

Université de Limoges

Ecole Doctorale Sciences Biologiques et Santé

Laboratoire PEIRENE

Thèse pour obtenir le grade de
Docteur de l'Université de Limoges

Génomique et Biologie Moléculaire

Présentée et soutenue par
Abdelmajid EL HOU

Le 10 décembre 2021

Identification de gènes impliqués dans la qualité des viandes chez les bovins allaitants français.

Thèse dirigée par Véronique Blanquet & Romain Philippe.

JURY :

Président du jury : Mme. Sandrine Lagarrigue, Professeur, AGROCAMPUS OUEST, UMR 1348 PEGASE, Rennes, France

Rapporteurs

Mme. Carole Charlier, Directeur de recherche, GIGA medical Genomics, Université de Liège, Belgique
Mme. Sandrine Lagarrigue, Professeur, AGROCAMPUS OUEST, UMR 1348 PEGASE, Rennes, France

Examineurs

Mme. Véronique Blanquet, Professeur, EA 7500 USC INRAE 1061, Université de Limoges, France
M. Romain Philippe, Chargé de recherche, EA 7500 USC INRAE 1061, Université de Limoges, France
M. Dominique Rocha, Directeur de recherche, UMR 1313 GABI, INRAE, Jouy-en-Josas, France
M. Yulixaxis Ramayo Caldas, Chargé de recherche, Institute of Agrifood Research and Technology (IRTA). Torre Marimon, Caldes de Montbui, Catalunya, Espagne



À mes parents, à ma famille et à tous ceux qui me sont chers. À mes chers professeurs :

Thomas HEAMS, Abdelhamid ROSSI, Omar CHAIR et Abdelaadim EL YAZIDI.

"Never give up on a dream just because of the time it will take to accomplish it. The time will pass anyway."
Earl Nightingale

Remerciements

Je tiens à saluer ici et à remercier très vivement les personnes qui, de près ou de loin, ont contribué à la concrétisation de ce travail de thèse de doctorat.

Merci à l'université de Limoges pour le financement de cette thèse.

Je tiens à remercier tous les membres du jury, en particulier Carole Charlier, Sandrine Lagarrigue et Yulixaxis Ramayo Caldas pour avoir accepté d'examiner ce travail. Merci pour la relecture de la thèse et pour les améliorations que vous aviez apporté à celle-ci.

Je tiens à remercier très chaleureusement mes directeurs de cette thèse, Romain Philippe et Véronique Blanquet, pour m'avoir fait confiance, puis pour m'avoir guidé, encouragé, conseillé tout au long des années de ma thèse. Merci pour le soutien moral, à chaque fois que j'en avais besoin. Mes remerciements vont également à Dominique Rocha, mon co-encadrant. Merci pour ta gentillesse et ta patience à mon égard durant cette thèse, tes réponses rapides aux mails, ainsi que pour tous tes conseils et encouragements.

Un grand merci à tout le personnel du laboratoire PEIRENE (Techniciens, ingénieurs... etc) et du département SV y compris le corps enseignant (maîtres de conférences, professeurs et doctorants contractuels). Merci à notre secrétaire Odile Duval pour sa disponibilité pour répondre à toute sorte de questions administratives et logistiques, mais surtout pour sa gentillesse et son encouragement.

Merci aux membres de l'équipe G2B de l'unité GABI (INRAE) à Jouy-en-Josas, en particulier Alexis Michnet pour la formation sur les analyses GWAS et Eric Venot pour sa formation au langage AWK.

Un grand merci à Bertrand Courtioux, le directeur de l'École Doctorale Sciences Biologiques et Santé (SBS - 615) pour tous ses efforts envers les doctorants afin de leur faciliter le suivi des formations de qualité, organiser des séminaires intéressants ainsi que des bons moments d'échanges entre les doctorants. Merci à toute l'équipe administrative de l'ED SBS en particulier l'assistante de direction Aurélie Lacombe pour ses réponses rapides pour toute question administrative.

Merci à mes amis Jordano-Mauritaniens (Montassir Al MANSI et SALEK), pour leurs conseils, encouragement et balades au bord de la vienne et au centre-ville de Limoges. Mes dernières pensées vont à mes parents qui ont su me donner une éducation de qualité et ont mis à ma disposition les maigres moyens financiers en leur possession pour que j'aie le plus loin possible dans mes études. Je pense à eux aujourd'hui, qu'ils trouvent ici toute ma gratitude et ma reconnaissance pour tous leurs efforts. Merci aussi à mes oncles (Omar, Mustapha et Mhamed EL HOU) pour leur soutien moral et financier pendant plusieurs années. Merci à mes frères et sœurs qui m'ont apporté leurs soutiens à leur manière. Et bien sûr un grand merci à mes chers professeurs Abdelhamid Rossi, Omar Chair, Abdeladim El Yazidi et Thomas Heams pour leur encouragement et merci à ma femme Loubna Hajamy pour son soutien moral en particulier durant les périodes de stress et pour sa patience vis-à-vis de mes nuits blanches.

Enfin, merci à celles et ceux dont j'ai omis les noms par oubli et qui se reconnaîtront dans ces quelques lignes.

Droits d'auteurs

Cette création est mise à disposition selon le Contrat :

« **Attribution-Pas d'Utilisation Commerciale-Pas de modification 3.0 France** »

disponible en ligne : <http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



Table des matières

Table des matières

Liste des abréviations.....	17
Introduction générale.....	19
Partie I. L'élevage des bovins allaitants en France et la qualité des viandes	22
I.1. Le cheptel bovin français	22
I.2. Production et consommation de viande bovine en France	23
I.2.1. La production de viande bovine en France	23
I.2.2. La consommation de viande en France	24
I.3. Une viande de qualité pourrait-elle regagner la confiance des consommateurs ?	24
I.3.1. Les qualités nutritionnelles de la viande	25
I.3.1.1. Les protéines	25
I.3.1.2. Teneur en matière grasse	25
I.3.1.3. Vitamines et minéraux.....	25
I.3.2. La qualité hygiénique de la viande.....	26
I.3.3. Les qualités technologiques de la viande	26
I.3.3.1. Le pH	26
I.3.3.2. La capacité de rétention d'eau	26
I.3.3.3. L'appétit de la viande à la conservation.....	26
I.3.4. Les qualités organoleptiques de la viande et ses facteurs de variation	26
I.3.4.1. La couleur	26
I.3.4.2. La flaveur	26
I.3.4.3. La jutosité	27
I.3.4.4. La tendreté.....	27
Partie II. Amélioration génétique des races à viande en France	28
II.1. Introduction	28
II.2. Rappel de notions de génétique.....	28
II.2.1. Les marqueurs génétiques, locus, allèles	28
II.3. Les méthodes de génotypage des SNP	29
II.3.1. La méthode TaqMan	29
II.3.2. Génotypage sur micropuce.....	30
II.4. Rappel historique de l'amélioration génétique des animaux d'élevage	32
II.5. Les méthodes de sélection et leurs limites	32
II.5.1. Sélection sur descendance.....	32
II.5.2. Sélection sur des gènes connus	32
II.5.3. Sélection assistée par marqueurs (SAM).....	33
II.5.3.1. La SAM1	34
II.5.3.2. La SAM2	34
II.5.4. Sélection génomique	34
II.6. L'organisation et les acteurs de la sélection des bovins allaitants en France	35
II.7. Les étapes d'un programme de sélection.....	36
II.7.1. Définition des objectifs de sélection	36
II.7.2. Recueil des informations	37
II.7.3. Evaluation génétique	38
II.7.4. Choix des reproducteurs.....	40
II.7.5. Diffusion des reproducteurs.....	40
II.7.6. Analyse des résultats de la sélection.....	40

II.8. Conclusion	41
Partie III. La cartographie des QTL chez les animaux.....	42
III.1. Introduction	42
III.2. La cartographie des QTL par analyse de liaison	42
III.2.1. Principe général	42
III.2.2. Les dispositifs utilisés pour la détection de QTL chez les animaux	43
III.2.2.1. Les populations « inbred »	43
III.2.2.2. Les populations « outbred »	43
III.2.3. Les méthodes statistiques utilisées	44
III.2.3.1. Analyse marqueur par marqueur	44
III.2.3.1.1. Analyse de variance	45
III.2.3.1.2. Par maximum de vraisemblance.....	45
III.2.3.2. Cartographie d'intervalle	45
III.2.3.2.1. Principe général.....	45
III.2.3.2.2. Le cas des populations « inbred » backcross.....	46
III.2.3.2.3. Le cas des populations « inbred » F2	47
III.2.3.2.4. Le cas des populations « outbred »	47
III.3. Etat des lieux sur les gènes et marqueurs associés à la qualité des viandes chez les bovins allaitants	47
III.3.1. Le programme Qualvigène	47
III.3.2. Gènes et marqueurs de la qualité des viandes	48
III.4. Conclusion	52
Partie IV. Le déséquilibre de liaison dans les populations animales	53
IV.1. Introduction.....	53
IV.2. Définition et propriétés du déséquilibre de liaison	53
IV.3. Mesures du déséquilibre de liaison.....	53
IV.3.1. Le calcul de DL en pratique.....	54
IV.4. Origine et facteurs impactant le DL.....	54
IV.4.1. Facteurs liés à l'histoire de la population	54
IV.4.1.1. La dérive génétique	54
IV.4.1.2. Les mutations	55
IV.4.1.3. Les goulots d'étranglement.....	55
IV.4.1.4. Le taux de recombinaison.....	55
IV.4.1.5. Le mélange des populations	56
IV.4.1.6. La sélection	57
IV.4.1.7. La conversion génique.....	57
IV.4.2. Facteurs liés à la qualité des données	58
IV.5. Les analyses de DL chez les bovins : état des lieux	58
IV.6. Conclusion.....	59
Chapitre 1: Le LRLD chez les bovins allaitants français	62
V.1. Introduction.....	62
V.2. Article	63
V.3. Discussion générale	77
V.4. Conclusion.....	79
Chapitre 2 : Détection de variants associés à la qualité des viandes bovines à l'aide des données de séquence génome imputé	82
VI.1. Introduction.....	82

VI.2. Matériel et méthodes	83
VI.2.1. Animaux et phénotypes	83
VI.2.2. Imputation pan-génomique	84
VI.2.3. Analyse d'association séquence imputée génome complet	84
VI.2.4. Annotation des variants et identification de mutations candidates.....	85
VI.3. Résultats	85
VI.3.1. Nombre de variants significatifs par race et par caractère.....	85
VI.3.2. Annotation des variants et identification de mutations candidates.....	90
VI.3.3. Comparaison avec des données bibliographiques.....	93
VI.4. Mutations candidates.....	96
VI.5. Discussion	96
VI.6. Conclusion.....	99
Discussion générale	101
Conclusion générale et perspectives	104
Références bibliographiques	106
Annexes	119

Table des figures

Table des figures

Figure 1 : Production de viande bovine (veaux, bœufs, taureaux, génisses et vaches, abattus par les abattoirs et à la ferme dont la viande est certifiée pour la consommation humaine) en Europe (1000 t équivalent carcasse) pour l'année 2019.....	23
Figure 2 : Type de viande produite et consommée en France en 2018.	24
Figure 3 : Exemple de polymorphisme SNP entre deux individus (image adaptée de icbf)...	29
Figure 4 : Principe du génotypage par la méthode TaqMan (source : applied-maths).....	30
Figure 5 : Puce à SNP 777k (BovineHD BeadChip) pour génotyper 8 individus (échantillons) pour 770 000 SNP (Illumina).	30
Figure 6: Principe de la méthode de génotypage d’Affymetrix (a) et d’Illumina (b).....	31
Figure 7 : Schéma représentatif des étapes d'un programme de sélection	37
Figure 8 : Progrès génétique estimé pour le poids-âge-type au sevrage entre 2000 et 2015 dans 9 races.....	40
Figure 9 : Principe de la détection d’un QTL. D’après (P. Le Roy & Elsen, 2000).....	43
Figure 10 : Exemple de profil de vraisemblance (P. Le Roy & Elsen, 2000).	46
Figure 11: Exemple de réaction induite par SCD pour convertir le stearoyl-CoA en oleoyl-CoA. D’après Uto, 2016.....	51
Figure 12: Exemple de mutation en DL transmise de génération en génération.	55
Figure 13 : Décroissance du DL en fonction du DL initial (D_0), du taux de recombinaison (r) et du temps (générations).....	56
Figure 14: Exemple de DL créé par fusion de deux populations en équilibre de liaison.	57
Figure 15: Exemple de conversion génique. Transfert unidirectionnel du fragment noir vers l’autre chromosome	57
Figure 16: Distribution des variants associés au caractère note de tendreté sur le BTA 29 du CHA	86
Figure 17: Distribution des variants associés au caractère taux de collagène insoluble sur le BTA2 du CHA.....	86
Figure 18: Diagramme de Venn des variants en commun entre races au seuil 10^{-7} . A : rendcar, B: poidgi, C: poidab, D : mcocisa, E: micoull, F: longjs, G : dsvif, H : gmq, I : epaicu, J : confca, K : dmvif, L : dap6c.....	89
Figure 19: Nombre de variants par pléiotropie dans les différentes races CHA, LIM et BLA.	90

Liste des tableaux

Table des tableaux

Tableau 1 : Les principaux cheptels bovins en 2018 (millions de têtes). Source : (GEB, 2019).....	22
Tableau 2 : Évolution du nombre de vaches et de bovins en France (en milliers de têtes) entre 1995 et 2019 ¹	22
Tableau 3 : Exemple de bases de données SNP.....	29
Tableau 4 : Exemple de puces commerciales existantes (Illumina & Affymetrix).	31
Tableau 5 : Les principales différences d'organisation entre la loi d'orientation agricole de 2006 et le règlement zootechnique de l'UE.....	36
Tableau 6 : Nombre de descendants nécessaires pour la détection d'un QTL dans un croisement entre lignées "inbred"	44
Tableau 7 : Nombre de descendants nécessaires pour une détection de QTL dans une population "outbred" (cas de familles de plein-frères BC ou F2)	44
Tableau 8: Nombre de variants significatifs au seuil p-valeur 10^{-7} (p-valeur 10^{-8}) par race et par caractère.	87
Tableau 9: Nombre de variants annotés par conséquence dans les 3 races.	91
Tableau 10: Liste des variants situés dans des régions d'épissage.	92
Tableau 11: Comparaison de nos variants avec les résultats de Allais et al, 2014	94
Tableau 12: Liste des variants significatifs associés à la note de tendreté chez le CHA et proches des QTL identifiés par (Leal-Gutiérrez et al., 2020a) sur le BTA29.	95
Tableau 13: Liste de mutations candidates pour les caractères d'aptitudes bouchères et qualité des viandes.....	96

Liste des abréviations

ADN : Acide DésoxyriboNucléique
AGMI: Acide Gras Mono-Insaturé
BLA: Blonde d'Aquitaine
BLUP: Best Linear Unbiased Predictor
CHA: Charolais
CIV : Centre d'Information des Viandes
cM : CentiMorgan
CNV: Copy Number Variation
CRÉDOC: Centre de Recherche pour l'Etude et l'Observation des Conditions de Vie
CTIG : Centre de Traitement de l'Information Génétique
EC : Equivalent Carcasse
GABI : Génétique Animale et Biologie Intégrative
G : Gramme
GWAS : Genome Wide Association Studie
HD : Haute densité
IA : Insémination Artificielle
IBOVAL : Indexation des Bovins Allaitants
IDELE : Institut de l'élevage
INRAE : Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement
JB : Jeunes Bovins
Kb : Kilo Base
Kgéc : Kilogramme équivalent carcasse
LIM : Limousin
MD : Moyen densité
Mg : milligramme
MN : Monte Naturelle
OMS : Organisation Mondiale de la Santé
OS : Organisme de Sélection
PCR : Polymerase Chaine Reaction
QTL : Quantitative Trait Locus
SAM : Sélection Assistée par Marqueur
SFFT : Site de fixation d'un facteur de transcription.
SG : Sélection Génomique
SNP : Single Nucleotide Polymorphisme
TEC : Tonne Equivalent Carcasse
UE : Union Européen

Introduction Générale

Introduction générale

La population mondiale de bovins en 2018 est d'environ 900 millions de têtes (GEB, 2019). L'Inde est le pays avec le plus grand troupeau au monde, suivi par le Brésil, la Chine et les états unis. Cependant, la France était le premier producteur de viande bovine en Europe. Elle possède le plus grand troupeau d'Europe avec un total de 18,5 millions de têtes et un troupeau de bovins à viande composé principalement de Charolaise, Limousin et Blonde d'Aquitaine ainsi que d'autres, mais avec des effectifs inférieurs à 200 000.

Les Français sont les plus gros consommateurs de viande bovine en Europe. D'autre part, la consommation de viande a diminué de façon continue au cours des 20 dernières années, en particulier la viande bovine, ovine et porcine. Alors que la part de la viande de volaille est passée de 25,8 % en 1998 à 34 % en 2018, et la consommation de viande est passée de 58 à 46 grammes par jour en 2007 et 2016 (France AgriMer, 2019). Une étude récente réalisée par le centre de recherche pour l'étude et l'observation des conditions de vie (CRÉDOC) sur la consommation des produits carnés révèle une diminution de 12 % sur les 10 dernières années (CREDOC, Tavoularis, 2018).

Pour regagner la confiance des consommateurs, les industriels adoptent une démarche qualité qui englobe toute la chaîne de production. Cependant, cette approche n'apporte aucune garantie quant à la qualité et aux caractéristiques organoleptiques de la viande.

La qualité de la viande est un caractère complexe et multifactoriel et ne peut pas être améliorée par les méthodes classiques de sélection animale (c.f. II.5) vu la nécessité de prendre des mesures en routine dans les abattoirs pour ces caractères. Ces mesures sont très lourdes à mettre en place techniquement et financièrement. Par conséquent, la filière s'est intéressée au déterminisme génétique de ces caractères, c'est-à-dire la recherche de gènes ou marqueurs responsables de la variabilité phénotypique chez ces animaux.

Le séquençage du génome complet de plusieurs espèces d'animaux domestiques et le progrès considérable des technologies de génotypage relativement peu coûteuses, ont permis d'entrevoir des perspectives beaucoup plus favorables pour l'identification de gènes impliqués dans les caractères d'intérêts économiques, y compris la qualité des viandes et le développement musculo-squelettique.

L'objectif de ma thèse est double. Premièrement, d'un point de vue fondamental, il s'agit de mieux comprendre la structure du génome bovin via l'analyse du déséquilibre de liaison à longue-distance (LRLD), et deuxièmement de rechercher les *Quantitative Trait Loci* (QTL) associés à la qualité des viandes et le développement musculaire chez les bovins allaitants français ; c'est-à-dire les régions du génome responsables de la variabilité phénotypique de ces caractères dans les populations bovines.

Dans la première partie de ce manuscrit de thèse, je fais une synthèse de la situation d'élevage bovin en France, un point sur la production et la consommation de la viande bovine en France et en Europe, ainsi qu'une synthèse de l'organisation de la sélection bovine en France, les méthodes de sélection et les approches de détection des QTL chez les bovins allaitants. Dans la deuxième partie bibliographique, je fais une introduction au déséquilibre de liaison, son importance en génétique animale, comment le mesurer, ainsi qu'un état des lieux sur les études faites à ce jour chez les bovins.

Le premier chapitre de ma thèse est consacré à l'étude du LRLD dans des populations de bovins allaitants français, sous forme d'un article publié dans *Genetics Selection Evolution*

(GSE). Le deuxième chapitre est consacré à la recherche de mutations impliquées dans la variabilité de la qualité des viandes et le développement musculaire chez les bovins allaitants français, en utilisant des données de séquences génome complet et des phénotypes issus du programme Qualvigène.

Étude bibliographique

Partie I. L'élevage des bovins allaitants en France et la qualité des viandes

I.1. Le cheptel bovin français

Depuis l'antiquité jusqu'à nos jours, les bovins ont été élevés pour fournir de la viande et des produits laitiers. Ils sont également utilisés comme animaux de trait pour labourer les champs ou transporter des objets lourds. Leurs peaux sont utilisées pour produire le cuir et leurs fumiers pour le carburant et les engrais agricoles.

La population mondiale de bovins comptait environ 900 millions de têtes en 2018 (GEB, 2019). L'Inde est le pays ayant le plus gros cheptel bovin du monde avec un total de 305 millions de têtes, suivi du Brésil (223 millions), la Chine (98 millions) et les États-Unis (94 millions de têtes) (Tableau 1).

Tableau 1 : Les principaux cheptels bovins en 2018 (millions de têtes). Source : (GEB, 2019).

	Total bovins	Dont vaches laitières
Inde	305	129
Brésil	223	24
Chine	98	7
États-Unis	94	9
Europe à 28	88	23
Argentine	55	2
Australie	27	2
Russie	20	7

La France est parmi les premiers producteurs de viande bovine en Europe. Elle dispose du cheptel le plus important d'Europe avec un total de 18,5 millions de têtes (Tableau 2). Le cheptel destiné à la production de viande est constitué principalement de races charolaise (1,4 million têtes), limousine (1,1 million têtes), blonde d'Aquitaine (0,5 million têtes), salers (0,2 million têtes) et aubrac (0,2 million têtes) (GEB, 2019). Le nombre de têtes a augmenté de 47 % pour la race aubrac entre 2009 et 2019, et de 9 % pour les races salers et limousine. En revanche, une diminution de 13 % est observée pour la race charolaise et de 2 % dans la Blonde d'Aquitaine sur la même période.

En décembre 2018, le nombre d'exploitations détenant des vaches allaitantes est d'environ 83 000. La majorité des étables sont composées de 30 à 99 vaches (47 %) ou de 5 à 29 vaches (43 %), le reste est constitué d'élevages intensifs avec plus de 100 vaches allaitantes (GEB, 2019).

Tableau 2 : Évolution du nombre de vaches et de bovins en France (en milliers de têtes) entre 1995 et 2019¹.

Au 1er janvier	1995	2000	2009	2017	2018	2019
Vaches laitières*	4760	4424	3939	3780	3747	3714
Vaches allaitantes	3983	4071	4133	4070	3977	3914
Total bovins	20507	20216	19902	19299	18894	18593

¹Source : Eurostat d'après SSP (1995 et 2000) et SPIE-BDNI, traitement Institut de l'Élevage. * Les vaches croisées de père laitier sont désormais comptabilisées comme laitières (2005 à 2018).

La région Nouvelle-Aquitaine a gardé sa place comme région avec l'effectif de vaches allaitantes le plus important (22 %), devant l'Auvergne-Rhône-Alpes (17 %) et l'Occitanie (13 %).

I.2. Production et consommation de viande bovine en France

I.2.1. La production de viande bovine en France

La France est le premier producteur de viande bovine en Europe avec un total de 1,43 millions tonnes équivalentes carcasse (téc) devant l'Allemagne (1,13 million téc), l'Italie (0,78 million téc) et l'Espagne (0,69 million téc) en 2019 (Figure 1). La viande bovine française est issue des troupeaux type « laitiers » (animaux issus de la production de lait) pour 32 % et des troupeaux allaitants (animaux élevés spécifiquement pour leur viande) pour 68 %. Cette viande provient des vaches (47 %), des taureaux (reproducteurs d'un troupeau) (31 %), des génisses (femelles jeunes engraisées jusqu'à l'âge de trois ans en moyenne) (17 %), et des bœufs (taureau castré) (5 %) (Figure 2) (GEB, 2019).

Une part importante de ces viandes bovines est destinée à l'exportation (y compris 86% des jeunes bovins (JB)) (Figure 2), ce qui représente un chiffre de 2,5 milliards d'euros d'exportations en 2018 (y compris les animaux vivants hors reproducteurs et les abats) (GEB, 2019). La principale destination de ces exportations est l'Italie, suivie par l'Espagne et l'Allemagne.

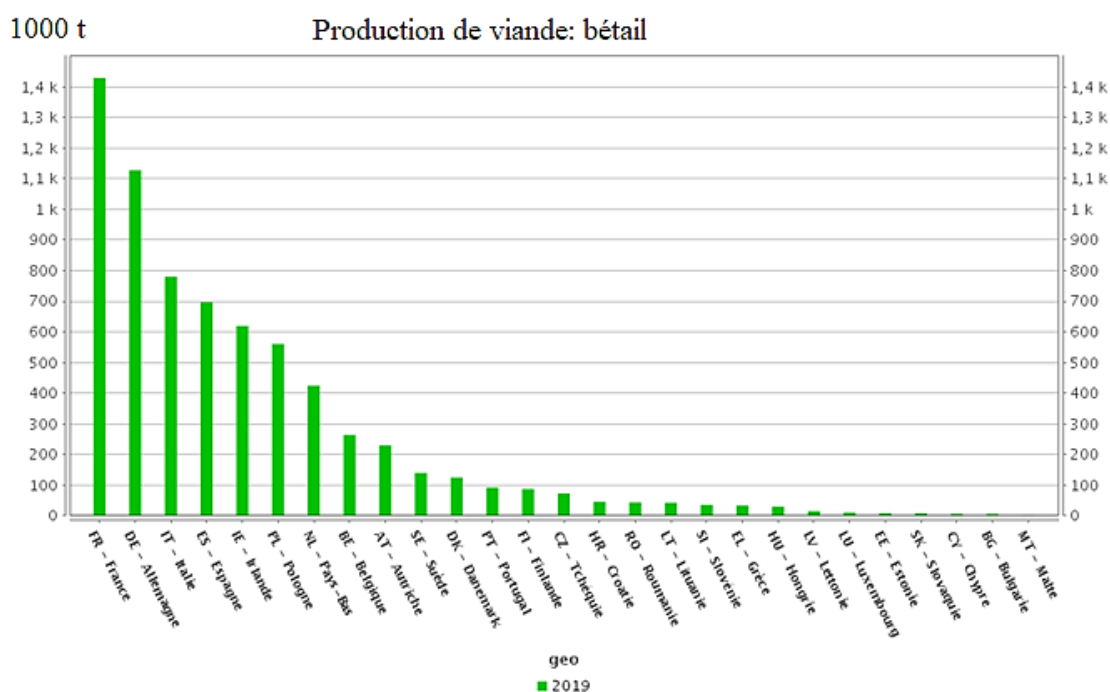


Figure 1 : Production de viande bovine (veaux, bœufs, taureaux, génisses et vaches, abattus par les abattoirs et à la ferme dont la viande est certifiée pour la consommation humaine) en Europe (1000 t équivalent carcasse) pour l'année 2019.

Source : Eurostat, 2019.

Production : 1,3 million de téc Consommation : 1,34 million de téc

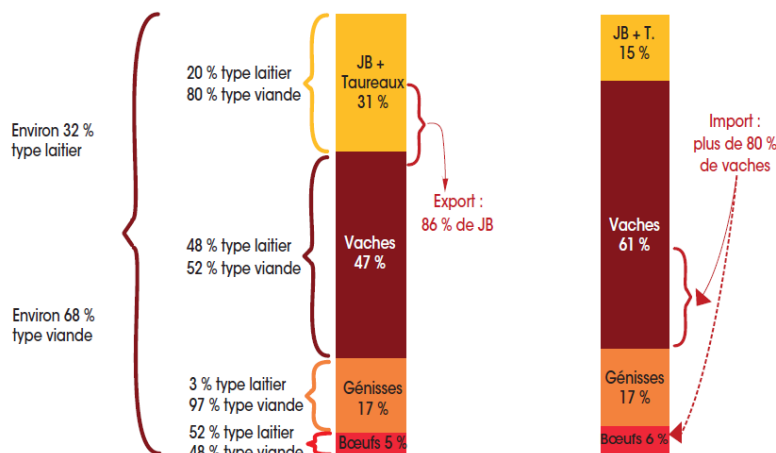


Figure 2 : Type de viande produite et consommée en France en 2018.
Source : Estimations GEB-institut d'élevage, % de tonnage de gros bovins.

