguinité estimé précédemment à l'échelle globale. Ainsi, la consanguinité n'a pas le même effet sur les performances pour les six caractères étudiés selon les régions du génome et les régions impliquées dans l'expression de la dépression de consanguinité sont nombreuses (Keller et al., 2012 ; Ayroles et al., 2009). En moyenne, ces effets sont proches de zéro dans cette étude mais la majorité des régions analysées présentent une pente négative.

L'importance du nombre de régions impliquées dans la dépression de consanguinité représente bien le fait que le fardeau génétique soit réparti le long du génome et que son expression dépend de l'expression de nombreux gènes à effets plus ou moins délétères.

Nous avons étudié si les effets fortement délétères de la consanguinité sur les caractères étaient liés au potentiel délétère de mutations dans les régions correspondantes du génome, en utilisant le score SIFT. Le faible nombre de SNPs avec un score SIFT potentiel de 0 n'est pas étonnant. D'une part, les prédictions SIFT ne sont pas disponibles pour la totalité des SNPs étudiés (seulement 3 % des SNPs appartenant aux régions « pics de pente »). D'autre part, une mutation délétère pour la fonction d'une protéine ne va pas forcément avoir un effet délétère pour un caractère. Enfin, le score SIFT évalue l'effet d'une substitution d'acide aminé au sein d'une protéine, or, une mutation au sein d'une région régulatrice du génome ne codant pas pour une protéine peut également avoir un impact sur les performances des animaux pour un caractère. Le score SIFT ne nous a ici pas aidés à mieux comprendre les mécanismes expliquant les effets de la consanguinité sur le caractère cellules et les caractères laitiers.

Nous avons voulu comprendre comment se comporte la dépression de consanguinité à proximité des régions du génome fortement sélectionnées car ayant un fort effet sur les caractères de production, comme le QTL de production laitière situé dans la région du gène *DGAT1* (Gautier et al., 2007 ; Kühn et al., 2004 ; Grisart et al., 2002 ; Cases et al., 1998).

Nous avons calculé les corrélations entre la proportion de vaches possédant un ROH à chaque position du génome et la pente observée pour chaque caractère. En effet, une forte proportion d'individus possédant un ROH dans une région du génome constitue une signature de sélection (Saravanan et al., 2020). Nous pouvons observer que de manière générale, la proportion de vaches possédant un ROH à chaque position du génome n'était pas généralement très fortement corrélée aux pentes. Les corrélations les plus élevées étaient de 0,22 pour le lait, 0,24 pour la MG et 0,27 pour la MP. Cela suggère que plus la proportion de ROH est élevée dans une région du génome, c'est-à-dire plus la région génomique est sous sélection, plus l'effet de la consanguinité dans cette région sera bénéfique pour les caractères mentionnés.

Le gène acyl CoA diacylglycérol acyltransférase 1 (*DGAT1*) est situé sur le chromosome 14, entre la 603 813 et la 612 791^{ème} paire de base. Il code pour une enzyme microsomale qui catalyse l'étape finale de la biosynthèse des triglycérides (Cases et al., 1998). Cette enzyme joue un rôle primordial dans l'absorption des graisses intestinales, l'assemblage des lipoprotéines, la formation du tissu adipeux et la lactation (Cases et al., 1998). Un QTL (*Quantitative Trait Locus*) de production laitière a été identifié dans la région du gène *DGAT1*, plus particulièrement pour le TB (Gautier et al., 2007 ; Kühn et al., 2004 ; Grisart et al., 2002).

Plus de 26 % des vaches du jeu de données possèdent un ROH au locus du gène *DGAT1*, c'est plus que 80 % des autres positions du génome, suggérant une forte sélection à ce locus. La MAF des SNPs inclus dans le gène *DGAT1* est très faible, aux alentours de 0,5 %. Chez la Montbéliarde, l'allèle A du gène *DGAT1* est presque complètement fixé (Gautier et al., 2007), ce qui est cohérent avec notre observation. Le premier SNP inclus dans le gène présente en revanche une MAF très élevée (plus de 45 %). Cette position code pour région 5' non traduite (5'-UTR) de l'ARNm de *DGAT1* (McLaren et al., 2016). Ce type de région est impliqué dans la régulation de l'expression du gène, selon sa longueur par exemple. En cas d'absence de cette région, la fixation des ribosomes à l'ARNm sera impossible, inhibant ainsi sa traduction (Pickering et Willis, 2005). Etant données les MAF observées, c'est probablement la consanguinité pour ce SNP qui joue un rôle dans l'expression de la dépression de consanguinité, en particulier pour le TB. Il est probable que la consanguinité ait un impact sur l'expression du gène *DGAT1* via la région 5'-UTR de son ARNm.

On peut observer des pentes significativement positives pour les six caractères étudiés, ce qui suggère que dans certaines régions du génome, la présence de ROH semble favorable aux performances (Keller et al., 2012). Nous pouvons émettre plusieurs hypothèses pour expliquer la présence de pentes positives :

- (i) Il s'agit de zones de déficit en homozygotes pour un haplotype létal ou purgées. Dans ce cas, l'un des deux homozygotes est peu voire pas présent dans la population car ce génotype n'est pas viable, tandis que l'autre l'est ;
- (ii) Il s'agit de régions du génome sous forte sélection (Kim et al., 2013). Dans ce cas, si la MAF dans ces régions est très faible, il n'existera pas d'homozygote de l'allèle rare pour des raisons de probabilités ;
- (iii) L'haplotype délétère est dominant ou partiellement dominant sur l'haplotype sain (voir Figure 2 page 28).

Ce dernier cas concerne probablement des allèles délétères à effet modéré ou faible. En effet, si l'effet était létal ou très fort, il y a peu de chances qu'un allèle délétère n'ait pas été éliminé rapidement après son apparition dans la population, par sélection négative.

Les régions génomiques pour lesquelles l'effet de la consanguinité sur les performances semble bénéfique peuvent donc correspondre à des régions portant des variants à effet délétère faible ou modéré sur les performances. Dans ces cas-là, maintenir la diversité dans ces régions tendrait à donner des valeurs de performances inférieures à la moyenne des deux types d'homozygotes, ce qui n'est pas favorable pour le progrès génétique. Il est donc nécessaire de ne pas mettre de côté les régions avec des pentes positives dans la gestion du fardeau génétique mais de trouver des façons de les gérer différemment des régions avec des pentes négatives. Dans certains cas, une région génomique présentera une pente positive pour un caractère et négative pour un autre. La gestion de ces cas n'est pas triviale et devra faire l'objet d'un travail approfondi.

S'il existait des cas de superdominance, les pentes auraient tendance à être négatives. Or, il n'est pas intéressant d'éliminer complètement les allèles délétères de la population dans les cas de superdominance si l'on veut conserver l'avantage des hétérozygotes. Ainsi, comme nous l'avons déjà suggéré, il semble bénéfique de chercher à maximiser la diversité dans ces régions à l'échelle de la population afin d'augmenter la proportion d'hétérozygotes et donc de l'association haplotypique la plus favorable.

On notera que, notre jeu de données étant constitué uniquement de vaches en production (ayant effectué au moins une lactation), nous risquons fortement de ne pas détecter les régions génomiques portant des allèles délétères à effet très fort car les animaux exprimant ces allèles exprimeraient une forte dépression de consanguinité, souvent précoce, qui les écarteraient des troupeaux. Les régions génomiques portant des allèles délétères à effet très fort (mort précoce, affections empêchant la croissance ou la production), s'apparentant à des anomalies génétiques, seraient plutôt détectées par des approches de détection des zones de déficit en homozygotes (Pausch et al., 2015 ; Fritz et al., 2013 ; VanRaden et al., 2011). La méthode des zones de déficit en

homozygotes et celle utilisée ici sont complémentaires pour repérer les régions du génome impliquées dans l'expression de la dépression de consanguinité.

Avec la méthode de détection des régions impliquées dans l'expression de la dépression de consanguinité utilisée ici, nous ne prenons pas en compte quels haplotypes sont responsables de l'effet de la consanguinité sur les performances. Cela ne pose pas de problème ici car l'objectif de cette étude n'est pas de détecter les variants alléliques délétères mais de savoir dans quelles régions du génome il est nécessaire de maintenir la diversité génétique pour éviter l'expression du fardeau génétique (Gómez-Romano et al., 2016).

Se focaliser sur les régions présentant les pentes les plus fortes seulement limiterait le risque de faux positifs. Le risque de faux négatifs (pente non significativement différente de 0) existe dans les cas d'additivité stricte, lorsque la moyenne des performances des homozygotes pour les deux haplotypes est égale aux performances des hétérozygotes.

On peut penser qu'il serait préférable de purger complètement les allèles délétères de la population. Or, la purge volontaire de nombreux allèles impliqués dans le fardeau (d'autant plus que le nombre de caractères pris en compte augmente) par une sélection drastique des individus non porteurs de ces allèles serait très dangereuse pour la diversité génétique de la population, en particulier lorsque la taille efficace de cette dernière est faible. De plus, l'édition du génome pour éliminer ces allèles (Johnsson et al., 2019) est une technique encore coûteuse et difficile à mettre en place en routine, surtout pour un grand nombre d'allèles et d'animaux. Enfin, les méthodes permettant de détecter les haplotypes impliqués dans le fardeau génétique sont particulièrement coûteuses en temps de calcul et en mémoire en comparaison avec la méthode utilisée ici, ce qui rend notre méthode d'autant plus avantageuse pour l'objectif poursuivi.

Des méthodes basées sur la détection des effets de dominance et leur association avec la présence de ROH ne donnent pas de résultats plus probants que notre méthode (Doekes, 2020), suggérant que notre méthode est pertinente dans l'étude de la dépression de consanguinité le long du génome.

En évitant la consanguinité dans ces régions, on évite la mise à l'état homozygote des allèles délétères récessifs et donc l'expression de la dépression de consanguinité (Gómez-Romano et al., 2016). Cela est possible soit en sélectionnant les reproducteurs pour être les moins apparentés possible entre eux pour ces régions, soit en concevant des plans d'accouplement prenant en compte l'apparentement génomique avec un poids plus important pour ces régions.

Pour chaque caractère, les régions dans lesquelles la consanguinité a un effet sur les performances sont nombreuses. Avec la multiplication des caractères à prendre en compte, il serait probable

que quasiment tout le génome soit couvert de régions impliquées dans l'expression de la dépression de consanguinité. Ainsi, il serait pertinent de limiter les régions retenues à celles avec les effets les plus forts. Cela permettrait également de ne pas augmenter trop fortement les contraintes de sélection des reproducteurs ou de plans d'accouplements, pour éviter d'être trop préjudiciable au progrès génétique.

V - Conclusion

Nous avons vu que la dépression de consanguinité s'exprime chez la Montbéliarde pour au moins cinq caractères de production et de santé sur six caractères étudiés. Selon les caractères, la dépression de consanguinité a fait perdre jusqu'à plus de 4 % du progrès génétique chez les vaches de race Montbéliarde nées entre 2012 et 2017. En estimant la consanguinité à l'échelle globale du génome pour la détection et l'estimation de la dépression de consanguinité, il n'est pas possible de prendre en compte l'hétérogénéité de la répartition de la consanguinité le long du génome.

Grâce aux ROH, nous avons estimé les effets de la consanguinité sur les performances des vaches de race Montbéliarde nées entre 2012 et 2017 pour six caractères de santé et de production laitière le long du génome. Nous avons vu que ces effets étaient très hétérogènes le long du génome, avec des régions pour lesquelles la consanguinité est favorable pour un caractère et d'autres pour lesquelles elle est défavorable.

La méthode d'estimation des effets de la consanguinité le long du génome présentée ici permettrait de définir simplement des régions du génome au sein desquelles le maintien de la diversité génétique est indispensable, par exemple en pondérant les apparentements entre individus pour donner plus d'importance à ces régions. Dans cette optique, il paraît indispensable d'évaluer l'efficacité de la prise en compte des régions impliquées dans la dépression de consanguinité dans les schémas de sélection et les plans d'accouplements, afin de s'assurer qu'elle limite la perte de progrès génétique due à la dépression de consanguinité. Notons que la détection des régions impliquées dans l'expression de la dépression de consanguinité devrait être renouvelée régulièrement et spécifiquement à chaque population ou race, car l'architecture génétique des caractères ainsi que du fardeau génétique dépend de la structure génétique de chaque race et évolue dans le temps.

Bien que l'identification des régions qui contribuent fortement à la dépression de consanguinité soit très importante pour la conservation et la gestion de la dépression de consanguinité dans la population étudiée, elle ne donne pas un aperçu complet de l'architecture génétique de la dépression de consanguinité. Pour une compréhension complète du mécanisme de la dépression de consanguinité, il serait nécessaire de connaître le nombre de locus impliqués, leur mode de

transmission et l'ampleur des effets estimés. Ainsi, le choix des stratégies de détection des régions du génome impliquées dans le fardeau génétique doit s'effectuer en fonction des objectifs poursuivis : gérer les régions impliquées dans le fardeau dans le cadre de la sélection dans les espèces de rente, ou augmenter la compréhension du mécanisme de la dépression de consanguinité.

Chapitre 5 Discussion générale et perspectives

I - Diversité génétique et schémas de sélection dans les races locales et/ou de petit effectif

Les principales races bovines laitières françaises de petit effectif sont l'Abondance (24 184 lactations enregistrées en 2019), la Brune (race de faible effectif en France, il s'agit toutefois d'une race internationale) (16 402 lactations enregistrées en 2019), la Tarentaise (ou Tarine) (8 602 lactations enregistrées en 2019) et la Vosgienne (1 413 lactations enregistrées en 2019) (Institut de l'Élevage, 2020b). Leurs effectifs efficaces pour les animaux nés entre 2016 et 2019 ont été estimés entre 54 et 96 selon les races (Danchin-Burge, 2020). Ils sont donc du même ordre de grandeur que ceux des races de grand effectif, malgré des effectifs réels très différents.

Dans les races de petit effectif, la diversité génétique est en danger à cause de la dérive génétique, couplée à la sélection, et d'autant plus importante que la taille des populations est faible. De plus, en cas de baisse trop importante de la diversité génétique dans ces races, la marge de manœuvre pour limiter cette perte est faible à cause de leurs petits effectifs réels et du faible nombre de taureaux disponibles. Ainsi, les efforts pour maintenir la diversité génétique dans ces races, la plupart du temps locales, sont généralement très importants et c'est l'une des préoccupations majeures des organismes et entreprises de sélection responsables de la sélection des taureaux au sein de ces races. Cela peut aller jusqu'à un relâchement de la pression de sélection dans les races en conservation présentant un intérêt patrimonial et culturel.

L'impact de la mise en œuvre récente, aux alentours de 2016, d'évaluations génomiques dans les races locales et/ou de petit effectif devrait être étudié dans les années à venir afin de garantir la durabilité des programmes de sélection dans ces races, en termes de progrès et de diversité génétique. Ce travail n'a pas été possible dans le cadre de ma thèse car en 2018, première année de ma thèse, la sélection génomique était encore trop récente dans les races de petit effectif et nous ne disposions donc pas d'un recul suffisant pour y effectuer une étude rétrospective. D'ici cinq ans (horizon 2025), il serait intéressant de réaliser ce travail pour ces races, afin de permettre aux entreprises de sélection d'avoir des retours sur les conséquences des modifications des schémas de sélection décidées lors de la mise en place de la sélection génomique.

Cependant, comme la marge de manœuvre est réduite dans ces races, il ne faudra pas attendre 2025 pour tester de nouvelles solutions de gestion de la diversité génétique. En particulier, d'autres modifications des schémas de sélection, l'utilisation du croisement ou de taureaux dont la semence est conservée en cryobanque permettraient de maintenir ou de réinjecter de la diversité génétique dans ces races. Les simulations semblent être une bonne façon d'anticiper

l'efficacité de ces différentes solutions pour maintenir voire augmenter la diversité et le progrès génétique dans les races locales et/ou de petit effectif.

Dans cette optique, des travaux de simulations ont débuté au sein d'Alice, en partenariat étroit avec les entreprises de sélection et dans le sillon de cette thèse, afin d'évaluer l'impact, par exemple, de l'utilisation ou non de taureaux confirmés, du nombre de jeunes taureaux nouvellement diffusés et de la durée de leur diffusion, sur le progrès et la diversité génétique. Ces travaux sont issus de nombreux échanges avec les entreprises de sélection, ayant permis d'identifier les problématiques spécifiques aux races locales, et sont amenés à devenir une prestation quasi de routine pour les entreprises de sélection souhaitant tester de nouveaux schémas de sélection.

II - Les leviers d'action de la gestion de la diversité à l'ère de la génomique : efficacité et faisabilité sur le terrain

La disponibilité croissante de données de génotypage de plus en plus dense, en particulier dans les races bovines laitières, est une opportunité formidable. En effet, les données de génotypage permettent une meilleure caractérisation et donc une meilleure gestion de la diversité génétique.

La sélection génomique est un outil permettant l'augmentation du progrès génétique sans nécessairement mettre en danger la diversité génétique des races bovines laitières françaises. Les évaluations génomiques permettent d'augmenter le nombre de candidats à la sélection à coût constant, voire le nombre de taureaux diffusés avec un renouvellement plus rapide du catalogue des taureaux diffusés. Cela devrait permettre de tirer parti des avantages des évaluations génomiques sur le plan du progrès génétique tout en ménageant la diversité génétique. En revanche, si le nombre de taureaux ou de candidats à la sélection n'augmente pas voire diminue, ou si les taureaux sélectionnés sont plus apparentés qu'auparavant, l'effet d'accélération de la perte de diversité génétique dû au raccourcissement important des intervalles de génération n'est plus compensé et la perte de diversité génétique s'accélère, comme nous l'avons observé dans la race Prim'Holstein.

Les conséquences sur la diversité génétique sont donc très différentes selon la manière d'utiliser l'évaluation génomique ou les technologies de reproduction comme le transfert embryonnaire. La conception et les modifications des schémas de sélection, tout comme la manière de diffuser les taureaux sélectionnés, jouent un rôle majeur dans la gestion de la diversité génétique. L'accroissement du taux de consanguinité annuel ou par génération n'est donc pas inéluctable et il est possible d'augmenter le progrès génétique tout en étant attentif à la diversité génétique.

Dans le cadre de cette thèse, j'ai profité d'un partenariat fort avec plusieurs entreprises de sélection, ce qui m'a permis de mener à bien une grande partie de mes travaux (en particulier les

simulations des schémas de sélection) mais aussi de me rendre compte du réel intérêt que les entreprises de sélection des races bovines laitières françaises, locales ou de grand effectif, portent au maintien de la diversité génétique dans les races qu'elles sélectionnent.

Ce lien fort entre la recherche et la filière bovins lait est à la fois enrichissant et indispensable, tant pour les chercheurs que pour les acteurs de la filière bovins lait. Il est en effet important pour les chercheurs de connaître les enjeux des professionnels de la filière, dans le cadre de la recherche appliquée, pour proposer des solutions de gestion de la diversité génétique ancrées dans le réel et qui auront une chance d'être appliquées efficacement, comme dans le cadre de la recherche fondamentale, pour organiser la collecte de données ou mettre en place des programmes de recherche en partenariat avec les éleveurs, les entreprises de sélection et les chercheurs.

Permettre aux entreprises de sélection de prendre des décisions en toute connaissance de cause est indispensable au vu de la responsabilité qu'elles ont dans le maintien de la diversité génétique mais aussi du progrès génétique dans les races qu'elles sélectionnent. Ainsi, les études a posteriori sur l'impact de changements de schémas de sélection participent à éclairer les prises de décision des entreprises de sélection, bien que leur aspect rétrospectif n'autorise que des remises en question tardives. C'est pourquoi il est indispensable d'associer ces études à des études prospectives, par simulations, comme cela avait été fait pour évaluer l'effet de différentes stratégies d'utilisation de la sélection génomique sur le progrès et la diversité génétique en race Montbéliarde (Colleau et al., 2015, 2009), et comme nous l'avons fait dans le cadre de l'utilisation du transfert embryonnaire.

En plus de la maîtrise des schémas de sélection, d'autres leviers d'action existent pour gérer la diversité génétique : les plans d'accouplement (Bérodier, 2020), les méthodes d'évaluation des animaux et la gestion du fardeau génétique. A mi-chemin entre les plans d'accouplement et la maîtrise des schémas de sélection, on peut inclure dans ces leviers la gestion du nombre de doses diffusées par taureau et la proportion maximale d'inséminations réalisées à partir de la semence de chaque taureau.

A - Les plans d'accouplement

Il est possible d'optimiser les plans d'accouplement pour limiter la perte de diversité génétique tout en maximisant le progrès génétique. L'utilisation des apparentements génomiques et d'optimisation linéaire (ou programmation linéaire) semble être l'une des meilleures manières d'optimiser cet équilibre entre progrès et diversité (Bérodier, 2020).

Si les plans d'accouplement au sein des entreprises de sélection peuvent être relativement facilement respectés, ce n'est pas forcément le cas dans les élevages, et ce pour deux raisons principales :

(i) **La difficulté à respecter scrupuleusement les plans d'accouplement sur le terrain**, pour des raisons de disponibilité des doses de semence de chaque taureau au moment de l'insémination (pour des raisons de stock global ou de stock dans la cuve de l'inséminateur qui s'est déplacé dans la ferme). De plus, il est souvent nécessaire de réaliser deux à trois inséminations, à trois semaines d'intervalle minimum, avant que la vache ne soit gestante. En effet, le taux de réussite de la première insémination chez les vaches laitières en France est de 51 % pour les vaches et 57 % pour les génisses (Institut de l'Élevage, 2020c), et est encore plus bas lors de l'utilisation de semence sexée. Les éleveurs sont donc généralement amenés à choisir des taureaux aux doses plus chères pour les premières inséminations et moins chères pour les suivantes. Dans ce cas, il arrive fréquemment que les vaches soient inséminées avec une race bien conformée pour valoriser les veaux à l'abattoir ou à l'engraissement, au lieu d'obtenir des génisses de renouvellement. Ainsi, il n'est pas garanti que le veau obtenu soit celui prévu par le plan d'accouplement, même si l'éleveur était décidé à le respecter au départ.

(ii) **L'indépendance des éleveurs**. En effet, les éleveurs sont responsables de leur renouvellement et ne sont pas obligés de suivre les recommandations des entreprises de sélection fournissant des doses. Chaque éleveur souhaitera généralement choisir les meilleurs taureaux au catalogue. Ne pas l'y autoriser, en limitant drastiquement le nombre de doses disponibles pour un taureau ou en imposant des inséminations avec des doses précises, reviendrait à brider cette liberté, voire à créer des inégalités de traitement entre plusieurs élevages. Un élevage pourrait en effet obtenir un très bon progrès génétique au détriment de l'élevage voisin, qui obtiendrait un progrès génétique inférieur à son potentiel, dans le cadre de plans d'accouplement optimisant le progrès et la diversité génétique à l'échelle de la race.

Dans le but de proposer des plans d'accouplement efficaces dans la gestion de la diversité génétique, il est donc nécessaire de prendre en compte l'égalité de traitement entre les élevages mais également la robustesse des plans d'accouplement, afin d'assurer leur efficacité en cas de déviation plus ou moins importante dans leur application.

La mise en place de limites de diffusion de chaque taureau nécessite une bonne collaboration entre les éleveurs et les entreprises de sélection, afin que tous les acteurs de la filière se sentent impliqués dans la gestion de la diversité génétique dans les races bovines laitières françaises, en en comprenant les enjeux.

B - Les méthodes d'évaluation des animaux

Ensuite, le choix des animaux sélectionnés est crucial dans la gestion de la diversité génétique au sein des races. En effet, si les reproducteurs sélectionnés sont très apparentés, leurs descendants présenteront une consanguinité d'autant plus élevée. Dès la fin des années 1990, des méthodes pour maximiser la réponse à la sélection tout en évitant l'accélération de la baisse de la diversité génétique ont été proposées et l'une des méthodes les plus connues est celle des contributions optimales (OC) (Meuwissen, 1997). Cette dernière consiste à maximiser le progrès génétique en limitant l'apparentement entre les reproducteurs sélectionnés pour maintenir le taux de consanguinité à un niveau prédéterminé, via un choix judicieux des reproducteurs mais également par le calcul des contributions de chacun d'entre eux (c'est-à-dire le nombre de descendants que devra avoir chaque reproducteur). La sélection des parents est donc effectuée sur des critères de valeur génétique mais aussi d'apparentement avec les autres reproducteurs. Cet apparentement peut être estimé à partir des pedigrees mais aussi des génotypes.

En cas d'utilisation des génotypes, l'inclusion dans le calcul des allèles rares peut modifier la liste des animaux sélectionnés sur la base de leur valeur génétique et de leur apparentement aux autres animaux sélectionnés (Eynard et al., 2015). L'utilisation de différents types de matrices d'apparentement génomiques peut aussi avoir un impact sur la variabilité génétique de la population (Wang et al., 2020). De plus, l'inclusion d'allèles rares permet de conserver plus de diversité génétique (Eynard et al., 2015). Dans le cadre de la sélection génomique, nous avons vu que les effets des allèles aux différents locus génotypés est estimé à partir d'une population de référence pour laquelle on dispose des génotypes mais aussi des performances. Il a été montré que des changements dans la constitution de la population de référence pouvait avoir un impact sur la liste des animaux sélectionnés (Eynard et al., 2018a). Si la population de référence est mise à jour au fil du temps grâce à des approches de type OC, il est possible de mieux optimiser le progrès génétique en limitant la perte de diversité génétique (Eynard et al., 2018a).

Il est également possible de prendre en compte l'originalité des animaux par rapport au reste de la population afin de pondérer leur valeur génétique. C'est déjà fait dans plusieurs races, en France et aux USA par exemple, avec des notes d'originalité pour les taureaux à destination des éleveurs ou pour les pères et mères à taureaux dans les schémas des entreprises de sélection. Dans ce cas, il serait intéressant de comparer plusieurs méthodes déjà existantes de calcul de l'originalité, et trouver la meilleure pondération à effectuer afin d'optimiser l'efficacité de ces valeurs génétiques prenant en compte l'originalité pour gérer le progrès et la diversité génétique.

De plus, de nouvelles méthodes d'évaluation font leur apparition, comme le *single-step*, qui consiste à évaluer en même temps et en une seule étape les animaux génotypés et non génotypés,

en se basant sur une matrice d'apparentement à la fois généalogique et génomique. L'avènement en cours de l'utilisation de la séquence complète au lieu du génotypage de plus ou moins haute densité pourrait également modifier fortement la manière d'évaluer les animaux, grâce à une connaissance plus fine du génome et de l'architecture génétique des caractères et du fardeau génétique. Ces nouvelles méthodes d'évaluation permettront de mieux prendre en compte l'apparentement entre les animaux y compris pour les allèles rares et pourront être plus précises que les précédentes, ce qui pourrait mener à un progrès génétique plus rapide, et peut-être à la possibilité de mieux gérer la diversité génétique (Thomasen et al., 2016).

Il est donc possible de modifier la manière d'évaluer et de sélectionner les animaux reproducteurs pour maintenir la baisse de diversité sous un seuil choisi selon le contexte de la race (conservation, race locale, maximisation du progrès génétique).

C - Le croisement et le matériel génétique issu de cryobanques

En plus de gérer la diversité au sein d'une race et seulement avec des animaux contemporains, il est possible de réinjecter de la diversité génétique dans une race à effectif génétique réduit, grâce au croisement ou à l'utilisation de matériel génétique conservé en cryobanque.

Deux types de matériel biologique sont concernés par la cryoconservation en cryobanques : le sperme et les embryons. Sur le long terme, il est possible d'utiliser la semence congelée d'un taureau ayant vécu plusieurs générations avant la population actuelle. Si le mâle est assez original génétiquement par rapport à la population actuelle, cela permet de réintroduire de la diversité génétique (Doekes et al., 2018; Eynard et al., 2018b). Cependant, cela présente un risque pour le progrès génétique si l'écart génétique est trop important. Le *single-step* pourrait être une méthode d'évaluation pertinente pour faciliter l'évaluation des animaux conservés en cryobanque en cas de difficulté pour les génotyper (pas assez de matériel génétique pour cela, que ce soit du sang ou du sperme). Il en va de même pour l'utilisation d'embryons cryoconservés. La conservation de semence et d'embryons permettent donc une conservation de gènes sur le long terme, comme une assurance en cas de disparition des gènes. Les outils génomiques ont permis d'aider à mieux prioriser les animaux dont le matériel génétique est le plus pertinent à cryoconserver.

Pour réintroduire de la diversité génétique dans une race, il est possible d'introgesser des allèles différents de ceux présents dans la race en la croisant avec une autre race ou en y introduisant des animaux issus de populations génétiquement distantes. Le risque lors de ce type d'introgession est la réduction du contenu génomique natif. Cela pose problème en particulier dans les cas où c'est la part de génome natif qui permet l'adaptation de la race à son milieu. La mesure dans laquelle l'évolution adaptative des allèles immigrants conduit à un remplacement complet ou

partiel du génome natif dépend des taux de recombinaison, de la taille du génome, de la taille efficace des populations et de la répartition des effets d'adaptation (Harris et al., 2019).

Dans les espèces de rente, la notion de race est plus difficile qu'il n'y paraît à délimiter. En général, une race est une race si suffisamment de gens considèrent que c'en est une. La FAO définit entre autres une race comme étant « un groupe d'animaux d'une espèce domestique possédant des caractéristiques externes identifiables et définissables qui permettent de les séparer visuellement des autres groupes de la même espèce définis de la même manière ». Ainsi, il semblerait que le croisement, s'il aboutit après quelques générations à des animaux identifiables à la race de départ, pourrait être accepté par la filière en général. Cependant, le remplacement d'une partie du génome d'une race par celui d'une autre race pose actuellement des difficultés dans l'évaluation génétique ou génomique des animaux croisés au-delà de la première génération de croisement.

De nombreux projets de recherche s'intéressent aux limites de l'utilisation du croisement et de matériel cryoconservé dans les races bovines laitières. L'efficacité dans la gestion de la diversité génétique de chacune de ces mesures doit être étudiée attentivement, dans le respect des contraintes de progrès génétique.

D - Quelles combinaisons de leviers actionner pour gérer au mieux la diversité génétique ?

Les leviers d'action pour mieux gérer la diversité génétique à l'ère de la génomique sont nombreux. Il est cependant difficile d'estimer leur efficacité relative, ainsi que leur acceptabilité pour les éleveurs comme pour les entreprises de sélection. Dans cette optique, il pourrait être intéressant d'étudier la perception de l'importance de la diversité génétique et de son aspect économique, ainsi que des différentes mesures de sa gestion chez les éleveurs et les entreprises de sélection. Le but serait d'identifier les freins techniques, économiques et sociétaux à une prise en compte plus forte de la diversité génétique dans le monde de l'élevage.

Pour les éleveurs et pour les entreprises de sélection, il est difficile de trouver le bon compromis entre progrès et diversité génétique. En effet, et bien que la diversité génétique soit indispensable au maintien du progrès génétique sur le long terme, il est difficile de sacrifier trop de progrès génétique au profit de la diversité génétique, à la fois pour rester :

- Compétitif : en diffusant des taureaux du même niveau que les concurrents nationaux ou internationaux pour les entreprises de sélection,
- Et rentable : en élevant des vaches qui produisent suffisamment de lait et/ou de qualité suffisante en restant en bonne santé pour que les éleveurs puissent vivre de leur activité, ce qui n'est pas si souvent le cas en France.

La problématique de la concurrence est particulièrement prégnante en race Prim'Holstein car il s'agit d'une race internationale, et les entreprises de sélection françaises sont en compétition avec de nombreuses entreprises de sélection à l'international. C'est probablement l'une des raisons pour lesquelles la race Prim'Holstein est celle qui présente le taux de consanguinité le plus élevé actuellement parmi les trois races bovines laitières françaises de grand effectif, au profit d'un progrès génétique élevé (Doublet et al., 2019).

Des approches d'aide à la décision multicritère pourraient être testées pour choisir quelles mesures mettre en œuvre pour atteindre les niveaux de progrès génétique et de maintien de la diversité génétique souhaités par les acteurs de la filière. Les outils d'aide à la décision multicritère, comme PROMETHEE (*preference ranking organisation method for enrichment evaluations*) (Brans et Vincke, 1985), sont généralement utilisés dans le domaine de la recherche opérationnelle et consistent à comparer plusieurs scénarios (dans notre cas : des combinaisons différentes entre toutes les méthodes de gestion de la diversité génétique) et de les classer au regard de différents critères de décision (qui pourraient être ici : le taux de consanguinité, le progrès génétique, le coût des scénarios, leur acceptabilité pour les éleveurs et les consommateurs, etc.). La hiérarchisation des différents scénarios de gestion de la diversité génétique se ferait en les comparant deux à deux selon l'ensemble de ces critères, avec pour chaque critère, un seuil d'indifférence et un seuil de préférence.

Prenons l'exemple du taux de consanguinité annuel. Il est souhaitable que celui-ci soit le plus bas possible. On pourrait définir un seuil d'indifférence de 0,05 % d'écart entre deux scénarios et un seuil de préférence de 0,15 % (ces valeurs sont données uniquement à titre d'exemple et ne correspondent pas à des seuils conseillés). Concrètement, si le scénario n°1 donne un taux de consanguinité annuel de 0,5 % et le scénario n°2 donne un taux de consanguinité annuel de 0,53 %, l'écart entre les deux sera de 0,03 %, valeur inférieure au seuil d'indifférence. Les deux scénarios seront donc considérés comme équivalents selon le critère du taux de consanguinité (le degré de préférence pour le scénario n°1 sera de 0). Si l'écart entre les deux scénarios est supérieur au seuil de préférence, par exemple un taux de consanguinité annuel du scénario n°2 de 0,68 % (écart de 0,18 %), on considérera que le scénario n°1 est absolument préférable à l'autre scénario selon le critère du taux de consanguinité (le degré de préférence pour le scénario n°1 sera égale à 1). Pour un écart compris entre le seuil d'indifférence et le seuil de préférence, le degré de préférence pour le scénario n°1 sera compris entre 0 et 1 et proportionnel à l'écart entre les scénarios.

Des seuils doivent être définis pour tous les critères. La définition des seuils devrait être réalisée suite à la consultation et à la discussion entre les acteurs de la filière et la communauté scientifique. Après prise en compte de tous les critères et les calculs des degrés de préférence de

tous les scénarios deux à deux, on obtient un classement des scénarios, soit complètement hiérarchisé, soit avec des ex aequo, selon la méthode utilisée. La robustesse du classement peut être testée en faisant varier plus ou moins fortement les seuils d'indifférence et de préférence pour tous les critères pour voir si le classement des scénarios change.

Ce type d'approche permettrait d'améliorer la prise de décision pour mieux gérer la diversité génétique, en essayant de trouver le meilleur compromis possible entre progrès, diversité, rentabilité et acceptabilité, et d'optimiser l'utilisation des méthodes de gestion de la diversité génétique. Afin de mener ce type d'étude à bien, des collaborations sont nécessaires entre les généticiens, qui travaillent à l'échelle individuelle ou de la population pour les animaux, et des chercheurs en sciences humaines et sociales ou en sciences de gestion, qui travaillent à l'échelle du système tout entier.

III - Comment gérer la dépression de consanguinité ?

Durant ma thèse, j'ai étudié la dépression de consanguinité chez les vaches de race Montbéliarde nées entre 2012 et 2017. Les effets de la consanguinité sur les performances le long du génome que nous avons estimés ne sont a priori valables que pour cette race et la période actuelle. En effet, dans une autre race ou à une autre époque, avec des objectifs de sélection différents, la structure génomique des caractères et donc de la dépression de consanguinité peut changer fortement. Si l'on peut sans trop de risque extrapoler des résultats tels que « la consanguinité a des effets très hétérogènes sur les performances le long du génome » à d'autres populations, on ne peut pas extrapoler les valeurs de ces effets.

Au-delà des apports scientifiques, ce type d'étude est pertinent dans le cadre de la sensibilisation des professionnels de la filière bovins lait à propos de l'importance du maintien de la diversité génétique. En effet, quantifier la portion de progrès génétique perdu chaque année à cause de la dépression de consanguinité permet de mieux visualiser l'impact réel de l'augmentation du taux de consanguinité sur les performances des animaux et donc des élevages. Cet aspect est loin d'être négligeable car des professionnels bien informés et sensibilisés sont plus à même de prendre des décisions éclairées sur la gestion de la diversité génétique dans les schémas de sélection, les plans d'accouplement, ou le choix des animaux sélectionnés, dans les fermes comme dans les entreprises de sélection.

Quand la consanguinité augmente, la dépression de consanguinité s'exprime. Toutefois, comme la consanguinité augmente, la fréquence de la mise à l'état homozygote des allèles récessifs délétères constituant le fardeau génétique augmente. On peut donc s'attendre à ce que la sélection contre ces allèles délétères soit également accrue quand la consanguinité augmente, ce qui donne

l'occasion à une population d'être "purgée" de son fardeau génétique (Barrett et Charlesworth, 1991; Lande et Schemske, 1985). C'est ce qu'on appelle la purge du fardeau génétique.

À la suite d'une purge du fardeau génétique, la valeur sélective d'une population peut revenir au niveau de celle d'une population dans laquelle le fardeau ne s'exprime pas, voire à un niveau plus élevé (en particulier si les hétérozygotes sont désavantagés par rapport aux homozygotes pour les allèles sains) (Hedrick, 1994; Lande et Schemske, 1985). La purge du fardeau génétique semble donc être très pertinente pour limiter la dépression de consanguinité. Cependant, la purge peut également causer une dépression de consanguinité sévère à court terme, résultant de l'expression transitoire du fardeau génétique de la population. Ainsi, pour pouvoir se servir volontairement de la purge de consanguinité dans un programme de sélection, il ne faut pas que la taille efficace de la population concernée soit trop faible (Caballero et al., 2017). Dans le cas contraire, réaliser des accouplements volontairement consanguins dans le but de purger le fardeau génétique n'est pas conseillé à cause d'un risque d'extinction très élevé, dû à la fixation des allèles délétères (Caballero et al., 2017 ; Leberg et Firmin, 2008 ; Hedrick, 1994) et pourrait être moins efficace que la réalisation d'accouplements au hasard ou de paires d'individus les moins apparentés possibles (de Cara et al., 2013a).

La purge du fardeau génétique via une augmentation élevée et volontaire de la consanguinité a été expérimentée dans une population de gazelles de Speke (*Gazella spekei*) au sein d'un zoo. On a constaté que le niveau de dépression de consanguinité était considérablement réduit après deux ou trois générations (Templeton et Read, 1984), ce qui a conduit à soutenir cette approche (Ralls et Ballou, 1986; Templeton et al., 1986). Cependant, une nouvelle analyse des mêmes données n'a pas permis de retrouver des preuves d'amélioration génétique dans la population (Kalinowski et al., 2000).

Il n'est donc pas certain que la purge du fardeau génétique ait bien lieu dans des populations réelles. Des études de simulations indiquent que la purge est plus susceptible de se produire lorsque les allèles délétères ont un effet important et lorsque l'augmentation de la consanguinité se produit lentement et sur plusieurs générations (Boakes et Wang, 2005; Hedrick, 1994). Il est donc possible que des générations successives de forte consanguinité dans une population réelle puissent ne purger que le fardeau génétique résultant des allèles létaux et sub-létaux, permettant aux allèles légèrement délétères de rester présents. Plusieurs études ont mis en évidence l'existence de purge du fardeau génétique dans des races bovines laitières, en particulier pour les caractères sélectionnés fortement et depuis longtemps comme les caractères de production laitière (Doekes et al., 2019; Mc Parland et al., 2009).

Cependant, étant données les tailles efficaces faibles voire très faibles dans les races bovines laitières, purger volontairement le fardeau génétique en laissant la consanguinité augmenter rapidement ne semble pas être une solution viable pour limiter les conséquences de la perte de diversité génétique (de Cara et al., 2013b).

A l'ère de la génomique, il est désormais possible de s'intéresser à la dépression de consanguinité à l'échelle du génome, et pas seulement à l'échelle globale. Il serait pertinent de tirer parti de cela en sélectionnant contre les haplotypes impliqués dans le fardeau génétique ; ou encore en maîtrisant l'évolution de la consanguinité, non plus globalement, mais différenciellement selon les régions du génome où elle a un effet défavorable sur les caractères d'intérêt.

Plusieurs méthodes de prise en compte de ces régions sont alors possibles, selon deux grands principes :

- (i) Intégrer ces informations aux évaluations génomiques, grâce à des méthodes de type BayesRC (MacLeod et al., 2016), qui permettent de donner plus de poids dans l'évaluation à certains SNPs par rapport aux autres, afin de mieux estimer et prendre en compte leur effet sur les performances des animaux.
- (ii) Utiliser des approches de type OC en pondérant les coefficients d'apparentement ou assimilés en fonction des régions du génome (Gómez-Romano et al., 2016), pour favoriser une plus grande diversité génétique au sein des régions du génome dans lesquelles la consanguinité a un effet délétère pour les performances.

Une fois de telles méthodes développées, il faudra évaluer leur efficacité par simulations, que ce soit leur efficacité propre, ou leur efficacité lorsqu'elles sont intégrées (ensemble ou seules) à d'autres stratégies de gestion de la diversité génétique dans les races bovines laitières françaises.

En complément de ces approches, travailler sur des évaluations génomiques prenant en compte les effets de dominance permettrait de mieux maîtriser l'expression de la dépression de consanguinité, mais aussi de mieux comprendre son mécanisme. L'identification des allèles délétères pourrait être une première étape avant d'envisager leur élimination. Cette élimination est possible en écartant de la reproduction les animaux qui en sont porteurs ou plus progressivement en évitant l'accouplement entre porteurs. Ces méthodes sont déjà employées pour limiter ou stopper la diffusion des anomalies génétiques dans les races bovines laitières, mais également dans d'autres espèces. L'édition du génome des taureaux est une autre manière d'éliminer les allèles délétères (Johnsson et al., 2019), mais ce type de méthode pose des questions d'acceptabilité sociétale, de responsabilité en cas d'erreur (allèle mal détecté, indispensable pour des fonctions biologiques vitales ou souhaitable pour de futurs objectifs de sélection) et de coût.

Conclusions

Ma thèse avait pour but de contribuer à rendre l'élevage plus durable dans le contexte de la sélection génomique. Son objectif principal était de comprendre et de prédire l'impact des modifications des schémas de sélection sur la diversité génétique des races bovines laitières françaises de grand effectif, afin de mieux la gérer et la maintenir, à l'ère de la génomique.

Les résultats de ma thèse soulignent le fait que les changements apportés aux schémas de sélection peuvent avoir un impact important sur la diversité génétique. La maîtrise des schémas de sélection, via leur dimension (nombre de pères et mères à taureaux, mode de sélection des reproducteurs, nombre de taureaux diffusés, etc.) ou l'intensité d'utilisation des technologies de reproduction, est un véritable levier d'action pour gérer la diversité génétique dans les races bovines laitières françaises. Il est donc important de continuer à surveiller la diversité génétique et d'étudier l'impact futur des différentes stratégies appliquées dans les schémas sélection sur la diversité génétique.

En plus de la maîtrise des schémas de sélection, d'autres leviers de gestion de la diversité génétique et des conséquences de son érosion existent, incluant les plans d'accouplement, les méthodes d'évaluation des animaux, la maîtrise de la diffusion des taureaux et la prise en compte, lors des évaluations, du fardeau génétique et/ou des régions du génome impliquées dans la dépression de consanguinité.

Il est nécessaire d'évaluer l'efficacité des mesures correspondant à chacun de ces leviers sur le terrain, ainsi que la faisabilité de leur mise en place et/ou de leur pérennisation. Les échanges entre les entreprises de sélection, les éleveurs et les scientifiques sont indispensables pour proposer des méthodes de gestion efficaces de la diversité génétique, permettant un bon compromis entre diversité et progrès génétique.