1.2.2. La consommation de viande en France

Les Français consomment environ 23,2 kgec/habitant de viande bovine par an sur un total moyen de 87,5 kgec/habitant (toute espèce confondue) (France AgriMer, 2019). La consommation de viande a diminué de manière continue sur ces 20 dernières années, en particulier la viande bovine, ovine et porcine. En revanche, la part de la viande de volaille est passée de 25,8 % en 1998 à 34 % en 2018. Le niveau de consommation des viandes de boucherie (bœuf, veau, agneau, porc frais et viande chevaline) est passé de 58 à 46 grammes par jour entre 2007 et 2016.

La diminution de la consommation est liée à plusieurs facteurs économiques et sociétaux. Une étude récente menée par le CRÉDOC sur la consommation des produits carnés révèle une diminution de 12 % sur les 10 dernières années. Les principales explications sont liées à l'évolution des comportements alimentaires chez le consommateur, l'augmentation des prix (surtout de la viande bovine), les inquiétudes pour la santé humaine, la sensibilisation à l'impact sur l'environnement, au bien-être animal, mais aussi à la qualité des produits par rapport aux prix élevés (CREDOC, Tavoularis, 2018).

1.3. Une viande de qualité pourrait-elle regagner la confiance des consommateurs ?

L'un des plus importants axes pour les industriels afin de regagner la confiance des consommateurs est de s'inscrire dans une démarche de qualité qui englobe toute la chaîne de production. En effet, des études ont montré qu'une partie importante de non-rachat de la viande est liée à la grande variabilité de la qualité des produits, et notamment la tendreté (Bonny et al., 2017; Gagaoua & Picard, 2020). La qualité de la viande est un caractère complexe et multifactoriel. Il résulte de l'effet combiné de plusieurs facteurs biologiques et

technologiques. Cette notion de qualité est très étendue et généralement, on distingue les qualités nutritionnelles, hygiéniques, technologiques et organoleptiques.

I.3.1. Les qualités nutritionnelles de la viande

La viande est une source de nutriments particulièrement concentrée et traditionnellement considérée comme essentielle à la croissance et au développement optimal de l'homme (Higgs, 2000). Elle est particulièrement riche en protéines de haute valeur biologique, ainsi qu'en micronutriments tels que le fer, le sélénium, le zinc et la vitamine B12 (Pereira & Vicente, 2013).

I.3.1.1. Les protéines

La teneur moyenne en protéines de la viande est comprise entre 20 et 24 % (Wyness, 2016), soit 30 à 50 % des apports en protéines conseillés pour un adulte¹, avec une grande variabilité selon les espèces et les morceaux de viande. De plus, les protéines de la viande se distinguent par leur richesse en acides aminés essentiels (isoleucine, leucine, lysine, méthionine, tryptophane, thréonine, valine et phénylalanine) qui doivent être fournis par l'alimentation (G. Wu, 2009). Les végétariens doivent combiner les céréales et les légumineuses pour obtenir tous les acides aminés essentiels, car les céréales sont pauvres en lysine, et les légumineuses sont faibles en méthionine (Elango, Ball, & Pencharz, 2009).

I.3.1.2. Teneur en matière grasse

La teneur en matière grasse de la viande diffère de manière significative selon l'espèce animale, son âge et la partie de la carcasse utilisée (Valsta, Tapanainen, & Männistö, 2005). Suite à 30 ans d'action de la filière pour améliorer la composition corporelle des animaux, une étude menée par l'INRAE et pilotée par le centre d'information des viandes (CIV) entre 2006 et 2009 a montré que la majorité des morceaux de viande crue analysés ont moins de 8 % des lipides (CIV & Inrae, 2009). La viande est riche en acides gras polyinsaturés, parmi ces acides gras, on trouve les oméga-3 qui ont un rôle protecteur dans les maladies cardiovasculaires (Calder & Yaqoob, 2009). De plus, la viande peut contribuer jusqu'à 20 % de l'apport en ces oméga-3 (Russo, 2009).

I.3.1.3. Vitamines et minéraux

La viande est aussi une source importante de minéraux et de vitamines. Chez le bœuf, la teneur en fer varie entre 2,3 et 3,4 mg/100g d'aliment cru ce qui représente 15 à 20 % des apports conseillés pour les femmes et jusqu'à 40 % pour les hommes. Le fer héminique représente 70 % du fer de la viande de bœuf, et son coefficient d'absorption est de 25 %, comparé avec le fer non-héminique des végétaux (moins de 10 %) (CIV & Inrae, 2009). De plus, la viande rouge fournit environ 25 % des apports alimentaires recommandés en riboflavine, niacine, vitamine B6 et acide pantothénique pour 100 g, et pratiquement deux tiers des besoins journaliers en vitamine B12 (Williams, 2007). Elle est également l'une des meilleures sources de zinc, de sélénium et de phosphore.

¹ 0.83g de protéines/kg de poids/jour (OMS)

I.3.2. La qualité hygiénique de la viande

La qualité hygiénique de la viande regroupe les caractéristiques microbiologiques et chimiques liées à la santé du consommateur. L'aliment ne doit pas contenir des microbes pathogènes ou des substances étrangères tels que les résidus de médicaments ou d'autres substances susceptibles d'altérer le goût ou l'odeur de la viande et surtout la santé du consommateur.

I.3.3. Les qualités technologiques de la viande

Les qualités technologiques de la viande se définissent par son aptitude à la conservation et à la transformation. Les principaux facteurs qui contribuent à la qualité technologique des viandes sont le pH, la capacité de rétention d'eau et l'aptitude à la conservation.

I.3.3.1. Le pH

Le pH a un effet sur la qualité de la viande et principalement sur la couleur. Son évolution post-mortem détermine l'aptitude de la viande à la conservation et à la transformation.

I.3.3.2. La capacité de rétention d'eau

La capacité de rétention d'eau permet d'apprécier la qualité de la viande et son rendement en chair. Ce pouvoir de rétention d'eau est en relation forte avec l'aspect que présentera la viande au moment de la vente, et donc sa mesure permet une prévision sur son mode d'utilisation.

I.3.3.3. L'aptitude de la viande à la conservation

L'aptitude de la viande à la conservation par réfrigération est influencée directement par le pH. Une viande de pH élevé (> 6) est considérée comme inapte à la conservation, car ce pH élevé favorise la dégradation des glucides par les microorganismes, ce qui induit des mauvaises odeurs (Newton & Gill, 1981).

I.3.4. Les qualités organoleptiques de la viande et ses facteurs de variation

Pour le consommateur, la qualité de la viande se traduit par sa couleur, sa flaveur, sa jutosité et sa tendreté. Ce sont donc, les qualités perçues par les sens du consommateur.

I.3.4.1. La couleur

La couleur est la première qualité perçue par le consommateur au moment de l'achat. C'est donc un premier critère qui détermine la décision d'achat de la viande. Elle dépend principalement de la quantité de pigment (myoglobine) présent dans le muscle. Les viandes rouges intenses sont très riches en myoglobine. L'intensité de la couleur dépend aussi du pH du muscle. Si le pH dépasse 6, la viande devient très foncée (Monin, 1991).

I.3.4.2. La flaveur

La flaveur de la viande se traduit par l'ensemble des sensations perçues par le goût et l'odorat lorsque le morceau de viande est en bouche. Elle est déterminée par les molécules

précurseurs d'arômes qui sont présentes dans le gras et qui se libèrent au moment de la cuisson.

I.3.4.3. La jutosité

La jutosité est un autre paramètre qui détermine la qualité organoleptique de la viande. C'est l'impression initiale d'humidité libérée à la surface de la viande pendant la mastication et le degré de la salivation induite (Muir, Deaker, & Bown, 1998). Elle dépend de la teneur en eau et en graisse, c'est pourquoi elle est liée à la capacité de rétention d'eau (Monin, 1991). La jutosité est accentuée par la stimulation de la salivation, due en particulier à la présence du gras intramusculaire. Ce dernier empêche la viande de se sécher pendant la cuisson.

I.3.4.4. La tendreté

C'est la qualité la plus appréciée par les consommateurs et sa variation affecte leur décision de rachat. La tendreté est définie comme la facilité avec laquelle la viande est découpée puis broyée lors de la mastication. Bien que l'étiologie précise ne soit pas entièrement comprise, un certain nombre de facteurs affectent la tendreté. Parmi ces facteurs impliqués dans la variation de la tendreté, l'âge des animaux, la durée et la température de maturation, la baisse du pH post-mortem, la longueur du sarcomère et la protéolyse, qui affectent la conversion du muscle en viande. Cependant, il est désormais évident que la variation d'autres facteurs tels que la composition génétique, la composition du muscle et le type des fibres ainsi que la race et la nutrition des animaux peut également contribuer à la variation observée de la tendreté de la viande (Muir et al., 1998 ; Mwangi et al., 2019 ; Špehar, Vincek, & Žgur, 2008 ; Wood et al., 1999).

Partie II. Amélioration génétique des races à viande en France

II.1. Introduction

Nous avons vu dans la première partie que les qualités des viandes sont gouvernées par des facteurs génétiques et environnementaux, notamment les procédés de transformation, les conditions de conservation et les méthodes de cuisson qui ont un rôle majeur dans la qualité des produits. Plus en amont, les conditions de production au sein des élevages, au travers de l'hygiène et de la santé des animaux, l'alimentation et la gestion des troupeaux jouent aussi sur cette qualité des produits. Enfin, tout à fait, en amont, le type d'animal qu'on élève ou sa génétique influe également sur la qualité des produits. La filière s'est concentrée sur l'optimisation de ces facteurs afin d'obtenir un produit final de qualité. Cependant, le déterminisme génétique de la qualité des viandes reste largement inexplicé.

Le séquençage du génome complet de plusieurs espèces d'animaux domestiques et le progrès considérable des technologies de génotypage relativement peu coûteuses, ont permis d'entrevoir des perspectives beaucoup plus favorables pour l'identification de gènes impliqués dans les caractères d'intérêts économiques, y compris la qualité des viandes. Ces avancées technologiques ont permis aussi, d'un point de vue fondamental, de mieux comprendre la relation génotype-phénotype et ses mécanismes impliqués, mais aussi d'exploiter ces outils moléculaires afin de sélectionner les animaux pour des caractères d'intérêt économiques.

Dans cette partie, nous allons faire un point sur l'amélioration génétique des bovins allaitants, son organisation, ses méthodes et leurs limites.

II.2. Rappel de notions de génétique

II.2.1. Les marqueurs génétiques, locus, allèles

Les marqueurs génétiques sont des variations (polymorphismes) de la séquence d'acide désoxyribonucléique (ADN) entre individus à un locus donné (un emplacement sur le génome) dans une population. Les différentes formes à un locus donné sont appelées allèles. Il existe de nombreux types de marqueurs, qui vont d'un simple changement de paire de bases (Polymorphisme d'un seul nucléotide ou SNP : *Single Nucleotide Polymorphism*) à des microsatellites (répétition de 1 à 4 paires de bases), des CNV (*Copy Number Variation*) ou encore des insertions/délétions. Ces dernières années, les chercheurs en génétique animale se sont concentrés sur l'identification de marqueurs génétiques pour les caractères de production et de la qualité des produits comme la viande. Les avantages économiques de cette amélioration génétique de la production ou de la qualité sont importants et peuvent être obtenus en démêlant la relation entre le génome et ces caractères (appelés phénotypes) (J. P. Kerry & Ledward, 2009).

Les SNP sont aujourd'hui les marqueurs les plus utilisés chez le bovin et correspondent à la substitution d'une seule paire de base entre individus de la même espèce. Ces formes de polymorphismes sont très souvent bi-alléliques et nombreuses sur le génome (ex. Figure 3). Les SNPs peuvent se trouver dans des régions codantes (exons), régulatrices (ex. Régions promotrices) ou introniques (introns) du génome, et peuvent affecter l'expression génique (en modifiant les sites de fixation des facteurs de transcription) la conformation des protéines, les sites d'épissage ou autres mécanismes de la régulation génique.

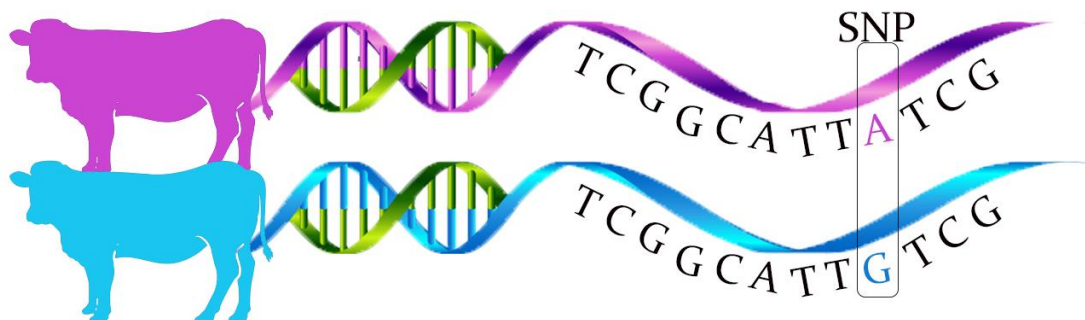


Figure 3 : Exemple de polymorphisme SNP entre deux individus (image adaptée de icbf²).

II.3. Les méthodes de génotypage des SNP

Le génotypage des SNPs a une grande importance dans l'étude de la structure génétique des populations, la cartographie génétique et les études de phylogénie. Il consiste à déterminer les locus SNP à l'échelle du génome ou de la région génomique d'intérêt. C'est un outil de base pour les analyses d'association pan-génome et la cartographie fine.

Les nouvelles technologies basées sur la robotisation ainsi, la réduction des volumes réactionnels ont permis de réduire considérablement les coûts de génotypage. Le développement rapide de ces technologies a permis d'établir plusieurs bases de données publiques pour la collecte et l'identification des SNPs comme par exemple, dbSNP et Ensembl (Tableau 3).

Tableau 3 : Exemple de bases de données SNP

Base de données	Site web
dbSNP	https://www.ncbi.nlm.nih.gov/snp/
Ensembl	https://www.ensembl.org/Bos_taurus/Info/Index
Santa Cruz	http://hgdownload.soe.ucsc.edu/downloads.html#cow
EMBL-EBI	https://www.ebi.ac.uk/eva/

Il existe plusieurs méthodes de génotypage des SNPs. Certaines méthodes permettent le génotypage à un seul locus et d'autres (méthodes à haut débit) permettent le génotypage de plusieurs milliers voire millions de SNPs simultanément. En génétique animale il est très souvent nécessaire de génotyper plusieurs marqueurs allant de quelques dizaines pour le suivi des QTLs en sélection assistée par marqueur (SAM) jusqu'à des dizaines de milliers, voire des millions pour les analyses d'association (GWAS : *Genome-Wide Association Studies*). Ci-dessous nous allons présenter quelques exemples des principales technologies :

II.3.1. La méthode TaqMan

La technologie « 5' nuclease assay » (Livak, 1999) connue sous le nom commercial TaqMan est une technique qui offre un moyen rapide pour l'analyse des polymorphismes SNP. Le principe de la méthode est illustré dans la Figure 4.

² Source : <https://www.icbf.com/?p=6092>

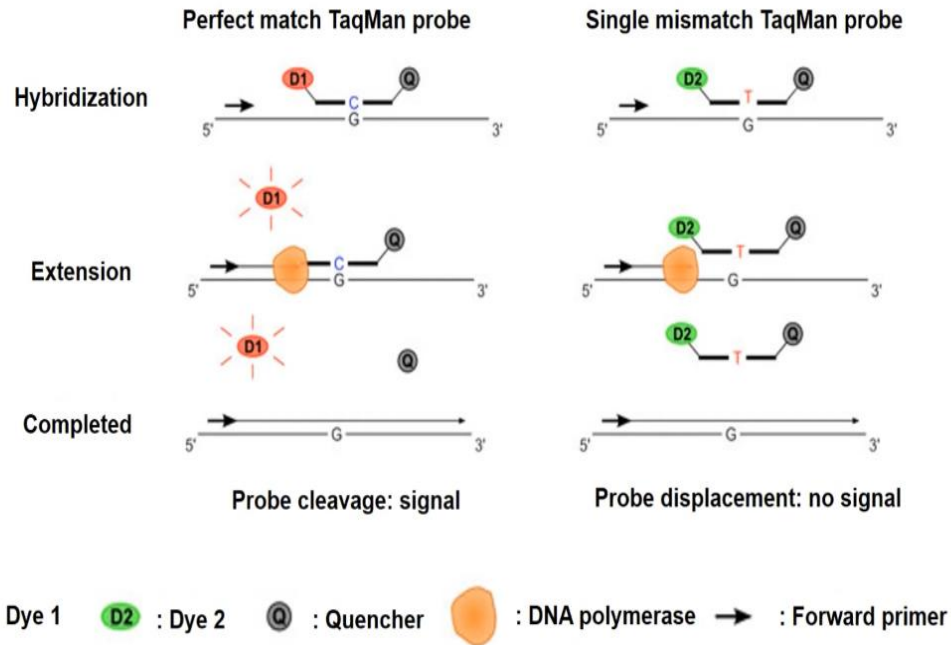


Figure 4 : Principe du génotypage par la méthode TaqMan (source : applied-maths).

Une réaction PCR (*Polymerase Chain Reaction*) avec le SNP à génotyper est réalisée en ajoutant une sonde marquée par un fluorochrome (quencher (Q) et rapporteur (D)) qui se fixe au niveau du SNP. Pendant l'étape d'élongation de la PCR, l'ADN polymérase dégrade l'extrémité 5' de la sonde, ce qui libère le rapporteur, qui devient fluorescent (Figure 4 à gauche). Si l'allèle au SNP n'est pas complémentaire avec le nucléotide de la sonde, alors cette dernière ne s'hybride pas et par conséquent, elle ne sera pas dégradée par la polymérase (Figure 4 à droite) et donc la fluorescence n'est pas décrochée. En pratique, deux sondes (avec des fluorochromes différents), chacune complémentaire de l'un des allèles au SNP sont utilisées pour génotyper un échantillon.

II.3.2. Génotypage sur micropuce

Les puces à ADN sont des supports solides divisés en plusieurs milliers voire millions de cellules, sur lesquels sont fixés des fragments d'ADN (ex. Figure 5).

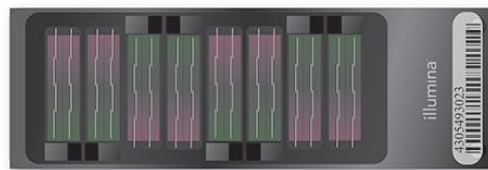


Figure 5 : Puce à SNP 777k (BovineHD BeadChip) pour génotyper 8 individus (échantillons) pour 770 000 SNP (Illumina).

Les puces commerciales disponibles aujourd'hui pour le génotypage des SNPs sont issues des plateformes Affymetrix et Illumina. Ces puces sont très variées en terme de densité de marqueurs (Tableau 4). Des puces à haute densité (HD) sont disponibles aujourd'hui pour l'humain et plusieurs espèces animales et végétales.

Tableau 4 : Exemple de puces commerciales existantes (Illumina & Affymetrix).

Nom de la puce	Société	Nombre de SNP	Nombre d'échantillons maximum par puce
Bovine LD	Illumina	7 931	24
Bovine MD	Illumina	53 218	24
Bovine HD	Illumina	777 962	8
OvineSNP50	Illumina	54 241	12
Genome-Wide BOS 1	Affymetrix	648 855	1
Genome-Wide chicken	Affymetrix	580 961	1

Pour les puces Affymetrix, chaque SNP est représenté par une paire de sondes dans deux puits différents, représentant les deux génotypes possibles au locus. La présence ou l'absence de signal dans chaque puits permet de déterminer le génotype pour ce SNP (Figure 6a).

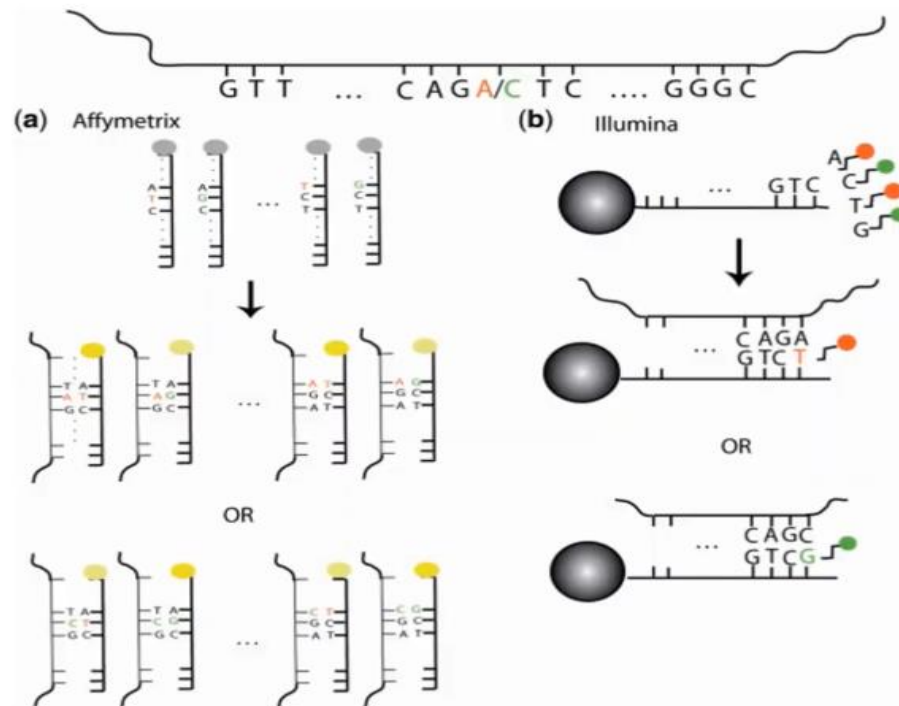


Figure 6: Principe de la méthode de génotypage d'Affymetrix (a) et d'Illumina (b).

Pour les puces Illumina, les sondes présentes dans chaque puits correspondent à la séquence d'ADN génomique précédant le SNP à génotyper, auxquelles l'ADN de l'individu s'hybride. Dans une phase d'élongation et en présence d'ADN polymérase et de nucléotides marqués, le signal émis permet de déterminer le nucléotide incorporé, et par conséquent le génotype de l'individu pour ce SNP (Figure 6b).

II.4. Rappel historique de l'amélioration génétique des animaux d'élevage

L'amélioration génétique est une activité économique qui implique plusieurs acteurs. Elle cible le secteur tout à fait en amont des filières pour mettre à disposition des éleveurs et des filières un type d'animal adapté à leurs besoins en tirant parti des différences d'origine génétique intra ou inter-populations, dans un objectif de sélection bien défini à l'avance.

La sélection animale a démarré de façon empirique dès les débuts de la domestication qui date de 10 500 ans (avant notre ère) selon les études (Bollongino et al., 2012). Certains caractères ont probablement fait objet de sélection empirique dès le début de la domestication et notamment ceux qui favorisaient la domestication elle-même comme la docilité des animaux d'élevage.

À partir des années 1850, des éleveurs se sont réunis pour choisir des groupes d'animaux selon certaines caractéristiques, ce qu'on appelle aujourd'hui une race animale, qui est un phénomène important dans l'homogénéisation de l'aspect extérieur des populations d'élevage. Des « livres généalogiques » (*herdbooks*) ont été élaborés pour enregistrer les généalogies des animaux et suivre leurs pedigrees dans un but de sélection (Pierre Sellier, Didier Boichard, 2019).

La redécouverte des lois de Mendel en 1900 et le développement de la génétique et ses branches (génétique des populations et génétique quantitative) ont permis le développement des concepts et des méthodes qui ont trouvé leur pleine application en amélioration génétique des animaux à partir du début des années 1950. Le développement des techniques de biologie moléculaire et leurs applications en amélioration génétique ont eu lieu à partir de la fin des années 1980, grâce aux techniques de typage moléculaire de gènes connus.

II.5. Les méthodes de sélection et leurs limites

II.5.1. Sélection sur descendance

La sélection sur descendance est une méthode très largement utilisée en France pendant plusieurs années. Elle se repose sur les performances des descendants (mesurées en ferme ou en station) pour sélectionner les reproducteurs mâles qui seront utilisés à grande échelle par insémination artificielle (IA), notamment dans les filières où les caractères majeurs de sélection sont exprimés que dans un seul sexe (production laitière par exemple). C'est une méthode très fiable pour la sélection des mâles et elle donne de bons résultats pour les caractères à faible héritabilité.

La limite de ces méthodes est l'intervalle de générations qui dure plusieurs années chez les bovins afin d'obtenir des performances mesurables (comme la mesure du poids à l'abattage, rendement en carcasse, ... etc.), mais aussi le coût élevé de sa mise en œuvre (Dudouet, 2010). Certains caractères d'intérêt économique comme la qualité des viandes sont très difficiles à améliorer par ces méthodes comme par exemple la tendreté ou le persillé, car les mesures de ces caractères sont très onéreuses et ne peuvent pas être mesurées en routine dans les abattoirs.

II.5.2. Sélection sur des gènes connus

La sélection génique repose sur des gènes connus à effet majeur sur la variabilité d'un caractère. En effet, le génotypage des individus peut se faire dès la naissance chez les mâles et les femelles et ne nécessite pas de mesure de performances. Cette approche permet un gain de temps en réduisant l'intervalle de génération, ainsi d'élargir le champ de sélection en différenciant les individus pour lesquels certains gènes contrôlent des caractères exprimés uniquement chez le sexe opposé. La sélection sur gène identifié permet aussi de réduire les

coûts de phénotypage pour certains caractères lourds à mettre en place ou nécessitant l'abattage de l'animal (ex. Mesures de la composition corporelle), et ainsi permet de réduire les risques dans le cas de sélection sur la résistance à des maladies.

En pratique, dans les populations animales, cinq types d'utilisations de gène à effet majeur sur un caractère ont été cités par (Elsen, 2011) :

- Élimination des défauts génétiques (appelés aussi tares génétiques) comme la syndactylie (Duchesne et al (2006)) ou BLAD (*Bovine Leucocyte Adhesion Deficiency*) (Shuster et al (1992)) chez les bovins ; une maladie génétique induisant une déficience du système immunitaire conduisant à la mort de l'animal au bout de quelques mois.
- Sélection d'allèles de résistance : une approche rapide et de faibles risques par rapport à la méthode classique qui nécessite des phénotypes via des tests d'infection dans le cas des maladies infectieuses ou bien nécessitant beaucoup de temps dans le cas des maladies à durée d'incubation longue (ex. Encéphalopathies). La résistance à la tremblante est un bon exemple de ce type de sélection chez les ovins en France (ex. Elsen et al 1999).
- Sélection sur un caractère non-mesurable en ferme comme l'aptitude au désaisonnement de la reproduction chez les ovins.
- Présélection d'animaux avec de bons allèles dans les schémas de sélection où l'effectif d'animaux candidats à la sélection est faible, ouvrant la voie à l'effet de la dérive qui peut éliminer les bons allèles par le fait du hasard.
- L'introgression d'allèles favorables identifiés chez une race donnée dans d'autres races par des séries de croisement en retour comme par exemple, le gène sans cornes chez le Charolais.