La mondialisation de l'élevage bovin laitier, via les échanges internationaux de semence ou d'embryons, crée une concurrence forte dans la filière pour les races internationales comme la Holstein, et tend à favoriser l'accroissement du progrès génétique au détriment de la diversité génétique. Cela va à l'encontre des préconisations et fait peser une menace sur la durabilité de l'élevage. A l'avenir, la gestion de la diversité devrait être envisagée à l'échelle internationale, grâce à la coopération des entreprises de sélection et des éleveurs du monde entier.

Notre utilisation du vivant via l'élevage nous donne des responsabilités. Dans le cadre de ma thèse, nous avons vu que cette responsabilité se situe en partie dans les enjeux liés à la diversité génétique et à son importance pour l'avenir de l'élevage, mais aussi des solutions pour mieux gérer la diversité génétique. Il est en effet nécessaire de s'assurer de l'efficacité des solutions

proposées afin d'éviter des erreurs irréversibles, en particulier dans la gestion du fardeau génétique ou son élimination. De plus, de nouvelles métriques de la diversité génétique, plus précises, pourraient voir le jour à l'avenir et modifier l'efficacité des mesures de gestion de la diversité génétique.

Ces questionnements sur la diversité génétique et sa gestion s'inscrivent dans ceux de la transition agroécologique, qui motive de nombreux projets de recherche au sein des instituts de recherche français comme l'INRAE. Ces problématiques existent dans les autres espèces domestiques, qu'elles soient animales ou végétales, bien que le contexte de chacune d'entre elles soit différent de celui des races bovines laitières, dans l'organisation de la sélection, de l'élevage, de la filière ou la structure génétique des populations. Des solutions adaptées doivent donc être mises en place pour améliorer la gestion de la diversité génétique de toutes ces espèces et races, si possible de manière intégrée entre espèces interdépendantes (par exemple entre une espèce animale et les espèces végétales consommées par cette espèce animale) et en partenariat avec les sciences humaines et sociales, pour une agriculture plus durable.

Chapitre Annexe 1 - Les mécanismes génétiques impliqués dans la dépression de consanguinité

Trois hypothèses ont été proposées pour expliquer les mécanismes génétiques impliqués dans la dépression de consanguinité, et donc dans l'expression du fardeau génétique :

- (i) **L'hypothèse de dominance partielle**, soit l'expression accrue des allèles récessifs délétères (Curik et al., 2017; Davenport, 1908) ;
- (ii) **L'hypothèse de superdominance**, soit la supériorité des hétérozygotes sur les deux types d'homozygotes (Curik et al., 2017 ; Charlesworth et Charlesworth, 1999 ; Shull, 1908). Un exemple très connu de superdominance est celui de la drépanocytose (maladie causée par un allèle récessif au locus du gène responsable de la structure de l'hémoglobine) : les individus homozygotes pour cet allèle meurent prématurément alors que les individus hétérozygotes présentent une résistance au paludisme supérieure aux homozygotes pour l'allèle sain ;
- (iii) **L'hypothèse d'épistasie**, soit la probabilité accrue de combinaisons de gènes favorables pour les hétérozygotes (Curik et al., 2001; Jain et Allard, 1966; Kempthorne, 1957).

Les deux premières hypothèses, celle de la dominance partielle et celle de superdominance, constituent les deux hypothèses historiques pour expliquer la dépression de consanguinité. Dans l'hypothèse de dominance partielle, la valeur génétique des hétérozygotes est supposée être supérieure à la moyenne des homozygotes alors qu'avec la superdominance, elle est supposée être supérieure aux deux types d'homozygotes (Kristensen et Sørensen, 2005).

L'une des différences primordiales entre ces deux hypothèses est que l'élimination des allèles délétères par la purge n'est possible et n'a de sens que dans le cadre de l'hypothèse de dominance partielle (Benesh et al., 2014).

Pour tenter de comprendre si une hypothèse est plus valable qu'une autre parmi celles de la dominance partielle et de la superdominance, certaines approches consistent à croiser des lignées consanguines pour savoir si la valeur des individus issus des croisements terminaux est égale (hypothèse de superdominance) ou supérieure (hypothèse de dominance partielle avec purge des allèles délétères) à celle de lignées de croisement (Roff, 2002). D'autres approches envisagent l'identification de mutations délétères grâce à l'utilisation des données issues du génotypage (Kristensen et al., 2010). Ces approches ont eu tendance à montrer que le rôle de la superdominance était mineur par rapport à celui de la dominance partielle (Roff, 2002 ; Charlesworth et Charlesworth, 1999 ; Lacy et al., 1996). Dans certains cas, ce qui apparaît comme

de la superdominance s'avère être plutôt de la pseudo-superdominance : on aura alors une combinaison de mutations délétères récessives à des locus en fort déséquilibre de liaison (Charlesworth et Willis, 2009; Graham et al., 1997).

Des schémas de superdominance ont par exemple été observés parmi des lignées consanguines de *Drosophila melanogaster*. Dans ces lignées, 567 gènes ont été identifiés comme étant impliqués dans la dépression de consanguinité, dont 75 % de gènes à effets additifs ou partiellement dominants et 25 % de gènes exprimant des schémas de superdominance (Ayroles et al., 2009). Ces expériences ont été possibles grâce à la facilité d'élevage des insectes *Drosophila melanogaster*, dont le cycle de vie est très court (environ deux semaines) et qui produisent beaucoup de descendants (environ 400 œufs par femelle).

Dans les espèces de rente, il est plus difficile de réaliser de telles expériences pour évaluer quelle hypothèse est la plus plausible, ou si les deux cohabitent et dans quelles proportions, pour différentes raisons (coût, éthique, intervalles de génération longs, peu de descendants) (Kristensen et Sørensen, 2005). Il a toutefois été observé dans la race bovine laitière Holstein, à partir de l'étude de 14 caractères, que la dépression de consanguinité était d'autant plus forte que la variance de dominance était élevée (Misztal et al., 1997). En bref, il est assez difficile de conclure quant à l'importance comparée des rôles de la dominance partielle et de la superdominance sur la dépression de consanguinité.

La troisième hypothèse proposée pour expliquer les mécanismes génétiques impliqués dans la dépression de consanguinité est l'épistasie. En effet, les individus hétérozygotes ont plus de chances de combiner des interactions géniques favorables que les individus homozygotes. Les mécanismes génomiques de l'épistasie et son importance dans la dépression de consanguinité sont encore mal compris (Kristensen et Sørensen, 2005).

Si les mécanismes impliqués dans la dépression de consanguinité sont encore mal connus, de nombreuses méthodes de détection et d'évaluation de l'amplitude de la dépression de consanguinité et du fardeau génétique ont été proposées.

Chapitre Annexe 2 - Mesurer la diversité génétique au sein d'une race

I - Les différentes métriques permettant d'évaluer la diversité génétique

Nous allons présenter les principales métriques permettant d'évaluer la diversité génétique au sein d'une race, et en particulier celles s'appuyant sur la **consanguinité**, l'**apparentement** et la **taille efficace de la population** considérée.

Comme vu précédemment, 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 Figure 17).

Soit l'individu I, produit de l'accouplement entre A et B, deux individus apparentés. Alors l'individu I sera consanguin. Pour estimer le niveau de consanguinité de l'individu I ou le niveau d'apparentement des individus A et B, on peut calculer le coefficient de consanguinité de l'individu I, qui sera égal au coefficient de parenté de ses parents A et B. Le coefficient de parenté de A et B est défini comme la probabilité qu'un allèle tiré au hasard chez l'individu A soit identique par descendance (ou IBD) à un allèle du même locus tiré au hasard chez l'individu B (voir Figure 18). On note ce coefficient de parenté f_{AB} .

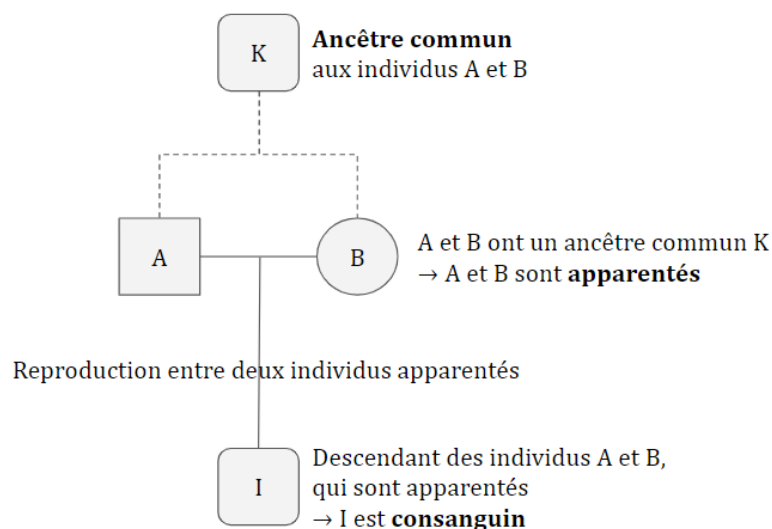


Figure 17 Consanguinité et apparentement - définitions

Le coefficient de consanguinité de l'individu I correspond par conséquent à la probabilité que ses deux allèles à un locus tiré au hasard soient identiques par ascendance (ou IBD), c'est-à-dire qu'un locus tiré au hasard soit autozygote (voir Figure 18). On note ce coefficient de consanguinité F_I .

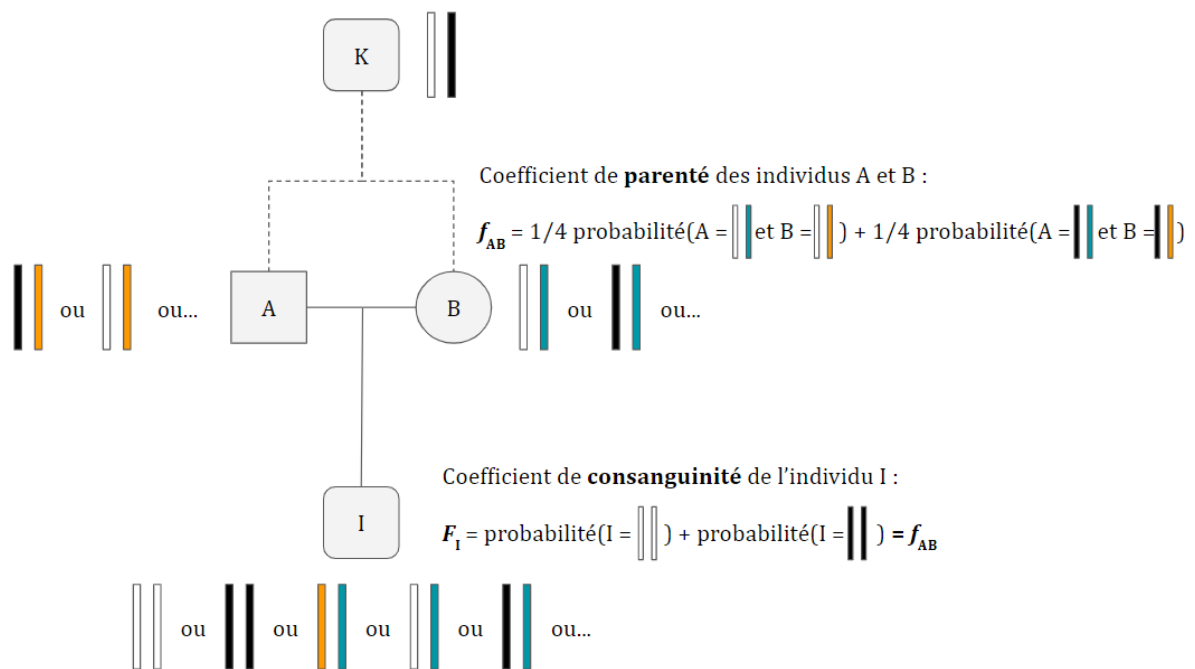


Figure 18 Coefficients de consanguinité et de parenté – définitions dans le cas de la présence d'un seul ancêtre commun

Cette approche probabiliste a été développée dans le cadre de l'hypothèse d'absence de sélection (Malécot, 1948). L'apparentement est transmissible, c'est-à-dire que si deux individus sont apparentés, leurs descendants le seront également : les ancêtres communs aux descendants seront leurs parents ainsi que tous les ancêtres communs à leurs parents. En revanche, la consanguinité n'est pas transmissible. En effet, si deux individus consanguins mais non apparentés se reproduisent, leurs descendants auront un coefficient de consanguinité nul car le coefficient de parenté de leurs parents est nul.

Les deux métriques sont donc complémentaires pour estimer la diversité d'une race. En effet, les animaux d'une race peuvent être très consanguins mais être élevés au sein de lignées très distinctes et donc avec un très faible apparentement entre elles mais très fort au sein de chaque lignée. Dans ce cas, le calcul de coefficients de parenté ou d'autres estimateurs de l'apparentement est nécessaire pour interpréter correctement les niveaux de consanguinité observés via les coefficients de consanguinité et donc pour avoir une idée plus juste de la structure de la population.

Une autre façon d'estimer la diversité au sein d'une population est de calculer sa taille efficace, notée N_e . La **taille efficace d'une population** dépend toujours d'un paramètre d'intérêt (consanguinité par exemple). On se place dans une population idéale (accouplements au hasard,

naissances simultanées pour chaque génération c'est-à-dire générations non chevauchantes, taille de population constante et même nombre de descendants par individu) et on cherche quelle devrait être la taille de cette population idéale pour que le paramètre d'intérêt soit égal à celui observé dans la population réelle (N), sous l'effet de la dérive seulement (Wright, 1931). Cette taille correspond à la taille efficace de la population. Etant donné qu'elle dépend du paramètre à partir duquel on l'estime, il n'est pas possible de comparer des estimateurs de la taille efficace obtenus par des méthodes différentes.

Ainsi, dans le cadre de la conservation de la diversité génétique, on s'intéresse le plus souvent à son évolution dans le temps et au ratio entre taille efficace et taille réelle de la population N_e/N . On peut remarquer que, dans des populations de grand effectif, la dérive génétique aura un effet similaire à celui observé dans une population idéale d'un effectif correspondant à la taille efficace de la population. Par conséquent, dans ces populations de grand effectif, si le ratio N_e/N est très faible, la dérive peut être beaucoup plus forte qu'attendu au regard du seul effectif réel de la population. La fixation des allèles due à la dérive sera alors plus rapide, ce qui participera à une perte de diversité plus rapide qu'attendue. A titre d'exemple, la taille efficace de la race bovine laitière Prim'Holstein en France est estimée à 96 individus pour une taille de population analysée de 3 033 721 individus, avec un ratio N_e/N de 0,003 % environ, quand la race ovine Roussin possède la même taille efficace pour un effectif réel de la population analysée de 3 450 individus, portant le ratio N_e/N à 2,78 % dans cette race (Danchin-Burge, 2020, 2018). Dans ce cas, la force de la dérive génétique dans ces deux races car leur taille efficace est identique. On comprend alors que la race ovine est bien plus diversifiée que la Prim'Holstein car sa taille efficace est plus proche de sa taille réelle.

En résumé, les indicateurs fondés sur la consanguinité, l'apparentement ou la taille efficace de la population sont complémentaires et permettent d'apprécier la diversité génétique au sein d'une race. Nous allons voir qu'il est possible d'estimer ces indicateurs grâce à plusieurs types de données :

- Les données généalogiques via les pedigrees,
- Les génotypes des animaux.

II - Estimer la diversité génétique à partir des pedigrees

A - Coefficients de consanguinité et de parenté

- **La méthode des chemins** (Wright, 1923)

La méthode des chemins est utilisée dans le calcul des coefficients de parenté et de consanguinité à la main. Nous allons suivre l'exemple de la Figure 19 ci-dessous, avec un pedigree très simple, afin de présenter les principes de cette méthode.

Pour calculer le coefficient de parenté de deux individus A et B ou le coefficient de consanguinité d'un individu I dont les parents sont A et B, qui sont égaux, il faut commencer par identifier tous les ancêtres communs à A et B. Dans le pedigree présenté dans la Figure 19, le seul ancêtre commun aux individus A et B est l'individu K. Il est nécessaire de calculer le coefficient de consanguinité des ancêtres communs identifiés. On notera F_K le coefficient de consanguinité de l'individu K.

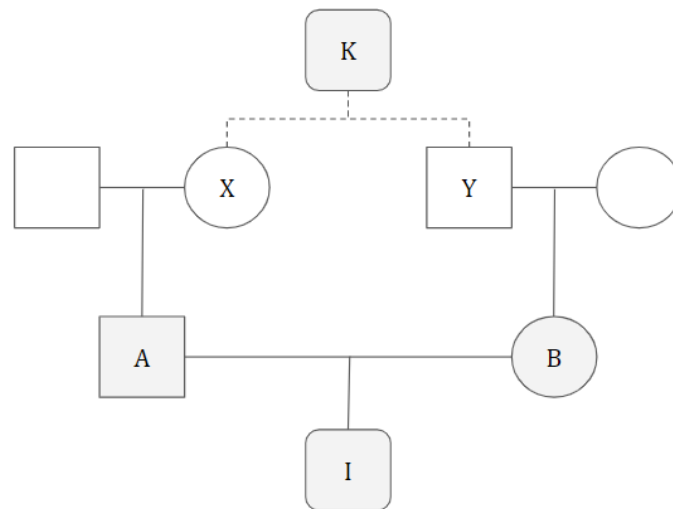


Figure 19 Exemple de pedigree simple pour le calcul de coefficients de consanguinité et de parenté selon la méthode des chemins

Pour chaque ancêtre commun à A et B, ici l'individu K, on identifie tous les chemins par lesquels A et B sont apparentés en passant par K dans le pedigree. Ces chemins représentent les voies par lesquelles les deux copies d'un même allèle ont pu être transmises de K à A et de K à B. Il existe deux règles pour définir un chemin de parenté :

- (i) Un chemin de parenté est représenté par une suite d'individus possédant des liens parent-descendant ;
- (ii) Un chemin de parenté remonte les générations depuis l'un des individus de la paire considérée jusqu'à l'ancêtre commun choisi puis redescend vers l'autre individu de la paire. Un individu ne peut être inclus qu'une seule fois dans un chemin de parenté.

Dans l'exemple de la Figure 19, le seul chemin de parenté reliant A à B en passant par K est : A - X - K - Y - B.

Pour chaque chemin de parenté, on compte le nombre d'individus n faisant partie du chemin, en incluant les individus A et B. Ici, le chemin présente une longueur de 5 individus. On définit la probabilité $P(A = B \text{ via } K)$ qu'un allèle pris au hasard chez l'individu A soit identique à un allèle du même locus pris au hasard chez l'individu B, et ce par ascendance en provenant de l'individu K. Avec un chemin de parenté de longueur n , cette probabilité est exprimée comme suit :

$$P(A = B \text{ via } K) = (1 + F_K) \left(\frac{1}{2}\right)^n$$

Enfin, on prend en compte l'ensemble des ancêtres communs et des chemins possibles reliant les individus A et B pour le calcul du coefficient de parenté entre ces deux individus f_{AB} , égal au coefficient de consanguinité de l'individu I ayant pour parents A et B F_I :

$$f_{AB} = F_I = \sum_K \sum_{\text{chemins}} (1 + F_K) \left(\frac{1}{2}\right)^n$$

Dans le cas particulier où l'on cherche à calculer le coefficient de parenté entre un individu et l'un de ses ancêtres, l'ancêtre est alors considéré comme son propre ancêtre et donc comme l'ancêtre commun à lui-même et son descendant.

- **La méthode tabulaire** (Cruden, 1949)

Si la méthode des chemins est simple à mettre en œuvre à la main et dans des cas simples, la méthode tabulaire et les méthodes qui en sont dérivées sont plus adaptées pour calculer des coefficients de parenté et de consanguinité dans des grands jeux de données.

Pour présenter les principes de la méthode tabulaire, nous prendrons l'exemple du pedigree présenté dans la Figure 20.

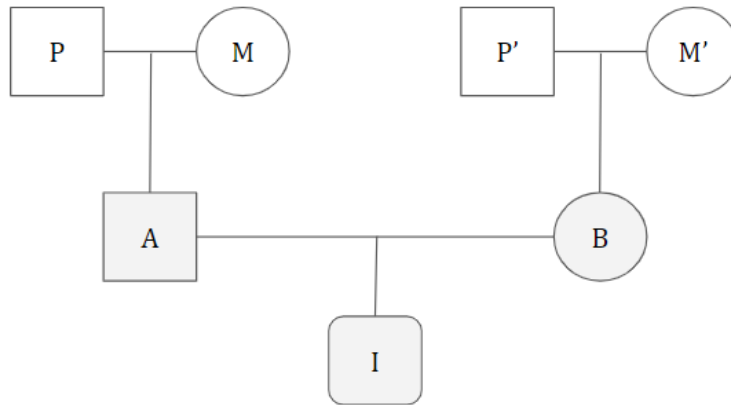


Figure 20 Exemple de pedigree simple pour le calcul de coefficients de consanguinité et de parenté selon la méthode tabulaire

La méthode tabulaire consiste à considérer que les allèles présents chez une paire d'individus, A et B dans l'exemple de la Figure 20, proviennent d'un échantillonnage de ceux de leurs parents. L'allèle paternel ou maternel d'un individu donné peut être considéré comme le résultat du tirage d'un allèle au hasard dans le génotype du parent concerné de l'individu. On considère deux individus A et B, sans relation parent-descendant. Soient P le père et M la mère de l'individu A, et P' le père et M' la mère de l'individu B. Lorsque l'on tire au hasard un allèle chez l'individu A, l'allèle obtenu peut être l'allèle paternel ou l'allèle maternel avec une probabilité de $1/2$. Il en est de même pour l'individu B. On obtient ainsi l'arbre des probabilités d'origine des allèles tirés au sort de la Figure 21.

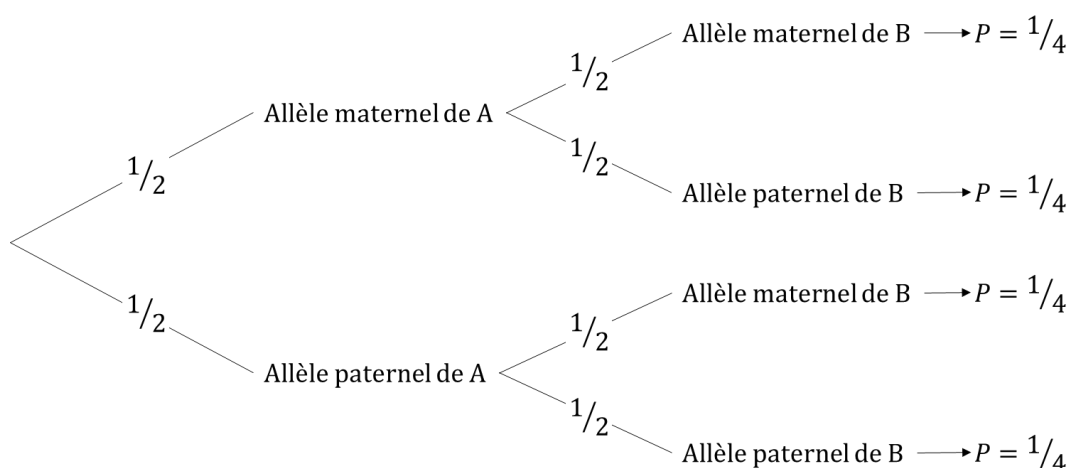


Figure 21 Arbre des probabilités d'origine des allèles tirés au sort à un même locus pour les individus A et B

Il y a donc quatre possibilités d'origine lors du tirage aléatoire des allèles des individus A et B à un même locus, équiprobables de probabilité 1/4.

Les probabilités d'identité des gènes entre deux individus donnés peuvent donc se déduire des probabilités d'identité à l'échelon parental. A un locus donné, la probabilité que l'allèle pris au hasard chez l'individu A soit identique à l'allèle tiré chez l'individu B, notée f_{AB} et égale au coefficient de parenté de A et B, correspond donc à :

$$f_{AB} = \frac{1}{4}f_{MM'} + \frac{1}{4}f_{MP'} + \frac{1}{4}f_{PP'} + \frac{1}{4}f_{PM'}$$

Avec $f_{MM'}$ la probabilité que l'allèle tiré au sort chez l'individu M (mère de A) soit identique à l'allèle tiré au sort chez l'individu M' (mère de A), c'est-à-dire le coefficient de parenté de M et M'. $f_{MP'}$, $f_{PP'}$ et $f_{PM'}$ sont définis de la même façon.

Le coefficient de consanguinité de I, noté F_I , est alors égal à f_{AB} , le coefficient de parenté de ses parents A et B.

Dans le cas particulier où l'on veut calculer le coefficient de parenté entre un individu et l'un de ses parents, cette équation n'est plus valable. En appliquant le même principe à l'individu B et à son père P', on obtient :

$$f_{P'B} = \frac{1}{2}f_{P'P'} + \frac{1}{2}f_{P'M'} = \frac{1}{2}(1 + F_{P'}) + \frac{1}{2}f_{P'M'}$$

Avec $F_{P'}$ le coefficient de consanguinité de l'individu P', égal au coefficient de parenté de ses propres parents. En effet, une fois qu'on a tiré un allèle chez l'individu P', on a une chance sur deux de tirer à nouveau ce même allèle et une chance sur deux de tirer le second allèle. Or la probabilité que le second allèle soit identique au premier correspond au coefficient de consanguinité de P'.

La méthode tabulaire est plus rapide que la méthode des chemins, notamment car elle fonctionne de manière séquentielle et ne nécessite pas de remonter entièrement le pedigree de chaque individu pour calculer son coefficient consanguinité.

Remarque : Les coefficients de consanguinité et de parenté calculés à partir des pedigrees seront toujours compris entre 0 et 1.

B - Taille efficace

Comme précisé précédemment, la taille efficace d'une population dépend toujours d'un paramètre d'intérêt. Ainsi, il existe de nombreuses méthodes pour calculer la taille efficace d'une population grâce aux données généalogiques. Nous allons en détailler deux seulement.

- **Estimation de la taille efficace à partir du nombre de mâles et de femelles reproducteurs** (Wright, 1931)

Il est possible de calculer la taille efficace d'une population à partir du nombre de mâles et de femelles reproducteurs qui s'y trouvent à un instant donné. Sous l'hypothèse de contribution équilibrée des reproducteurs de chaque sexe, la taille efficace de la population peut être calculée comme suit :

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

Avec N_f le nombre de femelles reproductrices présentes dans la population et N_m le nombre de mâles reproducteurs présents dans la population. D'autres formules existent afin de prendre en compte le déséquilibre de contribution des reproducteurs, qui peut être très fort dans les populations animales sélectionnées. Ces formules prennent en compte la variance du nombre de gamètes que les individus de la génération t apportent à la génération $t + 1$.

- **Estimation de la taille efficace à partir des coefficients de consanguinité** (Kempthorne, 1957)

On peut définir la taille efficace de la population N_e en notant comment le coefficient de consanguinité moyen change d'une génération à l'autre, puis en définissant N_e comme la taille de la population idéalisée qui présente la même variation du coefficient de consanguinité moyen que la population considérée.

Pour une population idéale, les coefficients de consanguinité suivent cette relation de récurrence :

$$F_t = \frac{1}{N} \left(\frac{1 + F_{t-2}}{2} \right) + \left(1 - \frac{1}{N} \right) F_{t-1}$$

Avec F_t la consanguinité moyenne de la population à une génération t et N la taille de la population idéale. On peut déduire de cette relation la relation explicite entre la taille de la population idéale et la consanguinité à une génération donnée t :

$$1 - F_t = (1 - F_0) \left(1 - \frac{1}{2N} \right)^t$$

On peut alors en déduire que :

$$\frac{1 - F_{t+1}}{1 - F_t} = 1 - \frac{1}{2N}$$

En considérant que la seule force évolutive en présence est la dérive (inclus dans l'hypothèse de population idéale), N doit être égal à N_e pour que l'équation précédente soit vérifiée. Cela donne :

$$N_e = \frac{1 - F_t}{2(F_{t+1} - F_t)}$$

On peut alors calculer la taille efficace d'une population en utilisant les coefficients de consanguinité calculés à partir des informations généalogiques fournies par les pedigrees.

C - Limites de ces méthodes

Dans le calcul des coefficients de consanguinité et de parenté à partir des pedigrees, on définit une population de base constituée par les individus dont l'ascendance est inconnue. On considère que les individus de la population de base ont un coefficient de consanguinité nul et ne sont pas apparentés entre eux. Ainsi, la qualité du calcul des coefficients de parenté et de consanguinité à partir des pedigrees dépend fortement de la qualité des données généalogiques disponibles, en termes de profondeur (nombre de générations d'ancêtres connus pour chaque individu) mais aussi d'erreurs possibles lors de l'enregistrement des liens de parenté.

Les erreurs ou les ascendances incomplètes entraînent souvent une sous-estimation des coefficients de parenté et donc des coefficients de consanguinité (Engelsma et al., 2011 ; Boichard et al., 1997). Or, il est rare de disposer de pedigrees exhaustifs et dans certaines espèces animales d'élevage, les pedigrees sont difficiles voire impossibles à enregistrer. Il suffit que plusieurs mâles reproducteurs soient mélangés avec les femelles reproductrices pendant la période de reproduction pour qu'il devienne difficile d'attribuer avec certitude un jeune au bon père, voire à la bonne mère, par exemple pour les espèces de poissons, pour lesquelles les jeunes ne restent pas près de leur mère dans les premières phases de vie.

Au-delà des difficultés liées à l'exactitude et à l'exhaustivité des pedigrees, nous avons vu que le calcul des coefficients de consanguinité à partir des pedigrees représentaient des probabilités d'identité par descendance des allèles à un locus tiré au sort. Ces coefficients correspondent donc à l'espérance de la proportion de génome autozygote chez un individu, si tous les marqueurs sont indépendants. Ils ne prennent donc pas en compte la variance liée aux recombinaisons inter et intra-chromosomiques pendant la méiose, qui entraînent des écarts à l'espérance pouvant s'accumuler génération après génération. Ainsi, les coefficients de consanguinité estimés à partir des pedigrees peuvent être considérés comme des mesures de la consanguinité attendue. Par conséquent, deux membres de la même fratrie (même père, même mère) auront le même coefficient de consanguinité à partir des pedigrees, alors que leurs niveaux individuels réels d'autozygotie ne seront pas identiques. Ces écarts entre la consanguinité attendue et la

consanguinité réalisée sont intrinsèques à l'utilisation des pedigrees, quelles que soient leur exactitude et leur exhaustivité (Hill et Weir, 2011).

Les mesures de la taille efficace de la population à partir de la consanguinité estimée sur les pedigrees présentent également des limites liées à la qualité des pedigrees. De plus, la prise en compte des contraintes liées aux populations réelles (taille non constante, déséquilibre de contribution, présence de sélection, ratio mâles/femelles différent de 1) peut rendre les calculs de la taille efficace des populations complexes et les résultats difficiles à interpréter.

III - Estimer la diversité génétique à partir des génotypes

Le génotypage haut débit a été utile non seulement pour améliorer l'efficacité de la sélection mais également pour donner accès à de nouvelles façons d'estimer la diversité génétique des populations dans lesquelles une part importante d'individus sont génotypés.

Ces données permettent de s'émanciper des limites imposées par les mesures de la diversité à partir de données généalogiques. Les données de génotypes donnent accès à un nouveau niveau d'informations car elles permettent d'étudier la diversité tout le long du génome et pas seulement à l'échelle des individus.

A - Hétérozygotie, consanguinité, apparentement, taille efficace : mesures moléculaires de la diversité

A l'échelle de la population, l'hétérozygotie est une mesure de la diversité génétique corrélée à la survie (Reed et Frankham, 2003). Ainsi, pour gérer la diversité génétique d'une population, il est intéressant d'estimer son hétérozygotie globale et d'observer son évolution dans le temps.

A l'échelle individuelle, nous avons vu que le niveau de consanguinité réelle correspondait à la proportion de génome autozygote, c'est-à-dire homozygote et provenant du même ancêtre commun. Les individus consanguins présentent donc une hétérozygotie globale plus basse que les individus non consanguins. Suivant ce principe, les mesures de la consanguinité à partir des marqueurs moléculaires consistent en général à comparer l'hétérozygotie (ou l'homozygotie) observée à l'échelle du génome d'un individu à l'hétérozygotie (ou à l'homozygotie) attendue sous les hypothèses de panmixie et d'absence de force évolutive, étant données les fréquences alléliques dans la population pour tous les marqueurs étudiés.

En particulier, la **méthode des moments** permet de calculer un **coefficient de consanguinité** F_1 pour un individu I (Ritland, 1996). Dans le cas de locus à deux allèles, cette méthode s'applique comme suit :

$$F_1 = \frac{1}{L} \sum_{l=1}^L \left(\frac{S_{1l} - p_l^2}{p_l} + \frac{S_{2l} - q_l^2}{q_l} \right)$$

Avec :

- L le nombre total de locus pris en compte ;
- S_{1l} et S_{2l} des indicateurs prenant la valeur 1 si le marqueur du locus l est homozygote pour l'allèle 1 ou pour l'allèle 2 respectivement, et la valeur 0 si le marqueur du locus l est hétérozygote pour l'allèle 1 ou pour l'allèle 2 respectivement ;
- p_l et q_l les fréquences des allèles 1 et 2 respectivement dans la population et donc p_l^2 et q_l^2 les probabilités d'être homozygote pour l'allèle 1 ou l'allèle 2 respectivement dans une population avec p_l et q_l pour fréquences alléliques au locus l ;
- On notera que pour des locus à deux allèles, quel que soit le locus l , $p_l + q_l = 1$.

Une seconde méthode simple des moments pour L locus indépendants à deux allèles chacun permet également de calculer le coefficient de consanguinité d'un individu I comme suit (Ritland, 1996) :

$$F_1 = \frac{S - J}{1 - J}$$

Avec :

- S la proportion de locus homozygotes chez l'individu I ;
- J la proportion attendue de locus homozygotes pour un individu de la population à laquelle I appartient, telle que :

$$J = \frac{1}{L} \sum_{l=1}^L (p_l^2 + q_l^2)$$

Il est également possible d'estimer les relations entre individus (et donc leur apparentement) en calculant des **matrices d'apparentement génomique**. Les principes d'obtention de ces matrices sont décrits ci-après (VanRaden, 2007).

Soit $\mathbf{M} = (m_{ij})$ la matrice d'incidence spécifiant les allèles présents chez les individus génotypés et pour lesquels on veut obtenir les relations d'apparentement. La matrice \mathbf{M} est de dimension N par L avec N le nombre d'individus considérés et L le nombre de locus génotypés. On se place à

nouveau dans le cas de génotypes définis à partir de SNPs, avec deux allèles par locus. Pour chaque locus j , tel que $1 \leq j \leq L$, on notera p_j la fréquence de l'allèle majoritaire à ce locus, c'est-à-dire le plus fréquent des deux allèles possibles ($p_j \geq 0,5$). La fréquence de l'allèle minoritaire, aussi appelée MAF pour *minor allele frequency*, est alors égale à $1 - p_j$.

Pour tous $1 \leq i \leq N$ et $1 \leq j \leq L$, m_{ij} est égal au nombre d'allèles majoritaires portés par l'individu i au locus j (0 : homozygote pour l'allèle minoritaire, 1 : hétérozygote, ou 2 : homozygote pour l'allèle majoritaire).

On note \mathbf{M}' la matrice transposée de la matrice \mathbf{M} . Les éléments m''_{ij} de la matrice carrée \mathbf{MM}' qui ne sont pas sur la diagonale correspondent au niveau de partage de l'allèle majoritaire entre l'individu i et l'individu j , et les éléments m''_{ii} de la diagonale sont le reflet de la consanguinité de l'individu i .

On note \mathbf{P} la matrice de dimension N par L telle que la colonne j , $1 \leq j \leq L$, de la matrice \mathbf{P} est égale à $2p_j \times \mathbf{1}_N$. On peut alors obtenir la matrice $\mathbf{Z} = \mathbf{M} - \mathbf{P}$. Cette soustraction permet de prendre en compte les fréquences des allèles dans le calcul de l'apparentement. En effet, deux individus ont plus de chances de partager des allèles rares s'ils sont apparentés, alors que la probabilité pour deux individus de partager des allèles très fréquents est élevée même par hasard. On note \mathbf{Z}' la transposée de \mathbf{Z} .

On peut alors calculer \mathbf{G} la matrice d'apparentement génomique comme suit :

$$\mathbf{G} = \frac{1}{2 \sum_{j=1}^L p_j (1 - p_j)} \times \mathbf{ZZ}'$$

Les éléments g_{ij} de \mathbf{G} ne se situant pas sur la diagonale correspondent alors au coefficient d'apparentement entre les individus i et j .

Remarque : Les coefficients d'apparentement calculés à partir de ces méthodes (comparaison des niveaux d'hétérozygotie et matrices d'apparentement génomique) seront compris entre -1 et 1. Il arrive que les individus d'une population présentent un excès de marqueurs hétérozygotes observés par rapport à ce qui est attendu à partir des fréquences alléliques de la population (Curie-Cohen, 1982). Dans ce cas la consanguinité est définie comme l'accouplement d'individus plus apparentés que la moyenne de la population (Keller et al., 2011). Des accouplements non aléatoires qui évitent les accouplements les plus consanguins vont donc tendre à maintenir le niveau d'hétérozygotie individuelle constant d'une génération à la suivante. L'hétérozygotie sera alors plus importante que ce qui serait attendu au hasard, donnant des estimations négatives de consanguinité par exemple.

On notera que ces méthodes consistent toutes à mettre en relation le génotype d'un individu ou d'une paire d'individus avec les fréquences alléliques de la population de laquelle ils sont issus. Elles nécessitent donc l'estimation préalable des fréquences alléliques de la population. Les fréquences alléliques doivent correspondre à celles de la population composée d'individus fondateurs non sélectionnées, et pas aux fréquences observées à un instant donné. Des méthodes ont été développées pour obtenir ces fréquences alléliques (Gengler et al., 2007).

Il est enfin possible d'estimer N_e la **taille efficace d'une population** grâce aux génotypes. Dans les espèces diploïdes, on peut partir du principe que, dans une population idéale de taille N_e , si l'on tire au hasard deux allèles parmi les individus de la génération $t + 1$, la probabilité qu'ils proviennent du même individu de la génération t est égale à $1/2N_e$. La probabilité qu'ils proviennent de deux individus différents est donc $1 - \frac{1}{2N_e}$. Le niveau d'hétérozygotie observée à la génération $t + 1$, noté H_{t+1} , peut alors s'exprimer en fonction de N_e et du niveau d'hétérozygotie observée à la génération t , noté H_t , comme suit :

$$H_{t+1} = \left(1 - \frac{1}{2N_e}\right)H_t$$

On peut donc en déduire que N_e , la taille efficace de la population à partir de laquelle on calcule H_{t+1} et H_t , s'exprime comme suit :

$$N_e = -\frac{1}{2} \times \frac{H_t}{H_{t+1} - H_t}$$

Il est possible d'utiliser non pas les génotypes mais les séquences génomiques complètes des individus pour estimer la diversité d'une population. Cela permet de limiter la perte d'informations sur les variants rares dans une population (Eynard et al., 2016) tout en améliorant la précision des prédictions génomiques. C'est le cas chez les bovins par exemple (Druet et al., 2014; Pérez-Enciso, 2014). Cependant, l'utilisation des SNPs, s'ils sont assez denses, est moins coûteuse que celle des séquences, tant en coûts de séquençage qu'en coûts de stockage et de traitement des données. De plus, l'utilisation des SNPs permet déjà de calculer des coefficients de consanguinité et de parenté d'une précision très satisfaisante, en particulier en comparaison avec les pedigrees (Wang, 2016 ; Kardos et al., 2015).

B - Limites des mesures fondées sur les données moléculaires

Si les mesures de la diversité à partir des génotypes permettent d'éviter les inconvénients de l'utilisation des pedigrees évoqués plus haut (exactitude et exhaustivité des pedigrees, impossibilité d'enregistrer les généalogies, mesure de l'espérance de la consanguinité ou de

l'apparement plutôt que des valeurs réalisées), l'utilisation de ces données présentent également des limites.

Tout d'abord, l'utilisation des génotypes ne permet pas de calculer les différentes mesures de la diversité de populations peu ou pas génotypées. Bien que les coûts de génotypage aient chuté depuis quelques années et que le génotypage soit devenu très courant dans certaines espèces comme les bovins, toutes les filières n'ont pas les mêmes moyens financiers ou techniques pour génotyper largement les individus de leurs populations.

De plus, des erreurs dans l'estimation des fréquences alléliques de référence de la population peuvent changer drastiquement les valeurs d'apparement et de consanguinité calculées à partir des génotypes. Le génotypage non systématique des animaux et l'absence de génotypes d'animaux anciens peuvent participer à rendre difficile l'estimation des fréquences alléliques dans une population.

C - Mesurer la consanguinité à partir des Runs of Homozygosity (ROH)

A partir de l'utilisation des génotypes, une autre méthode existe pour estimer la consanguinité des animaux. Elle consiste à identifier des blocs supposés autozygotes dans le génome, appelés Runs Of Homozygosity (ROH) ou îlots d'homozygotie (McQuillan et al., 2008).

Les accouplements entre individus apparentés A et B causent la transmission à leurs descendants de longs segments de chromatides homologues d'ADN provenant du même ancêtre commun K (voir Figure 22). Les individus consanguins, comme l'individu I, présentent donc des segments

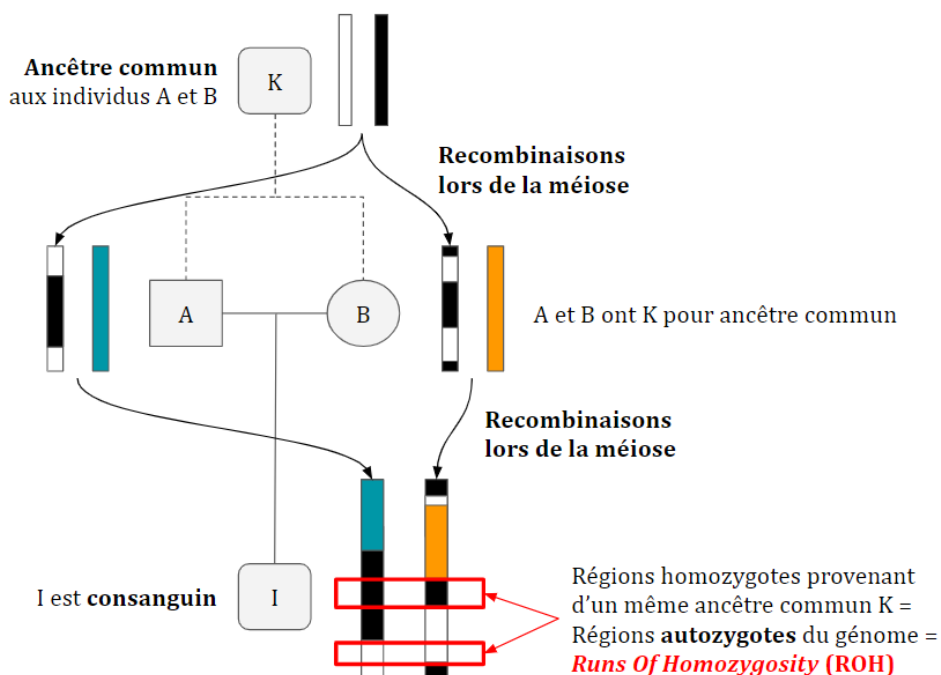


Figure 22 Origine des *Runs Of Homozygosity* (ROH)

génomiques autozygotes constitués de locus homozygotes consécutifs de longueur variable en fonction des recombinaisons ayant eu lieu lors de la méiose. Si ces segments sont suffisamment longs, il est probable qu'ils soient en effet autozygotes, et pas seulement présents par hasard.

Après identification des ROH, il est possible de calculer un coefficient de consanguinité. Les ROH représentant les portions génomiques autozygotes, on peut définir $F_{ROH,i}$ le coefficient de consanguinité à partir des ROH de l'individu i comme étant la proportion de son génome qui est inclus dans des ROH. On a alors :

$$F_{ROH,i} = \frac{\sum L_{ROH,i}}{L_{\text{génomique}}}$$

Avec $\sum L_{ROH,i}$ la somme de la longueur de tous les ROH présents chez l'individu i , c'est-à-dire la longueur totale de son génome inclus dans des ROH, et $L_{\text{génomique}}$ la longueur du génome sur laquelle on a effectué la détection des ROH (McQuillan et al., 2008).