L'introgression d'allèle favorable dans une race receveuse peut avoir des effets négatifs au moment de l'introgression, car l'expression de l'allèle n'est pas toujours maîtrisée tant que le processus d'introgression n'est pas fini. Il peut y avoir des effets d'épistasie défavorable entre l'allèle introduit et les autres allèles de la race receveuse. Une autre limite de ces méthodes de sélection est l'effet défavorable possible sur d'autres caractères d'intérêt économique ou zootechnique principalement pour les loci pléotropiques. Finalement, dans le cas de gènes de résistance à des maladies, il peut apparaître des phénomènes de résistance suite à des mutations chez l'agent pathogène, rendant la sélection inefficace.

II.5.3. Sélection assistée par marqueurs (SAM)

Suite au développement des technologies de génotypage et à l'arrivée des marqueurs génétiques, de grands programmes de recherche ont été développés pour détecter des régions chromosomiques induisant des variations sur les caractères chez les animaux d'élevage, ce qu'on appelle des QTL (pour *Quantitative Trait Loci*) (c.f. partie III). À partir des informations obtenues, il est possible de génotyper les animaux sur les marqueurs identifiés par l'approche QTL et qui sont à proximité des gènes fonctionnels (Sélection assistée par marqueurs : SAM).

La SAM comme la sélection génique permet de réduire les intervalles de génération en génotypant les animaux à la naissance, et permet donc de diminuer les coûts de phénotypage et de réduire les risques. Cependant, la SAM diffère de la sélection génique par le fait que les données portent sur des marqueurs et non sur des génotypes actifs, les QTLs sont très souvent nombreux et polygéniques expliquant une part de la variabilité phénotypique. Des méthodes permettent d'estimer la valeur génétique assistée par marqueur (MAEBV) à partir

de la combinaison des informations génotypiques et phénotypiques des individus candidats à la sélection.

On distingue deux types de SAM en fonction de la précision des informations disponibles.

II.5.3.1. La SAM1

Mise en place en 2000 dans les trois principales races laitières françaises (Holstein, Montbéliarde et Normande) (Didier Boichard et al., 2002). Une dizaine de milliers d'animaux génotypés pour 43 microsatellites et marquant 14 QTL liés à la quantité et qualité du lait, la résistance aux mammites et à la fertilité femelle ont été utilisés, afin d'estimer la valeur génétique des jeunes mâles candidats au testage sur descendance et les jeunes femelles candidates au noyau de sélection. La mise en route de la SAM en bovins laitiers en France entre 2001 et 2007 a permis d'augmenter la précision des index des jeunes animaux comparés aux index obtenus par sélection classique.

II.5.3.2. La SAM2

Le développement des puces de génotypage à moyenne densité (54 000 SNPs) a permis de faire des localisations QTL beaucoup plus fines. Plusieurs régions QTL avec des intervalles de confiance beaucoup plus petits ont été identifiées par caractère à l'aide de ces marqueurs en exploitant le déséquilibre de liaison, ce qui rend la SAM2 proche d'une sélection multi-génique (François Guillaume et al., 2008). La SAM2 a été mise en place en 2008 dans les mêmes races de bovins laitiers en France, mais au même moment, la sélection génomique (SG) a émergé, et par conséquent, la majorité des opérateurs ont choisi de se lancer directement dans la nouvelle approche de sélection (Elsen, 2011). En France, le choix a été fait sur la combinaison de la SAM2 et la SG vu la disponibilité des informations (François Guillaume et al., 2011).

Les limites de la SAM sont similaires à celles de la sélection génique. En effet, l'apparition d'un déséquilibre de liaison avec le fond polygénique induit un retard de progrès à long terme si l'allèle au QTL est fixé. La localisation des QTL n'est pas aussi précise et nécessite le génotypage de plusieurs apparentés afin d'estimer les effets des QTLs. Enfin, le testage sur descendance reste toujours nécessaire dans ces approches, car la SAM ne permet pas de capturer toute la variabilité génétique vu le caractère polygénique des phénotypes d'intérêt économique.

II.5.4. Sélection génomique

La facilité de génotypage à haut débit des marqueurs SNPs et l'arrivée des puces de génotypage pan-génomiques de type BovineSNP50 (Illumina), ainsi le coût relativement faible a permis de génotyper des animaux sur des dizaines de milliers de marqueurs.

Mise en place à partir de 2009 chez les bovins laitiers en France (Holstein, Montbéliarde et Normande), et en 2015 chez les 3 principales races allaitantes (Charolaise, Limousine et Blonde d'Aquitaine), l'évaluation génomique est un calcul d'index qui se repose sur l'estimation de la valeur génétique des candidats à la reproduction à partir de leur génotype à des dizaines de milliers de marqueurs suite à l'établissement des équations de prédiction sur une population de référence (Meuwissen, Hayes, & Goddard, 2001). La SG a l'avantage d'élargir la gamme des pères à taureaux, et par conséquent de maintenir la variabilité génétique chez les animaux sélectionnés. Elle permet aussi d'évaluer de nouveaux critères avec plus de précision et sur

des intervalles de temps très réduits, car le génotypage et l'estimation de la valeur génétique (GEBV) de l'animal peut se faire à la naissance dans les deux sexes, sans passer par des testages sur descendance. Par conséquent, la SG réduit les coûts de schémas de sélection.

L'une des principales limites de la mise en place d'une SG est la nécessité d'avoir une grande population de référence (génotypée et phénotypée avec plusieurs milliers individus) et d'une densité de marqueurs importante, ainsi que des capacités de calcul puissantes.

Le développement des méthodes d'amélioration génétique est appuyé sur le développement d'autres techniques de reproduction comme l'insémination artificielle qui permet la diffusion de semences améliorées à plusieurs milliers de vaches à l'échelle nationale et internationale, et le transfert d'embryons qui permet d'avoir un gain important sur le progrès génétique. L'amélioration génétique est aussi basée sur le développement des capacités de calculs et l'élaboration de méthodes statistiques sophistiquées pour traiter les masses d'information dans une phase appelée évaluation génétique (c.f. II.7).

II.6. L'organisation et les acteurs de la sélection des bovins allaitants en France

L'amélioration génétique est une activité très réglementée à l'échelle nationale et internationale. En France, un cadre législatif qui s'appelait, « la loi d'orientation agricole³ » votée en 2006 a remplacé « la loi sur l'élevage⁴ » de 1966 qui restait depuis une quarantaine d'années le cadre légal des activités d'amélioration génétique. Le règlement zootechnique de l'Union européenne (UE)⁵ voté en 2016 et mise en application en novembre 2018 a bouleversé le dispositif génétique en France pour la plupart des espèces d'élevage y compris les bovins. Les changements apportés par ce nouveau dispositif concernent principalement la gestion des livres généalogiques, la mise en œuvre des programmes de sélection, le contrôle des performances et l'évaluation génétique des animaux. La mission des organismes de sélection (OS) est élargie dans ce nouveau dispositif pour prendre le rôle de certains acteurs définis dans la loi d'orientation agricole. Dans le cadre du règlement zootechnique, les OS prennent en charge la tenue du livre généalogique, le contrôle des performances (recueil des informations), la conduite des programmes de sélection et le calcul (indexation) et la publication des valeurs génétiques des animaux qui était pris en charge conjointement par l'Institut de l'Élevage (IDELE) et l'unité de Génétique Animale et Biologie Intégrative (GABI, UMR 1313 INRAE) et le Centre de Traitement de l'Information Génétique (CTIG) du département de génétique animale de l'INRAE à Jouy-en-Josas.

D'autres acteurs, notamment les opérateurs de base comme les entreprises de sélection (ES) (coopératives d'IA et leurs unions, groupements d'éleveurs et firmes privées) prennent en charge l'élaboration des schémas de sélection, le choix ou la procréation des futures reproducteurs et la mise sur le marché des reproducteurs sélectionnés. Le dispositif d'amélioration génétique actuel en France est donc passé d'un système administré à un système basé sur la contractualisation, les principaux changements sont illustrés dans le tableau suivant (Tableau 5) selon le ministère de l'Agriculture de l'agroalimentaire et de la forêt (2016).

³ L. n° 2006-11

⁴ L. n° 66-1005

⁵ Règlement (UE) 2016/1012 du Parlement européen et du Conseil du 8 juin 2016 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.

Tableau 5 : Les principales différences d'organisation entre la loi d'orientation agricole de 2006 et le règlement zootechnique de l'UE.

Loi d'orientation agricole	Règlement zootechnique de l'UE
<ul style="list-style-type: none"> - Les OS agréés par l'état ont la responsabilité de la définition des orientations raciales et de la tenue du livre généalogique. - Les établissements de l'élevage (EDE) effectuent l'enregistrement des parentés et l'identification des animaux. - Les organismes de contrôle de performances ont la responsabilité d'enregistrement des performances. - Les entreprises de sélection (ES) élaborent les schémas de sélection et s'occupent de la mise sur le marché des reproducteurs. 	<ul style="list-style-type: none"> - Les OS de l'UE (OSUE) agréés par l'état et reconnue par les états membres de l'UE prennent en charge l'organisation et les décisions de sélection. - Les EDE conservent leurs missions d'enregistrement et de certification de la parenté des ruminants. - Les organismes de contrôle des performances gardent aussi leurs rôles de réalisation des contrôles de performances par délégation des OS. - GeneEval est une équipe d'ingénieurs en génétique et en informatique créée en 2018 par Alice (union de coopératives d'élevage) et race de France a pour mission d'assurer le calcul des index des ruminants pour les caractères choisis par chaque OS. - Les ES gardent leurs rôles de mise en œuvre des schémas de sélection.

II.7. Les étapes d'un programme de sélection

L'amélioration génétique fait porter les efforts de sélection ou de gestion sur l'ensemble d'une population, même quand elle est répartie au sein de nombreux élevages. C'est un travail collectif entre plusieurs opérateurs comme on l'a vu précédemment. C'est une activité qui se passe en différentes étapes (Figure 7) dans un cadre appelé programme de sélection.

II.7.1. Définition des objectifs de sélection

La première étape dans un programme de sélection est de définir les objectifs de sélection, qui dépendent largement du contexte économique et social de chaque pays ou région, des contraintes de la filière et de l'élevage et de l'avenir du système d'élevage et des orientations raciales (ex. (Phocas et al., 1998)). Par exemple chez les bovins allaitants en France, les caractères relatifs aux facilités de naissance, les aptitudes bouchères (comme la croissance, le développement musculaire et squelettique, l'efficacité alimentaire et le rendement carcasse) et les qualités maternelles (comme la longévité, la fertilité, l'aptitude au vêlage, l'aptitude à l'allaitement et l'efficacité reproductive) sont intégrés dans les objectifs de sélection. Une analyse du contexte est donc nécessaire pour bien définir ces objectifs de sélection.

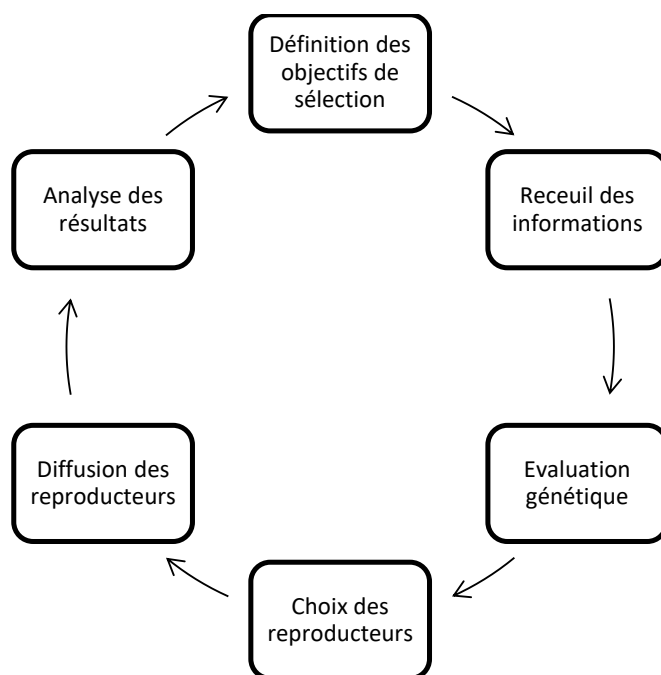


Figure 7 : Schéma représentatif des étapes d'un programme de sélection

En France, contrairement à d'autres pays européens comme l'Allemagne (troupeau à dominante laitier) ou l'Irlande (animaux croisés), la production de viande bovine est issue essentiellement de races pures (71 % en 2018) (institut de l'élevage, 2018). Cette particularité a permis la mise en place des évaluations génétiques dans la majorité des races allaitantes françaises (Griffon et al., 2017). Les objectifs de sélection diffèrent d'une race à l'autre. Chez le charolais par exemple, les objectifs prioritaires sont de maintenir une croissance naturelle et une meilleure qualité bouchère, de parvenir à la production d'un veau par vache et par an, sans assistance au vêlage et de produire des carcasses de bonne qualité. Chez la limousine, les objectifs actuels sont d'obtenir des vaches produisant un veau sevré par an sans assistance au vêlage et des taurillons de 15 à 17 mois donnant des carcasses haut de gamme (un poids de 400 kg de meilleure conformation et sans excès de gras). Idem pour la Blonde d'Aquitaine où la priorité est donnée au développement musculaire (production de viande), améliorer la capacité d'allaitement, et conserver le format, la vitesse de croissance et les facilités de vêlage de la race (France Génétique Élevage).

Malgré les orientations de sélection différentes pour chaque race, les éleveurs préfèrent utiliser globalement des taureaux avec l'index le plus élevé pour le développement musculaire, la conformation (sevrage, post-sevrage et jeunes bovins) et le développement squelettique au sevrage. Alors que les index pour la facilité de naissance, la finesse d'os, les aptitudes au vêlage et à l'allaitement, ou la réussite à l'IA ne semblent globalement pas être des critères pris en compte par les éleveurs lors du choix de leurs reproducteurs, même si leurs valeurs génétiques sont moyennes ou mauvaises pour ces objectifs. Cela montre peut-être la vision différente des besoins entre les organismes de sélection et les éleveurs (Griffon et al., 2017).

II.7.2. Recueil des informations

C'est une phase très importante dans tout programme de sélection, car sans informations, on ne peut pas évaluer le potentiel génétique des animaux. Différents types d'informations sont nécessaires dans un programme de sélection :

- **L'identification de l'animal** avec un « passeport » qui l'accompagne jusqu'à la fin de sa vie. Ce passeport contient plusieurs informations sur l'animal (son numéro unique d'identité, l'attestation sanitaire, sa race et son sexe).
- **L'état-civil de l'animal** qui consiste à enregistrer les parents de l'individu à sa naissance. C'est une activité importante dans les programmes de sélection et qui est établie sur la base de déclarations des éleveurs et assurée par des contrôles de filiation.
- **Le contrôle de performances** qui est la base de la sélection animale. C'est l'ensemble des mesures de caractères sur l'animal qui se réalisent en stations de contrôle ou en ferme par des spécialistes. Ce sont généralement des caractères quantitatifs qui prennent des valeurs numériques et qui sont économiquement importants, mais aussi, ils peuvent être valorisés au-delà de la sélection dans la conduite des élevages.
- **Le génotypage des animaux** consiste à déterminer les allèles à chaque locus sur le génome ou une partie du génome. On distingue différents types de typages selon les marqueurs ou les gènes à typer :
 - o Le génotypage pour des marqueurs sur des gènes majeurs à fonction connue impliqués par exemple dans des anomalies génétiques (ex. *BLAD*) ou des gènes d'intérêt (sans cornes, coloration, ... etc.) ou même pour le suivi de filiation (exemple de l'allèle *A^{br}* chez la Normande (Girardot et al., 2006)).
 - o Les marqueurs moléculaires de type SNP qui sont faciles à génotyper et de façon automatisée par les nouvelles technologies de génotypage (c.f. II.3). Des puces à SNP permettent de génotyper plusieurs individus pour des milliers voir des dizaines de milliers de marqueurs (Craig & Stephan, 2005; Gunderson, Steemers, Lee, Mendoza, & Chee, 2005; Matukumalli et al., 2009).

II.7.3. Evaluation génétique

L'évaluation génétique est une prédiction de la valeur génétique (index) des animaux pour un certain nombre de caractères. La mesure des performances montre une variabilité phénotypique entre les animaux d'élevage. Cette variabilité peut être due à une variabilité génétique ou à des différences du milieu (facteurs non-génétiques). On distingue deux types de milieu :

- Milieu contrôlé (macro-milieu) : rassemble les facteurs que nous pouvons identifier et enregistrer comme l'année et le lieu de naissance d'un animal.
- Milieu non contrôlé (micro-milieu) : c'est l'ensemble des facteurs environnementaux sur lesquels nous ne pouvons pas agir ou contrôler.

L'analyse de cette variabilité se base sur des modèles statistiques et mathématiques qui consistent à mesurer ou à estimer la part de la variabilité expliquée par chaque facteur de variation. On peut donc écrire le modèle suivant :

$$y = \mu + m + g + e$$

Avec **y** la performance, **μ** la moyenne, **m** l'effet du milieu contrôlé, **g** la valeur génétique (effet moyen du génotype) et **e** la résiduelle environnementale. La valeur génétique globale (**g**) peut se décomposer en une valeur génétique additive (**a**) et une valeur génétique non additive (**d**) :

$$g = a + d$$

Un paramètre très important en évaluation génétique est l'héritabilité (**h^2**) du caractère. Elle décrit dans quelle proportion un caractère quantitatif est transmis à la descendance. Par

définition c'est le rapport de la variance génétique (σ_A^2) sur la variance phénotypique (σ_P^2) (Gerald Wiener, 2009) :

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2}$$

Autrement dit, c'est la part de la variation additive sur la variation totale observée. Les valeurs de l'héritabilité varient entre 0 (pas d'effet génétique additif qui induit des variations sur le caractère) et 1 (toute la variabilité est d'origine génétique additive).

Le modèle génétique peut s'écrire à l'échelle de la population comme suit :

$$y = X\beta + Zu + e$$

Avec y est le vecteur des performances ($y \sim N(X\beta, V)$), X et Z sont des matrices d'incidence reliant les effets fixes et aléatoires (respectivement) aux performances. β est le vecteur des effets fixes (macro-milieu), u est le vecteur des effets aléatoires non résiduels (les valeurs génétiques) tel que $u \sim N(0, G)$ et e le vecteur des effets aléatoires résiduels ($e \sim N(0, R)$). $G = A\sigma_a^2$ est la matrice de variances-covariances entre effets aléatoires dans le cas du modèle génétique, A est la matrice des relations de parenté et σ_a^2 la variance des valeurs génétiques. $R = I\sigma_e^2$ est la matrice de variances-covariances entre effet résiduelles avec I la matrice d'identité et σ_e^2 , la variance résiduelle.

La méthode BLUP (*Best Linear unbiased Predictor*) (Henderson, 1975) permet d'estimer conjointement les effets fixes du milieu et les valeurs génétiques de tous les animaux. Cette méthode est basée sur une combinaison linéaire des données de l'individu et ses apparenté.

Henderson (1963) a montré qu'on peut estimer les valeurs génétiques en résolvant le système d'équations suivant :

$$\hat{u} = (Z'MZ + G^{-1})^{-1}Z'My$$

avec $M = R^{-1} - R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}$

On peut donc montrer que l'index BLUP (\hat{a}) d'un individu est une combinaison d'informations de son ascendance, lui-même et sa descendance (cité par (Sophie Allais, 2011b)), soit :

$$\hat{a} = x_1 * \hat{a}_{asc} + x_2 * (P - f^{\wedge}) + x_3 * \hat{a}_{desc}$$

avec x_1 le poids relatif à l'ascendance, \hat{a}_{asc} l'index sur ascendance, x_2 le poids relatif à la performance propre de l'animal (P), f^{\wedge} est l'estimation des effets fixes, x_3 le poids relatif à la descendance et \hat{a}_{desc} l'index sur descendance.

Les évaluations génétiques sont faites en ferme (Indexation des Bovins Allaitants : IBOVAL), en station ou sur descendance :

- En ferme sur des reproducteurs mâles et femelles pour les caractères relatifs aux conditions de naissance, les aptitudes bouchères, la croissance jusqu'au sevrage et les qualités maternelles. Les performances enregistrées sont le poids à la naissance, les conditions de naissance, le poids à 210 jours, le développement musculaire, le développement squelettique, la grosseur des canons (le diamètre de l'os canon de l'animal), la note de comportement au pointage et le nombre de mouvements lors de la pesée au sevrage (Boulesteix, 2018).
- En station de contrôle des performances pour les taureaux de monte naturelle (MN) ou d'IA sur les caractères d'aptitudes bouchères. Le but est de contrôler la conformation des animaux par pointage et mensurations et mesure de la composition corporelle de plusieurs individus dans des conditions homogènes.
- Sur descendance pour les aptitudes bouchères et/ou qualités maternelles de leurs descendants (Sophie Allais, 2011b).

II.7.4. Choix des reproducteurs

Une fois les animaux évalués sur la base de ces évaluations (calcul des index pour chaque caractère), l'éleveur peut choisir (sélectionner) ses reproducteurs pour les générations suivantes en fonction de la combinaison des aptitudes souhaitées pour son exploitation.

II.7.5. Diffusion des reproducteurs

Les reproducteurs choisis sur la base de leur valeur génétique vont être diffusés soit dans le cadre de MN soit dans le cadre d'IA. En élevage allaitant, la diffusion se fait majoritairement par MN (88 % en 2015) (Griffon et al., 2017) puisque la détection des chaleurs est difficile chez les vaches lors de la période de reproduction et surtout en présence de veau qui peut perturber le comportement de la mère, alors que chez les bovins laitiers, la diffusion se fait majoritairement par IA, car la détection de chaleur chez les femelles est beaucoup plus facile en bâtiment sans présence de veau.

II.7.6. Analyse des résultats de la sélection

La dernière étape d'une évaluation génétique consiste à analyser les résultats de sélection dans la population et voir si les moyennes des caractères sélectionnés ont évolué dans le sens souhaité, et aussi voir s'il n'y a pas d'effet sur d'autres caractères. Ces résultats permettent de revisiter les objectifs de sélection. Un exemple de progrès génétique estimé pour le poids-âge-type à 210 jours (PAT 210) au sevrage pour 9 races allaitantes est présenté dans la Figure 8.

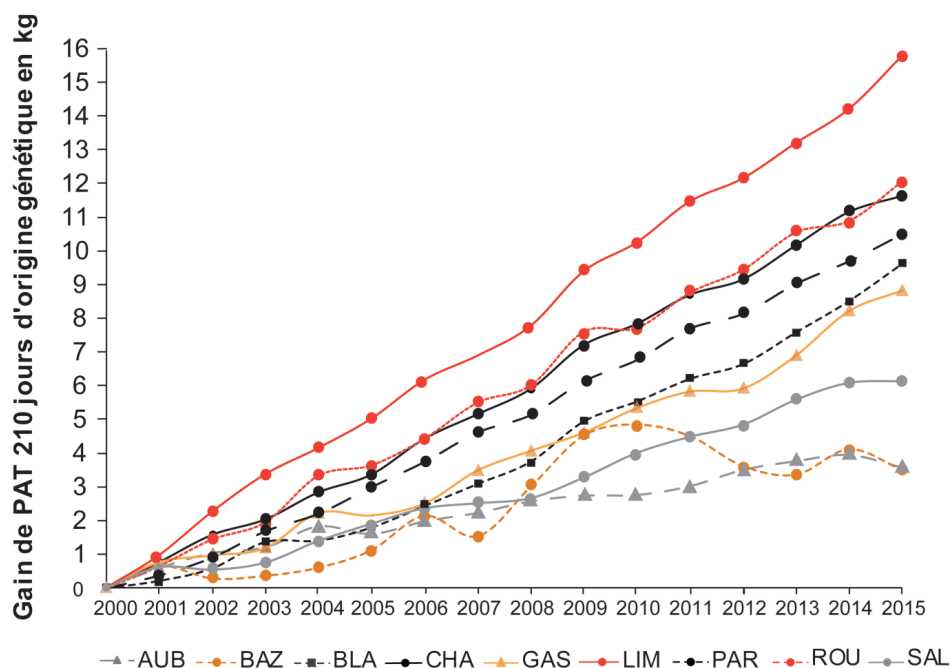


Figure 8 : Progrès génétique estimé pour le poids-âge-type au sevrage entre 2000 et 2015 dans 9 races. Moyenne par année de naissance des valeurs génétiques directes- index croissance au sevrage - référence IBOVAL 2017_01. Source : (Griffon et al., 2017). **AUB** = Aubrac, **BAZ** = Bazadaise, **BLA** = Blonde d'Aquitaine, **CHA** = Charolaise, **GAS** = Gasconne, **LIM** = Limousine, **PAR** = Parthenaise, **ROU** = Rouge des Prés, **SAL** = Salers. Toutes les races sont mises artificiellement à 0 kg pour l'année 2000. Seule la pente est comparable entre les races.

On remarque une évolution positive du PAT 210 dans toutes les races, avec une amplitude variable, allant de 3,5 kg (Aubrac, Bazadaise) jusqu'à 16 kg (Limousine). Cette sélection sur la croissance au sevrage a entraîné une augmentation défavorable du poids de naissance des veaux (entre 0,2 et 1 kg) dans toutes les races (à l'exception de la race Rouge des Prés) et une dégradation des conditions de naissance pour certaines races (Griffon et al., 2017).

II.8. Conclusion

Nous avons vu dans cette partie que l'organisation législative de l'amélioration génétique des animaux d'élevage, en particulier les bovins en France a subi des changements ces dernières années suite à l'évolution des technologies et des connaissances sur la génétique bovine, dans l'objectif de répondre à des défis de compétitivité à l'échelle européenne et internationale. Nous avons vu aussi que certains caractères comme la qualité des viandes bovines sont difficiles à améliorer avec les méthodes classiques de sélection, car ils ne sont pas mesurés en routine dans les abattoirs, le phénotypage de ces caractères restant très coûteux. La génomique est une piste intéressante à explorer pour essayer de comprendre le déterminisme génétique de la qualité des viandes et de trouver des marqueurs intéressants pour la sélection des meilleurs reproducteurs. En effet, plusieurs gènes et marqueurs ont été identifiés, certains tests à ADN pour la qualité des viandes sont déjà disponibles sur le marché, mais leur efficacité n'est pas la même quelle que soit la race (Hocquette, Lehnert, Barendse, & Picard, 2007).

Partie III. La cartographie des QTL chez les animaux

III.1. Introduction

La variabilité des valeurs génétiques entre individus d'une population résulte des effets des polymorphismes génétiques. Certains gènes ont des effets plus importants que d'autres. On parle de gène majeur quand la variabilité phénotypique du caractère d'intérêt étudié est due à l'effet d'un seul gène. Dans la plupart des cas chez les animaux d'élevage, la variabilité phénotypique des caractères quantitatifs est contrôlée par plusieurs gènes à faible effet.

La cartographie consiste à localiser le locus (ou l'ensemble de loci) dont la variation allélique est associée à la variation d'un caractère quantitatif (ou performance). Ce locus (ou loci) s'appelle locus de caractère quantitatif (*Quantitative Trait Locus* ou *QTL*). Généralement, les caractères quantitatifs sont gouvernés par plusieurs gènes et par des effets de l'environnement de sorte qu'un ou plusieurs QTL peuvent affecter un même caractère (Abiola et al., 2003).

On parle de présence d'un QTL quand on observe une différence significative de performances moyennes entre individus ayant reçu différents allèles de marqueurs de leurs parents. L'idée de la cartographie est de tester si, pour un ensemble de marqueurs à une position donnée sur le génome, la distribution des performances est la même quel que soit leur génotype.

Dans cette partie, nous allons faire un point sur les principales méthodes de détection des QTL dans les populations animales, en passant par le principe et les méthodes de l'approche classique d'analyse de liaison, les méthodes de cartographie fine des QTL et un état des lieux sur les QTL et gènes associés à la qualité des viandes et au développement musculaire chez les bovins.