L'exemple du calcul du coefficient de consanguinité à partir des ROH pour l'individu I de la Figure 22 est présenté dans la Figure 23.

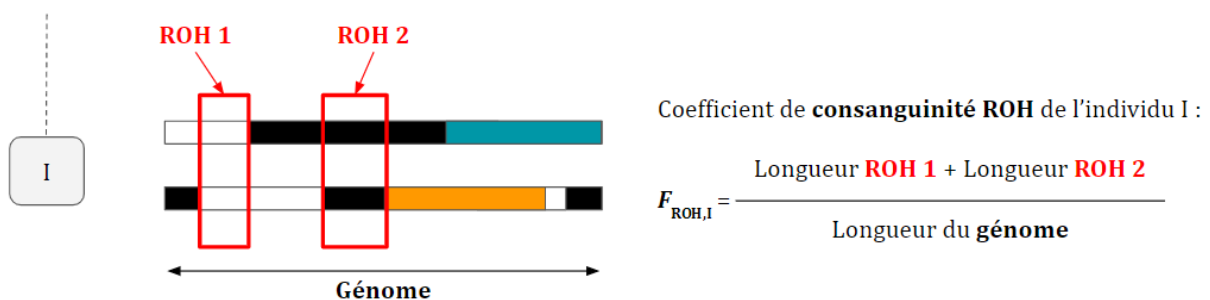


Figure 23 Exemple de calcul du coefficient de consanguinité d'un individu I à partir des *Runs of Homozygosity* (ROH)

Les coefficients de consanguinité estimés à partir des ROH, à condition que les ROH soient correctement détectés, fournissent une estimation de la consanguinité plus proche de la consanguinité (ou autozygotie) réalisée que les estimations à partir des pedigrees voire que celles effectuées à partir des méthodes nécessitant l'estimation des fréquences alléliques de la population (Forutan et al., 2018). On n'a donc pas besoin des informations sur la population d'origine d'un individu pour estimer son coefficient de consanguinité.

On peut également calculer un coefficient de consanguinité (ou d'autozygotie) localement au niveau du génome afin de savoir où la consanguinité est la plus élevée par exemple. Dans ce cas, on s'intéresse seulement à une fenêtre du génome de longueur $L_{\text{fenêtre}} < L_{\text{génomique}}$ et on

comptabilise uniquement les ROH ou portions de ROH qui chevauchent cette fenêtre du génome dans le calcul de $\sum L_{ROH,i}$.

L'exemple du calcul du coefficient de consanguinité local à partir des ROH pour l'individu I de la Figure 22 est présenté dans la Figure 24.

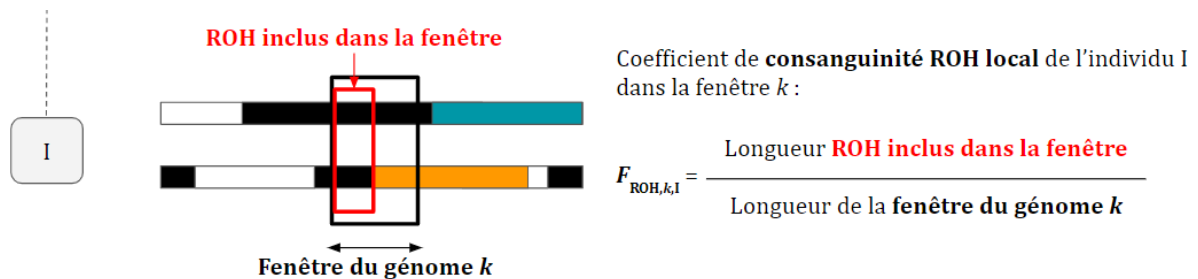


Figure 24 Exemple du calcul d'un coefficient de consanguinité local à partir des *Runs Of Homozygosity* (ROH) pour l'individu I

Une caractéristique particulièrement intéressante des ROH est le lien entre la longueur d'un ROH et l'ancienneté de l'ancêtre commun duquel il est issu. Les ROH longs correspondent à une consanguinité récente et les ROH courts à une consanguinité plus ancienne (Ferenčaković et al., 2013a; Purfield et al., 2012). Ceci est lié aux recombinaisons intra-chromosomiques ayant lieu durant la méiose, recoupant les segments autozygotes au fil des générations (Thompson, 2008). Cela permet par exemple d'estimer uniquement la consanguinité récente en calculant des coefficients de consanguinité à partir de ROH plus longs qu'un seuil donné ou la consanguinité ancienne en prenant en compte uniquement les ROH les plus courts. Il est également possible d'étudier la distribution de la longueur des ROH chez un individu ou un ensemble d'individus d'une population. Cette relation entre longueur et ancienneté de la consanguinité permet donc d'accéder à l'histoire de la population.

Les ROH permettent également d'analyser la structure du génome à l'échelle de la population. En particulier, lorsque de nombreux individus au sein d'une population partagent des ROH à une même position du génome, cette région pourrait avoir été sous sélection (Gibson et al., 2006). De plus, les ROH peuvent rester intacts sur plusieurs générations lorsqu'ils sont situés dans des régions possédant des faibles taux de recombinaison. Les ROH sont donc un outil intéressant pour comprendre comment la consanguinité et la sélection affectent le génome et quelles régions génomiques en particulier à l'échelle de la population.

D - Méthodes de détection et limites des mesures fondées sur les ROH

Plusieurs approches ont été proposées pour identifier les segments autozygotes du génome à partir de régions génomiques constituées de marqueurs homozygotes consécutifs.

Les procédures empiriques basées sur des règles visant à caractériser les ROH reposent sur la définition préalable de seuils spécifiques pour leur nombre minimal de marqueurs homozygotes et la longueur minimale des segments ainsi que la proportion maximale de marqueurs hétérozygotes autorisés (pour tenir compte des erreurs de génotypage) (Chang et al., 2015; Purcell et Chang).

Une approche statistique formelle consistant à évaluer l'autozygotie de fenêtres du génome a ensuite été proposée pour tenir compte des fréquences alléliques de la population et des taux d'erreur de génotypage (Kardos et al., 2017; Pemberton et al., 2012 ; Broman et Weber, 1999). Parallèlement à ces approches, une modélisation probabiliste a été proposée pour identifier les zones autozygotes le long des chromosomes, via un modèle de Markov caché à deux états : présence ou absence de ROH (Leutenegger et al., 2003). Les probabilités de transition entre ces deux états des marqueurs successifs dépendent de la distance génétique entre les marqueurs, d'un paramètre contrôlant le taux de changement par unité de distance génétique et du coefficient individuel de consanguinité. Cette approche à deux états revient à supposer que tous les ROH d'un individu ont la même longueur attendue et donc que tous les ancêtres communs à ses parents appartiennent à la même génération avec des chemins généalogiques de même longueur. Pour prendre en compte des sources de consanguinité plus variées, l'approche à deux états a été étendue pour différencier les ROH en plusieurs classes, chaque classe correspondant à un âge de consanguinité donné (Druet et Gautier, 2017).

Les méthodes basées sur un modèle de Markov caché donnent de bons résultats pour estimer la consanguinité des animaux (Druet et Gautier, 2017) mais sont coûteux en temps de calcul et en mémoire. De plus, elles nécessitent une bonne connaissance des fréquences alléliques de la population et nous avons déjà vu qu'estimer correctement les fréquences alléliques d'une population était compliqué.

Les méthodes de détection basées sur des règles, comme celle proposée par le logiciel PLINK (Chang et al., 2015; Purcell et Chang), nécessitent un choix minutieux des paramètres de détection comme la longueur minimale des ROH, la densité minimale de SNPs au sein d'un ROH, l'écart maximal autorisé entre deux marqueurs au sein d'un ROH, etc. Il est primordial d'ajuster ces paramètres en fonction de la population étudiée et de la densité du génotypage (Meyermans et al., 2020 ; Ferenčaković et al., 2013b ; Gibson et al., 2006). Ainsi, les longueurs et les positions des ROH peuvent varier en fonction des paramètres utilisés dans les algorithmes de détection (Gibson

et al., 2006). Une fois ces paramètres ajustés, ces méthodes donnent de bons résultats pour estimer la consanguinité des animaux (Forutan et al., 2018). De plus, la détection des ROH basée sur des règles par l'algorithme de PLINK (Chang et al., 2015; Purcell et Chang) est très rapide (Howrigan et al., 2011), ce qui en fait une méthode pratique pour l'estimation de coefficients de consanguinité de populations de grande taille.

IV - Bilan

Nous avons vu que les coefficients de consanguinité et de parenté ainsi que la taille efficace d'une population (par exemple l'ensemble des animaux d'une race donnée) étaient de bons indicateurs de la diversité au sein de ces populations. Ces indicateurs peuvent être estimés à partir de données généalogiques ou des génotypes des animaux (grâce à la comparaison entre le génotype d'un individu et les fréquences alléliques au sein de sa population, ou à la détection de ROH).

La précision des estimations de la diversité à partir des pedigrees dépend de l'exactitude et de l'exhaustivité des pedigrees. De plus, les coefficients de consanguinité ou de parenté estimés à partir des pedigrees correspondent à la valeur attendue de ces paramètres.

Utiliser les génotypes des animaux permet d'obtenir des estimations plus précises de la diversité au sein d'une population (race par exemple), en donnant accès à la consanguinité réalisée. Certaines méthodes basées sur les génotypes nécessitent une bonne estimation des fréquences alléliques de la population au risque de fausser les résultats. Le calcul de coefficients de consanguinité à partir des ROH est précis, efficace et tend à devenir une méthode standard d'estimation de la consanguinité.

On ne génotype pas tous les individus d'une population alors qu'on peut avoir accès plus facilement aux pedigrees dans la plupart des espèces de rente. Dans le cadre de la sélection génomique, des approches *single-step* sont développées pour prendre en compte l'apparentement des individus d'une population qu'ils soient génotypés ou non, par la construction d'une matrice d'apparentement commune, génomique et généalogique, notée **H** (Aguilar et al., 2010; Legarra et al., 2009; Misztal et al., 2009). L'utilisation des pedigrees pour estimer les coefficients de consanguinité et de parenté peut être combinée à celle des données moléculaires. En effet, les pedigrees peuvent parfois porter des informations sur des relations de parenté éloignées qui ne sont plus détectables au niveau génétique (Wang, 2016). De plus, les estimations de la consanguinité basées sur une combinaison des pedigrees et des données de génotypage permettent d'inclure plus d'individus dans les estimations que l'utilisation seule des données de génotypage, mais surtout de comparer les valeurs de consanguinité des individus avec et sans génotypes car ces dernières sont à la même échelle.

Liste des publications

Communications orales en congrès internationaux

Anna-Charlotte Doublet (oratrice), Pascal Croiseau, Sébastien Fritz, Chris Hozé, Alexis Michenet, Denis Laloë, Gwendal Restoux, 2018. *The impact of genomic selection on genetic diversity and genetic gain in French dairy cattle breeds*, EAAP, Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science, Volume No 24, page 199

Anna-Charlotte Doublet (oratrice), Gwendal Restoux, Sébastien Fritz, Alexis Michenet, Chris Hozé, Denis Laloë, Pascal Croiseau, 2019. *Impact of multiple ovulation and embryo transfer on genetic gain and diversity in dairy cattle*, EAAP, Book of Abstracts of the 70th Annual Meeting of the European Federation of Animal Science, Volume No 25, page 610

Katy Paul, **Anna-Charlotte Doublet**, Denis Laloë, Pascal Croiseau, Gwendal Restoux (orateur), 2019. *Evolution of ROH's distribution along the genome over a decade of genomic selection in dairy cattle*, EAAP, Book of Abstracts of the 70th Annual Meeting of the European Federation of Animal Science, Volume No 25, page 208

Articles scientifiques

Anna-Charlotte Doublet, Pascal Croiseau, Sébastien Fritz, Alexis Michenet, Chris Hozé, Coralie Danchin-Burge, Denis Laloë, Gwendal Restoux, 2019. *The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds*. *Genetics Selection Evolution*, Volume 51, Article number 52, <https://doi.org/10.1186/s12711-019-0495-1>. **Publié.**

Anna-Charlotte Doublet, Gwendal Restoux, Sébastien Fritz, Laura Balberini, Guillaume Fayolle, Chris Hozé, Denis Laloë, Pascal Croiseau, 2020. *Intensifying the use of reproductive technologies and reducing the dimensions of breeding schemes put genetic diversity at risk in dairy cattle breeds*. **Soumis.**

Activités autour de la recherche

Co-encadrement d'un stage de M2 de 6 mois (Katy Paul)

- L'évolution de la distribution des ROH le long du génome au moment du passage à la sélection génomique dans les races bovines laitières françaises de grand effectif.

Organisation du séminaire des doctorants du Département de Génétique Animale de l'INRAE

- Organisation et création des sessions thématiques,
- Création du livret d'accueil et du logo,
- Demandes de devis,
- Rédaction d'un document de retour d'expérience.

Cette activité m'a permis d'acquérir des compétences en gestion, en organisation d'évènement, en résolution de problèmes, en créativité, en prise de décision.

Trésorière de l'association des doctorants de l'INRAE de Jouy-en-Josas, Doc'J

- Organisation de réunions,
- Organisation d'un évènement sponsorisé : le colloque Doc'J,
 - Avec un concours de présentations en 180 secondes et un concours de posters,
 - Une table ronde « Intégrité scientifique, éthique de la science, lien entre sciences et société »,
- Gestion administrative et des comptes.

Cette activité m'a permis d'acquérir des compétences en gestion, en organisation de réunions et d'évènements.

Pint of Science, festival de science

- Expérience de médiation scientifique lors d'une soirée sur le thème « L'Homme au cœur de la sélection animale ».

Projet de communication avec les étudiants de première année de l'IUT de Cachan

- Echanges avec les étudiants et tournage d'une vidéo avec objectif de vulgarisation du sujet de thèse.

Participation au colloque Doc'J

- Participation aux concours de présentation de la thèse en 180 secondes et de posters, deux années consécutives.

Séminaire des doctorants du Département de Génétique Animale de l'INRAE

- Ma thèse en 180 secondes (en première année de thèse),
- Présentation de 10 minutes (en deuxième année de thèse).

Journées des doctorants ABIES

- Participation au concours de posters des journées des doctorants de l'école doctorale ABIES,
 - Obtention d'un prix de 250 € pour mon poster.

Ces activités, tout comme les communications en congrès internationaux, m'ont permis d'acquérir des compétences en communication orale et médiation scientifique.

Formations suivies pendant la thèse

Type de formation	Intitulé de la formation	Autres informations	Durée de la formation
Outils et méthodes pour bien exercer son métier de doctorant	Management de projets (modules 1 et 2)	Formation organisée par l'Anféia (organisme de formation de l'employeur)	21 heures
	Initiation à la gestion de projet, conduire sa thèse comme un projet	Formation organisée par ABIES et suivie les 13 et 14 février 2018	14 heures
	Journée de formation à la rédaction du manuscrit	Formation organisée par ABIES (suivie en visio pendant le confinement)	8 heures
Séminaires scientifiques – Communication et médiation scientifique	Colloque Doc'J - Colloque des jeunes chercheurs – Edition 2018	Colloque organisé par l'association Doc'J	8 heures
	Colloque Doc'J - Colloque des jeunes chercheurs - Edition 2019	Colloque organisé par l'association Doc'J	8 heures
	69th annual EAAP meeting - Conventional and traditional livestock production systems - new challenges	Congrès international à Dubrovnik, Croatie (2018)	35 heures
	70th Annual EAAP meeting of the European Federation of Animal Science	Congrès international à Ghent, Belgique (2019)	35 heures
	Journées des doctorants ABIES 2018	Organisées par l'école doctorale ABIES en avril 2018 à AgroParisTech	16 heures
	Journées des doctorants ABIES 2019	Organisées par l'école doctorale ABIES en avril 2019 à AgroParisTech	16 heures
Formations scientifiques pluridisciplinaires proposées par une école doctorale	Journal Club	Formation organisée par l'école doctorale ABIES à AgroParisTech Sessions réparties entre le 5 février 2018 et le 16 avril 2018	18 heures
Autre	Journée d'accueil des nouveaux doctorants	Organisée par l'école doctorale ABIES en novembre 2017 à AgroParisTech	7 heures

Formations scientifiques disciplinaires	Interactive post-graduate course on characterization, management and exploitation of genomic diversity in animals	Organisée par l'Institute of Animal Sciences, à Wageningen, Pays-Bas en octobre 2018	35 heures
Aide à l'insertion professionnelle – Définir son projet professionnel	Forum professionnel des doctorants et docteurs – Doc'Avenir 2018	Organisé par l'école doctorale ABIES en juillet 2018 à AgroParisTech	7 heures
Connaissance de l'entreprise	MOOC de la Coopération Agricole AgroParisTech, plateforme FUN (cours en ligne) - 6 heures	Formation organisée par AgroParisTech sur la plateforme FUN MOOC (cours en ligne)	6 heures
Ethique et intégrité scientifique	Intégrité scientifique dans les métiers de la recherche	Formation organisée par l'Université de Bordeaux sur la plateforme FUN MOOC (cours en ligne)	15 heures
TOTAL			249 heures

Bibliographie

Abney, M., McPeck, M.S., et Ober, C. (2000). Estimation of variance components of quantitative traits in inbred populations. *Am. J. Hum. Genet.* 66, 629–650.

AETE (2019). Commercial Embryo Transfer Activity in Europe 2018 [cité le 10 août 2020]. Disponible : <https://www.aete.eu/publications/statistics/>

Agreste (2019). Enquête annuelle laitière 2018 (Ministère de l'Agriculture et de l'Alimentation - Service de la Statistique et de la Prospective) [cité le 10 août 2020]. Disponible : agreste.agriculture.gouv.fr/agreste-web/download/publication/publie/Chd1913/Chd1913.pdf

Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A., Tsuruta, S., et Lawlor, T.J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score1. *J. Dairy Sci.* 93, 743–752.

Alice (2020). Les statistiques nationales [cité le 10 août 2020]. Disponible : <https://www.alice.fr/documents-chiffres-cles/les-donnees-cles-en-france/les-statistiques-nationales.html>

Ayroles, J.F., Hughes, K.A., Rowe, K.C., Reedy, M.M., Rodriguez-Zas, S.L., Drnevich, J.M., Cáceres, C.E., et Paige, K.N. (2009). A genomewide assessment of inbreeding depression: gene number, function, and mode of action. *Conserv. Biol. J. Soc. Conserv. Biol.* 23, 920–930.

Barrett, S.C.H., et Charlesworth, D. (1991). Effects of a change in the level of inbreeding on the genetic load. *Nature* 352, 522–524.

Benesh, D.P., Weinreich, F., Kalbe, M., et Milinski, M. (2014). Lifetime inbreeding depression, purging, and mating system evolution in a simultaneous hermaphrodite tapeworm. *Evol. Int. J. Org. Evol.* 68, 1762–1774.

Bérodier, M. (2020). Utilisation en ferme des données de génotypage pour une gestion optimisée et durable de l'élevage laitier. Université Paris-Saclay.

Bérodier, M., Berg, P., Meuwissen, T., Brochard, M., et Ducrocq, V. (2019). Improving mating plans at herd level using genomic information. (Gand, Belgique), p. 210.

Bjelland, D.W., Weigel, K.A., Vukasinovic, N., et Nkrumah, J.D. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *J. Dairy Sci.* 96, 4697–4706.

Boakes, E., et Wang, J. (2005). A simulation study on detecting purging of inbreeding depression in captive populations. *Genet. Res.* 86, 139–148.

Boichard, D., Maignel, L., et Verrier, É. (1997). The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol. GSE* 29, 5–23.

Boichard, D., Guillaume, F., Baur, A., Croiseau, P., Rossignol, M.N., Boscher, M.Y., Druet, T., Genestout, L., Colleau, J.J., Journaux, L., et al. (2012). Genomic selection in French dairy cattle. *Anim. Prod. Sci.* 52, 115.

Bouquet, A., et Juga, J. (2013). Integrating genomic selection into dairy cattle breeding programmes: a review. *Animal* 7, 705–713.

- Bouquet, A., Sørensen, A.C., et Juga, J. (2015). Genomic selection strategies to optimize the use of multiple ovulation and embryo transfer schemes in dairy cattle breeding programs. *Livest. Sci.* *174*, 18–25.
- Brans, J.P., et Vincke, Ph. (1985). Note—A Preference Ranking Organisation Method. *Manag. Sci.* *31*, 647–656.
- Broman, K.W., et Weber, J.L. (1999). Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. *Am. J. Hum. Genet.* *65*, 1493–1500.
- Bulmer, M.G. (1971). The Effect of Selection on Genetic Variability. *Am. Nat.* *105*, 201–211.
- Burdon, R.D., et Russell, J.H. (1998). Inbreeding depression in selfing experiments: statistical issues. *Forrest Genet.* *5*, 179–189.
- Caballero, A., Bravo, I., et Wang, J. (2017). Inbreeding load and purging: implications for the short-term survival and the conservation management of small populations. *Heredity* *118*, 177–185.
- de Cara, M.Á.R., Villanueva, B., Toro, M.Á., et Fernández, J. (2013a). Using genomic tools to maintain diversity and fitness in conservation programmes. *Mol. Ecol.* *22*, 6091–6099.
- de Cara, M.Á.R., Villanueva, B., Toro, M.Á., et Fernández, J. (2013b). Purging deleterious mutations in conservation programmes: combining optimal contributions with inbred matings. *Heredity* *110*, 530–537.
- Carrillo, J.A., et Siewerdt, F. (2010). Consequences of long-term inbreeding accumulation on preweaning traits in a closed nucleus Angus herd. *J. Anim. Sci.* *88*, 87–95.
- Cases, S., Smith, S.J., Zheng, Y.-W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusi, A.J., et al. (1998). Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 13018–13023.
- Cassell, B.G., Adamec, V., et Pearson, R.E. (2003). Effect of incomplete pedigrees on estimates of inbreeding and inbreeding depression for days to first service and summit milk yield in Holsteins and Jerseys. *J. Dairy Sci.* *86*, 2967–2976.
- Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., et Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* *4*, 7.
- Charlesworth, B., et Charlesworth, D. (1999). The genetic basis of inbreeding depression. *Genet. Res.* *74*, 329–340.
- Charlesworth, D., et Willis, J.H. (2009). The genetics of inbreeding depression. *Nat. Rev. Genet.* *10*, 783–796.
- CNIEL (2020). L'économie laitière en chiffres - Édition 2020 (CNIEL) [cité le 10 août 2020]. Disponible : <https://www.calameo.com/read/0022300516d0430c4176c>
- Colleau, J.J., Fritz, S., Guillaume, F., Baur, A., DUPASSIEUX, D., Boscher, M.Y., Journaux, L., Eggen, A., et Boichard, D. (2009). Simulation des potentialités de la sélection génomique chez les bovins laitiers. In 16èmes Rencontres Recherches Ruminants, (Paris, France), p. np.

Colleau, J.J., Fritz, S., Guillaume, F., Baur, A., Dupassieux, D., Boscher, M.Y., Journaux, L., Eggen, A., et Boichard, D. (2015). Simulation des potentialités de la sélection génomique chez les bovins laitiers. *INRA Prod Anim* 28, 251–258.

Croquet, C., Mayeres, P., Gillon, A., Vanderick, S., et Gengler, N. (2006). Inbreeding depression for global and partial economic indexes, production, type, and functional traits. *J. Dairy Sci.* 89, 2257–2267.

Croquet, C., Mayeres, P., Gillon, A., Hammami, H., Soyeurt, Hélène, Vanderick, S., et Gengler, N. (2007). Linear and Curvilinear Effects of Inbreeding on Production Traits for Walloon Holstein Cows. *J. Dairy Sci.* 90, 465–471.

Cruden, D. (1949). The Computation of Inbreeding Coefficients for Closed Populations. *J. Hered.* 40, 248–251.

Curie-Cohen, M. (1982). Estimates of Inbreeding in a Natural Population: A Comparison of Sampling Properties. *Genetics* 100, 339–358.

Curik, Sölkner, et Stipic (2001). The influence of selection and epistasis on inbreeding depression estimates. *J. Anim. Breed. Genet.* 118, 247–262.

Curik, I., Ferenčaković, M., Sölkner, J., Curik, I., Ferenčaković, M., et Sölkner, J. (2017). Genomic dissection of inbreeding depression: a gate to new opportunities. *Rev. Bras. Zootec.* 46, 773–782.

Danchin-Burge, C. (2018). Indicateurs de variabilité génétique - races ovines allaitantes - Edition 2018.

Danchin-Burge, C. (2020). Indicateurs de variabilité génétique - races bovines - Edition 2020.

Darwin, C. (1876). *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom* (London: J. Murray).

Darwin, C. (1868). *The variation of animals and plants under domestication. Volume 2* (London: J. Murray).

Davenport, C.B. (1908). Degeneration, Albinism and Inbreeding. *Science* 28, 454–455.

de Roos, A.P.W., Schrooten, C., Veerkamp, R.F., et van Arendonk, J.A.M. (2011). Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. *J. Dairy Sci.* 94, 1559–1567.

Dickerson, G.E., et Hazel, L.N. (1944). Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J. Agric. Res.* 69, 459–476.

Doekes, H.P. (2020). Genomic characterization and conservation of genetic diversity in cattle. Wageningen University [cité le 29 septembre 2020]. Disponible : <https://research.wur.nl/en/publications/0b015e68-d5b0-47d8-a2f9-ffff721b27cb>

Doekes, H.P., Veerkamp, R.F., Bijma, P., Hiemstra, S.J., et Windig, J. (2018). Value of the Dutch Holstein Friesian germplasm collection to increase genetic variability and improve genetic merit. *J. Dairy Sci.* 101, 10022–10033.

Doekes, H.P., Veerkamp, R.F., Bijma, P., de Jong, G., Hiemstra, S.J., et Windig, J.J. (2019). Inbreeding depression due to recent and ancient inbreeding in Dutch Holstein–Friesian dairy cattle. *Genet. Sel. Evol.* 51, 54.

- Doublet, A.-C., Croiseau, P., Fritz, S., Michenet, A., Hozé, C., Danchin-Burge, C., Laloë, D., et Restoux, G. (2019). The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genet. Sel. Evol.* *51*, 52.
- Druet, T., et Gautier, M. (2017). A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Mol. Ecol.* *26*, 5820–5841.
- Druet, T., Macleod, I.M., et Hayes, B.J. (2014). Toward genomic prediction from whole-genome sequence data: impact of sequencing design on genotype imputation and accuracy of predictions. *Heredity* *112*, 39–47.
- Dubois, S. (2016). Le transfert embryonnaire chez les bovins. Thèse pour obtenir le grade de Docteur Vétérinaire. Vetagro Sup - Campus vétérinaire de Lyon.
- Dumont, B., Fortun-Lamothe, L., Jouven, M., Thomas, M., et Tichit, M. (2013). Prospects from agroecology and industrial ecology for animal production in the 21st century. *Anim. Int. J. Anim. Biosci.* *7*, 1028–1043.
- Engelsma, K.A., Veerkamp, R.F., Calus, M.P.L., et Windig, J.J. (2011). Consequences for diversity when prioritizing animals for conservation with pedigree or genomic information: Prioritization with pedigree or genomic data. *J. Anim. Breed. Genet.* *128*, 473–481.
- Ercanbrack, S.K., et Knight, A.D. (1991). Effects of inbreeding on reproduction and wool production of Rambouillet, Targhee, and Columbia ewes. *J. Anim. Sci.* *69*, 4734–4744.
- Eynard, S.E., Windig, J.J., Leroy, G., van Binsbergen, R., et Calus, M. (2015). The effect of rare alleles on estimated genomic relationships from whole genome sequence data. *BMC Genet.* *16*, 24.
- Eynard, S.E., Windig, J.J., Hiemstra, S.J., et Calus, M.P.L. (2016). Whole-genome sequence data uncover loss of genetic diversity due to selection. *Genet. Sel. Evol.* *48*.
- Eynard, S.E., Croiseau, P., Laloë, D., Fritz, S., Calus, M.P.L., et Restoux, G. (2018a). Which Individuals To Choose To Update the Reference Population? Minimizing the Loss of Genetic Diversity in Animal Genomic Selection Programs. *G3 Bethesda Md* *8*, 113–121.
- Eynard, S.E., Windig, J.J., Hulsege, I., Hiemstra, S.-J., et Calus, M.P.L. (2018b). The impact of using old germplasm on genetic merit and diversity—A cattle breed case study. *J. Anim. Breed. Genet.* *135*, 311–322.
- FAO (2007). The state of the world's animal genetic resources for food and agriculture (Rome: Commission on Genetic Resources for Food and Agriculture, Food and Agriculture Organization of the United Nations) [cité le 17 juin 2020]. Disponible : <http://www.fao.org/3/a-a1250e.pdf>
- Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T.R., Klemetsdal, G., Curik, I., et Sölkner, J. (2013a). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J. Anim. Breed. Genet.* *130*, 286–293.
- Ferenčaković, M., Sölkner, J., et Curik, I. (2013b). Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genet. Sel. Evol.* *45*, 42.
- Ferenčaković, M., Sölkner, J., Kapš, M., et Curik, I. (2017). Genome-wide mapping and estimation of inbreeding depression of semen quality traits in a cattle population. *J. Dairy Sci.* *100*, 4721–4730.

- Ferket, P.R., van Heugten, E., van Kempen, T. a. T.G., et Angel, R. (2002). Nutritional strategies to reduce environmental emissions from nonruminants. *J. Anim. Sci.* *80*, E168–E182.
- Fisher, R.A. (1918). The correlation between relatives on the supposition of mendelian inheritance. *Trans. Roy. Soc.* *52*, 399–433.
- Forutan, M., Ansari Mahyari, S., Baes, C., Melzer, N., Schenkel, F.S., et Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genomics* *19*, 98.
- France Génétique Elevage. Races bovines laitières [cité le 17 juin 2020]. Disponible : <http://fr.france-genetique-elevage.org/-Races-bovines-laitieres-.html>
- FranceAgriMer (2020). Les marchés des produits laitiers, carnés et avicoles - Bilan 2019 - Perspectives 2020 [cité le 10 août 2020]. Disponible : <https://www.franceagrimer.fr/content/download/63165/document/BIL-VIA-LAI-Bilan2019-Perspectives2020.pdf>
- Frantz, L.A.F., Mullin, V.E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., Linderholm, A., Mattiangeli, V., Teasdale, M.D., Dimopoulos, E.A., et al. (2016). Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* *352*, 1228–1231.
- Fritz, S., Capitan, A., Djari, A., Rodriguez, S.C., Barbat, A., Baur, A., Grohs, C., Weiss, B., Boussaha, M., Esquerré, D., et al. (2013). Detection of Haplotypes Associated with Prenatal Death in Dairy Cattle and Identification of Deleterious Mutations in GART, SHBG and SLC37A2. *PLOS ONE* *8*, e65550.
- Galli, C., Crotti, G., Notari, C., Turini, P., Duchi, R., et Lazzari, G. (2001). Embryo production by ovum pick up from live donors. *Theriogenology* *55*, 1341–1357.
- Galli, C., Duchi, R., Crotti, G., Turini, P., Ponderato, N., Colleoni, S., Lagutina, I., et Lazzari, G. (2003). Bovine embryo technologies. *Theriogenology* *59*, 599–616.
- Gama, L.T., et Smith, C. (1993). The role of inbreeding depression in livestock production systems. *Livest. Prod. Sci.* *36*, 203–211.
- Gautier, M., Capitan, A., Fritz, S., Eggen, A., Boichard, D., et Druet, T. (2007). Characterization of the *DGAT1* K232A and Variable Number of Tandem Repeat Polymorphisms in French Dairy Cattle. *J. Dairy Sci.* *90*, 2980–2988.
- Gengler, N., Mayeres, P., et Szydlowski, M. (2007). A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal* *1*, 21–28.
- Gibson, J., Morton, N.E., et Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* *15*, 789–795.
- Gómez, M.D., Valera, M., Molina, A., Gutiérrez, J.P., et Goyache, F. (2009). Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. *Livest. Sci.* *122*, 149–155.
- Gómez-Romano, F., Villanueva, B., Fernández, J., Woolliams, J.A., et Pong-Wong, R. (2016). The use of genomic coancestry matrices in the optimisation of contributions to maintain genetic diversity at specific regions of the genome. *Genet. Sel. Evol.* *48*, 2.

- González-Recio, O., López de Maturana, E., et Gutiérrez, J.P. (2007). Inbreeding depression on female fertility and calving ease in Spanish dairy cattle. *J. Dairy Sci.* 90, 5744–5752.
- Gouyon, P.-H. (2002). L'introduction de la génétique en France. (Colloque : L'amélioration des plantes, continuités et ruptures, Montpellier), p. 6.
- Graham, G.I., Wolff, D.W., et Stuber, C.W. (1997). Characterization of a Yield Quantitative Trait Locus on Chromosome Five of Maize by Fine Mapping. *Crop Sci.* 37, crops1997.0011183X003700050033x.
- Granleese, T., Clark, S.A., Kinghorn, B.P., et Werf, J.H.J. van der (2019). Optimizing female allocation to reproductive technologies considering merit, inbreeding and cost in nucleus breeding programmes with genomic selection. *J. Anim. Breed. Genet.* 136, 79–90.
- Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., et al. (2002). Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine *DGAT1* Gene with Major Effect on Milk Yield and Composition. *Genome Res.* 12, 222–231.
- Groupe Economie du Bétail (2018). Chiffres-clés Bovins 2018 - Productions lait et viande [cité le 22 novembre 2018]. Disponible : <https://abcis.com/wp-content/uploads/2019/07/2018-Chiffres-cles-Bovins-par-page-A4.pdf>
- Groupe Economie du Bétail (2019). Les chiffres clés du GEB - bovins 2019 - Productions lait et viande (Institut de l'Élevage et Confédération Nationale de l'Élevage) [cité le 17 juin 2020]. Disponible : <http://idele.fr/filieres/bovin-viande/publication/idelesolr/recommends/chiffres-cles-bovins-2019.html>
- Guillou, M., Guyomard, H., Huyghe, C., Peyraud, J.-L., Vert, J., et Claquin, P. (2013). Le projet agro-écologique : Vers des agricultures doublement performantes pour concilier compétitivité et respect de l'environnement [cité le 15 juin 2020]. Disponible : <https://agriculture.gouv.fr/telecharger/42709?token=3967e95baa153da2cdc469d2c44bd697>
- Gutiérrez, J.P., Cervantes, I., et Goyache, F. (2009). Improving the estimation of realized effective population sizes in farm animals. *J. Anim. Breed. Genet.* 126, 327–332.
- Harris, K., Zhang, Y., et Nielsen, R. (2019). Genetic rescue and the maintenance of native ancestry. *Conserv. Genet.* 20, 59–64.
- Hayes, B.J., Bowman, P.J., Chamberlain, A.J., et Goddard, M.E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92, 433–443.
- Hedrick, P.W. (1994). Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity* 73, 363–372.
- Herrero, M., Havlik, P., Valin, H., Notenbaert, A., Rufino, M.C., Thornton, P.K., Blummel, M., Weiss, F., Grace, D., et Obersteiner, M. (2013). Biomass use, production, feed efficiencies, and greenhouse gas emissions from global livestock systems. *Proc. Natl. Acad. Sci.* 110, 20888–20893.
- Hill, W.G., et Weir, B.S. (2011). Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet. Res.* 93, 47–64.

Howard, J.T., Haile-Mariam, M., Pryce, J.E., et Maltecca, C. (2015). Investigation of regions impacting inbreeding depression and their association with the additive genetic effect for United States and Australia Jersey dairy cattle. *BMC Genomics* 16.

Howard, J.T., Pryce, J.E., Baes, C., et Maltecca, C. (2017). Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability. *J. Dairy Sci.* 100, 6009–6024.

Howrigan, D.P., Simonson, M.A., et Keller, M.C. (2011). Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics* 12, 460.

IDF (2019). Bulletin of the IDF N° 501/ 2019: The World Dairy Situation 2019 (International Dairy Federation).

Insee (2019). Le compte prévisionnel de l'agriculture pour 2019 (Insee) [cité le 10 août 2020]. Disponible : <https://www.insee.fr/fr/statistiques/4267450>

Institut de l'Élevage (2020a). BGIAP 2019 Bilan Génétique de l'Insémination en races bovines laitières [cité le 10 août 2020]. Disponible : http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/bilan-genetique-de-linsemination-en-races-bovines-laitieres-3.html

Institut de l'Élevage (2020b). Résultats de contrôle laitier - Espèce bovine - France 2019 (Institut de l'Élevage) [cité le 10 août 2020]. Disponible : http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/resultats-de-contrôle-laitier-france-2019.html

Institut de l'Élevage (2020c). Forte dégradation de la fécondité bovine en 2017-2018 [cité le 18 septembre 2020]. Disponible : http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/reproscope-performances-2017-2018-une-forte-degradation-de-la-fecondite.html

Jain, S.K., et Allard, R.W. (1966). The Effects of Linkage, Epistasis, and Inbreeding on Population Changes Under Selection. *Genetics* 53, 633–659.

Jenko, J., McClure, M.C., Matthews, D., McClure, J., Johnsson, M., Gorjanc, G., et Hickey, J.M. (2019). Analysis of a large dataset reveals haplotypes carrying putatively recessive lethal and semi-lethal alleles with pleiotropic effects on economically important traits in beef cattle. *Genet. Sel. Evol.* 51, 9.

Johnsson, M., Gaynor, R.C., Jenko, J., Gorjanc, G., de Koning, D.-J., et Hickey, J.M. (2019). Removal of alleles by genome editing (RAGE) against deleterious load. *Genet. Sel. Evol.* 51.

Kalinowski, S.T., Hedrick, P.W., et Miller, P.S. (2000). Inbreeding depression in the Speke's gazelle captive breeding program. *Conserv. Biol.* 14, 1375–1384.

Kardos, M., Luikart, G., et Allendorf, F.W. (2015). Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity* 115, 63–72.

Kardos, M., Qvarnström, A., et Ellegren, H. (2017). Inferring Individual Inbreeding and Demographic History from Segments of Identity by Descent in *Ficedula* Flycatcher Genome Sequences. *Genetics* 205, 1319–1334.

- Keller, M.C., Visscher, P.M., et Goddard, M.E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* 189, 237–249.
- Keller, M.C., Simonson, M.A., Ripke, S., Neale, B.M., Gejman, P.V., Howrigan, D.P., Lee, S.H., Lencz, T., Levinson, D.F., Sullivan, P.F., et al. (2012). Runs of Homozygosity Implicate Autozygosity as a Schizophrenia Risk Factor. *PLoS Genet.* 8, e1002656.
- Kempthorne, O. (1957). An introduction to genetic statistics (Iowa State University Press).
- Kim, E.-S., Cole, J.B., Huson, H., Wiggans, G.R., Tassell, C.P.V., Crooker, B.A., Liu, G., Da, Y., et Sonstegard, T.S. (2013). Effect of Artificial Selection on Runs of Homozygosity in U.S. Holstein Cattle. *PLOS ONE* 8, e80813.
- Kristensen, T.N., et Sørensen, A.C. (2005). Inbreeding – lessons from animal breeding, evolutionary biology and conservation genetics. *Anim. Sci.* 80, 121–133.
- Kristensen, T.N., Pedersen, K.S., Vermeulen, C.J., et Loeschcke, V. (2010). Research on inbreeding in the “omic” era. *Trends Ecol. Evol.* 25, 44–52.
- Kühn, C., Thaller, G., Winter, A., Bininda-Emonds, O.R.P., Kaupe, B., Erhardt, G., Bennewitz, J., Schwerin, M., et Fries, R. (2004). Evidence for Multiple Alleles at the *DGAT1* Locus Better Explains a Quantitative Trait Locus With Major Effect on Milk Fat Content in Cattle. *Genetics* 167, 1873–1881.
- Lacy, R.C., Alaks, G., et Walsh, A. (1996). Hierarchical analyses of inbreeding depression in *Peromyscus polionotus*. *Evol. Int. J. Org. Evol.* 50, 2187–2200.
- Laloë, D. (1993). Precision and information in linear models of genetic evaluation. *Genet. Sel. Evol.* 25, 557.
- Laloë, D. (2011). La genèse et le développement des concepts de l'évaluation génétique classique. *INRAE Prod. Anim.* 24, 323–330.
- Lamberson, W.R., et Thomas, D.L. (1984). Effects of inbreeding in sheep: a review. *Anim. Breed. Abstr.* 52, 287–297.
- Land, R.B., et Hill, W.G. (1975). The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Anim. Sci.* 21, 1–12.
- Lande, R., et Schemske, D.W. (1985). The Evolution of Self-Fertilization and Inbreeding Depression in Plants. I. Genetic Models. *Evolution* 39, 24–40.
- Launay, A. (2015). Évaluation génétique - Méthodes et Définitions (Institut de l'Élevage).
- Leberg, P.L., et Firmin, B.D. (2008). Role of inbreeding depression and purging in captive breeding and restoration programmes. *Mol. Ecol.* 17, 334–343.
- Legarra, A., Aguilar, I., et Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92, 4656–4663.
- Leroy, G. (2011). Genetic diversity, inbreeding and breeding practices in dogs: results from pedigree analyses. *Vet. J. Lond. Engl.* 1997 189, 177–182.

Leroy, G. (2014). Inbreeding depression in livestock species: review and meta-analysis. *Anim. Genet.* *45*, 618–628.

Leroy, G., Danchin-Burge, C., Palhiere, I., Baumung, R., Fritz, S., Mériaux, J.C., et Gautier, M. (2012). An ABC estimate of pedigree error rate: application in dog, sheep and cattle breeds. *Anim. Genet.* *43*, 309–314.

Leutenegger, A.-L., Prum, B., Génin, E., Verny, C., Lemainque, A., Clerget-Darpoux, F., et Thompson, E.A. (2003). Estimation of the inbreeding coefficient through use of genomic data. *Am. J. Hum. Genet.* *73*, 516–523.

Loi sur l'élevage (1966). Loi n°66-1005 du 28 décembre 1966 sur l'élevage - Légifrance.

Løvendahl, P., Difford, G.F., Li, B., Chagunda, M.G.G., Huhtanen, P., Lidauer, M.H., Lassen, J., et Lund, P. (2018). Review: Selecting for improved feed efficiency and reduced methane emissions in dairy cattle. *Animal* *12*, s336–s349.

Lynch, M. (1988). Design and Analysis of Experiments on Random Drift and Inbreeding Depression. *Genetics* *120*, 791–807.

MacLeod, I.M., Bowman, P.J., Vander Jagt, C.J., Haile-Mariam, M., Kemper, K.E., Chamberlain, A.J., Schrooten, C., Hayes, B.J., et Goddard, M.E. (2016). Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* *17*, 144.

Malécot, G. (1948). *Les Mathématiques de l'Hérédité* (Paris, France: Masson).

Mandal, A., Pant, K.P., Rout, P.K., et Roy, R. (2004). Effects of Inbreeding on Lamb Survival in a Flock of Muzaffarnagari Sheep. *Asian-Australas. J. Anim. Sci.* *17*, 594–597.

Martin, P., Taussat, S., Vinet, A., Krauss, D., Maupetit, D., et Renand, G. (2019). Genetic parameters and genome-wide association study regarding feed efficiency and slaughter traits in Charolais cows. *J. Anim. Sci.* *97*, 3684–3698.

Masel, J. (2011). Genetic drift. *Curr. Biol.* *CB 21*, R837-838.

Mc Parland, S., Kearney, F., et Berry, D.P. (2009). Purging of inbreeding depression within the Irish Holstein-Friesian population. *Genet. Sel. Evol.* *41*, 16.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., et Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biol.* *17*, 122.

McQuillan, R., Leutenegger, A.-L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., et al. (2008). Runs of homozygosity in European populations. *Am. J. Hum. Genet.* *83*, 359–372.