III.2. La cartographie des QTL par analyse de liaison

III.2.1. Principe général

Le principe général d'une détection de QTL consiste à observer dans la descendance d'un parent hétérozygote M1/M2 (M1 et M2 étant les allèles au marqueur M) s'il existe une différence de performance moyenne selon l'allèle transmis (M1 ou M2) (P. Le Roy & Elsen, 2000) (Figure 9). Cette différence peut s'expliquer par la ségrégation d'allèles (Q1 ou Q2) en un QTL génétiquement lié au marqueur M.

Sous l'hypothèse nulle (absence de QTL), la valeur de la performance moyenne est indépendante du génotype au marqueur. En revanche, la présence d'une différence significative entre les valeurs moyennes du caractère quantitatif des génotypes au marqueur montre la possibilité d'une liaison du marqueur avec un QTL.

L'hypothèse peut être testée avec une statistique de student (*test-t*) quand les comparaisons ne concernent que deux classes (ex. Protocoles backcross ou demi-frères) ou dans un modèle linéaire d'analyse de variance dans le cas de populations « outbred » (Knott, Elsen, & Haley, 1996). Le modèle peut s'écrire sous cette forme :

$$Y_{ij} = s_i + (2p_{ij} - 1)a_i + e_i$$

Avec Y_{ij} la performance du $j^{\text{ème}}$ descendant du père i , s_i est l'effet père i , p_{ij} est la probabilité que le descendant j ait reçu l'allèle 1 du père i conditionnellement à l'information aux marqueurs. a_i est l'effet de l'allèle 1 du père i au QTL supposé et e_i est la résiduelle du modèle supposée de moyenne nulle et de variance σ^2 .

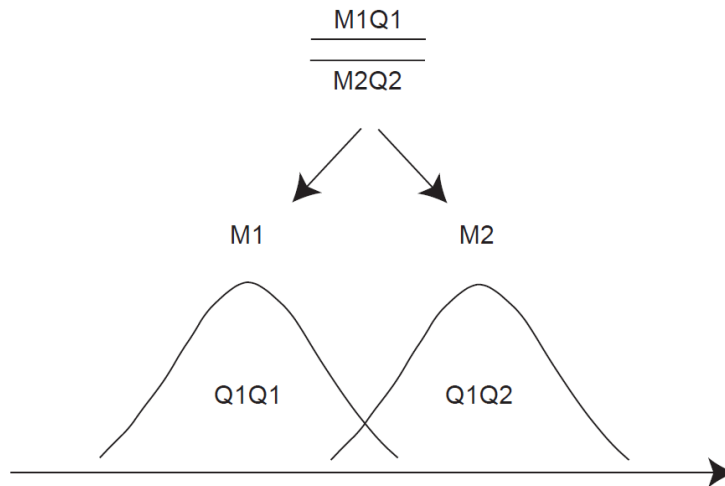


Figure 9 : Principe de la détection d'un QTL. D'après (P. Le Roy & Elsen, 2000)

III.2.2. Les dispositifs utilisés pour la détection de QTL chez les animaux

La détection de QTL implique la création d'une descendance en ségrégation afin de suivre la transmission des allèles de génération en génération. De nombreux dispositifs familiaux sont utilisés pour la cartographie des QTL chez les animaux.

III.2.2.1. Les populations « inbred »

Ce sont des populations composées d'individus consanguins, où la majorité des locus est espérée être homozygote. C'est le cas par exemple chez les végétaux et certaines lignées d'animaux de laboratoire.

Les protocoles les plus utilisés avec ce type de population consistent à croiser deux lignées « inbred » afin d'obtenir une génération F1 constitué d'individus hybrides (supposés hétérozygotes pour tous les locus). La génération F1 est croisée entre elle pour donner une descendance F2 qui sera analysée pour la détection des QTL. La génération F1 peut aussi être croisée avec l'une des lignées parentales afin d'obtenir une génération rétro-croisée (RC ou *backcross* : BC) qui sera ensuite analysée pour la détection des QTL. Dans le cas de la génération rétro-croisée (ou BC), la moitié des individus sont homozygotes et la moitié hétérozygote. Une comparaison de moyenne de performances entre ces deux sous-groupes permet de cartographier les QTL. En revanche, $\frac{1}{4}$ des individus de la population F2 sont homozygotes pour le premier allèle du marqueur, $\frac{1}{4}$ d'homozygotes au deuxième allèle du marqueur et $\frac{1}{2}$ d'individus de la population sont hétérozygotes. Une comparaison de moyenne de performances des deux groupes homozygotes permet de cartographier les QTL.

III.2.2.2. Les populations « outbred »

La cartographie des QTL dans les populations « outbred » nécessite un nombre important d'individus comparé avec les populations « inbred », pour palier le problème de seuil de significativité des tests statistiques. Une comparaison d'effectifs nécessaires pour cartographier un QTL avec une puissance de 90 % et pour un risque α de 5 % dans une population « inbred » (BC ou F2) et une population « outbred » (BC ou F2) est illustrée dans les tableaux 6 et 7 ((Soller, Brody, & Genizi, 1976; Soller & Genizi, 1978), cité par Mohamed Kileh Wais 2012).

Tableau 6 : Nombre de descendants nécessaires pour la détection d'un QTL dans un croisement entre lignées "inbred"

a*	0,125	0,25	0,5
Backcross	2672	668	168
F2	1336	334	84

* : a est l'effet additif du QTL. Pour simplifier les tableaux, on supposera que le degré de dominance au QTL est nul. Les fréquences des allèles au marqueur (M1 et M2) et au QTL (Q1 et Q2) sont égales à ½.

Tableau 7 : Nombre de descendants nécessaires pour une détection de QTL dans une population "outbred" (cas de familles de plein-frères BC ou F2)

a*	0,125		0,25		0,5	
	BC	F2	BC	F2	BC	F2
4	13200	14356	3300	3590	824	898
5	12390	13620	3098	3404	774	852
6	12134	12332	3034	3334	758	834
10	12518	13362	3130	3340	782	834
20	14540	15154	3634	3790	910	948
100	26060	26060	6514	6514	1630	1630

* : a est l'effet additif du QTL. Pour simplifier les tableaux, on supposera que le degré de dominance au QTL est nul. Les fréquences des allèles au marqueur (M1 et M2) et au QTL (Q1 et Q2) sont égales à ½.

Dans les populations « outbred » et particulièrement chez les animaux d'élevage (exemple de bovins), le nombre de descendants affiché dans le tableau 7 est très rarement atteignable. Ce qui nécessite parfois d'utiliser des croisements de type BC ou F2 entre races différentes (exemple entre Holstein et Normand chez les bovins laitiers) afin d'augmenter la puissance de détection des QTL.

Les dispositifs familiaux les plus utilisés dans ces populations pour la détection des QTL sont des dispositifs sur deux générations d'individus (dispositif « filles » ou « père ») ou des dispositifs nécessitant trois générations (dispositifs « petites-filles » ou « grand-père »).

III.2.3. Les méthodes statistiques utilisées

De nombreuses méthodes pour détecter et localiser les QTL chez les humains et les animaux sont décrites dans la littérature. Une revue bibliographique des méthodes classiques est présentée par Wais (2012). Un résumé de ces méthodes est décrit dans les paragraphes suivants.

III.2.3.1. Analyse marqueur par marqueur

L'approche marqueur par marqueur consiste à rechercher le QTL impliqué dans le caractère quantitatif étudié, autour de chaque marqueur testé. Deux approches sont proposées pour cette méthode, l'une est basée sur l'analyse de variance (ANOVA), proposée par Soller *et al.*,

(1976), et l'autre basée sur le maximum de vraisemblance (Luo & Kearsey, 1989 ; Rebai, Goffinet, & Mangin, 1995).

III.2.3.1.1. Analyse de variance

Sous l'hypothèse de liaison entre le marqueur et le QTL (taux de recombinaison nul), l'approche par analyse de variance consiste à appliquer une ANOVA à un seul facteur qui est le génotype au marqueur.

Dans le cas des descendants BC d'une population « inbred », l'analyse de variance consiste à comparer la moyenne des performances entre les individus homozygotes et les individus hétérozygotes au marqueur testé.

En revanche, chez les populations « outbred » et en particulier les animaux d'élevage, le taux de recombinaison entre le marqueur et le QTL n'est pas toujours nul, par conséquent le parent transmettant un allèle au marqueur ne transmet pas forcément l'allèle au QTL au même descendant. Cette situation nécessite donc d'analyser les données intra-famille, soit par mesure de l'effet du marqueur intra-famille père, pour les dispositifs parentaux (Soller et genizi, 1978) ou intra-famille grand-père, pour les dispositifs grand-parentaux (Weller *et al.*, 1990).

Le modèle linéaire pour un dispositif père s'écrit :

$$Y_{ijk} = \mu + s_i + a_{ij} + e_{ijk}$$

Avec Y_{ijk} la performance du $k^{\text{ème}}$ descendant du $i^{\text{ème}}$ père possédant le génotype j au marqueur. μ la moyenne générale des performances, s_i l'effet du père i , a_{ij} est l'effet du marqueur j intra-famille de père i ($j=1$ si l'allèle reçu du père est le premier et $j=2$ si l'allèle reçu est le second) et e_{ijk} la résiduelle intra-marqueur intra-famille.

La différence estimée (D), entre performances moyennes des deux groupes de descendants ayant reçu l'un ou l'autre allèle de leur père, est un effet apparent du marqueur :

$$D = 2a(1 - 2r)$$

Avec $2a$ (différence entre les génotypes homozygotes Q1/Q1 et Q2/Q2) l'effet du QTL et r le taux de recombinaison entre le marqueur et le QTL. Par cette méthode, il y a donc confusion entre la position et l'effet du QTL.

III.2.3.1.2. Par maximum de vraisemblance

L'approche par maximum de vraisemblance permet d'estimer les deux paramètres r et a simultanément. Cependant, une estimation précise du taux de recombinaison nécessite une grande population.

L'approche marqueur par marqueur ne nous donne pas la position du QTL par rapport au marqueur. La cartographie d'intervalle est une méthode alternative qui permet d'estimer séparément les effets du QTL et sa position.

III.2.3.2. Cartographie d'intervalle

III.2.3.2.1. Principe général

La disponibilité des cartes génétiques pour les animaux d'élevage a permis de rechercher les QTL en balayant tout le génome. La cartographie d'intervalle, proposée par Lander et Botstein (1989) est une approche basée sur l'analyse des données par groupes de liaison par chromosome. Elle permet d'estimer l'effet du QTL et sa position la plus probable sur le génome. Le principe de la méthode consiste à scanner l'ensemble du génome en testant l'hypothèse d'absence de QTL en chaque position x , entourée par deux marqueurs.

III.2.3.2.2. Le cas des populations « inbred » backcross

Dans le cas des populations BC, la cartographie d'intervalle considère les marqueurs deux à deux définissant des intervalles successifs sur le génome. L'idée est de tester si un QTL existe entre deux marqueurs liés, et donc déterminer sa position la plus probable, son effet et sa significativité. Cette démarche multi-marqueurs est beaucoup plus puissante et robuste que les approches uni-marqueur (Haley, Knott, & Elsen, 1994 ; P. Le Roy & Elsen, 2000; Wais, 2012). La cartographie des QTL est basée sur la méthode de vraisemblance dans ce type de populations. La vraisemblance est fonction de la position du QTL (x), de la variance (σ^2) et des moyennes des performances conditionnelles (μ et $\mu + \Delta$), avec Δ l'effet de substitution au QTL :

$$L(\mu, \sigma^2, \Delta) = \prod \left\{ \left\{ P \left(d_i^x = \frac{1}{M_i N_j} \right) \cdot f \left(\frac{y_i}{d_i^x} = 1 \right) \right\} + \left\{ P \left(d_i^x = \frac{2}{M_i N_j} \right) \cdot f \left(\frac{y_i}{d_i^x} = 2 \right) \right\} \right\}$$

$d_i^x = q$ (avec $q = 1$ ou 2 dans le cas d'un BC), c'est-à-dire que l'individu i a reçu de son père F1, en position x , un segment chromosomique de la population grand-parentale q et, aux marqueurs flanquants M_i et N_j . Les fonctions $f \left(\frac{y_i}{d_i^x} = 1 \right)$ [de moyenne μ et variance σ^2] et $f \left(\frac{y_i}{d_i^x} = 2 \right)$ [de moyenne $\mu + \Delta$ et variance σ^2] sont les fonctions de densité supposées normales de variable aléatoire y_i , dans les populations de génotype Q1Q1 et Q1Q2 au QTL respectivement.

L'hypothèse nulle (absence de QTL ou $\Delta = 0$) à la position x est testée contre l'hypothèse H1 (existence de QTL ou $\Delta \neq 0$) et le rapport de vraisemblance ($LRT = 2 \ln \left(L(\hat{\mu}, \hat{\Delta}, \hat{\sigma}^2) / L(\hat{\mu}, 0, \hat{\sigma}^2) \right)$) ou le test de Lod Score ($LOD = \log \left(L(\hat{\mu}, \hat{\Delta}, \hat{\sigma}^2) / L(\hat{\mu}, 0, \hat{\sigma}^2) \right)$) peut être calculé en toute position, afin d'obtenir un profil de vraisemblance (ex. Figure 10).

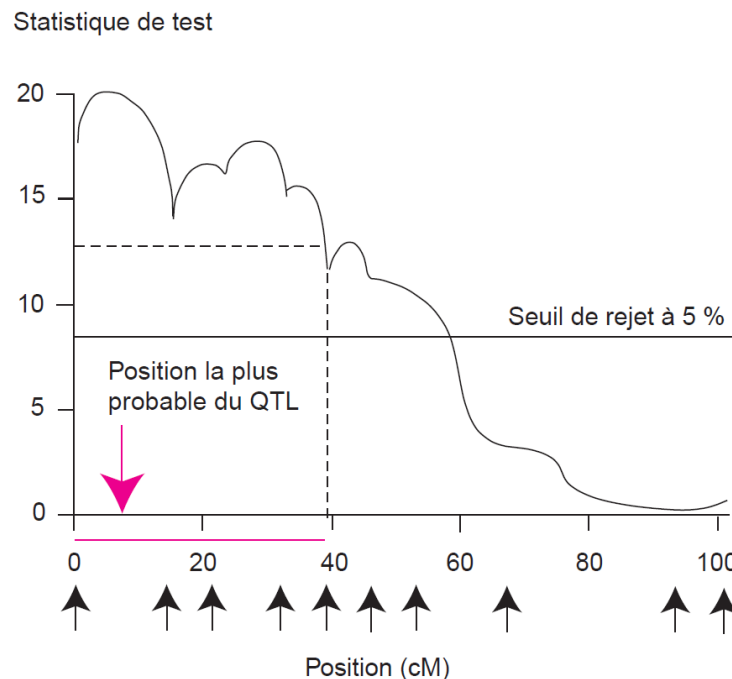


Figure 10 : Exemple de profil de vraisemblance (P. Le Roy & Elsen, 2000).

III.2.3.2.3. Le cas des populations « inbred » F2

Dans le cas des populations F2, la méthode utilisée pour la détection des QTL est dérivée des approches de vraisemblance. C'est une approche par régression linéaire multiple, proposée par Haley et Knott (1992) et Martinez et Curnow (1992). Une illustration de la méthode est décrite par Wais, 2012.

III.2.3.2.4. Le cas des populations « outbred »

La cartographie d'intervalle dans les populations « outbred » devient beaucoup plus complexe du fait de la nécessité de prendre en compte les structures familiales et la non-informativité des marqueurs dans le modèle.

Une approche basée sur le test de rapport de vraisemblance telle que décrit par Lander et Botstein (1989) est proposée par Le Roy *et al.* (1998), dans une population de plein-frères et demi-frères. Prenant un exemple d'une population de n familles de père ($i = 1, \dots, n$) avec n_i accouplements par père i ($j = 1, \dots, n_i$) et n_{ij} descendants par mère ij . Considérant les pères et les mères non apparentés et chaque mère est accouplée avec un seul père. Le modèle de vraisemblance linéarisée intra-famille de plein-frères, pour une position donnée (x) (Le Roy *et al.*, 1998) s'écrit comme suivant, en prenant la phase la plus probable du père :

$$V^x = \prod_i \prod_j \sum_{hd_{ij}} P\left(\frac{hd_{ij}}{M_i}, \hat{hs}_i\right) \prod_k f\left(\frac{yp_{ijk}}{\hat{hs}_i}, \hat{hd}_{ij}, M_i\right)$$

Avec M_i l'information marqueur pour la famille du père i , $hs_i = \{hs_{ip}, hs_{im}\}$, est la phase du père i , où hs_{ip} étant l'haplotype du segment chromosomique reçu par le père i par la voie paternelle et hs_{im} l'haplotype reçu par la voie maternelle. \hat{hs}_i est la phase la plus probable des marqueurs génétiques pour le père i , hd_{ij} la phase la plus probable des marqueurs génétiques pour la mère ij et yp_{ijk} est l'observation phénotypique du descendant ijk ($k = 1, \dots, n$) pour le caractère en question.

III.3. Etat des lieux sur les gènes et marqueurs associés à la qualité des viandes chez les bovins allaitants

Plusieurs gènes et marqueurs SNPs associés à la qualité de la viande dans différentes races bovines ont été identifiés dans la littérature (*c.f.* III.3.2). Le programme Qualvigène était un programme important de recherche français pour identifier les polymorphismes associés à la qualité des viandes et le développement musculaire dans les principales races à viande bovine françaises.

III.3.1. Le programme Qualvigène

Mis en place en 2003 par l'INRAE, l'UNCEIA, l'IDELE et les unités de sélection, le programme Qualvigène (décrit en détail par (Allais, Levéziel, Payet-Duprat, *et al.*, 2010)) visait à constituer un outil puissant d'analyse de déterminisme génétique des qualités de la viande dans les trois principales races à viande en France : Charolaise, Limousine et Blonde d'Aquitaine (Payet N., Malafosse A., Renand G., Menissier F., 2006).

Il est basé sur 3 années de contrôle sur descendance sur des animaux destinés à l'insémination artificielle et évalués sur aptitudes bouchères sur des jeunes bovins (JB) en station de contrôle.

La population Qualvigène est constituée de 3349 JB (1114 CHA, 1254 LIM et 981 BLA) abattus entre 2003 et 2006 et issus de 114 pères (48 CHA, 36 LIM et 30 BLA, respectivement). Les JB CHA sont abattus à un poids vif de 730 kg, tandis que les JB LIM et BLA sont abattus à un âge constant de 15,8 et 13,9 mois, respectivement.

La population Qualvigène est phénotypée pour des caractères liés au muscle, les aptitudes bouchères et la qualité de la viande, ainsi que les performances mesurées en routine pendant le testage sur descendance. Les mesures ont été effectuées par des techniciens de l'institut de l'élevage (IDELE).

Les caractères liés aux aptitudes bouchères ont été mesurés tout au long de la chaîne sur les animaux vivants et après abattage. Les animaux sont pointés durant les 4 semaines avant l'abattage et une note sur la morphologie a été donnée pour le développement musculaire (dmvif) et le développement squelettique (dsvif). Ensuite, le poids (poidab) et l'âge à l'abattage ont (ageaba) été enregistrés. Le poids de carcasse a été enregistré à chaud (PC) sur la chaîne d'abattage et le rendement en carcasse (rendcar) a été déduit, ensuite le poids des gras internes (poidgi) (gras de bassin et de rognon) a été pesé. Une note a été donnée pour la conformation SEUROP (confca) et le gras interne (NGI) au moment du ressuage. La longueur de la carcasse de la symphyse pubienne au milieu de la première côte (LCO), la longueur jarret-symphyse (longjs) et l'épaisseur de cuisse (epaicu) ont été enregistrées pour chaque demi-carcasse. La surface du muscle *Longissimus dorsi* (surfld) a été mesurée 24 heures après l'abattage, et le pourcentage de gras dans la 6ème côte (dap6c) a également été mesuré (le rapport du poids de dépôts adipeux sur la somme du poids du muscle *Longissimus dors*, du poids des autres muscles et des dépôts adipeux de cette côte), ainsi que des mesures de VOS (Velocity Of Sound) ont aussi été réalisées.

Pour les caractères liés au muscle, les mesures ont été réalisées sur un échantillon de viande de la 7ème côte du muscle longissimus dorsi (LD). La surface moyenne des fibres musculaires a été calculée par analyse d'images (mtafibr), ensuite le nombre moyen de fibres du muscle LD (nbfibr) a été déduit en utilisant la surface du muscle LD et la surface moyenne des fibres musculaires. La teneur en collagène total a été mesurée par la méthode décrite dans les normes NF V 04-415, basée sur la méthode de BERGMAN et LOXLEY (1963) qui utilise la concentration en acide hydroxyproline ($C_{collagène} = 7,5 * C_{hydroxyproline}$) (AFNOR, 2002). La teneur en collagène insoluble (mtxcoli) a été mesurée en triplicata par dosage après un traitement de l'échantillon libérant la fraction thermolabile du collagène soluble (BONNET & KOPP, 1986). La solubilité du collagène (solcol) a ensuite été calculée. La teneur en lipides intramusculaires (mtxlipi) a été mesurée par la méthode d'extraction à l'éther de pétrole sur des échantillons séchés (cité par (Sophie Allais, 2011a)).

Concernant les caractères de la qualité de la viande, la couleur (micoull) a été mesurée 5 fois à l'aide d'un spectrophotomètre (cité par (Sophie Allais, 2011a)). La force de cisaillement (mcocisa) (10 répétitions) et les qualités sensorielles ont été mesurées sur des steaks ayant 14 jours de maturation et grillés à 55°C. Les qualités sensorielles sont des notes de tendreté (mnotend), de jutosité (mnojuto) et de flaveur (mnoflav) de la viande qui ont été obtenues par un jury de dégustation (composée de 12 dégustateurs) spécifique de la race (et souvent identique pour les 3 années du programme), en moyennant les notes des membres du jury.

III.3.2. Gènes et marqueurs de la qualité des viandes

Le premier volet du programme Qualvigène consistait en primo-localisation des QTL par analyse de liaison à l'aide des marqueurs microsatellites disponibles. Des analyses plus fines avec des puces SNP de moyenne et haute densité ont été ainsi appliquées suite aux avancées

des technologies de génotypage. Dans le cadre de cette thèse, une analyse d'association dans les 3 races (CHA, LIM et BLA) en utilisant des données de séquence imputées génome complet a été effectuée (cf. chapitre 2).

Le deuxième volet du programme Qulavigène s'intéressait à la validation de la pertinence de certains marqueurs SNP génotypés sur des gènes candidats pour la qualité de la viande, issus des travaux de recherche français ou étrangers.

La tendreté de la viande a fait objet de plusieurs études dans l'espoir d'identifier des gènes impliqués dans ce caractère. La *calpaïne 1* (*CAPN1*), la *calpastatine* (*CAST*) et la *lysyl-oxydase* (*LOX*) sont les principaux gènes candidats étudiés.

- ***Calpaïne 1 (CAPN1)***

Le gène de la *calpaïne 1* (ou μ -*calpaïne*) code pour la grande sous-unité de l'enzyme μ -*calpaïne* qui est une protéase à cystéine liée au processus d'attendrissement post-mortem de la viande. En 2011, sur 4 SNPs génotypés au niveau du gène *CAPN1*, Allais et al., 2011 ont identifiés 2 marqueurs SNP (en positions 45 221 250 et 45 241 089 sur l'assemblage btou 4.0) ayant un effet significatif sur le caractère force de cisaillement et note de tendreté dans la race CHA (S. Allais et al., 2011a). Page et al., 2002 ont identifié sur le chromosome 29 des polymorphismes SNP associés à la variation de la tendreté dans des races américaines croisées (Piémontaise x Angus et Jersey x Limousin) (Page et al., 2002). D'autres études ont confirmé l'association des polymorphismes SNP de ce gène avec la tendreté de la viande dans d'autres races allaitantes (ex. (H. J. Lee et al., 2019b; S. H. Lee et al., 2014 ; Sun et al., 2018)). Par conséquent, la mutation causale n'est pas encore identifiée. De plus, une analyse d'expression des gènes *CAPN1* et *CAPN2* dans une population brésilienne de la race Nellore suggère que la tendreté de la viande dans cette race ne dépend pas directement de l'expression de ces deux gènes candidats, mais est associée à l'expression d'autres gènes tels que *CAST2*, *HSP90AA1*, *DNAJA1* et *HSPB1* (Malheiros et al., 2018).

- ***Calpastatine (CAST)***

La protéase *calpastatine* dont le gène est situé sur le chromosome 7 est l'inhibiteur de la *calpaïne 1* (Koohmaraie, 1996). Il est également identifié comme facteur majeur affectant l'attendrissement post-mortem de la viande (Enriquez-Valencia et al., 2017). Plusieurs études d'association ont montré l'effet de polymorphismes de ce gène sur la tendreté de la viande (ex. (H. J. Lee et al., 2019b; S. H. Lee et al., 2014; Sun et al., 2018)).

D'autres marqueurs associés aux gènes *CAPN1* et *CAST* ont été suggérés comme étant associés à la tendreté et aux autres qualités de la viande (Enriquez-Valencia et al., 2017; Koohmaraie, 1996 ; Y. X. Li et al., 2013 ; Schenkel et al., 2006; Sevane, Armstrong, Wiener, Pong Wong, & Dunner, 2014; Sun et al., 2018).

- ***Lysyl-oxydase (LOX)***

Le collagène intramusculaire peut affecter la qualité de la viande en limitant sa tendreté et sa facilité de cuisson. Une étude australienne publiée en 2006 a identifié un QTL associé à l'adhésion du muscle semi-tendineux proche du gène codant la lysyl oxydase (Drinkwater et al., 2006).

Le persillé de la viande, ou le gras intramusculaire est l'un des caractères les plus appréciés par les consommateurs. Plusieurs gènes candidats ayant un rôle dans la variabilité de la

teneur en graisse et la composition en acides gras de la viande bovine ont été identifiés dans la littérature comme la *leptine* (Shin & Chung, 2007), *thyroglobuline* (*TG*), la *synthase des acides gras* (*Fatty acide synthase : FASN*), diacylglycérol-O-acyltransférase (*DGAT*), l'*hormone de croissance 1* (*Growth hormone : GH*) (Barendse, Bunch, Harrison, & Thomas, 2006), *sterol regulatory element-binding protein 1* (*SREBF1*) (Bhuiyan et al., 2009), la *stéaroyl-CoA désaturase* (*SCD*) et la *protéine 4 de liaison aux acides gras* (*Fatty acide binding protein 4 : FABP4*) (Dujková, Ranganathan, Dufek, Macák, & Bezdíček, 2015 ; Kaplanová et al., 2013; Matsushashi et al., 2011a ; S. Zhang, Knight, Reecy, & Beitz, 2008).

- **Le gène de la leptine (LEP)**

Le gène de la *leptine*, appelé aussi gène de l'obésité est situé sur le chromosome 4 chez le bovin et est constitué de 3 exons (assemblage *ARS-UCD1.2*). Ce gène code la leptine qui est une hormone régulatrice des réserves de graisses et modulatrice de l'appétit. Ce gène, situé à proximité d'un microsatellite (*BM1500*) a été considéré comme gène candidat influençant les caractéristiques liées à la graisse de la viande chez les bovins (Anton et al., 2011 ; Daix et al., 2008). Des études ont montré l'association des polymorphismes SNP du gène de la *leptine* avec des caractères de poids, taille corporelle, carcasse et qualité de la viande chez différentes races bovines (Schenkel et al., 2005 ; Shin & Chung, 2007 ; Tian et al., 2013).