Merton, J.S., de Roos, A.P.W., Mullaart, E., de Ruigh, L., Kaal, L., Vos, P.L.A.M., et Dieleman, S.J. (2003). Factors affecting oocyte quality and quantity in commercial application of embryo technologies in the cattle breeding industry. *Theriogenology* *59*, 651–674.

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., Hayes, B.J., et Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* *157*, 1819–1829.

Meyermans, R., Gorssen, W., Buys, N., et Janssens, S. (2020). How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics* 21, 94.

Ministère de l'agriculture, de l'agroalimentaire et de la forêt (2016). Le règlement zootechnique de l'Union Européenne (Ministère de l'agriculture, de l'agroalimentaire et de la forêt) [cité le 10 août 2020]. Disponible : <https://agriculture.gouv.fr/le-reglement-zootechnique-europeen-un-nouveau-cadre-reglementaire-europeen-pour-la-genetique-animale>

Misztal, I., Lawlor, T., et Gengler, N. (1997). Relationships among estimates of inbreeding depression, dominance and additive variance for linear traits in Holsteins. *Genet. Sel. Evol.* 29, 319–326.

Misztal, I., Legarra, A., et Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J. Dairy Sci.* 92, 4648–4655.

Montbéliarde Association L'ISU, l'outil de pilotage de la sélection montbéliarde [cité le 10 août 2020]. Disponible : <https://www.montbeliarde.org/objectifs-de-selection.html>

Neto, A.S.C., Sanches, B.V., Binelli, M., Seneda, M.M., Perri, S.H., et Garcia, J.F. (2005). Improvement in embryo recovery using double uterine flushing. *Theriogenology* 63, 1249–1255.

O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W., et Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biol. Conserv.* 133, 42–51.

Organisme de Sélection en race Normande (2018). Objectifs de sélection - OS Race Normande [cité le 10 août 2020]. Disponible : https://www.lanormande.com/les_objectifs_de_la_selection.html

Pausch, H., Schwarzenbacher, H., Burgstaller, J., Flisikowski, K., Wurmser, C., Jansen, S., Jung, S., Schnieke, A., Wittek, T., et Fries, R. (2015). Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. *BMC Genomics* 16, 312.

Pedersen, L.D., Kargo, M., Berg, P., Voergaard, J., Buch, L.H., et Sørensen, A.C. (2012). Genomic selection strategies in dairy cattle breeding programmes: Sexed semen cannot replace multiple ovulation and embryo transfer as superior reproductive technology. *J. Anim. Breed. Genet.* 129, 152–163.

Pemberton, T.J., Absher, D., Feldman, M.W., Myers, R.M., Rosenberg, N.A., et Li, J.Z. (2012). Genomic patterns of homozygosity in worldwide human populations. *Am. J. Hum. Genet.* 91, 275–292.

Pérez-Enciso, M. (2014). Genomic relationships computed from either next-generation sequence or array SNP data. *J. Anim. Breed. Genet. Z. Tierzucht Zuchtungsbiologie* 131, 85–96.

Phocas, F., Belloc, C., Bidanel, J., Delaby, L., Dourmad, J.-Y., Dumont, B., Ezanno, P., Fortun-Lamothe, L., Foucras, G., Frappat, B., et al. (2017). Quels programmes d'amélioration génétique des animaux pour des systèmes d'élevage agro-écologiques ? *INRAE Prod. Anim.* 30, 31–46.

Pickering, B.M., et Willis, A.E. (2005). The implications of structured 5' untranslated regions on translation and disease. *Semin. Cell Dev. Biol.* 16, 39–47.

Prim'Holstein France (2020). Programme de sélection en race Prim'holstein | Prim'Holstein France [cité le 10 août 2020]. Disponible : <https://primholstein.com/la-prim-holstein/programme-de-selection/>

- Pryce, J.E., Goddard, M.E., Raadsma, H.W., et Hayes, B.J. (2010). Deterministic models of breeding scheme designs that incorporate genomic selection. *J. Dairy Sci.* *93*, 5455–5466.
- Pryce, J.E., Haile-Mariam, M., Goddard, M.E., et Hayes, B.J. (2014). Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet. Sel. Evol.* *46*, 71.
- Purcell, S., et Chang, C. PLINK 1.9 [cité le 3 octobre 2018]. Disponible : www.cog-genomics.org/plink/1.9/
- Purfield, D.C., Berry, D.P., McParland, S., et Bradley, D.G. (2012). Runs of homozygosity and population history in cattle. *BMC Genet.* *13*, 70.
- Qi, M., Yao, Y., Ma, H., Wang, J., Zhao, X., Liu, L., Tang, X., Zhang, L., Zhang, S., et Sun, F. (2013). Transvaginal Ultrasound-guided Ovum Pick-up(OPU) in Cattle. *J. Biomim. Biomater. Tissue Eng.* *18*.
- Races de France (2014a). Race Bovine Montbéliarde - Races de France [cité le 10 août 2020]. Disponible : <https://www.racesdefrance.fr/bovins?id=108>
- Races de France (2014b). Race Bovine Normande - Races de France [cité le 10 août 2020]. Disponible : <https://www.racesdefrance.fr/bovins?id=109>
- Races de France (2014c). Race Bovine Prim'Holstein - Races de France [cité le 10 août 2020]. Disponible : <https://www.racesdefrance.fr/bovins/races-laitieres?id=111>
- Ralls, K., et Ballou, J. (1986). Captive breeding programs for populations with a small number of founders. *Trends Ecol. Evol.* *1*, 19–22.
- Ramayo-Caldas, Y., Zingaretti, L., Popova, M., Estellé, J., Bernard, A., Pons, N., Bellot, P., Mach, N., Rau, A., Roume, H., et al. (2020). Identification of rumen microbial biomarkers linked to methane emission in Holstein dairy cows. *J. Anim. Breed. Genet.* *137*, 49–59.
- Reed, D.H., et Frankham, R. (2003). Correlation between fitness et genetic diversity. *Conserv. Biol.* *17*, 230–237.
- Renand, G., Vinet, A., Decruyenaere, V., Maupetit, D., et Dozias, D. (2019). Methane and Carbon Dioxide Emission of Beef Heifers in Relation with Growth and Feed Efficiency. *Animals* *9*, 1136.
- Restoux, G., Huot de Longchamp, P., Fady, B., et Klein, E.K. (2012). Variability of individual genetic load: consequences for the detection of inbreeding depression. *Genetica* *140*, 39–51.
- Ritland, K. (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genet. Res.* *67*, 175–185.
- Robertson, I., et Nelson, R.E. (2010). Certification and identification of embryos (Chapter 9). In *Manual of the International Transfer Society*, pp. 86–105.
- Roff, D.A. (2002). Inbreeding Depression: Tests of the Overdominance and Partial Dominance Hypotheses. *Evolution* *56*, 768–775.
- Rosen, B.D., Bickhart, D.M., Schnabel, R.D., Koren, S., Elsik, C.G., Tseng, E., Rowan, T.N., Low, W.Y., Zimin, A., Couldrey, C., et al. (2020). De novo assembly of the cattle reference genome with single-molecule sequencing. *GigaScience* *9*.

- Santana, M.L., Oliveira, P.S., Eler, J.P., Gutiérrez, J.P., et Ferraz, J.B.S. (2012). Pedigree analysis and inbreeding depression on growth traits in Brazilian Marchigiana and Bonsmara breeds. *J. Anim. Sci.* *90*, 99–108.
- Saravanan, K.A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., et Mishra, B.P. (2020). Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livest. Sci.* *241*, 104257.
- Shull, G.H. (1908). The Composition of a Field of Maize. *J. Hered.* *os-4*, 296–301.
- Silva, M.H.M.A. da, Malhado, C.H.M., Kern, E.L., Daltro, D. dos S., Cobuci, J.A., Carneiro, P.L.S., Silva, M.H.M.A. da, Malhado, C.H.M., Kern, E.L., Daltro, D. dos S., et al. (2019). Inbreeding depression in Holstein cattle in Brazil. *Rev. Bras. Zootec.* *48*.
- Smith, J.M., et Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet. Res.* *23*, 23–35.
- Smith, L.A., Cassell, B.G., et Pearson, R.E. (1998). The effects of inbreeding on the lifetime performance of dairy cattle. *J. Dairy Sci.* *81*, 2729–2737.
- Templeton, A.R., et Read, B. (1984). Factors eliminating inbreeding depression in a captive herd of speke's gazelle (*Gazella spekei*). *Zoo Biol.* *3*, 177–199.
- Templeton, A.R., Hemmer, H., Mace, G., Seal, U.S., Shields, W.M., et Woodruff, D.S. (1986). Local adaptation, coadaptation, and population boundaries. *Zoo Biol.* *5*, 115–125.
- Thomas, G., Bourrigan, X., et Astruc, J.-M. (2018). Résultats de Contrôle Laitier - France 2017.
- Thomas, M., Fortun-Lamothe, L., Jouven, M., Tichit, M., González-García, E., Dourmad, J.-Y., et Dumont, B. (2014). Agro-écologie et écologie industrielle : deux alternatives complémentaires pour les systèmes d'élevage de demain. *Prod. Anim.* *27*, 89–100.
- Thomasen, J.R., Willam, A., Egger-Danner, C., et Sørensen, A.C. (2016). Reproductive technologies combine well with genomic selection in dairy breeding programs. *J. Dairy Sci.* *99*, 1331–1340.
- Thompson, E.A. (2008). The IBD process along four chromosomes. *Theor. Popul. Biol.* *73*, 369–373.
- VanRaden, P.M. (2007). Genomic measures of relationship and inbreeding. *Interbull Bull.* 33–33.
- VanRaden, P.M., Olson, K.M., Null, D.J., et Hutchison, J.L. (2011). Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* *94*, 6153–6161.
- van Wyk, J.B., Erasmus, G.J., et Konstantinov, K.V. (1993). Inbreeding in the Elsenburg Dormer sheep stud. *South Afr. J. Anim. Sci.* *23*, 77–80.
- van Wyk, J.B., Fair, M.D., et Cloete, S.W.P. (2009). Case study: The effect of inbreeding on the production and reproduction traits in the Elsenburg Dormer sheep stud. *Livest. Sci.* *120*, 218–224.
- Wang, J. (2016). Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? *Theor. Popul. Biol.* *107*, 4–13.
- Wang, L., Janss, L.L., Madsen, P., Henshall, J., Huang, C.-H., Marois, D., Alemu, S., Sørensen, A., et Jensen, J. (2020). Effect of genomic selection and genotyping strategy on estimation of variance components in animal models using different relationship matrices. *Genet. Sel. Evol.* *52*, 31.

WHFF (2019). Statistics - WHFF • World Holstein Friesian Federation [cité le 10 août 2020]. Disponible : <https://www.whff.info/documentation/statistics.php>

Wiener, G., Lee, G.J., et Woolliams, J.A. (1992). Effects of rapid inbreeding and of crossing of inbred lines on the body weight growth of sheep. *Anim. Sci.* 55, 89-99.

Willems, O.W., Miller, S.P., et Wood, B.J. (2013). Aspects of selection for feed efficiency in meat producing poultry. *Worlds Poult. Sci. J.* 69, 77-88.

Wright, S. (1921). Systems of Mating. I. the Biometric Relations between Parent and Offspring. *Genetics* 6, 111-123.

Wright, S. (1923). Mendelian Analysis of the Pure Breeds of Livestock. The Measurement of Inbreeding and Relationship. *J. Hered.* 14, 339-348.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics* 16, 97-159.

Wright, S. (1968). *Evolution and the Genetics of Populations, Volume 3: Experimental Results and Evolutionary Deductions* (University of Chicago Press).

Annexes

Annexe 1 Positions and effects of additive quantitative trait loci (QTLs) along the simulated genome for the simulated embryo transfer French dairy cattle breeding scheme and its different variations.

Chromosome	SNP number on the chromosome	Effect of homozygosity at allele 0	Effect of heterozygosity	Effect of homozygosity at allele 1
1	13	2.65000902261731	0	-2.65000902261731
1	46	4.04321501613447	0	-4.04321501613447
1	279	5.16256017458212	0	-5.16256017458212
1	381	4.95766225521855	0	-4.95766225521855
1	399	0.0418184298480472	0	-0.0418184298480472
1	433	1.04085002231946	0	-1.04085002231946
1	440	4.54842993036214	0	-4.54842993036214
1	591	2.49129481738992	0	-2.49129481738992
1	611	0.811873097958928	0	-0.811873097958928
1	726	0.9182183924319	0	-0.9182183924319
1	743	6.2473360237838	0	-6.2473360237838
1	770	0.459709759477787	0	-0.459709759477787
1	944	0.0200970738521899	0	-0.0200970738521899
1	1264	3.86063998210712	0	-3.86063998210712
1	1323	6.09864767954884	0	-6.09864767954884
1	1394	0.0422957321416645	0	-0.0422957321416645
1	1408	0.228635452610289	0	-0.228635452610289
1	1517	0.0499054738620643	0	-0.0499054738620643
1	1545	0.0881619630581057	0	-0.0881619630581057
1	1565	1.28046952754945	0	-1.28046952754945
1	1634	0.718151704678496	0	-0.718151704678496
1	1805	1.85132620529803	0	-1.85132620529803
1	1856	17.428609811555	0	-17.428609811555
1	1997	0.282781534070361	0	-0.282781534070361
1	2002	0.0808089461934074	0	-0.0808089461934074
1	2028	6.05886114934598	0	-6.05886114934598
1	2034	3.02889878730205	0	-3.02889878730205
1	2083	2.34276279349919	0	-2.34276279349919
1	2099	0.42214600848486	0	-0.42214600848486
1	2131	1.97810897613828	0	-1.97810897613828
1	2145	0.00341147639856452	0	-0.00341147639856452
1	2343	2.6242107816887	0	-2.6242107816887
1	2348	0.157514416891878	0	-0.157514416891878
1	2600	2.95081964413072	0	-2.95081964413072
1	2629	1.10848650942877	0	-1.10848650942877
1	2633	0.02575662432519	0	-0.02575662432519
1	2673	4.71005052924318	0	-4.71005052924318
2	72	0.112548079640944	0	-0.112548079640944
2	254	3.01219887852302	0	-3.01219887852302
2	496	0.0432386510826053	0	-0.0432386510826053

2	529	0.339047188957978	0	-0.339047188957978
2	629	0.294409596941935	0	-0.294409596941935
2	657	4.26081724317158	0	-4.26081724317158
2	673	1.34017301322324	0	-1.34017301322324
2	732	0.86520950016583	0	-0.86520950016583
2	745	0.103120484448188	0	-0.103120484448188
2	790	0.00386900994385823	0	-0.00386900994385823
2	890	1.05796007916797	0	-1.05796007916797
2	926	8.04779536384494	0	-8.04779536384494
2	1091	0.0127882733632495	0	-0.0127882733632495
2	1250	0.000253233285360206	0	-0.000253233285360206
2	1309	14.3239440254155	0	-14.3239440254155
2	1469	1.27012316852972	0	-1.27012316852972
2	1481	0.814216538263928	0	-0.814216538263928
2	1483	0.318317111037001	0	-0.318317111037001
2	1532	3.40786120417134	0	-3.40786120417134
2	1561	0.279661843364689	0	-0.279661843364689
2	1571	6.12149094875317	0	-6.12149094875317
2	1659	0.208311217867508	0	-0.208311217867508
2	1716	3.25591835179511	0	-3.25591835179511
2	1718	1.49145201918693	0	-1.49145201918693
2	2002	1.46403222054004	0	-1.46403222054004
2	2129	0.00390754731534007	0	-0.00390754731534007
2	2138	0.7630421798414	0	-0.7630421798414
3	70	0.00316231083825047	0	-0.00316231083825047
3	124	0.401302787045247	0	-0.401302787045247
3	218	2.7680357869227	0	-2.7680357869227
3	255	2.75369737576732	0	-2.75369737576732
3	373	0.245428282711005	0	-0.245428282711005
3	380	0.591404689731204	0	-0.591404689731204
3	477	0.0516356899931795	0	-0.0516356899931795
3	564	3.1159181902466	0	-3.1159181902466
3	602	0.0189630716233304	0	-0.0189630716233304
3	627	1.66870275457167	0	-1.66870275457167
3	658	1.5353369423066	0	-1.5353369423066
3	713	2.97991699755012	0	-2.97991699755012
3	1089	0.400195899865274	0	-0.400195899865274
3	1133	0.449583702328462	0	-0.449583702328462
3	1141	1.90409099141674	0	-1.90409099141674
3	1459	14.0800168561054	0	-14.0800168561054
3	1474	0.00586962719115667	0	-0.00586962719115667
3	1699	0.450097069595996	0	-0.450097069595996
3	1721	0.427682545877889	0	-0.427682545877889
3	1728	11.1298900298365	0	-11.1298900298365
3	1789	0.77952623747326	0	-0.77952623747326
3	2018	0.0747149354397176	0	-0.0747149354397176

4	67	2.46064092805983	0	-2.46064092805983
4	72	0.235449886992034	0	-0.235449886992034
4	159	0.33267114799173	0	-0.33267114799173
4	176	1.05345761361398	0	-1.05345761361398
4	196	4.33515686763372	0	-4.33515686763372
4	199	0.0563831393653673	0	-0.0563831393653673
4	306	1.46299120677573	0	-1.46299120677573
4	578	7.2904766737671	0	-7.2904766737671
4	776	0.0454919948686085	0	-0.0454919948686085
4	794	0.378067178388327	0	-0.378067178388327
4	937	0.218779145034492	0	-0.218779145034492
4	1034	0.37302530271992	0	-0.37302530271992
4	1114	2.95150086578935	0	-2.95150086578935
4	1153	1.79095867706455	0	-1.79095867706455
4	1163	5.71559965337402	0	-5.71559965337402
4	1179	8.15497297429141	0	-8.15497297429141
4	1345	4.97314566613102	0	-4.97314566613102
4	1580	1.40872258435826	0	-1.40872258435826
4	1597	0.158197734668382	0	-0.158197734668382
4	1626	2.69433677925975	0	-2.69433677925975
4	1942	4.59101421949357	0	-4.59101421949357
4	1956	3.50032155667046	0	-3.50032155667046
5	103	4.81026795004191	0	-4.81026795004191
5	208	0.484244377055198	0	-0.484244377055198
5	213	1.17442353656431	0	-1.17442353656431
5	346	6.13235209441725	0	-6.13235209441725
5	360	0.781247251899001	0	-0.781247251899001
5	408	4.6610422679705	0	-4.6610422679705
5	428	0.425782685697564	0	-0.425782685697564
5	539	0.407679481803175	0	-0.407679481803175
5	563	13.1701116034831	0	-13.1701116034831
5	634	0.416844168565608	0	-0.416844168565608
5	810	5.75569798355976	0	-5.75569798355976
5	833	0.259319349082704	0	-0.259319349082704
5	840	0.0325293669941772	0	-0.0325293669941772
5	876	3.44731400388234	0	-3.44731400388234
5	927	1.57311495807102	0	-1.57311495807102
5	936	1.23673278432409	0	-1.23673278432409
5	955	31.2756183700343	0	-31.2756183700343
5	997	1.38221925214264	0	-1.38221925214264
5	1180	0.497147818994634	0	-0.497147818994634
5	1409	6.58349669023016	0	-6.58349669023016
5	1432	0.00296450656010982	0	-0.00296450656010982
5	1557	0.611011252802201	0	-0.611011252802201
5	1687	2.89832963740228	0	-2.89832963740228
6	7	0.0183712099278186	0	-0.0183712099278186

6	72	0.955557419218076	0	-0.955557419218076
6	146	0.0320804761638862	0	-0.0320804761638862
6	170	5.32294524026195	0	-5.32294524026195
6	241	0.0732567212990728	0	-0.0732567212990728
6	281	0.296621932386459	0	-0.296621932386459
6	335	0.0292093633661658	0	-0.0292093633661658
6	385	3.8211857620203	0	-3.8211857620203
6	528	0.0142090077727275	0	-0.0142090077727275
6	543	1.87536045313126	0	-1.87536045313126
6	612	14.9293336188041	0	-14.9293336188041
6	622	0.000184800528313824	0	-0.000184800528313824
6	782	0.932351084036338	0	-0.932351084036338
6	900	0.124953674802091	0	-0.124953674802091
6	943	4.9946913316459	0	-4.9946913316459
6	1075	0.425550819893919	0	-0.425550819893919
6	1120	0.785369962641444	0	-0.785369962641444
6	1306	8.03604024067557	0	-8.03604024067557
6	1525	4.80506134527242	0	-4.80506134527242
6	1687	6.02902561696551	0	-6.02902561696551
6	1743	0.255118538071965	0	-0.255118538071965
6	1987	1.25211383584995	0	-1.25211383584995
7	112	0.925545853364116	0	-0.925545853364116
7	249	0.113231401677636	0	-0.113231401677636
7	520	10.159558450378	0	-10.159558450378
7	592	1.91701379974145	0	-1.91701379974145
7	694	0.0291392134480839	0	-0.0291392134480839
7	704	0.155451785142413	0	-0.155451785142413
7	752	0.230216501971448	0	-0.230216501971448
7	775	1.17815924750931	0	-1.17815924750931
7	823	0.0068408412256474	0	-0.0068408412256474
7	1000	1.07257012966517	0	-1.07257012966517
7	1017	0.864197174916518	0	-0.864197174916518
7	1063	0.0259774254292891	0	-0.0259774254292891
7	1089	0.0484242237950545	0	-0.0484242237950545
7	1145	0.0419292322042091	0	-0.0419292322042091
7	1258	0.0918515301519485	0	-0.0918515301519485
7	1261	2.17797979549803	0	-2.17797979549803
7	1385	0.753502820544887	0	-0.753502820544887
7	1387	9.11799955668958	0	-9.11799955668958
7	1532	1.48435738714947	0	-1.48435738714947
7	1533	6.3783760276671	0	-6.3783760276671
7	1658	0.0605754695577753	0	-0.0605754695577753
7	1697	0.111470300007206	0	-0.111470300007206
7	1716	2.37823580406007	0	-2.37823580406007
8	32	0.0207727438509778	0	-0.0207727438509778
8	113	1.19157499453618	0	-1.19157499453618

8	122	2.14706717184446	0	-2.14706717184446
8	156	1.34095208526841	0	-1.34095208526841
8	294	0.0463761315834323	0	-0.0463761315834323
8	349	0.396318343614877	0	-0.396318343614877
8	605	20.7021026142894	0	-20.7021026142894
8	650	0.021838365700851	0	-0.021838365700851
8	775	0.0351682044447958	0	-0.0351682044447958
8	881	0.414617823366633	0	-0.414617823366633
8	887	0.800627737359919	0	-0.800627737359919
8	890	12.6424793135293	0	-12.6424793135293
8	1032	3.74711541620842	0	-3.74711541620842
8	1055	0.0368931009536479	0	-0.0368931009536479
8	1061	0.000980756697841534	0	-0.000980756697841534
8	1080	0.0000243985194126174	0	-0.0000243985194126174
8	1416	0.6100870789162	0	-0.6100870789162
8	1462	0.379917015934536	0	-0.379917015934536
8	1488	1.53729408125147	0	-1.53729408125147
8	1560	3.27007367166137	0	-3.27007367166137
8	1645	0.00223543713862751	0	-0.00223543713862751
8	1716	2.88699467524608	0	-2.88699467524608
8	1809	4.20232678043201	0	-4.20232678043201
8	1873	1.02285205450531	0	-1.02285205450531
9	83	0.000934901754726674	0	-0.000934901754726674
9	154	3.27843238825659	0	-3.27843238825659
9	175	11.4470461435565	0	-11.4470461435565
9	187	1.2890270738519	0	-1.2890270738519
9	242	0.498330620691537	0	-0.498330620691537
9	274	0.0513442628099305	0	-0.0513442628099305
9	336	0.283448507095982	0	-0.283448507095982
9	939	0.810077802342505	0	-0.810077802342505
9	941	3.33926972273436	0	-3.33926972273436
9	1006	2.07287292146562	0	-2.07287292146562
9	1038	2.51692474752031	0	-2.51692474752031
9	1131	4.55166196581934	0	-4.55166196581934
9	1205	3.8637426682132	0	-3.8637426682132
9	1385	0.360679195882796	0	-0.360679195882796
9	1448	4.84572716290276	0	-4.84572716290276
9	1449	0.271852358002135	0	-0.271852358002135
9	1477	0.648274680887237	0	-0.648274680887237
9	1500	0.0267726491154285	0	-0.0267726491154285
9	1581	13.2997163026906	0	-13.2997163026906
10	6	1.31123093378595	0	-1.31123093378595
10	94	4.91275564919893	0	-4.91275564919893
10	250	3.49958171541755	0	-3.49958171541755
10	253	0.0000236558645152314	0	-0.0000236558645152314
10	312	0.44218638981211	0	-0.44218638981211

10	470	0.22885880663574	0	-0.22885880663574
10	487	0.075565454961235	0	-0.075565454961235
10	540	0.103977620148944	0	-0.103977620148944
10	640	0.0158715667154235	0	-0.0158715667154235
10	820	0.788401727335902	0	-0.788401727335902
10	855	0.00907077879146933	0	-0.00907077879146933
10	902	1.56447038818966	0	-1.56447038818966
10	961	0.45278731453108	0	-0.45278731453108
10	1078	5.97030216706863	0	-5.97030216706863
10	1267	3.38627814031319	0	-3.38627814031319
10	1288	0.229119778918629	0	-0.229119778918629
10	1382	1.09055403103352	0	-1.09055403103352
10	1421	2.10440054821501	0	-2.10440054821501
10	1428	2.69568318976628	0	-2.69568318976628
10	1553	0.403944384844653	0	-0.403944384844653
10	1610	1.36155464600794	0	-1.36155464600794
10	1710	5.32915078698872	0	-5.32915078698872
10	1714	2.31302575585875	0	-2.31302575585875
11	41	3.36618818752868	0	-3.36618818752868
11	175	1.29823350280969	0	-1.29823350280969
11	203	0.242983006796773	0	-0.242983006796773
11	466	2.96852388750107	0	-2.96852388750107
11	551	1.3312180766908	0	-1.3312180766908
11	603	0.0163911341226562	0	-0.0163911341226562
11	609	0.659808976485047	0	-0.659808976485047
11	740	0.177662996210996	0	-0.177662996210996
11	796	0.0000302272522060445	0	-0.0000302272522060445
11	834	0.195058221033838	0	-0.195058221033838
11	934	1.39193506657434	0	-1.39193506657434
11	1024	0.00303069409823028	0	-0.00303069409823028
11	1179	4.55475514257826	0	-4.55475514257826
11	1401	0.938603358335636	0	-0.938603358335636
11	1435	3.22212617691665	0	-3.22212617691665
11	1462	1.97489070598612	0	-1.97489070598612
11	1518	0.507723357931737	0	-0.507723357931737
11	1544	0.0295024644572521	0	-0.0295024644572521
11	1692	0.104598017505904	0	-0.104598017505904
11	1709	2.46276880645702	0	-2.46276880645702
11	1734	0.000281569629453598	0	-0.000281569629453598
11	1802	0.0468089933182939	0	-0.0468089933182939
12	113	0.00298114454910012	0	-0.00298114454910012
12	129	0.493380472055754	0	-0.493380472055754
12	376	1.65077776953625	0	-1.65077776953625
12	399	5.52112892814638	0	-5.52112892814638
12	510	1.03088443576638	0	-1.03088443576638
12	632	2.73600218478296	0	-2.73600218478296

12	666	0.271461139984032	0	-0.271461139984032
12	725	8.08499266201889	0	-8.08499266201889
12	743	0.204754813393301	0	-0.204754813393301
12	756	0.544526086947914	0	-0.544526086947914
12	776	0.0411330819744266	0	-0.0411330819744266
12	907	3.49458836074395	0	-3.49458836074395
12	954	0.590351262011454	0	-0.590351262011454
12	966	0.623149519530033	0	-0.623149519530033
12	1020	0.0985917533490924	0	-0.0985917533490924
12	1323	0.0349352001738338	0	-0.0349352001738338
12	1371	1.88613876862803	0	-1.88613876862803
13	62	0.272666983075603	0	-0.272666983075603
13	280	0.00463945113462966	0	-0.00463945113462966
13	334	4.09141708812438	0	-4.09141708812438
13	339	0.322121488793178	0	-0.322121488793178
13	389	0.319759688684078	0	-0.319759688684078
13	419	2.49107655787161	0	-2.49107655787161
13	430	0.655839523327001	0	-0.655839523327001
13	544	0.900942974846742	0	-0.900942974846742
13	587	0.569278560288013	0	-0.569278560288013
13	734	1.77790539511879	0	-1.77790539511879
13	787	0.105889467682368	0	-0.105889467682368
13	1068	0.181052840656451	0	-0.181052840656451
13	1244	0.0119589179917611	0	-0.0119589179917611
13	1252	1.0154910048134	0	-1.0154910048134
14	26	0.154081599559553	0	-0.154081599559553
14	110	1.15142130037306	0	-1.15142130037306
14	159	0.217920179100303	0	-0.217920179100303
14	168	0.213977952918569	0	-0.213977952918569
14	188	0.34579499918446	0	-0.34579499918446
14	219	1.84135951136885	0	-1.84135951136885
14	231	4.71737264072939	0	-4.71737264072939
14	271	0.00016213419674367	0	-0.00016213419674367
14	326	3.46618973371208	0	-3.46618973371208
14	555	0.197808360656539	0	-0.197808360656539
14	603	5.1071423189273	0	-5.1071423189273
14	668	0.00320151543142691	0	-0.00320151543142691
14	681	0.437377530724633	0	-0.437377530724633
14	811	0.00146741178038403	0	-0.00146741178038403
14	838	1.39830855057411	0	-1.39830855057411
14	841	5.7889525587425	0	-5.7889525587425
14	1028	1.19577307428644	0	-1.19577307428644
14	1040	3.14607401030792	0	-3.14607401030792
14	1090	3.13788844453219	0	-3.13788844453219
14	1294	9.33818598216415	0	-9.33818598216415
14	1356	0.894421047731119	0	-0.894421047731119

15	41	0.598137649066498	0	-0.598137649066498
15	268	3.01619196995498	0	-3.01619196995498
15	521	4.14194444215983	0	-4.14194444215983
15	606	0.476951638061875	0	-0.476951638061875
15	653	0.00572499578136585	0	-0.00572499578136585
15	763	0.16702019739681	0	-0.16702019739681
15	962	8.79707012148461	0	-8.79707012148461
15	1217	0.00229810466684213	0	-0.00229810466684213
15	1258	4.54979984024964	0	-4.54979984024964
15	1294	0.789578385901744	0	-0.789578385901744
15	1310	1.02439251881043	0	-1.02439251881043
15	1345	0.0436929082594098	0	-0.0436929082594098
16	5	0.0836735189306747	0	-0.0836735189306747
16	61	1.32255356096415	0	-1.32255356096415
16	81	0.0128703070320303	0	-0.0128703070320303
16	125	0.237324600938968	0	-0.237324600938968
16	235	0.000217729607972298	0	-0.000217729607972298
16	257	0.0277387895802905	0	-0.0277387895802905
16	475	2.82891989534052	0	-2.82891989534052
16	663	0.00971917005113528	0	-0.00971917005113528
16	690	1.6090556332915	0	-1.6090556332915
16	793	0.699155408425609	0	-0.699155408425609
16	839	0.00764631480558219	0	-0.00764631480558219
16	873	0.0287396592970931	0	-0.0287396592970931
16	975	0.0159065651273975	0	-0.0159065651273975
16	1077	0.412222084481946	0	-0.412222084481946
16	1145	0.236735910381229	0	-0.236735910381229
16	1161	2.44962146855568	0	-2.44962146855568
17	8	0.231054015026556	0	-0.231054015026556
17	45	0.00741606200846408	0	-0.00741606200846408
17	118	3.23019251697124	0	-3.23019251697124
17	297	0.0609599604894334	0	-0.0609599604894334
17	339	0.0202190955092424	0	-0.0202190955092424
17	487	9.57951708122193	0	-9.57951708122193
17	592	0.00612057645372645	0	-0.00612057645372645
17	603	0.177600720085055	0	-0.177600720085055
17	627	0.399779724330074	0	-0.399779724330074
17	778	0.643102198550672	0	-0.643102198550672
17	961	0.00291530808336672	0	-0.00291530808336672
17	971	0.0777818694581076	0	-0.0777818694581076
17	1086	0.320993696516304	0	-0.320993696516304
18	29	0.24755088343876	0	-0.24755088343876
18	292	0.690493071020109	0	-0.690493071020109
18	331	0.587838607703245	0	-0.587838607703245
18	374	0.369527591920317	0	-0.369527591920317
18	378	1.10771404953978	0	-1.10771404953978

18	453	0.368230125485861	0	-0.368230125485861
18	594	0.03722484206735	0	-0.03722484206735
18	603	0.00248295264344436	0	-0.00248295264344436
18	690	0.471161803610511	0	-0.471161803610511
18	780	13.1665693061094	0	-13.1665693061094
18	810	0.00909188606790862	0	-0.00909188606790862
18	845	0.767940932660066	0	-0.767940932660066
18	1002	0.00689854030611686	0	-0.00689854030611686
19	231	6.95980087916802	0	-6.95980087916802
19	338	0.711357544285564	0	-0.711357544285564
19	376	0.00292705233309753	0	-0.00292705233309753
19	383	4.25780028789597	0	-4.25780028789597
19	481	0.0145408292813196	0	-0.0145408292813196
19	654	22.8375911832387	0	-22.8375911832387
19	759	0.0127027350462078	0	-0.0127027350462078
19	809	0.00302890023704722	0	-0.00302890023704722
19	823	0.0755424933941261	0	-0.0755424933941261
19	980	1.22490940598896	0	-1.22490940598896
19	1072	3.64839253125785	0	-3.64839253125785
19	1076	1.7149387462508	0	-1.7149387462508
20	40	0.586121698991782	0	-0.586121698991782
20	73	2.31333409336828	0	-2.31333409336828
20	103	1.72553677662717	0	-1.72553677662717
20	146	0.642046125189751	0	-0.642046125189751
20	301	1.67930445781057	0	-1.67930445781057
20	307	9.02789651008662	0	-9.02789651008662
20	382	1.54034684015055	0	-1.54034684015055
20	719	0.790660787069067	0	-0.790660787069067
20	908	0.219992228876332	0	-0.219992228876332
20	959	0.0013780063314942	0	-0.0013780063314942
20	986	1.98539077960982	0	-1.98539077960982
20	1028	0.239585475217005	0	-0.239585475217005
20	1065	4.03394378491708	0	-4.03394378491708
21	145	0.237212046694967	0	-0.237212046694967
21	226	0.86181965534167	0	-0.86181965534167
21	248	0.777648195081131	0	-0.777648195081131
21	251	4.80029607651833	0	-4.80029607651833
21	280	0.445705937734944	0	-0.445705937734944
21	331	0.00972802116611879	0	-0.00972802116611879
21	404	0.994202508193548	0	-0.994202508193548
21	418	0.0606629836183235	0	-0.0606629836183235
21	616	0.0296185802539273	0	-0.0296185802539273
21	632	3.49198969903896	0	-3.49198969903896
21	663	4.99710972168143	0	-4.99710972168143
21	695	0.00359587695640669	0	-0.00359587695640669
21	765	0.888643237700599	0	-0.888643237700599

21	774	1.32945080188294	0	-1.32945080188294
21	788	0.0743983302736663	0	-0.0743983302736663
21	887	1.66576672223036	0	-1.66576672223036
21	898	0.0342716249607919	0	-0.0342716249607919
22	26	0.71232720525948	0	-0.71232720525948
22	129	0.00145498871844077	0	-0.00145498871844077
22	216	2.71429601814115	0	-2.71429601814115
22	539	0.00168766441636739	0	-0.00168766441636739
22	597	0.346107977387235	0	-0.346107977387235
22	733	5.82082143862715	0	-5.82082143862715
22	799	0.266116746503377	0	-0.266116746503377
22	975	0.115823540884182	0	-0.115823540884182
23	184	0.589268811334705	0	-0.589268811334705
23	187	0.881467725435667	0	-0.881467725435667
23	311	0.10553968806056	0	-0.10553968806056
23	326	0.046795797866759	0	-0.046795797866759
23	345	1.97099171410478	0	-1.97099171410478
23	386	0.363307110647019	0	-0.363307110647019
23	431	0.281029252665731	0	-0.281029252665731
23	466	0.0683282165695164	0	-0.0683282165695164
23	509	0.728582417782906	0	-0.728582417782906
23	510	1.92143860118914	0	-1.92143860118914
23	542	0.543251880796052	0	-0.543251880796052
23	669	0.0196093050899671	0	-0.0196093050899671
23	710	0.0218208027805475	0	-0.0218208027805475
23	714	0.0000398667908215711	0	-0.0000398667908215711
24	173	13.3290548453825	0	-13.3290548453825
24	269	0.391912178065114	0	-0.391912178065114
24	362	1.64636211704708	0	-1.64636211704708
24	374	0.425669235873767	0	-0.425669235873767
24	427	3.80228512963397	0	-3.80228512963397
24	513	0.427296855593771	0	-0.427296855593771
24	515	7.37106134596548	0	-7.37106134596548
24	714	4.85321689373218	0	-4.85321689373218
24	857	0.0196688612128603	0	-0.0196688612128603
24	980	0.00233659136550531	0	-0.00233659136550531
25	114	0.251746759936434	0	-0.251746759936434
25	216	0.795409085891022	0	-0.795409085891022
25	243	0.168567621948314	0	-0.168567621948314
25	293	0.497869263803566	0	-0.497869263803566
25	298	0.160217611106504	0	-0.160217611106504
25	347	0.0226758301929089	0	-0.0226758301929089
25	484	1.01001608152995	0	-1.01001608152995
25	732	0.121643523766693	0	-0.121643523766693
25	797	0.219994591180191	0	-0.219994591180191
26	56	0.0951671642505996	0	-0.0951671642505996

26	238	3.048044044432	0	-3.048044044432
26	506	1.58834167040544	0	-1.58834167040544
26	563	0.0300996673760749	0	-0.0300996673760749
26	789	0.315741010100177	0	-0.315741010100177
26	820	1.10905400105909	0	-1.10905400105909
27	204	0.0148133261255233	0	-0.0148133261255233
27	349	0.0809532499942724	0	-0.0809532499942724
27	358	1.17953125972273	0	-1.17953125972273
27	404	0.270059564608199	0	-0.270059564608199
27	482	2.69262225386995	0	-2.69262225386995
27	522	8.87151098453958	0	-8.87151098453958
27	633	2.17450378000628	0	-2.17450378000628
27	650	1.14758580284192	0	-1.14758580284192
27	716	7.70154374024647	0	-7.70154374024647
27	722	1.55334442533382	0	-1.55334442533382
27	735	0.979613360466115	0	-0.979613360466115
27	749	6.91962588237768	0	-6.91962588237768
28	191	0.28638054372896	0	-0.28638054372896
28	206	0.392799815599769	0	-0.392799815599769
28	235	0.319821231298604	0	-0.319821231298604
28	240	0.000026723467645674	0	-0.000026723467645674
28	258	0.367790246196497	0	-0.367790246196497
28	266	0.0129119134240494	0	-0.0129119134240494
28	312	0.000165198258569376	0	-0.000165198258569376
28	327	6.04069657936331	0	-6.04069657936331
28	365	2.01758754409084	0	-2.01758754409084
28	373	0.0197708450673093	0	-0.0197708450673093
28	485	0.890221610376744	0	-0.890221610376744
28	492	0.000256507646847296	0	-0.000256507646847296
28	504	2.17810563374081	0	-2.17810563374081
28	615	1.04376961096772	0	-1.04376961096772
28	633	0.0322077167315722	0	-0.0322077167315722
29	26	0.0236940315119136	0	-0.0236940315119136
29	139	1.75089430069001	0	-1.75089430069001
29	307	1.36836496958951	0	-1.36836496958951
29	313	1.70072962667128	0	-1.70072962667128
29	317	0.0546339228086695	0	-0.0546339228086695
29	401	2.13181157483351	0	-2.13181157483351
29	445	2.01562690339242	0	-2.01562690339242
29	464	11.2232983711204	0	-11.2232983711204
29	554	0.150817244900762	0	-0.150817244900762
29	560	1.62051760242648	0	-1.62051760242648
29	679	0.0520005860367456	0	-0.0520005860367456
29	683	0.28851690993435	0	-0.28851690993435
29	751	0.235282977262047	0	-0.235282977262047
29	782	7.5681538902783	0	-7.5681538902783

Titre : La diversité génétique à l'ère de la génomique : évolution de la consanguinité et ses conséquences dans trois races bovines laitières françaises

Mots clés : sélection génomique, transfert embryonnaire, consanguinité, diversité génétique, bovins laitiers, dépression de consanguinité

Résumé : La sélection génomique, mise en place depuis 2009 en France, a drastiquement modifié les schémas de sélection dans les races bovines laitières françaises de grand effectif en augmentant le nombre de candidats à la sélection et en réduisant les intervalles de génération. L'essor de l'utilisation du transfert embryonnaire a également remanié les schémas de sélection, intensifiant la sélection des mères à taureaux et réduisant encore les intervalles de génération. Ces technologies permettent une augmentation du progrès génétique.

Cependant, si leur utilisation n'est pas maîtrisée, elles mettent en danger la diversité génétique de ces races. Cette perte de diversité, caractérisée par une augmentation de la consanguinité, peut s'accompagner de dépression de consanguinité, d'une baisse de l'efficacité de la sélection et une perte de potentiel adaptatif. Il est donc primordial de comprendre et de prédire l'impact de ces changements sur la diversité génétique à l'ère de la génomique, afin de mieux la maintenir.

Un bilan de la mise en place de la sélection génomique a montré que cette technologie a permis une augmentation

du progrès génétique annuel chez la Montbéliarde, la Normande et la Prim'Holstein mais que la perte de diversité génétique s'est accélérée seulement chez la Prim'Holstein.

Une étude par simulations a montré que, dans le cadre de la sélection génomique, l'utilisation intensive du transfert embryonnaire dans un schéma de sélection de type Montbéliarde présentait un risque pour la diversité génétique, et que la réduction du nombre de taureaux dans ces schémas causait à la fois une baisse du progrès génétique et une accélération de la perte de diversité génétique.

Dans la race Montbéliarde, la dépression de consanguinité a fait baisser les performances pour cinq caractères de production et de santé sur les six caractères étudiés. Les effets de la consanguinité sur ces performances sont très hétérogènes le long du génome. Gérer la consanguinité et ses conséquences dans les races bovines laitières françaises nécessite la mise en œuvre de mesures efficaces, éthiques et applicables sur le terrain.

Title : Genetic diversity in the genomic era: evolution of inbreeding and its consequences in three French dairy cattle breeds

Keywords : genomic selection, embryo transfer, inbreeding, genetic diversity, dairy cattle, inbreeding depression

Abstract: Genomic selection, implemented in France since 2009, has drastically modified breeding schemes in large French dairy cattle breeds by increasing the number of candidates for selection and reducing generation intervals. The growing use of embryo transfer has also reshaped breeding schemes, intensifying the selection of dams of bulls and further reducing generation intervals. This resulted in an increased genetic gain.

However, if their use is not controlled, they put the genetic diversity of these breeds at risk. This loss of diversity, characterized by an increased inbreeding, can be accompanied by inbreeding depression, a decrease in selection efficiency and a loss of adaptive potential. It is therefore essential to understand and predict the impact of these changes on genetic diversity. We studied the implementation of genomic selection and showed that this technology has led to an increase in annual genetic gain in

Montbéliarde, Normande and Prim'Holstein while the loss of genetic diversity has only been accelerated in Prim'Holstein.

We showed by simulations that, in the context of genomic selection, the intensive use of embryo transfer in a Montbéliarde-type breeding scheme presented a risk for genetic diversity, and that the reduction in the number of bulls in these schemes caused both a decrease in genetic gain and an acceleration of the loss of genetic diversity.

In the Montbéliarde breed, inbreeding depression resulted in lower performance for five production and health traits out of six traits studied. The effects of inbreeding on these performances were very heterogeneous along the genome.

Managing inbreeding and its consequences in French dairy cattle breeds requires the implementation of effective, ethical and feasible actions in the field.