- **Le gène de la thyroglobuline (TG)**

Les hormones thyroïdiennes jouent un rôle important dans la régulation du métabolisme et peuvent affecter les processus de dépôt de graisse. Le gène codant pour la thyroglobuline (*TG*) produit le précurseur des hormones thyroïdiennes. Des marqueurs SNP significativement associés au persillé de la viande ont été identifiés sur la régions 3' de ce gène (G.-Y. Hou et al., 2011).

- **Le gène de la Fatty Acid Synthase (FASN)**

Ce gène a un rôle important dans la lipogenèse. Il code pour la synthase des acides gras, un enzyme qui régule la biosynthèse des acides gras à longue chaîne. Des études ont montré des polymorphismes et des niveaux d'expression du gène *FASN* associés au métabolisme des acides gras et à l'obésité chez les bovins (Rempel, Casas, Shackelford, & Wheeler, 2012; R. Roy, Taourit, Zaragoza, Eggen, & Rodellar, 2005). Une étude récente chinoise sur une population de 723 bovins de la race Simmental génotypés sur la puce Illumina BovineHD BeadChip a identifié une région significativement associée au groupe d'acides gras saturés (C14 :0) et mono-insaturés (C14 :1 cis-9) expliquant 10 et 6,5 % de la variance total, respectivement. Le gène *FASN* en position 51 Mb sur le chromosome 19 chevauche avec cette région, ainsi d'autres marqueurs SNP autour de ce gène sont associés avec le C14:0 (Zhu et al., 2017).

- **Les gènes Stearoyl-CoA Desaturase (SCD) et diacylglycérol-O-acyltransférase (DGAT1)**

Le gène *SCD* code pour l'enzyme $\Delta 9$ désaturase, qui introduit une double liaison dans les acides gras saturés (AS) pour les convertir en acides gras mono-insaturés (AGMI) (Figure 11) (Uto, 2016).



Figure 11: Exemple de réaction induite par SCD pour convertir le stearoyl-CoA en oleoyl-CoA. D'après Uto, 2016.

La forte activité de l'enzyme $\Delta 9$ désaturase est associée au persillé de la viande bovine (Smith, Gill, Lunt, & Brooks, 2009; Westerling & Hedrick, 1979), mais l'expression de ce gène varie d'une race à l'autre et au sein des individus de la même race (Dujková et al., 2015; Taniguchi et al., 2004).

L'enzyme codé par le gène *DGAT1* catalyse la formation des triglycérides à partir du diacylglycerole et l'Acyl-CoA. Des études ont montré des polymorphismes SNPs au niveau du gène *DGAT1* et *SCD1* associés avec le caractère persillé de la viande dans des races suédoises (X. Li, Ekerljung, Lundström, & Lundén, 2013) et chinoises (X. X. Wu et al., 2012).

- **Fatty Acid Binding Protein 4 (FABP4)**

Le gène *FABP4* est exprimé dans le tissu adipeux et code la protéine de liaison aux acides gras 4 de la famille des FABPs. Ces protéines jouent un rôle important dans l'absorption, le transport et le métabolisme des acides gras (Michal, Zhang, Gaskins, & Jiang, 2006). Des études ont montré l'association des polymorphismes SNPs du gène *FABP4* au phénotype gras intramusculaire chez la race Aberdeen Angus et Blonde d'Aquitaine (Dujková et al., 2015). D'autres ont montré l'association des polymorphismes de ce gène avec la variabilité du persillé (Michal et al., 2006) et l'épaisseur du gras dorsal chez des bovins Coréens (Cho et al., 2008).

- **Autres gènes**

Un effet significatif des polymorphismes au niveau du gène codant l'hormone de croissance bovine (*GH*) sur la composition en acides gras du muscle dans une population de 480 bovins Japonais de la race Japanese black a été révélé par (Matsushashi et al., 2011b). Le récepteur de l'hormone de croissance bovine (*GHR*) et le gène *RORC* (retinoic acid receptor-related orphan receptor C) situé sur le chromosome 3 et codant un récepteur à la vitamine A ont été également identifiés comme gènes candidats associés aux qualités de la viande chez certaines races bovine (Di Stasio, Destefanis, Brugiapaglia, Albera, & Rolando, 2005; Gorlov, Fedunin, Randelin, & Sulimova, 2014).

Les protéines découplantes (*uncoupling proteins : UCPs*) appartiennent à la famille des protéines de la membrane interne mitochondriale. Ces protéines sont considérées comme régulateur essentiel de l'homéostasie énergétique (Krauss, Zhang, & Lowell, 2005). Le gène *UCP2* est exprimé dans tous les tissus des mammifères et régule l'homéostasie énergétique à plusieurs niveaux d'expression (transcription et post-traduction) (Donadelli, Dando, Fiorini, & Palmieri, 2014). Le gène *UCP3* est exprimé principalement dans le muscle, tissu adipeux et le cœur (Azzu & Brand, 2010). Récemment, une étude chinoise a montré l'association des polymorphismes SNPs du gène *UCP2* et *UCP3* avec des caractères de mesures corporelles (hauteur au garrot et à la profondeur de la poitrine) et qualité des viandes (teneur en gras intramusculaire) chez les bovins Qinchuan (Yaning Wang, Yang, Gui, Wang, & Zan, 2016).

Enfin, des mutations au niveau du gène de la *myostatine* (*MSTN*) situé sur le chromosome 2 chez le bovin, sont responsables – à son état homozygote – de l'hypertrophie musculaire (phénotype culard) trouvé par exemple chez la race Blanc Bleu Belge. Une étude réalisée en 2010 par Sophie Allais et ses collègues (Allais, Levéziel, Payet-Duprat, *et al.*, 2010) dans le cadre du projet Qualvigène a montré que les jeunes bovins CHA et LIM portant une seule copie de l'allèle inactif de la *MSTN* ont un meilleur rendement de carcasse et score musculaire. Dans la race CHA, les animaux hétérozygotes avaient une meilleure carcasse et une meilleure qualité de viande.

III.4. Conclusion

Nous avons vu dans cette partie que la qualité de la viande est gouvernée par plusieurs gènes. Plusieurs polymorphismes associés à la qualité des viandes ont été identifiés dans la littérature. En revanche, l'effet d'un polymorphisme dans une race donnée n'a pas forcément le même effet dans une autre race. La taille de la population et la densité de marqueurs jouent aussi sur la puissance de détection des loci impliqués dans les caractères quantitatifs. Une analyse QTL sur les données Qualvigène en utilisant les données de séquence pourrait identifier de nouveaux QTL associés à la qualité de la viande, c'est l'objectif du deuxième chapitre de cette thèse.

Partie IV. Le déséquilibre de liaison dans les populations animales

IV.1. Introduction

Le concept et les premières études du déséquilibre de liaison remontent aux années 1917-1918 où les chercheurs Jennings (1917) et Robbins (1918) étudièrent indépendamment la transmission simultanée de deux loci dans une population (Jennings, 1917; Robbins, 1918). Les deux scientifiques ont montré mathématiquement que si pour chacun des loci, l'équilibre des fréquences alléliques et génotypiques d'une génération à l'autre dans une population « idéale » défini par Hardy-Weinberg est atteint en une génération, il n'en est pas de même si l'on considère deux loci simultanés. En effet, même si le concept du DL est très ancien, les études à l'échelle de la population n'ont lieu qu'à partir des années 2000 suite à l'émergence des marqueurs moléculaires répartis sur le génome.

Cette partie vise à présenter le déséquilibre de liaison, les méthodes et les statistiques utilisées pour l'estimer, les facteurs qui impactent sa variation dans les populations et un bilan des études faites à ce jour sur le DL dans les populations bovines.

IV.2. Définition et propriétés du déséquilibre de liaison

Le déséquilibre de liaison ou le déséquilibre de phase gamétique se définit comme l'association non-aléatoire (ou préférentielle) d'allèles à des locus différents dans une population (Jennings, 1917). Prenant un exemple de deux loci bialléliques (A et B), avec les allèles A/a au premier locus et B/b au second. Le déséquilibre de liaison (D) se calcul comme suit :

$$D_{AB} = p_{AB} - p_A q_B \quad (1)$$

Avec p_{AB} la fréquence de l'haplotype porteur des allèles A et B, p_A et q_B la fréquence de l'allèle A et B dans la population, respectivement. Si $D = 0$, on parle de l'équilibre de liaison (absence de déséquilibre de liaison) entre loci; autrement dit, la transmission indépendante des deux locus A et B. Si $D \neq 0$, on dit qu'il y a déséquilibre de liaison entre les deux locus considérés. Les valeurs de D varient entre $-0,25$ et $0,25$ et sa valeur minimale (D_{min}) ou maximale (D_{max}) dépend des fréquences alléliques. Si $D > 0$, sa valeur maximale est $q_B (1 - p_A)$ si $p_A > q_B$ ou $p_A (1 - q_B)$ si $p_A < q_B$. En revanche, si $D < 0$, $|D_{max}|$ est la plus petite des valeurs de $p_A q_B$ et $p_a q_b$.

IV.3. Mesures du déséquilibre de liaison

D'après la formule précédente (1), on constate que la mesure de DL entre deux loci (D) est dépendante des fréquences alléliques et haplotypiques. Pour pouvoir comparer ces valeurs pour différentes paires de loci, la mesure idéale de DL doit être normalisée et indépendante de ces fréquences (alléliques et haplotypiques). Pour pallier à ce problème, plusieurs statistiques qui permettent d'estimer le DL dans une population ont été proposées (Revue Cierco-Ayrolles et al. (2004)). Les plus utilisées dans les populations animales sont le r^2 et le D' . Le coefficient de corrélation (r^2) proposé par (Hill & Roberston, 1968) comme mesure de DL entre deux locus bi-alléliques est défini comme suit :

$$r^2 = \frac{D^2}{p_A p_a q_B q_b} \quad (2)$$

Cette mesure est préférée dans les populations animales, car elle est moins sensible à la taille de la population que le D' (Bohmanova, Sargolzaei, & Schenkel, 2010 ; H. Zhao, Nettleton, &

Dekkers, 2007). Elle est comprise entre 0 (équilibre de liaison) et 1 (déséquilibre de liaison complet). Cette mesure est généralisée pour les locus multi-alléliques sous cette forme :

$$r^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{D^2_{ij}}{p_i q_j}$$

Avec I le nombre d'allèles pour le locus A et J le nombre d'allèles pour le locus B.

Le coefficient D' proposé par Lewontin (Lewontin, 1964), appelé aussi déséquilibre normalisé est défini pour les loci bialléliques comme suivant :

$$D' = \frac{D}{D_{max}} \quad \text{si } D > 0 \text{ et } D' = \frac{D}{D_{min}} \text{ si } D < 0.$$

Ce calcul est généralisé en 1986 pour les loci multi-alléliques par Hedrick & Thomson (1986) sous cette formule :

$$D' = \sum_{i=1}^I \sum_{j=1}^J p_i q_j |D'_{ij}|$$

Avec $D'_{ij} = \frac{D_{ij}}{D_{max}}$ où $D_{max} = \min[p_i q_j, (1 - p_i) * (1 - q_j)]$ si $D_{ij} < 0$ et

$D_{max} = \min[p_i * (1 - q_j), (1 - p_i) * q_j]$ si $D_{ij} > 0$.

Les valeurs de D' varient entre -1 et 1. Le DL calculé par cette mesure peut être surestimé dans le cas de petites populations ou en cas de faibles fréquences alléliques (Bohmanova et al., 2010 ; H. Zhao et al., 2007).

IV.3.1. Le calcul de DL en pratique

En pratique, on utilise des logiciels informatiques pour estimer le DL sur des grands jeux de données. Des logiciels standards de génétique des populations, tel que Haploview (Barrett, Fry, Maller, & Daly, 2005), GOLD (Abecasis & Cookson, 2000) et Arlequin (Excoffier, Laval, & Schneider, 2005), ainsi que plusieurs packages R permettent d'estimer les statistiques de DL. PLINK est une boîte à outils très utilisée pour l'analyse des données génétiques et compte parmi les outils les plus efficaces en matière de calcul pour l'estimation du DL (Chang et al., 2015). VCFtools est un autre logiciel largement utilisé pour la manipulation et l'analyse des données génétiques qui fournit des utilitaires pour estimer le DL à partir des données génotypiques sous format Variant Call Format (VCF) (Danecek et al., 2011).

IV.4. Origine et facteurs impactant le DL

Plusieurs facteurs sont à l'origine du DL. Certains de ces facteurs impactent son évolution dans les populations animales. Les paragraphes suivants présentent les facteurs liés à l'histoire de la population et les facteurs liés à la qualité des données utilisés pour estimer le DL.

IV.4.1. Facteurs liés à l'histoire de la population

IV.4.1.1. La dérive génétique

La dérive génétique désigne les changements de fréquence allélique qui se produisent par l'échantillonnage aléatoire de gamètes des parents à la progéniture dans une population. Dans les populations de petite taille efficace (N_e), comme par exemple chez les bovins laitiers

français où la taille efficace de la population est estimée à environ une quarantaine d'individus (< 50) (Didier Boichard et al., 1996) la dérive génétique est forte. Chez les bovins allaitants, la taille efficace est plus importante. Elle est estimée à 247 pour la Blonde d'Aquitaine, 601 pour le Charolais et plus de 1000 pour la Limousine (Bouquet, Renand, & Phocas, 2009). En général, la dérive dans les petites populations stables (non en croissance) a tendance à augmenter le DL, car les haplotypes peuvent être perdus dans la population (Ardlie, Kruglyak, & Seielstad, 2002). Cette réduction du nombre d'haplotypes dans la population freine le déclin du DL au cours des générations.

IV.4.1.2. Les mutations

Le déséquilibre de liaison est généralement créé dans une population suite à une mutation qui se transmet de génération en génération sous forme de blocs avec les autres loci voisins (Figure 12). Si la mutation est létale ou perdue par la dérive génétique, le DL est ponctuel. En revanche, la mutation peut se transmettre de génération en génération et sa fréquence peut augmenter par la sélection ou par la dérive ; les deux forces les plus actives dans les populations animales d'élevage.

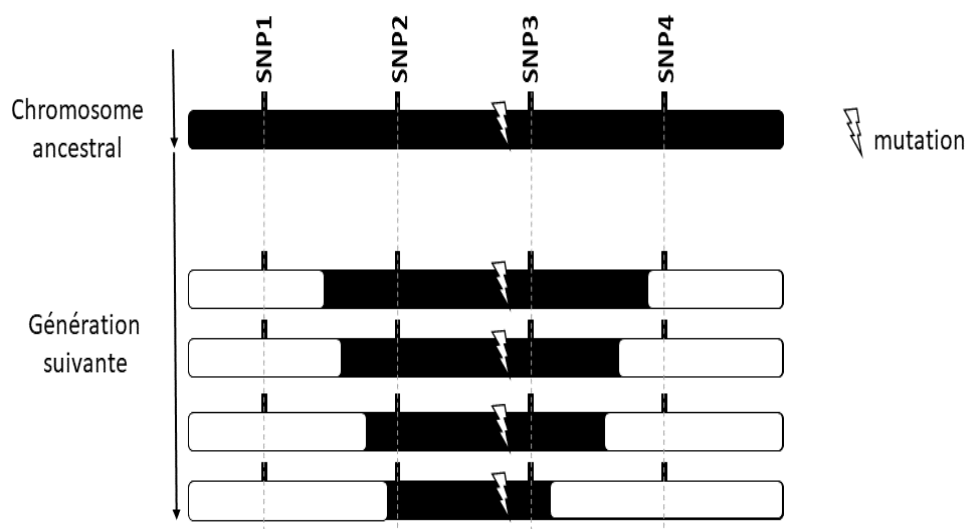


Figure 12: Exemple de mutation en DL transmise de génération en génération. La mutation est en fort DL avec les SNPs 2 et 3.

IV.4.1.3. Les goulots d'étranglement

Les goulots d'étranglement (*bottlenecks*) correspondent à une réduction sévère de la taille de la population. Une étude réalisée par Bouquet et ses collaborateurs (2009) (Bouquet et al., 2009) a montré que la population Blonde d'Aquitaine (race bovine allaitante Française) connaît un goulot d'étranglement depuis les années 1970 suite à l'utilisation intensive de l'insémination artificielle (IA) par les éleveurs français et qui a entraîné une diminution de la taille efficace de la population. Des goulots d'étranglement ont également été suggérés comme source de DL importante dans le génome humain (Claudia, Josef, Walther, & Günter, 2005).

IV.4.1.4. Le taux de recombinaison

La recombinaison (*crossing-over*) est le principal mécanisme qui va diminuer le niveau du DL dans une population au cours des générations. En effet, supposant une population idéale et panmictique. Si à la génération g le déséquilibre initial entre deux loci est D_0 et r le taux de

recombinaison entre ces locus (exprimé en %), le DL entre les locus diminue d'un facteur $(1-r)$ par génération. On peut donc estimer le DL à la génération g (D_g) par cette formule :

$$D_g = (1 - r)^g D_0$$

Les courbes de décroissance de DL en fonction du temps, du taux de recombinaison et du DL initial peuvent alors être tracées (Figure 3).

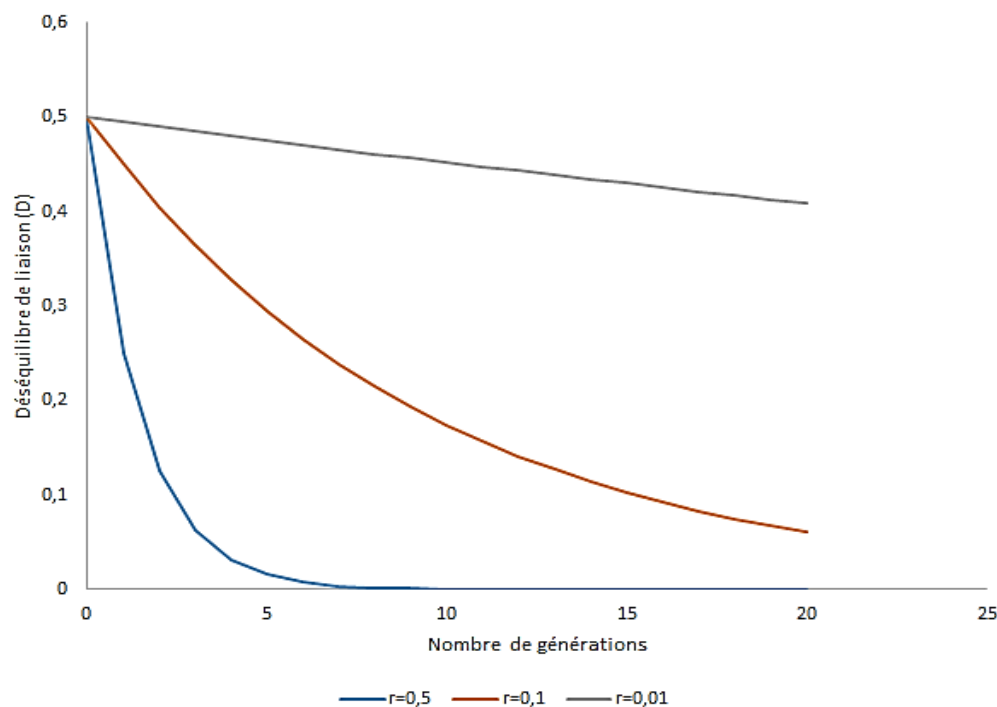


Figure 13 : Décroissance du DL en fonction du DL initial (D_0), du taux de recombinaison (r) et du temps (générations).

En observant la Figure 13, on remarque que pour les locus proches génétiquement (taux de recombinaison faible), le DL est maintenu sur plusieurs générations. Du DL résiduel ($D < 0,1$) est observé pour des loci éloignés ($r = 0,5$) si le DL initial est fort. Par exemple, pour une distance de 10 centiMorgan (cM), avec un déséquilibre initial de 0,5, le DL à la génération 14 (D_{14}) vaut 0,12, tandis que pour les locus indépendants ($r=0,5$) le DL s'approche de la valeur nulle après 3 à 4 générations.

IV.4.1.5. Le mélange des populations

Le DL peut être la conséquence d'un mélange (*admixture*) de populations génétiquement distinctes ayant des fréquences haplotypiques différentes (Nei & Li, 1973) (Figure 14). Un exemple de la race Blonde d'Aquitaine qui est formée par fusion de trois populations de Blonde du Sud-Ouest français (Quercy, Garonnais et Blonde des Pyrénées) dans les années 1960 (Beghain et al., 2013). Il est important de noter que des phénomènes de « stratification » peuvent exister dans les populations ce qui induit des faux positifs dans les analyses d'associations.

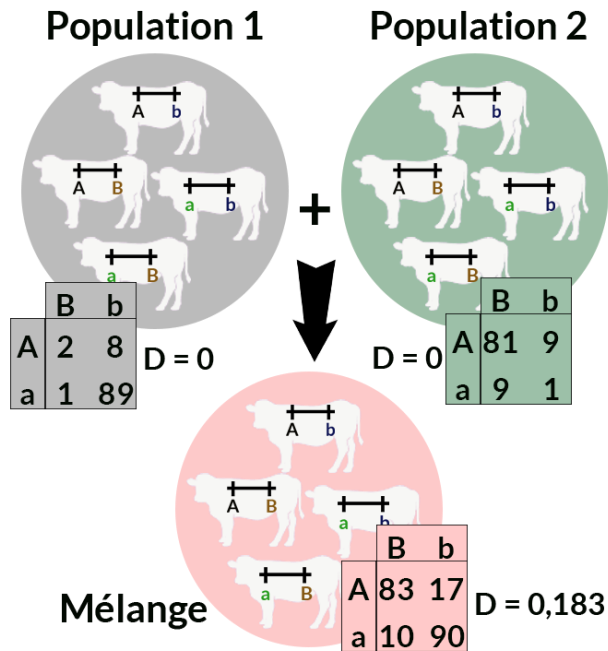


Figure 14: Exemple de DL créé par fusion de deux populations en équilibre de liaison.

IV.4.1.6. La sélection

La sélection peut être à l'origine du déséquilibre de liaison si elle favorise un haplotype sur l'autre dans une population. Le choix de quelques individus pour la reproduction peut induire un effet similaire à la dérive génétique, c'est-à-dire éliminer certains allèles ou haplotypes dans la population et par conséquent créer du DL. Ainsi, la sélection maintient le DL sur plusieurs générations. La sélection peut aussi affecter le DL par effet d'auto-stop génétique, par lequel un haplotype entier qui flanque une région positivement sélectionnée peut rapidement être balayé à haute fréquence ou même être fixé dans la population (Ardlie et al., 2002). Ainsi, la sélection contre des variants délétères peut augmenter le DL par élimination des haplotypes, mais l'effet est généralement plus faible. Une autre façon dont la sélection peut affecter le DL dans les populations est la sélection épistasique pour des combinaisons d'allèles à des loci éloignés sur le même chromosome.

IV.4.1.7. La conversion génique

Lors d'une conversion génique, une copie d'un fragment chromosomique parental est transférée à l'autre chromosome pendant la méiose (Figure 15). C'est un effet équivalent à deux événements de recombinaison très rapprochés, et peut décomposer le DL d'une manière similaire à la recombinaison ou à la mutation récurrente (Ardlie et al., 2002).



Figure 15: Exemple de conversion génique. Transfert unidirectionnel du fragment noir vers l'autre chromosome

IV.4.2. Facteurs liés à la qualité des données

Au-delà des facteurs liés à l'histoire des populations, d'autres facteurs liés à la méthodologie et à la qualité des données peuvent impacter les estimations du DL à l'échelle de la population. Selon la statistique utilisée pour estimer le DL, les conclusions peuvent être différentes. Par exemple, le calcul du DL à l'aide du coefficient de corrélation r^2 est moins sensible aux variations de fréquences alléliques comparé au D' , qui tend à surestimer le DL dans les populations de petites tailles et de faible MAF (*Minor Allele Frequency*) (Ardlie et al., 2002; Bohmanova et al., 2010 ; Lu et al., 2012; H. Zhao et al., 2007). La densité de marqueurs et la taille de la population jouent aussi sur l'estimation du DL. Une faible densité de marqueurs ou une répartition non-homogène sur le génome peut induire des biais dans l'estimation du DL.

IV.5. Les analyses de DL chez les bovins : état des lieux

Le développement des technologies de génotypage à haut débit (Craig & Stephan, 2005; Gunderson et al., 2005 ; Matukumalli et al., 2009) et la disponibilité de panel de polymorphismes SNP à haute densité ont permis de réaliser des études de DL approfondies sur le génome de la plupart des races bovines. Les premiers travaux d'analyse de DL à l'échelle des populations animales ont été initiés en 2000 avec l'étude de Farnier et ses collaborateurs (Farnier et al., 2000). L'équipe a analysé le DL dans une population de bovins laitiers (race Black-and-White néerlandaise) avec un panel de 284 microsatellites en utilisant la statistique D' . La conclusion tirée de cette étude est que le DL décroît rapidement avec la distance physique entre locus, mais aussi du DL non nul peut être observé entre loci très éloignés (plusieurs dizaines de centimorgans) voire même entre loci non synténiques (situés sur des chromosomes différents). D'autres études ont abouti à la même conclusion comme par exemple l'étude de Vallejo et ses collaborateurs portée sur 23 taureaux d'élite de la race Holstein nord-américaine, faiblement apparentés et génotypé pour 54 microsatellites (R. L. Vallejo, Li, Rogers, & Ashwell, 2003) ou l'étude de Khatkar et ses collaborateurs sur une population de 433 taureaux australiens de la race Holstein–Friesian génotypés pour 220 SNPs sur le chromosome 6 (Khatkar et al., 2006).

Le DL calculé par la statistique D' peut être surestimé dans les populations de petite taille et/ou pour des marqueurs SNPs avec de faibles MAF (Bohmanova et al., 2010; H. Zhao et al., 2007) c'est pour cela, la statistique r^2 est préférée dans les études avec marqueurs SNPs bialléliques. En effet, quel que soit le nombre ou le type de marqueurs utilisés, la relation entre les valeurs de DL et la distance entre loci est inversement proportionnelle, c'est-à-dire que, plus la distance entre marqueurs est longue, plus les valeurs de DL sont petites.

La plupart des analyses du DL avaient focalisé sur une fenêtre de distance de 500 kb entre marqueurs. Par exemple McKay et ses collaborateurs, dans une étude sur 8 races bovines (Angus, Brahman, Charolaise, Dutch Black et White Dairy, Holstein, Japanese Black, Limousine et Nelore) en utilisant une faible densité de marqueurs (2670 SNP) ont trouvé qu'au-delà de 500 kb, le DL moyen est proche de zéro (McKay et al., 2007). Une autre étude portant sur 1546 taureaux Holstein-Friesian génotypés pour 15 036 SNP a montré que les valeurs moyennes de r^2 sont inférieures à 0,1 pour les marqueurs ayant des distances comprises entre 200 et 500 kb (Khatkar et al., 2008). Cependant, Mokry et ses collègues (Mokry et al., 2014) ont montré dans une étude portant sur 395 bovins allaitants génotypés sur la puce HD (BovineHD BeadChip) que la valeur moyenne de r^2 variait de 0,05 à 0,07 à des distances comprises entre 400 et 500 kb, mais que des valeurs de DL pouvait persister jusqu'à 0,66 à ces distances. Beghain et ses collaborateurs ont analysé le DL dans 30 bovins non-apparentés de la race française Blonde d'Aquitaine génotypés sur des puces 54 k (Beghain et al., 2013). Les résultats sont concordants avec les études précédentes en termes d'entendu de DL et sa

diminution avec la distance physique, ainsi du DL usuel (r^2 supérieur à 0,2) est observé entre marqueurs ayant des distances allant jusqu'à 724 kb. Ce DL à longue distance (appelé dans ce manuscrit *LRLD* : pour *Long Range Linkage Disequilibrium*) fait l'objet du chapitre 1. En effet, des études récentes ont montré l'existence d'un DL à longue distance entre des paires de variants distants au sein du génome humain. Koch et ses collaborateurs (Koch, Ristroph, & Kirkpatrick, 2013), en analysant 60 individus Yoruba du Nigeria obtenus à partir de données HapMap, ont montré la présence de LRLD entre des paires de SNP à des distances > 250 kb parmi les 22 autosomes humains. Park (2019) a montré des interactions spécifiques de LRLD dans trois populations humaines différentes d'Afrique, d'Europe et d'Asie de l'Est du projet 1000 génomes (Park, 2019). Dans une autre étude sur la truite arc-en-ciel, le LRLD a été observé pour des distances > 1 Mb, mais avec un panel de SNP de faible densité (32k SNP), et la population présente une structure sub-familiale, ce qui peut expliquer ce LRLD (Roger L Vallejo et al., 2018).

IV.6. Conclusion

Plusieurs facteurs peuvent influencer l'étude du DL à l'échelle de la population animale, ce qui implique la nécessité de faire des contrôles de qualité sur les échantillons à analyser. L'échantillon doit être composé d'individus les moins apparentés pour représenter au mieux la variabilité haplotypique dans la population, ce qui n'est souvent pas le cas dans les populations d'élevage vu leur taille efficace réduite et leur forme en structures familiales. Il est aussi nécessaire d'utiliser une densité de marqueurs assez élevée pour couvrir la totalité du génome.

L'analyse du DL génome complet à l'échelle des populations est quasiment fait dans toutes les races bovines avec des densités de marqueurs différentes, allant de quelques microsatellites à plusieurs centaines de milliers de SNPs (Puces HD : 700 k). En revanche, l'étude de DL à longue distance n'a pas été abordée jusqu'à aujourd'hui dans les populations bovines. Il est donc important de vérifier l'existence de ce LRLD chez les bovins allaitants français. C'est l'objectif du premier chapitre de cette thèse.

Chapitre 1 : Le LRLD chez les bovins allaitants français

Chapitre 1: Le LRLD chez les bovins allaitants français

V.1. Introduction

Le déséquilibre de liaison est un outil essentiel pour les généticiens. Les informations sur son étendue et ses caractéristiques peuvent fournir des indications importantes afin de concevoir des stratégies visant à identifier le déterminisme génétique des caractères complexes ou pour développer des méthodes de sélection. Il permet de déterminer le nombre de marqueurs requis pour une analyse d'association efficace ou une sélection assistée par marqueurs. L'analyse du DL permet aussi d'identifier les éléments d'épistasie entre régions du génome très éloignées sur le même chromosome, ainsi que la détection de réarrangements chromosomiques. Le DL fournit également des informations sur l'histoire des populations. C'est une alternative utile pour estimer la taille efficace de la population lorsque les informations généalogiques ne sont pas disponibles. En outre, les informations sur le DL peuvent être utilisées pour détecter une sélection positive récente.

De nombreuses études de DL génome entier basées sur des données SNP de densité moyenne ou haute densité ont été réalisées dans la plupart des populations bovines laitières et allaitantes (*cf.* IV .5). Certaines de ces études ont montré l'existence de déséquilibre de liaison à longue distance (*Long-range linkage disequilibrium : LRLD*) entre variants dans certaines races bovines (ex. Beghain et al., 2013; Farnir et al., 2000), cependant, à ce jour, aucune étude détaillée sur le LRLD sur le génome bovin n'a été faite alors que sur le génome humain, le LRLD a été étudié récemment (Koch et al., 2013; Park, 2019).

Dans ce chapitre, nous présenterons la première étude complète sur le LRLD chez les bovins allaitants français, en analysant les données de 672 Charolais (CHA), 462 Limousin (LIM) et 326 Blonde d'Aquitaine (BLA) génotypés sur la puce Illumina BovineHD Beadchip. De plus, le déséquilibre de liaison génome entier et la structure des blocs haplotypiques ont été analysés dans ces trois races à viande.

RESEARCH ARTICLE

Open Access

Long-range linkage disequilibrium in French beef cattle breeds



Abdelmajid El Hou¹, Dominique Rocha², Eric Venot², Véronique Blanquet^{1†} and Romain Philippe^{1*†}

Abstract

Background: Linkage disequilibrium (LD) is a key parameter to study the history of populations and to identify and fine map quantitative trait loci (QTL) and it has been studied for many years in animal populations. The advent of new genotyping technologies has allowed whole-genome LD studies in most cattle populations. However, to date, long-range LD (LRLD) between distant variants on the genome has not been investigated in detail in cattle. Here, we present the first comprehensive study of LRLD in French beef cattle by analysing data on 672 Charolais (CHA), 462 Limousine (LIM) and 326 Blonde d'Aquitaine (BLA) individuals that were genotyped on the Illumina BovineHD Beadchip. Furthermore, whole-genome LD and haplotype block structure were analysed in these three breeds.

Results: We computed linkage disequilibrium (r^2) values for 5.9, 5.6 and 6.0 billion pairs of SNPs on the 29 autosomes of CHA, LIM and BLA, respectively. Mean r^2 values drop to less than 0.1 for distances between SNPs greater than 120 kb. However, for the first time, we detected the existence of LRLD in the three main French beef breeds. In total, 598, 266, and 795 LRLD events ($r^2 \geq 0.6$) were detected in CHA, LIM and BLA, respectively. Each breed had predominantly population-specific LRLD interactions, although shared LRLD events occurred in a number of regions (55 LRLD events were shared between two breeds and nine between the three breeds). Examples of possible functional gene interactions and QTL co-location were observed with some of these LRLD events, which suggests epistatic selection.

Conclusions: We identified long-range linkage disequilibrium for the first time in French beef cattle populations. Epistatic selection may be the main source of the observed LRLD events, but other forces may also be involved. LRLD information should be accounted for in genome-wide association studies.

Background

Linkage disequilibrium (LD), or non-random association of alleles between loci, is important for identifying and fine mapping quantitative trait loci (QTL) [1]. The number of markers required for a successful association analysis and marker-assisted selection depends on the extent of LD across the genome. LD information also provides information on the history of populations. It is a useful alternative for estimating effective population size (N_e) when pedigree information is not available [2,

3]. Furthermore, LD information is used to detect recent positive selection (e.g. in humans [4] and *Drosophila* [5]). In addition, information about the extent and patterns of LD can provide important insights for the design of strategies to identify the genetic basis of complex phenotypes or to develop genomic selection methods [6].

The development of high-throughput genotyping technologies [7–9] and the availability of high-density (HD) single nucleotide polymorphism (SNP) panels have made it possible to carry out detailed studies on LD across the whole genome in cattle (e.g. [10, 11]). Different measures of LD have been published in the literature (e.g. [12]). The most frequently used measures in animal populations are the square correlation coefficient (r^2) [13] and the normalized D' [14]. r^2

*Correspondence: romain.philippe@inrae.fr

[†]Véronique Blanquet and Romain Philippe share last authorship

¹ INRAE, PEIRENE EA7500, USC1061 GAMAA, Université de Limoges, 87060 Limoges, France

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



ranges from 0 (no-disequilibrium) to 1 (complete disequilibrium), and D' from -1 to 1 . r^2 is the preferred measure of LD in animal populations, because it is less sensitive to population size than D' , and D' tends to be inflated with small sample sizes and/or low allele frequencies [15, 16].

LD values decrease as the distance between markers on the genome increases. Most of the bovine studies using SNP data have shown that the average LD was close to zero for distances between markers greater than 500 kb. For example, using a low density of markers (2670 SNPs) but a very large number of animals from eight breeds, McKay et al. [17] found that for a distance between SNPs greater than 500 kb, the average LD was close to zero. Another study on 1546 Holstein–Friesian bulls that were genotyped for 15,036 SNPs showed that the mean r^2 values fall below 0.1 for distances between 200 and 500 kb [10]. However, in a study on 395 beef cattle animals genotyped on the BovineHD BeadChip, Mokry et al. [18] found that the average r^2 ranged from 0.05 to 0.07 at distances between 400 and 500 kb, but that, at these distances, the LD phase could persist to 0.66.

Recent studies have shown the existence of long-range LD (LRLD) between pairs of distant variants within the human genome. By analysing 60 Yoruba individuals from Nigeria obtained from the HapMap data, Koch et al. [19] showed the presence of LRLD between pairs of SNPs at distances greater than 250 kb among the 22 human autosomes. Park [20] showed the occurrence of specific LRLD interactions in African, European and East-Asian human populations from the 1000 Genomes Project. Another study on rainbow trout using a low-density SNP panel (31,788 SNPs) revealed LRLD for distances greater than 1 Mb, but the sub-family structure that exists in the population analysed explains this LRLD [21].

Several factors may be at the origin of LRLD events, such as population admixture [22], genetic drift or epistatic selection [23], recurrent bottlenecks [24] or chromosome structural variations (e.g. [25]). Genome assembly errors can also be the source of false observed LRLD events [19].

Numerous whole-genome LD studies based on medium-density (MD) or high-density (HD) SNP data have been performed in most dairy and beef cattle populations, and several have already shown the existence of LRLD in some bovine breeds (e.g. Beghain et al. [11]); however, to date, LRLD has not been investigated in detail in cattle. In these previous studies, the distribution of LRLD events along the cattle genome and the potential functional interactions between regions on LRLD have not been analyzed. In our study, we studied for the first time the extent of LRLD in three French beef cattle breeds and we explored the hypothesis that epistatic

selection could explain the LRLD events by searching for functional interactions between genes in LRLD.

Methods

Animals and genotyping

For this study, since we did not perform any experiments on animals, no ethical approval was required. We used SNP genotyping data that were obtained from the GEM-BAL (multi-breed genomics of beef and dairy cattle) research project [26–28] and a large population of 672 Charolais (CHA), 462 Limousine (LIM) and 326 Blonde d'Aquitaine (BLA) animals that were genotyped with the BovineHD BeadChip (Illumina Inc., San Diego, CA).

SNP quality control

SNPs located on sex chromosomes or without an assigned position in the ARS-UCD1.2 genome assembly (GenBank assembly Accession GCA 002263795.2) were discarded, as well as SNPs and animals with a low call rate ($<2\%$). SNP quality control (QC) was carried out based on a minor allele frequency (MAF) lower than 0.05 and on Hardy–Weinberg equilibrium test ($<10e-6$). Individuals that deviated by more than ± 3 standard deviations from the mean of the heterozygosity rate were removed from the analysis, as well as individuals with cryptic relatedness (π -hat threshold >0.125 : third degree relatives), based on a subset of pruned SNPs, using the *PLINK* v1.9 software [29], as recommended by Marees et al. [30].

Principal component analysis (PCA) was performed to determine the family structure of each population because it can have a strong impact on the LD pattern. PCA on SNP genotypes was performed using the *snpGdsPCA* function of the *SNPRelate* R package [31] based on pruned SNPs which were in approximate LD. The genotype matrix (individuals*SNPs) was used to calculate a correlation matrix by individuals, and then the eigenvector of this matrix was calculated. These eigenvectors were used to describe the population structure. The PCA analysis was followed by a cluster analysis on a matrix of genome-wide identity by state (IBS) pairwise distances using the *snpGdsHCluster* and *snpGdsIBS* functions of the *SNPRelate* package. Groups were determined by a permutation score using the same R package (*SNPRelate*).

Linkage disequilibrium analysis

For each pair of SNPs, we calculated the square correlation coefficient (r^2) as a measure of LD using the *PLINK* v1.9 software [29] and for all syntenic pairs of SNPs on each autosome, we also calculated the r^2 between two loci [13]. Background LD measured as the r^2 between non-syntenic SNPs was estimated among a subset of non-syntenic pairs of SNPs, which were selected using

the *-indep-pairwise* option of the *PLINK* software with the default parameters: 50 SNPs per window, a shift of five SNPs between windows at each step, and a pairwise r^2 threshold of 0.2.

Haplotype block analysis

Haplotype block patterns were estimated using the *-blocks* option based on the *PLINK* v1.9 software [29]. The same QC filters were applied to the data of the three breeds. *PLINK* v1.9 uses the haplotype block definition suggested by Gabriel et al. [32], by default, with blocks of a maximum size of 5000 kb.

Identification of long-range linkage disequilibrium

We defined long-range LD (henceforth LRLD event) between two haplotype blocks as follows: (1) at least two markers per haplotype block with an average r^2 higher than 0.4, 0.6 or 0.8; and (2) a haplotype block distance longer than or equal to 1 Mb on the same chromosome. We used the *Circos* software v0.69-6 [33] to visualize whole-genome LRLD events. By analysing the variance (one-way ANOVA), we checked whether the number of LRLD events was on average statistically different in the three breeds, and we used Spearman's test to check the correlation between the number of LRLD events and chromosome size.

LRLD and functional interactions

To check for the existence of functional interactions in the identified LRLD paired blocks, all the genes with in each block were retrieved from the Ensembl database (release 101) [34]. The Search Tool for the Retrieval of Interacting Genes (STRING) database [35] was queried for possible functional interactions between proteins encoded by genes within LRLD events. In addition, we used the cattle quantitative trait locus (QTL) database (Cattle QTLdb) [36] to investigate the presence of QTL associated with the same phenotype in both blocks, for each LRLD event. All 425 phenotypes of the Cattle QTLdb, including the 266 phenotypes related to meat/carcass and production traits were used. A common LRLD event between breeds was defined as the strict intersection of these LRLD events in each pair of breeds or in all three breeds. We checked if the proportion of

LRLD paired blocks showing functional interactions (STRING interaction and QTL) was not simply due to sampling. Then, we used the χ^2 test between the proportion of LRLD events with STRING interactions or QTL and the same number of randomly chosen pairs of blocks.

Results

Quality control for LD analysis

To evaluate the extent of LD in the three main French beef cattle breeds, 672 CHA, 462 LIM and 326 BLA animals were genotyped using the BovineHD BeadChip. However, to eliminate bias in LD related to family structure, only the least related (third-degree relatives) animals were selected based on the π -hat values. In total, 145 CHA, 106 LIM and 49 BLA animals passed QC, yielding 559,260, 541,319, and 563,740 SNPs for the CHA, LIM, and BLA animals, respectively.

These SNPs cover a total length of 2.48 Gb on the genome (Table 1) and (see Additional file 1: Tables S1–S3). SNPs were generally homogeneously distributed along the 29 autosomes, with fluctuations in some chromosomal regions (see Additional file 2: Figures S1–S3). Adjacent SNPs were separated by an average distance of 4.4 ± 7.0 kb, 4.6 ± 7.3 kb, 4.4 ± 6.9 kb for CHA, LIM and BLA, respectively and a median spacing of ~ 2.6 kb. All adjacent SNPs had an inter-distance shorter than 1 Mb, except *Bos taurus* (BTA) chromosomes 8 and 10 on which some adjacent SNPs had an inter-distance longer than 1 Mb.

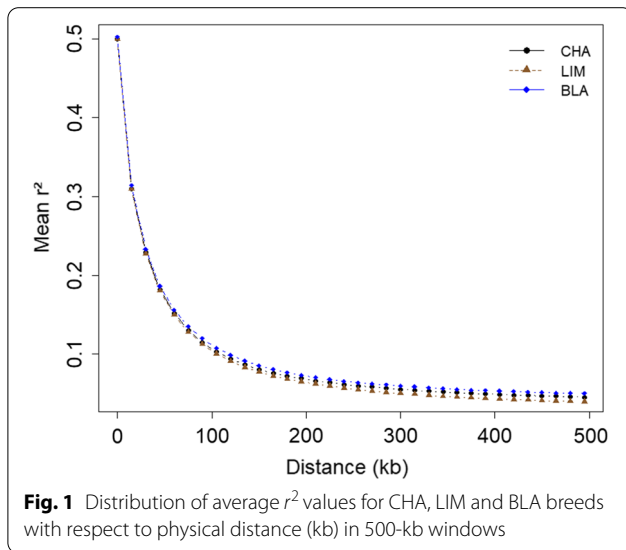
Principal component analysis in the three breeds showed that each population was homogeneous. These results were confirmed by a clustering method using a matrix of genome-wide identity by state (IBS) pairwise distances (see Additional file 3: Figures S4 and S5).

Genome-wide linkage disequilibrium analysis

We analysed the LD decay for SNPs within 500-kb windows. The mean values of r^2 , pooled over autosomes for an inter-SNP distance of 15 kb are summarized in Additional file 4: Tables S4–S6. The distribution of the r^2 values according to the physical distance between loci is shown in Fig. 1. As expected, there was an inversely proportional relationship between the mean r^2 and

Table 1 Distribution of SNPs along the autosomes of the Charolaise, Limousine and Blonde d'Aquitaine breeds

Breed	Number of SNPs	Size (Gb)	Average inter-distance (\pm sd) in kb	Median	Max (kb)
Charolaise	559,260	2.48	4.4 ± 6.9	2.6	1636.2
Limousine	541,319	2.48	4.6 ± 7.1	2.7	1636.7
Blonde d'Aquitaine	563,740	2.48	4.4 ± 6.7	2.6	1636.2



the physical distance between SNPs in the three breeds (Fig. 1 and see Additional file 4: Tables S4–S6). The average r^2 values dropped below 0.1 at a distance greater than 120 kb. Background LD was estimated on a set of 50,083, 51,301, and 46,925 non-syntenic SNPs for CHA, LIM and BLA, respectively, and resulted in values of 0.009 for CHA, 0.010 for LIM and 0.024 for BLA.

The average values of LD (r^2) varied between 0.5 at distances smaller than 15 kb, to less than 0.1 at distances greater than 120 kb (Fig. 1 and see Additional file 4: Tables S4–S6). The mean r^2 (\pm SD) values between pairs of SNPs ranged from 0.079 (\pm 0.154) to 0.121 (\pm 0.203) for CHA, LIM and BLA (see Additional file 5: Tables S7–S9). In contrast to the LIM and BLA breeds, BTA14 of the CHA breed showed a lower LD decay with physical distance than the other chromosomes (see Additional file 6: Figures S6–S8).

Haplotype block structure

We used the method defined by Gabriel et al. [32] based on genotyping data to identify the haplotype blocks on each autosome; the haplotype blocks that included only two SNPs (16,624 for CHA, 16,131 for LIM and 14,763

for BLA) were discarded to avoid formation of spurious blocks. In total, 460,224, 433,950 and 422,469 SNPs were clustered into haplotype blocks, which represent 82.29, 80.17, and 74.94% of all the SNPs for CHA (see Additional file 7: Table S10), LIM (see Additional file 7: Tables S11) and BLA (see Additional file 7: Tables S12), respectively. These haplotype blocks covered 1.55, 1.49, and 1.35 Gb of the total genome size for CHA, LIM and BLA, respectively (Table 2), and the chromosome coverage ranged from 45.87% (BTA23, BLA) to 67.42% (BTA7, CHA) in the three breeds (see Additional file 7: Tables S10–S12).

In the three breeds, we observed larger haplotypes on BTA6, 7, 12 and 23 than on the other chromosomes and we found small or no haplotype blocks on BTA10, 12 and 23 because of the low density of SNPs for these chromosomes (see Additional file 8: Figure S9 and Additional file 11: Table S13) for CHA, (see Additional file 9: Figure S10 and Additional file 11: Table S14) for LIM, and (see Additional file 10: Figure S11 and Additional file 11: Table S15) for BLA. In addition, large haplotype blocks were found at the extreme ends of BTA15, 21 and 23 in the three breeds.

Interestingly, large haplotype blocks on BTA14 (size range from 800 to 1267 kb) were found for CHA, although block sizes did not exceed 500 kb in the other two breeds (Fig. 2), which suggests a selection pressure on this chromosome. The largest block on BTA14 hosts 12 annotated genes (*XKR4*, *TMEM68*, *TGS1*, *LYN*, *RP620*, *UI*, *MOS*, *PLAG1*, *CHCHD7*, *SDR16C5*, *SDR16C6*, and *PENK*), with *PLAG1* known to be associated with stature and carcass yield in cattle (Fig. 3) [37].

Long-range linkage disequilibrium

In order to investigate the existence of LRLD in the three populations, the (r^2) measures of LD were calculated for each pair-wise combination of SNPs on each autosome (syntenic markers). In total, 5.9, 5.6, and 6.0 billion pairs were analysed for all the autosomes in the CHA, LIM and BLA breeds, respectively. In this study, we called LRLD between two haplotype blocks, if at least two SNPs in each block were in high LD ($r^2 \geq 0.4, 0.6$ or 0.8)

Table 2 Descriptive summary of the haplotype block analysis in the Charolais, Limousine and Blonde d’Aquitaine breeds

Breed	Number of SNPs	Chr size (Mb)	Number of blocks	Block coverage length (Mb)	Block coverage length (%)	Number of SNPs in blocks	% SNPs in blocks
Charolaise	559,260	2480.92	52,664	1551.78	62.55	460,224	82.29
Limousine	541,319	2481.03	50,553	1488.29	59.99	433,950	80.17
Blonde d’Aquitaine	563,740	2481.00	48,303	1352.07	54.50	422,469	74.94

Total number of SNPs, total chromosome (chr) size, number of blocks per breed, total length covered by haplotype blocks, block coverage length in percent, number of SNPs in blocks and percent of SNPs in blocks (% SNPs in blocks)

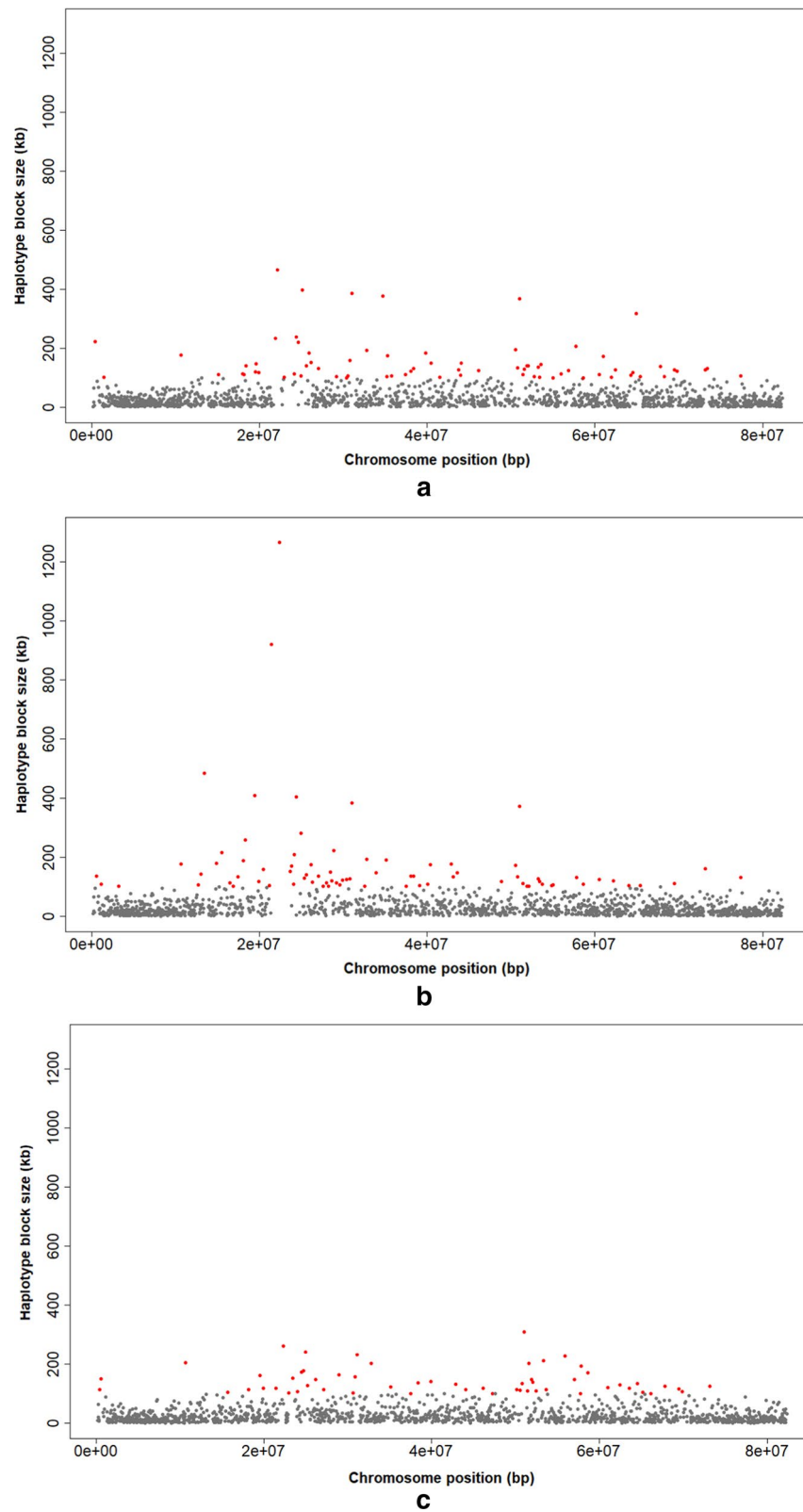
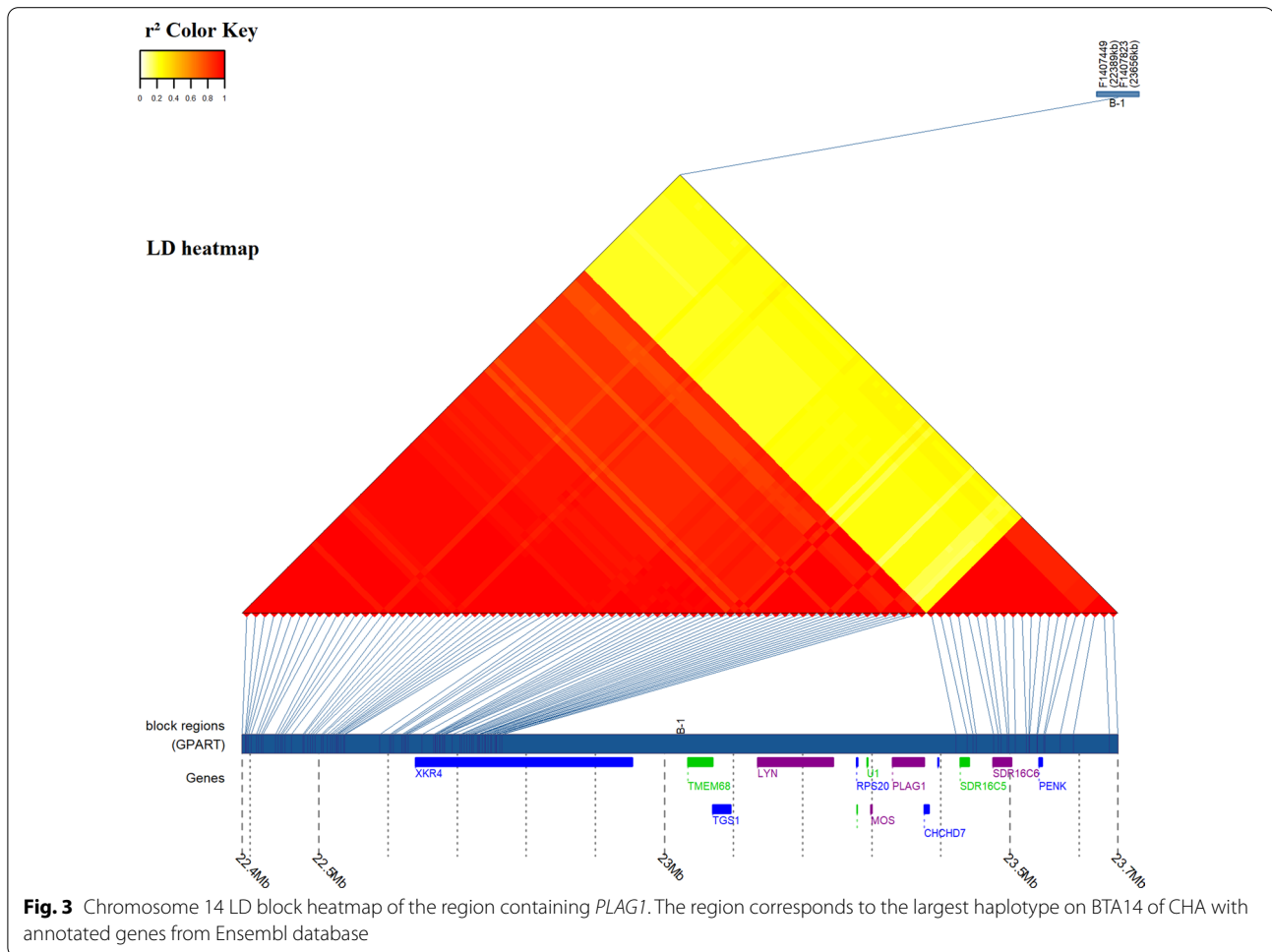


Fig. 2 Size and distribution of haplotype blocks on chromosome 14 of CHA (a), LIM (b) and BLA (c). Red points: block size ≥ 100 kb



with SNPs in the other block, and if the two blocks had an inter-distance greater than or equal to 1 Mb. In total, 15, 266 and 3288 LRLD events were found for $r^2 \geq 0.8$, 0.6 or 0.4, respectively, in LIM (see Additional file 12 Table S16), 29, 598 and 7840 LRLD events were found for $r^2 \geq 0.8$, 0.6 or 0.4, respectively, in CHA (see Additional file 12 Table S17), and 61, 795 and 22,517 LRLD events were found for $r^2 \geq 0.8$, 0.6 or 0.4, respectively in BLA (see Additional file 12 Table S18). A small number of LRLD events was observed for $r^2 \geq 0.8$ and a very large number for $r^2 \geq 0.4$ in the three breeds. In order to analyse a reasonable number of LRLD events, only LRLD events with a $r^2 \geq 0.6$ will be described in the next section. Table 3 shows the number of LRLD events per breed ($r^2 \geq 0.6$). In total, 598, 266 and 795 LRLD events were found for CHA, LIM and BLA, respectively, which indicates that LRLD occurs in these three beef cattle breeds. Each chromosome displayed a different number of LRLD events, with some chromosomes having few or no LRLD events depending on the population. The number of LRLD events in BLA was larger than in the

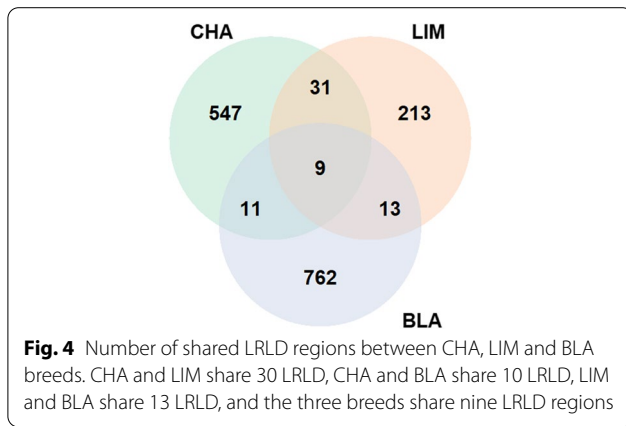
other two breeds (ANOVA, $p=0.001$) and the number of LRLD events was moderately correlated with chromosome size ($\rho=0.58$, $p=0.0010$; $\rho=0.49$, $p=0.0064$; $\rho=0.7300$, $p=8.6e-06$ for CHA, LIM and BLA, respectively). The three breeds shared nine common LRLD events, CHA shared 31 LRLD events with LIM only and 11 with BLA only, and LIM shared 13 LRLD events with BLA only (Fig. 4 and see Additional file 13 Table S19).

We looked at the specificity of chromosome-wide LRLD events in the three breeds (Fig. 5 and Additional file 14: Figure S12). Each breed had predominantly population-specific LRLD events, although common LRLD events also existed in a number of regions, as mentioned above. We were interested in the distance between pairs of LRLD blocks in each breed. The average distance between pairs of LRLD blocks was 1.97 ± 2.05 and 2.15 ± 1.79 Mb, for CHA and LIM, respectively (Table 3). However, this average distance on the whole genome was greater in the BLA breed (3.92 ± 4.96 Mb) and was the largest on almost all the chromosomes of the BLA breed.

Table 3 Descriptive summary of the LRLD events in the Charolaise, Limousine and Blonde d'Aquitaine breeds

BTA	Charolaise					Limousine					Blonde d'Aquitaine				
	N LRLD	Avg dist (Mb)	SD (Mb)	Max dist (Mb)	N LRLD	Avg dist (Mb)	SD (Mb)	Max dist (Mb)	N LRLD	Avg dist (Mb)	SD (Mb)	Max dist (Mb)	N LRLD	Avg dist (Mb)	SD (Mb)
1	30	1.88	0.83	3.73	27	1.99	0.73	3.87	77	3.77	4.85	34.27			
2	22	1.82	0.85	4.97	34	1.99	0.77	5.29	134	3.92	3.84	28.20			
3	17	1.67	0.57	2.54	2	1.32	0.45	1.64	18	6.00	7.03	25.41			
4	22	1.61	0.60	3.49	4	1.90	0.51	2.63	21	3.63	4.10	15.65			
5	43	2.00	0.75	3.91	12	1.42	0.58	3.11	40	2.88	4.62	29.18			
6	60	1.79	0.58	3.18	16	2.04	0.87	4.22	48	5.66	8.49	52.45			
7	74	2.54	1.11	4.39	13	2.47	1.28	4.29	35	3.13	3.75	20.11			
8	57	1.43	0.37	2.86	1	1.07	-	1.07	51	2.58	1.51	7.76			
9	45	1.80	0.73	4.19	15	2.39	1.31	4.37	25	3.10	2.97	14.59			
10	9	1.60	0.47	2.36	12	2.13	0.89	3.74	12	2.99	2.09	8.01			
11	12	1.55	0.45	2.42	21	3.36	1.70	5.58	44	3.27	2.99	12.22			
12	17	1.69	0.51	3.19	3	1.15	0.12	1.24	28	2.91	1.66	6.56			
13	5	1.56	0.20	1.83	0	-	-	-	29	4.63	4.01	22.00			
14	42	1.73	0.84	5.19	4	1.70	0.69	2.51	36	3.87	3.31	14.86			
15	61	1.77	0.75	3.60	46	2.27	1.37	6.64	34	3.95	2.75	10.45			
16	8	1.30	0.21	1.67	4	1.37	0.44	2.03	17	4.27	5.89	25.23			
17	5	1.67	0.34	1.99	8	1.45	0.33	1.96	35	6.24	11.10	36.61			
18	1	2.30	-	2.30	3	1.23	0.18	1.36	15	3.84	3.82	12.57			
19	13	1.37	0.35	2.27	4	2.84	0.14	3.01	0	-	-	-			
20	10	1.33	0.28	1.97	2	1.65	0.63	2.09	5	2.18	1.03	3.65			
21	3	1.23	0.17	1.40	6	1.71	0.58	2.42	20	3.69	3.42	12.87			
22	1	1.62	-	1.62	0	-	-	-	2	2.55	1.04	3.29			
23	6	3.33	0.42	4.08	13	1.61	0.44	2.31	20	1.80	1.15	5.04			
24	11	1.73	0.64	2.91	2	1.34	0.33	1.57	30	3.69	1.57	6.41			
25	0	-	-	-	1	1.07	-	1.07	2	4.84	5.21	8.53			
26	12	9.16	11.72	25.09	11	3.53	7.12	24.97	9	9.25	11.61	24.74			
27	5	1.06	0.07	1.19	2	1.57	0.31	1.79	3	1.30	0.12	1.42			
28	1	1.11	-	1.11	0	-	-	-	4	14.05	7.41	23.13			
29	6	1.49	0.36	1.99	0	-	-	-	1	6.73	-	6.73			
Total	598	1.97	2.05	25.09	266	2.15	1.79	24.97	795	3.92	4.96	52.45			

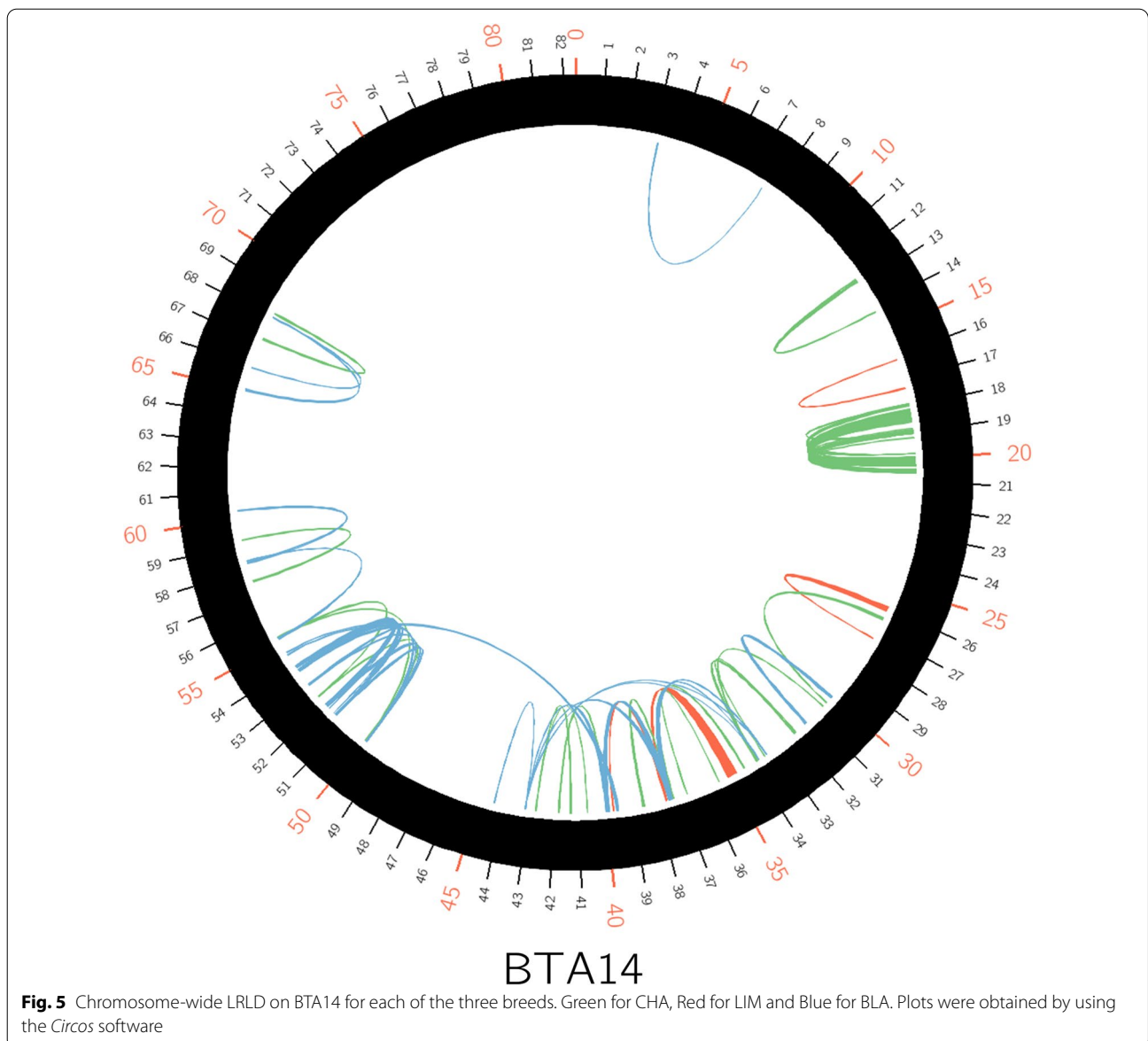
N LRLD: number of LRLD events, Avg dist: average distance in Mb; SD: standard deviation; Max dist: maximal distance between LRLD block pairs (in Mb)



Interestingly, in BLA, 60% of the LRLD events were separated by more than 2 Mb, whereas in CHA and LIM only 28.76 and 39.47% were separated by more than 2 Mb, respectively (Fig. 6).

LRLD functional interactions

Table 4 summarizes the number of LRLD events for which annotated genes were identified in both blocks and Fig. 7 shows the results of possible functional interactions (see Additional file 15: Tables S20–S22). Among the 598, 266 and 795 LRLD events in CHA, LIM and BLA, 270, 116 and 290 have annotated genes in both blocks, respectively. Fifteen of the shared LRLD events between two breeds have annotated genes in each of



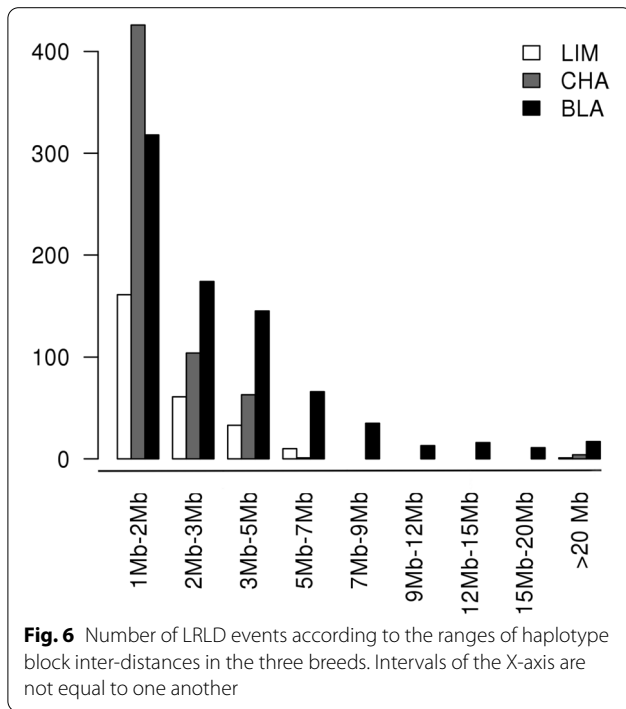


Table 4 Number of LRLD events with annotated genes in both blocks

Breed	Number of LRLD events with genes ^a
Charolaise	270
Limousine	116
Blonde d'Aquitaine	290
Common (2 breeds)	15
Common (3 breeds)	0

^a Number of LRLD events with annotated genes in both blocks

the blocks and no annotated genes were found in the nine regions shared between the three breeds.

The STRING database was queried to see if, within each LRLD event, there were any potential interactions between the proteins encoded by the genes present in their two blocks. Such interactions were found for 78, 38, and 56 LRLD events in CHA, LIM and BLA, respectively, but for only two LRLD events that were shared between two breeds (Table 5 and Fig. 7a). Furthermore, no common LRLD events were identified between the three breeds with STRING interactions (Fig. 7a and see Additional file 15: Tables S20–S22). In order to investigate if these numbers of interactions were due to chance only, the same numbers of pairs of blocks were randomly selected in the three breeds i.e. 598 in

CHA, 266 in LIM and 795 in BLA. In total, 16, 10 and 19 random pairs of blocks showed interactions in CHA, LIM and BLA, respectively. The Chi2 test between the random pairs of blocks and the LRLD events resulted in a p-value ranging from 1.204e–05 to 2.701e–11 (Table 5), which indicates that the interactions identified in the LRLD events were not due to sampling. Interestingly, we detected several examples of STRING-type text-mining interactions. For example, dentin matrix acidic phosphoprotein 1 encoded by *DMP1*, an extracellular protein involved in mineralization of dentin regulates dentin sialophosphoprotein encoded by *DSPP* [38]. The two genes, *DMP1* and *DSPP*, are located in two different blocks of the LRLD events 187–191 of in CHA (see Additional file 15: Table S20). Other examples are the *OR52W1* and *OR52B2* genes related to the olfactory system and the *CNGA4* (*cyclic nucleotide gated channel alpha 4*) gene, which are co-mentioned in PubMed Abstracts (text-mining) [39] and are in two different blocks of LRLD events 479–480 in CHA (see Additional file 15: Tables S20).

The cattle QTL database was used to check for the potential existence of QTL associated with the same phenotype in the two blocks of each LRLD event that could be an indicator of epistatic interactions. In CHA, LIM and BLA respectively, 38, 30 and seven LRLD events with QTL associated with the same phenotype (in each block) were found (Table 6). We checked for traits that are related to selection objectives in beef cattle (i.e. QTL associated with production and carcass or meat quality traits). We did not identify LRLD events with related QTL in LIM, but we found five and nine LRLD events with QTL of interest in CHA and BLA, respectively (Fig. 7b and see Additional file 15: Tables S20, S21 and S22). In CHA, four of the LRLD events may be associated with body weight and one with shear force. In BLA, five of the LRLD events may be associated with body weight, two with scrotal circumference, one with average daily gain and one with muscle anserine content. In the same way as for STRING interactions, the impact of sampling was tested (Table 6). Even if only two of the p-values of the Chi2 test between the random pairs of blocks and the LRLD events were significant (the other non-significant p-values were probably due to the small sample sizes), we observed a clear trend that more QTL were associated with beef-related traits in the LRLD events than with the randomly-chosen pairs of blocks.

Discussion

In this study, we investigated the existence of LRLD events in the three main French beef cattle breeds for the first time. The study covered a relatively large population of genotyped animals, with a very high density

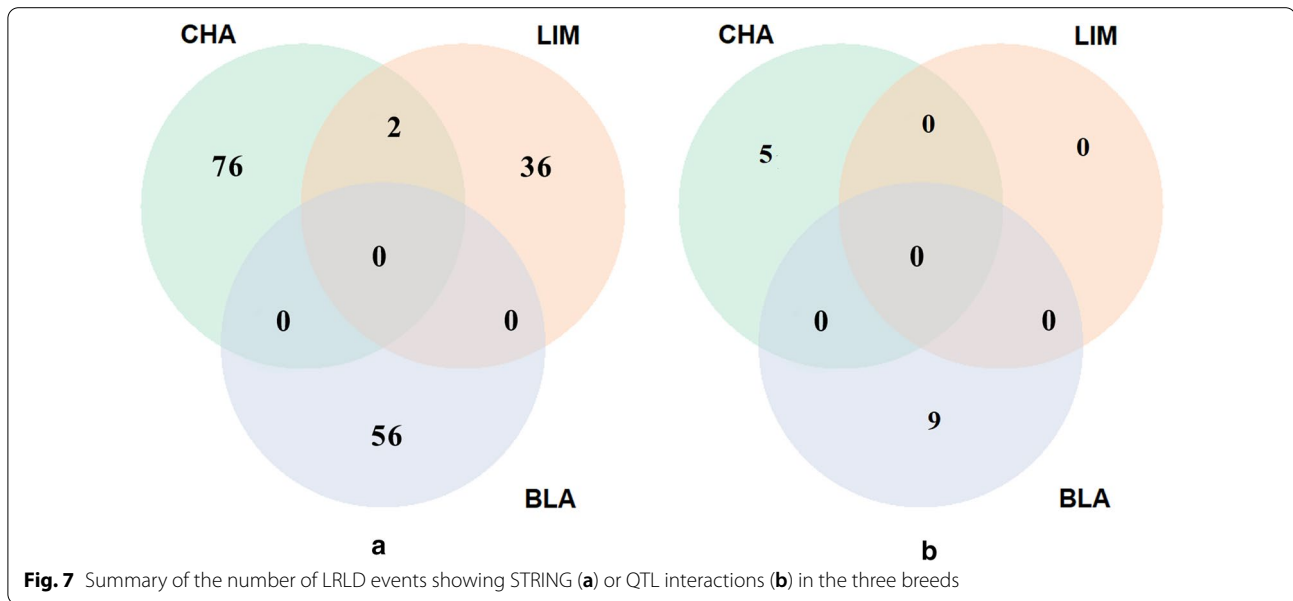


Table 5 Number of LRLD events and randomly-chosen pairs of blocks with functional interactions

Breed	Type of pairs of blocks	Number of pairs of blocks	Number of pairs of blocks with STRING interaction	P-value ^a
Charolaise	LRLD	598	78	2.701e-11
	Random	598	16	
Blonde d'Aquitaine	LRLD	795	56	1.204e-05
	Random	795	19	
Limousine	LRLD	266	38	2.668e-05
	Random	266	10	

^a P-value of a Chi² test between random pairs and LRLD pairs

Table 6 Number of LRLD events and random pairs of blocks with potential QTL interactions

Breed	Type of pairs of blocks	Number of pairs of blocks	Number of pairs of blocks with all QTL ^a	P-value ^b	Number of pairs of blocks with production-related QTL ^c	P-value ^b
Charolaise	LRLD	598	38	1.704e-09	5	0.101
	Random	598	1		1	
Blonde d'Aquitaine	LRLD	795	30	1.439e-09	9	0.002
	random	795	1		0	
Limousine	LRLD	266	7	0.032	0	NA
	Random	266	1		0	

^a Number of pairs of blocks with QTL for the same phenotype in the two blocks (all QTL of the Cattle QTLdb)

^b P-value of a Chi² test between random pairs and LRLD pairs

^c Number of pairs of blocks with QTL for the same phenotype in the two blocks (only QTL for production, carcass and meat quality traits)

of markers. Each population had a different number of LRLD events, which indicates that some long-range interactions are specific to each breed. Nonetheless, common LRLD regions also exist between the three breeds.

The average number of LRLD events was larger in BLA than in the other two breeds and the inter-distance between LRLD events reached up to ~52 Mb (BTA6) in BLA, but was not greater than 25.09 Mb in CHA and LIM (Table 3). There are several possible explanations for such

LRLD in the BLA breed, including population admixture. Indeed, the BLA breed was formed by merging three French South-West Blonde populations (Quercy, Garonnais and Blonde des Pyrénées) in the 1960s [11]. However, PCA and clustering analyses have shown that each of these three populations was genetically homogeneous. The occurrence of bottlenecks can also explain this long range LD in the three breeds. Bouquet et al. [40] showed that the BLA population experienced a bottleneck since the 1970s due to the extensive use of artificial insemination by French breeders, which resulted in a decrease in effective population size (estimated at 247), compared to CHA (601) and LIM (>1000). Bottlenecks have also been suggested as a source of large LD in the human genome [41]. In contrast, the unequal number of LRLD events per chromosome indicates that sub-populations and bottlenecks may not be the main source of the LRLD observed in this study, which suggests a role of selection in these populations.

Interestingly, in our study, we identified several functional interactions between genes from pairs of LRLD regions, which suggests epistatic selection. The functional interaction between genes in LRLD regions may explain some of the LD observed at long distances. Indeed, epistasis can contribute to LD between variants located at a long distance from each other on the same chromosome (e.g. [42]). For example, in CHA, SNPs in the *sin3A-associated protein 130 kDa (SAP130)* and *UDP-glucose glycoprotein glucosyltransferase 1 (UGGT1)* genes were in LRLD, see block number 31 in (Additional file 15: Table S20). These genes are known to be involved in the acetylation of the H3 histone and in protein metabolism [43], respectively. A genome-wide association study (GWAS) has shown that SNPs located 500 kb upstream of these two genes are associated with average daily gain phenotype [44]. The *MYL6B* and *MYL6* genes in CHA, see block number 108 in Additional file 15 Table S20 were identified in the Brown Hanwoo breed (Korean beef cattle) within a region under recent positive selection [45]. This region contains a QTL associated to a skeletal muscle generation phenotype [46]. Furthermore, the *CATSPER3* and *PITX1* genes in CHA, see block number 254 in Additional file 15 Table S20 were found in a candidate region under selection in European and African *Bos taurus* breeds [47]. In BLA, the *SNUPN*, *CSPG4*, and *PTPN9* genes, see block number 712 in Additional file 15 Table S22 were co-mentioned in PubMed Abstracts [48]. They have been identified in Korean cattle in selective sweep regions that are associated with marbling score [48]. In addition, the *OR52W1* and *OR52B2* genes related to the olfactory system, and the *cyclic nucleotide gated channel alpha 4 (CNGA4)* gene were co-mentioned in PubMed Abstracts because

they are involved in the olfactory transduction signal [39]. The *OR52W1* and *CNGA4* genes were identified in regions under selection in the Western Pyrénées sheep [39]. These genes are located in an LRLD event shared between CHA and LIM, see number 18 in Additional file 15: Tables S20 and S22. The orthologous genes of *FOXP2* and *MSANTD1* in CHA, see block number 82 in Additional file 15: Table S20, have been described as under selection in Moroccan Black and Northern goats, respectively [49]. The *DMP1* gene in CHA, see block number 187 in Additional file 15: Table S20, encodes an extracellular protein involved in the mineralization of dentin that regulates DSPP during dentinogenesis [38]. Interestingly, the two genes, *DMP1* and *DSPP*, are found in the same LRLD event. These findings could explain some of the relationships found between LRLD regions in the bovine genome.

The number of LRLD events that involve genes showing interactions was much smaller than the total number of LRLD events in the three breeds. The existence of interactions between genes and genetic regulatory elements on distant genomic regions might also explain some of the observed LRLD events. Indeed, many studies have shown long-distance interactions between non-coding elements of the genome, especially with Hi-C data (e.g. [50–52]). However, since regulatory regions are not well annotated in the bovine genome, it is currently difficult to identify this type of interaction.

We checked, in the cattle QTL database, the presence of QTL associated with the same phenotype in each block of LRLD events. We found five and nine LRLD events showing QTL associated with the same phenotype in CHA and BLA, respectively. Only traits that are selected by breeders in beef cattle were checked, which could explain the small number of overlaps observed between QTL and LRLD events. Indeed, phenotypes related to dairy cattle have been much more widely studied and are thus over-represented in the database. Nevertheless, LRLD events may be an indicator of epistasis between several regions and thus may impact GWAS. Epistasis is considered as one potential explanation of the “missing heritability” [53, 54]. However, computational requirements are a challenge for the detection of epistatic interactions. Using the LRLD information in a GWAS model could help to reduce the complexity by testing the interactions between SNPs in LRLD blocks.

Another possible explanation for some of the detected LRLD events is the presence of errors in the genome assembly. The new bovine genome assembly (ARS-UCD1.2) used in this study has an increased assembly accuracy [55]. However, assembly errors, such as duplications or inversions, can still remain and create false positive LRLD events.

In this study, we identified LRLD events in beef cattle and also confirmed the inversely proportional relationship between LD and the physical distance between markers. These results are consistent with the literature. A study conducted by Hozé et al. [27] on 16 dairy and beef cattle breeds showed that the average LD drops to around 0.1 within a distance of 100 kb. Khatkar et al. [10] estimated that LD values drop to less than 0.08 for distances between markers greater than 200 kb in the Australian Holstein–Friesian cattle. The same result was reported by McKay et al. [17] in a study on eight cattle populations, in which the extent of LD did not exceed 500 kb. Beghain et al. [11] reported useful LD values ($r^2 > 0.2$) up to 724 kb in a separate population of BLA animals and since family structure has been observed in this population, it could be the main source of LD observed at such a distance.

In the CHA breed, BTA14 showed a lower level of LD decay with physical distance than the other chromosomes. Several QTL associated with growth, carcass, meat quality and eating quality traits have been identified on this chromosome in cattle (e.g. [56]). Furthermore, Allais et al. [57] identified a QTL associated with tenderness score on BTA14 (at position 59.5 cM) in the French Charolaise breed. One possible explanation for this large extent of LD could be the presence of genes that have a selective advantage or neutral markers that segregate with QTL on this chromosome. A study based on sequence data detected some loci under selection on this chromosome [58] such as the *PLAG1* gene, which is associated with stature and carcass yield [37]. This gene was located in the largest haplotype block observed in the three breeds.

Conclusions

We conducted a linkage disequilibrium study on three populations of French beef cattle breeds genotyped on HD chips and report, for the first time, long-range linkage disequilibrium in these three breeds. Several hypotheses can explain this observed long-range linkage disequilibrium, such as bottlenecks, population admixture and epistatic selection. Our results should be taken into account for future genome-wide association studies.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12711-021-00657-8>.

Additional file 1: Tables S1–S3. Summary distribution and density of SNPs over the genome in the Charolaise (Table S1), Limousine (Table S2) and Blonde d'Aquitaine (Table S3) breeds. In each table are provided the number of SNPs per chromosome, the chromosome size (kb), average inter-distance (\pm SD) between markers (kb), and median and maximum distance between markers.

Additional file 2: Figures S1–S3. SNP density over autosomes. Each chromosome was divided in windows of 500 kb and SNP density was plotted. Summary distribution and density of SNPs over the genome in the Charolaise (Figure S1), Limousine (Figure S2) and Blonde d'Aquitaine (Figure S3) breeds. Each figure shows SNP density across the autosomes. Each chromosome is divided into windows of 500 kb and SNP density was plotted.

Additional file 3: Figures S4, S5. Family structure analysis in the Charolaise, Limousine and Blonde d'Aquitaine breeds using *SNPRelate R* package. **Figure S4.** Principal component analysis (PCA). **Figure S5.** Clustering analysis based on the matrix of genome-wide identity by state (IBS) pairwise distances.

Additional file 4: Tables S4–S6. Mean linkage disequilibrium (r^2) according to physical distance ranges in 500-kb windows in Charolaise (Table S4), Limousine (Table S5) and Blonde d'Aquitaine (Table S6).

Additional file 5: Tables S7–S9. Mean linkage disequilibrium (r^2) on each autosome (1–29) in 500-kb windows. Number of SNP pairs per chromosome, mean r^2 , standard deviation (sd), percent of SNP pairs with LD superior to 0.2 ($\% r^2 > 0.2$), and percent of SNP pairs with high LD ($\% r^2 > 0.8$) in Charolaise (Table S7), Limousine (Table S8) and Blonde d'Aquitaine (Table S9).

Additional file 6: Figures S6–S8. LD decay within physical distance on each autosome (1–29) in 500-kb windows in the Charolaise (Figure S6), Limousine (Figure S7) and Blonde d'Aquitaine breeds (Figure S8). Dotted horizontal lines are the LD background (computed on a subset of non-syntenic SNPs).

Additional file 7: Tables S10–S12. Summary of haplotype blocks for the Charolais (Table S10), Limousine (Table S11) and Blonde d'Aquitaine (Table S12) breeds. Each table provides the number of SNPs (Nb SNPs) per chromosome, chromosome size, number of haplotype blocks per autosome (1–29), block coverage length (Mb), chromosome block coverage (in percent), number of SNPs in blocks (Nb SNPs in Blocks) and percent of SNPs in blocks.

Additional file 8: Figure S9. Distribution of haplotype block sizes on each autosome (1–29) for the Charolaise breed. The red dots show haplotype blocks with size ≥ 100 kb. Blue circles show regions with larger haplotypes.

Additional file 9: Figure S10. Distribution of haplotype block sizes on each autosome (1–29) for the Limousine breed. The red dots show haplotype blocks with size ≥ 100 kb. Blue circles show regions with larger haplotypes.

Additional file 10: Figure S11. Distribution of haplotype block sizes on each autosome (1–29) for the Blonde d'Aquitaine breed. The red dots show haplotype blocks with size ≥ 100 kb. Blue circles show regions with larger haplotypes.

Additional file 11: Tables S13–S15. List of large haplotype blocks in the Charolaise (Table S13), Limousine (Table S14) and Blonde d'Aquitaine (Table S15) breeds.

Additional file 12: Tables S16–S18. List of LRLD events identified in the Limousine (Table S16), Charolaise (Table S17) and Blonde d'Aquitaine (Table S18) breeds. Each table show the r^2 threshold, the pair of blocks number (LRLD number), breed, chromosome (chr), start position of block 1 (start_block1), end position of block 1 (stop_block1), start position of block 2 (start_block2), end position of block 2 (stop_block2) and distance between blocks.

Additional file 13: Table S19. List of common LRLD events between two or all the three breeds.

Additional file 14: Figure S12. Chromosome-wide LRLD on the Charolaise (CHA), Limousine (LIM) and Blonde d'Aquitaine (BLA) autosomes. Green for CHA, Red for LIM and Blue for BLA. Plots were done using *Circos* software.

Additional file 15: Tables S20–S22. List of LRLD events with functional interactions in the Charolaise (Table S20), Limousine (Table S21) and Blonde d'Aquitaine (Table S22) breeds. STRINGdb shows the interacting proteins in each block in LRLD from the STRING database. Cattle QTL db shows the LRLD blocks with the same QTL name identified in the cattle QTL database.

Acknowledgements

The authors would like to acknowledge the GEMBAL consortium for providing data. GEMBAL project was funded by the Agence Nationale de la Recherche (ANR-10-GENM-0014), Apis-Gene and Races de France. The authors would also acknowledge the university of Limoges for funding this work, CTIG and Genotoul for providing computational facilities.

Authors' contributions

AEH performed data analysis. VB secured the funding and supervised the project with RP and DR. RP performed QTL and gene interaction analysis. AEH drafted the manuscript and RP, DR, VB and EV revised it. All authors read and approved the final manuscript.

Funding

This work was funded by the University of Limoges.

Availability of data and materials

The data that support the findings of this study are available from the GEMBAL consortium but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the GEMBAL consortium.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹INRAE, PEIRENE EA7500, USC1061 GAMAA, Université de Limoges, 87060 Limoges, France. ²INRAE, AgroParisTech, GABI, Université Paris-Saclay, 78350 Jouy-en-Josas, France.

Received: 20 May 2020 Accepted: 15 July 2021

Published online: 23 July 2021

References

- Terwilliger JD, Weiss KM. Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr Opin Biotechnol*. 1998;9:578–94.
- Flury C, Tapio M, Sonstegard T, Drögemüller C, Leeb T, Simianer H, et al. Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium. *J Anim Breed Genet*. 2010;127:339–47.
- Shin DH, Cho KH, Park KD, Lee HJ, Kim H. Accurate estimation of effective population size in the Korean dairy cattle based on linkage disequilibrium corrected by genomic relationship matrix. *Asian-Australas J Anim Sci*. 2013;26:1672–9.
- Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, et al. Detecting recent positive selection in the human genome from haplotype structure. *Nature*. 2002;419:832–7.
- Kim Y, Nielsen R. Linkage disequilibrium as a signature of selective sweeps. *Genetics*. 2004;167:1513–24.
- Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001;157:1819–29.
- Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet*. 2005;37:549–54.
- Craig DW, Stephan DA. Applications of whole-genome high-density SNP genotyping. *Expert Rev Mol Diagn*. 2005;5:159–70.
- Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, et al. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One*. 2009;4:e5350.
- Khatkar MS, Nicholas FW, Collins AR, Zenger KR, Cavanagh JAL, Barris W, et al. Extent of genome-wide linkage disequilibrium in Australian Holstein–Friesian cattle based on a high-density SNP panel. *BMC Genomics*. 2008;9:187.
- Beghain J, Boitard S, Weiss B, Boussaha M, Gut I, Rocha D. Genome-wide linkage disequilibrium in the Blonde d'Aquitaine cattle breed. *J Anim Breed Genet*. 2013;130:294–302.
- Hedrick PW. Gametic disequilibrium measures: proceed with caution. *Genetics*. 1987;117:331–41.
- Hill WG, Robertson A. Linkage disequilibrium in finite populations. *Theor Appl Genet*. 1968;38:226–31.
- Lewontin RC. The interaction of selection and linkage. I. General considerations: heterotic models. *Genetics*. 1964;49:49–67.
- Bohmanova J, Sargolzaei M, Schenkel FS. Characteristics of linkage disequilibrium in North American Holsteins. *BMC Genomics*. 2010;11:421.
- Zhao H, Nettleton D, Dekkers JCM. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms. *Genet Res*. 2007;89:1–6.
- McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J, Coppieters W, et al. Whole genome linkage disequilibrium maps in cattle. *BMC Genet*. 2007;8:74.
- Mokry FB, Buzanskas ME, Mudadu MA, Grossi DA, Higa RH, Ventura RV, et al. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics*. 2014;15:56.
- Koch E, Ristroph M, Kirkpatrick M. Long range linkage disequilibrium across the human genome. *PLoS One*. 2013;8:e80754.
- Park L. Population-specific long-range linkage disequilibrium in the human genome and its influence on identifying common disease variants. *Sci Rep*. 2019;9:11380.
- Vallejo RL, Silva RMO, Evenhuis JP, Gao G, Liu S, Parsons JE, et al. Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels: evidence that long-range LD is a major contributing factor. *J Anim Breed Genet*. 2018;135:263–74.
- Wilson JF, Goldstein DB. Consistent long-range linkage disequilibrium generated by admixture in a Bantu–Semitic hybrid population. *Am J Hum Genet*. 2000;67:926–35.
- Lewontin RC, Kojima K. The evolutionary dynamics of complex polymorphisms. *Evolution*. 1960;14:458–72.
- Schaper E, Eriksson A, Rafajlovic M, Sagitov S, Mehlig B. Linkage disequilibrium under recurrent bottlenecks. *Genetics*. 2012;190:217–29.
- Caceres A, Sindi SS, Raphael BJ, Caceres M, Gonzales JR. Identification of polymorphic inversions from genotypes. *BMC Bioinformatics*. 2012;13:28.
- Barbat M, Tribout T, Saintilan R, Venot E, Fouilloux M-N, Phocas F. Comparison of accuracies of genomic prediction in French Limousin cattle population according to the number of markers and to pedigree relationship between training and validation populations. In: *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production: 16–21 August 2014, Vancouver*; 2014.
- Hozé C, Fouilloux MN, Venot E, Guillaume F, Dassonneville R, Fritz S, et al. High-density marker imputation accuracy in sixteen French cattle breeds. *Genet Sel Evol*. 2013;45:33.
- Gunia M, Saintilan R, Venot E, Hozé C, Fouilloux MN, Phocas F. Genomic prediction in French Charolais beef cattle using high-density single nucleotide polymorphism markers. *J Anim Sci*. 2014;92:3258–69.
- Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7.
- Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *Int J Methods Psychiatr Res*. 2018;27:e1608.

31. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. 2012;28:3326–8.
32. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. *Science*. 2002;296:2225–9.
33. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19:1639–45.
34. Cunningham F, Achuthan P, Akanni W, Allen J, Amodio MR, Armean IM, et al. Ensembl 2019. *Nucleic Acids Res*. 2019;47:D745–51.
35. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47:D607–13.
36. Cattle Quantitative Trait Locus (QTL) Database. <https://www.animalgenome.org/cgi-bin/QTLdb/BT/index>. Accessed 14 Feb 2020.
37. Song Y, Xu L, Chen Y, Zhang L, Gao H, Zhu B, et al. Genome-wide association study reveals the *PLAG1* gene for knuckle, biceps and shank weight in Simmental Beef cattle. *PLoS One*. 2016;11:e0168316.
38. Gibson MP, Zhu Q, Wang S, Liu Q, Liu Y, Wang X, et al. The rescue of dentin matrix protein 1 (DMP1)-deficient tooth defects by the transgenic expression of dentin sialophosphoprotein (DSPP) indicates that DSPP is a downstream effector molecule of DMP1 in dentinogenesis. *J Biol Chem*. 2013;288:7204–14.
39. Ruiz-Larrañaga O, Langa J, Rendo F, Manzano C, Iriando M, Estonba A. Genomic selection signatures in sheep from the Western Pyrenees. *Genet Sel Evol*. 2018;50:9.
40. Bouquet A, Renand G, Phocas F. Evolution de la diversité génétique des populations françaises de bovins allaitants spécialisés de 1979 à 2008. *INRA Prod Anim*. 2009;22:317–30.
41. Schmegner C, Hoegel J, Vogel W, Assum G. Genetic variability in a genomic region with long-range linkage disequilibrium reveals traces of a bottleneck in the history of the European population. *Hum Genet*. 2005;118:276–86.
42. Hamon SC, Kardina SLR, Boerwinkle E, Liu K, Klos KLE, Clark AG, et al. Evidence for consistent intragenic and intergenic interactions between SNP effects in the *APOA1/C3/A4/A5* gene cluster. *Hum Hered*. 2006;61:87–96.
43. Consortium TGO. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res*. 2019;47:D330–8.
44. Serão NVL, González-Peña D, Beever JE, Bollero GA, Southey BR, Faulkner DB, et al. Bivariate genome-wide association analysis of the growth and intake components of feed efficiency. *PLoS One*. 2013;8:e78530.
45. Lim D, Strucken EM, Choi BH, Chai HH, Cho YM, Jang GW, et al. Genomic footprints in selected and unselected beef cattle breeds in Korea. *PLoS One*. 2016;11:e0151324.
46. Stella R, Biancotto G, Krogh M, Angeletti R, Pozza G, Sorgato MC, et al. Protein expression changes in skeletal muscle in response to growth promoter abuse in beef cattle. *J Proteome Res*. 2011;10:2744–57.
47. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. Composite selection signals for complex traits exemplified through bovine stature using multibreed cohorts of European and African *Bos taurus*. *G3 (Bethesda)*. 2015;5:1391–401.
48. Ryu J, Lee C. Genetic association of marbling score with intragenic nucleotide variants at selection signals of the bovine genome. *Animal*. 2016;10:566–70.
49. Benjelloun B, Alberto FJ, Streeter I, Boyer F, Coissac E, Stucki S, et al. Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (*Capra hircus*) using WGS data. *Front Genet*. 2015;6:107.
50. Ron G, Globerson Y, Moran D, Kaplan T. Promoter-enhancer interactions identified from Hi-C data using probabilistic models and hierarchical topological domains. *Nat Commun*. 2017;8:2237.
51. Choy M-K, Javierre BM, Williams SG, Baross SL, Liu Y, Wingett SW, et al. Promoter interactome of human embryonic stem cell-derived cardiomyocytes connects GWAS regions to cardiac gene networks. *Nat Commun*. 2018;9:2526.
52. Zhang M, Hill MC, Kadow ZA, Suh JH, Tucker NR, Hall AW, et al. Long-range Ptx2c enhancer promoter interactions prevent predisposition to atrial fibrillation. *Proc Natl Acad Sci USA*. 2019;116:22692–8.
53. Mackay TFC, Moore JH. Why epistasis is important for tackling complex human disease genetics. *Genome Med*. 2014;6:124.
54. Upton A, Trelles O, Cornejo-García JA, Perkins JR. Review: high-performance computing to detect epistasis in genome scale data sets. *Brief Bioinform*. 2016;17:368–79.
55. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *GigaScience*. 2020;9:giaa021.
56. Wibowo TA, Gaskins CT, Newberry RC, Thorgaard GH, Michal JJ, Jiang Z. Genome assembly anchored QTL map of bovine chromosome 14. *Int J Biol Sci*. 2008;4:406–14.
57. Allais S, Levéziel H, Hocquette JF, Rousset S, Denoyelle C, Journaux L, et al. Fine mapping of quantitative trait loci underlying sensory meat quality traits in three French beef cattle breeds. *J Anim Sci*. 2014;92:4329–41.
58. Boitard S, Boussaha M, Capitan A, Rocha D, Servin B. Uncovering adaptation from sequence data: lessons from genome resequencing of four cattle breeds. *Genetics*. 2016;203:433–50.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



V.3. Discussion générale

Dans cette étude, nous avons montré pour la première fois l'existence d'événements de déséquilibre de liaison à longue distance (LRLD) dans les trois principales races bovines françaises. L'étude a porté sur 3 populations relativement importantes d'animaux génotypés sur des puces à haute densité de marqueurs. Le nombre de LRLD identifiés est différent d'une race à l'autre, ce qui indique la spécificité des interactions à longue distance dans chacune des races. Néanmoins, des régions LRLD communes existent entre les trois races.

On a remarqué que le nombre moyen d'événements LRLD chez la Blonde d'Aquitaine est plus élevé que dans les deux autres races et l'inter-distance entre les blocs en LRLD atteint jusqu'à environ 52 Mb (sur le chromosome 6 par exemple), tandis que la distance maximale entre les paires de blocs en LRLD chez la race Charolaise et la Limousine ne dépasse pas 25,09 Mb. Ce nombre important de LRLD ainsi que les distances élevées entre blocs haplotypiques chez la BLA peuvent être expliquées par le mélange de populations (ou admixture) génétiquement distinctes et ayant des fréquences haplotypiques différentes (Nei & Li, 1973) (cf. IV.4.1.5). En effet, la race BLA est issue de la fusion de trois populations de Blonde du Sud-Ouest français (Quercy, Garonnais et Blonde des Pyrénées) dans les années 1960. Les goulets d'étranglement (*bottlenecks*) peuvent également expliquer ce LRLD dans les trois races. Les *bottlenecks* correspondent à une réduction sévère de la taille de la population suite à des phénomènes naturels (catastrophes naturelles, maladies, etc.) ou artificiels (comme l'utilisation intensive de l'IA dans l'élevage). Des travaux de recherche français ont montré que la population BLA connaît un goulet d'étranglement depuis les années 1970 en raison de l'utilisation intensive de l'insémination artificielle par les éleveurs français (Bouquet et al., 2009), qui a entraîné une diminution de la taille effective de la population (estimée à 247), par rapport à la race CHA (601) et à la LIM (> 1 000). En revanche, le nombre inégal d'événements LRLD par chromosome indique que les sous-populations et les goulets d'étranglement ne sont peut-être pas les principales sources de LRLD observée dans cette étude, ce qui suggère le rôle de la sélection dans ces populations d'élevage.

Il est aussi intéressant de noter que plusieurs interactions fonctionnelles entre les gènes des régions chromosomiques en LRLD ont été identifiées dans la présente étude, ce qui suggère une sélection épistatique probablement à l'origine de ce DL à longue-distance. Des exemples de gènes ou marqueurs interagissant à longue-distance ou impliqués dans le même phénotype ont été observés dans ces paires de blocs en LRLD. Par exemple chez le CHA, des SNP en LRLD ont été identifiés sur les gènes *SAP130* et *UGGT1*. Ces gènes sont connus respectivement pour leur rôle dans l'acétylation des histones H3 et le métabolisme des protéines (Carbon et al., 2019). Des polymorphismes SNP situés 500 kb en amont de ces gènes ont été identifiés dans des analyses d'association (GWAS) comme étant associés au phénotype gain moyen quotidien (Serão et al., 2013). Les gènes *MYL6B* et *MYL6* en LRLD ont été identifiés dans une région soumise à une sélection positive récente chez les Hanwoo (race bovine allaitante coréenne) (Lim et al., 2016a). Cette région contient des QTL associés au développement du muscle squelettique (Stella et al., 2011a).

En outre, les gènes *CATSPER3* et *PITX1* ont été trouvés dans une région candidate soumise à la de sélection dans des races bovines européennes et africaines (Randhawa, Khatkar, Thomson, & Raadsma, 2015). En BLA, les gènes *SNUPN*, *CSPG4* et *PTPN9* ont été mentionnés conjointement dans des résumés PubMed (Ryu & Lee, 2016). Ces gènes ont été identifiés chez des bovins coréens dans des régions soumises à la sélection et ils ont été associés au persillé de la viande (Ryu & Lee, 2016). De plus, les gènes *OR52W1* et *OR52B2*,

liés au système olfactif et le gène *CNGA4* ont été mentionnés conjointement dans des résumés Pubmed, car ces trois gènes sont impliqués dans la même voie de transduction de signal (Ruiz-Larrañaga et al., 2018). Ces gènes sont situés sur des blocs LRLD en commun entre le CHA et la LIM.

Le nombre d'événements LRLD présentant des interactions entre les gènes codants est beaucoup plus faible que le nombre total de LRLD dans les trois races. L'existence d'interactions entre les gènes et les éléments de régulation génétique sur des régions génomiques éloignées pourrait également expliquer une partie des événements LRLD observés. En effet, de nombreuses études ont montré des interactions à longue distance entre des éléments non-codants du génome (ex. (Choy et al., 2018; Ron, Globerson, Moran, & Kaplan, 2017; M. Zhang et al., 2019)). Les enhancers des vertébrés par exemple, peuvent être situés jusqu'à plus d'une mégabase des promoteurs sur lesquels ils agissent (Lettice et al., 2003; Sagai et al., 2004). Alors, bien que ces éléments régulateurs soient souvent éloignés de leurs gènes cibles, ils affectent l'expression des gènes en recrutant des facteurs de transcription dans des régions promotrices spécifiques (Vermunt, Zhang, & Blobel, 2019). En raison de cette action à longue distance, les polymorphismes régulateurs ne sont pas largement annotés sur le génome bovin et il est donc actuellement difficile d'identifier ce type d'interaction. La méthode 3C (chromosome conformation capture) (Dekker, Rippe, Dekker, & Kleckner, 2002) et ses dérivés (4C, 5C, capture-C et Hi-C) permet d'identifier la manière de communication entre les éléments régulateurs sur des grandes distances. La méthode 3C-seq est basée sur l'immobilisation chimique de la conformation de la chromatine et le séquençage des régions d'interaction (Sati & Cavalli, 2017).

Nous avons vérifié la présence de QTL associés au même phénotype dans chaque bloc LRLD dans la base de données des QTL bovins. Nous avons trouvé très peu d'événements LRLD montrant des QTL associés au même phénotype dans les deux races CHA et BLA. Mais seuls les caractères sélectionnés par les éleveurs de bovins allaitants ont été vérifiés, ce qui pourrait expliquer le faible nombre de chevauchements entre les QTL et les événements LRLD. En effet, les phénotypes liés aux bovins laitiers sont très étudiés et donc surreprésentés dans la base de données. Néanmoins, les événements LRLD peuvent être un indicateur d'épistasie entre plusieurs régions et peuvent donc avoir un impact sur les études GWAS. L'utilisation de ces informations LRLD dans des modèles GWAS détectant les interactions épistatiques pourrait aider à réduire la complexité en testant uniquement les interactions entre les SNP en LRLD.

Dans cette étude, nous avons également confirmé la relation inversement proportionnelle entre le DL et la distance physique entre les marqueurs. Ces résultats sont donc conformes à la littérature.

Le chromosome 14 de la race CHA a montré un déclin du DL plus faible en fonction de la distance physique par rapport aux autres chromosomes. Plusieurs QTL associés à des caractéristiques de croissance, de carcasse, de qualité de la viande ont été identifiés sur ce chromosome chez les bovins. L'une des explications possibles de cette large étendue du DL pourrait être due à des gènes qui ont un avantage sélectif sur ce chromosome, ou à des marqueurs neutres qui se ségrégent avec les QTL. La présence de loci adaptatifs sur ce chromosome, comme le gène *PLAG1* associé à la stature et au rendement en carcasse peut renforcer cette hypothèse (Song et al., 2016).

V.4. Conclusion

Dans cette étude réalisée sur trois populations de bovins allaitants français, et génotypées sur des puces HD, nous avons analysé dans le détail pour la première fois le déséquilibre de liaison à longue-distance. Plusieurs hypothèses sont possibles pour expliquer ce LRLD observé, telles que les goulots d'étranglement, le mélange des populations et la sélection épistatique. Les résultats de cette étude et l'approche suivie pourraient être pris en compte pour les futures études d'association à l'échelle du génome.

Chapitre 2 : Détection de variants associés à la qualité des viandes bovines à l'aide des données de séquence génome imputé

Chapitre 2 : Détection de variants associés à la qualité des viandes bovines à l'aide des données de séquence génome imputé

VI.1. Introduction

Les études d'association pan-génomiques (*GWAS : Genome-Wide Association Study*) ou études d'association à l'échelle du génome sont devenues un outil puissant et de plus en plus abordable pour découvrir les bases génétiques des maladies complexes chez l'homme (Gudmundsson et al., 2009 ; Hirschhorn & Daly, 2005) et les caractères économiquement importants chez les animaux d'élevage suite au développement du séquençage du génome et des technologies de génotypage à haut débit des polymorphismes SNP (Ramayo-Caldas, Renand, Ballester, Saintilan, & Rocha, 2016 ; Sanchez et al., 2020, 2019 ; L. Wang et al., 2014; L. Zhang et al., 2013).

Historiquement, plusieurs études d'association GWAS ont été réalisées chez les bovins et de nombreux gènes associés à des caractères d'intérêt économique ont été détectés (*ex. cf.* III.3.2).

Dans les races allaitantes, des analyses d'associations basées sur des données de moyenne et haute densité ont permis d'identifier des variants associés à la qualité de la viande et au développement musculo-squelettique chez les bovins allaitants français. Par exemple, une étude réalisée par Allais *et al.* (2014) sur 3 races allaitantes (Charolaise, Limousine et Blonde d'Aquitaine) génotypés sur les puces commerciales Illumina 50 k dans le cadre du programme Qualvigène a identifiée plus de 200 variants significatifs associés à la qualité de la viande chez le CHA et la BLA, et 78 variants significatifs chez la LIM (S. Allais et al., 2014).

Plus récemment, des analyses d'association basées sur la séquence du génome complet ont permis d'identifier de nouveaux gènes candidats dans les races bovines. En race laitière, Sanchez *et al.* (2019) par exemple ont identifié plusieurs gènes candidats associés à la composition et l'aptitude fromagère du lait chez la Montbéliarde (Sanchez et al., 2019). Une autre étude dans les races Holstein et Montbéliarde, basée sur les données de séquence du génome complet a identifié 3 gènes candidats (*ABCC4*, *IER3* et *CBFA2T2*) pour la résistance à la paratuberculose, une maladie infectieuse due à la bactérie *Mycobacterium paratuberculosis* (Sanchez et al., 2020).

D'autres études basées sur des données de séquence du génome complet ont été réalisées dans différentes races allaitantes et pour différents caractères d'intérêt économique, mais dans des races non-françaises. Par exemple, une étude canadienne menée dans une population de 7 573 animaux issus de plusieurs races bovines allaitantes, sur la base de 7 853 211 variants imputés et répartis sur tout le génome a identifié plusieurs variants significativement associés à l'efficacité alimentaire et ses composantes (gain moyen quotidien, apport journalier en matière sèche et le poids corporel métabolique) (F. Zhang et al., 2020). Une autre étude menée sur une population de 2110 animaux de la race Hanwoo (race allaitante) imputés génome complet a révélé 107 SNP significatifs associés aux caractères de la qualité de la viande (Bedhane et al., 2019). En outre, une étude américaine récente menée sur une population de 2285 animaux de race Angus a identifié plusieurs variants associés à la tendreté de la viande (Leal-Gutiérrez et al., 2020a). En revanche, aucune étude basée sur les données de séquence du génome complet n'est publiée à ce jour sur les caractères de la qualité des viandes et le développement musculaire dans les races bovines allaitantes françaises.

L'objectif de cette deuxième partie de la thèse est d'étudier le déterminisme génétique des caractères d'aptitudes bouchères et des qualités de la viande dans des races bovines allaitantes française (population Qualvigène précédemment décrite (cf. III.3.1)) grâce à une analyse GWAS en utilisant des données SNP imputées au niveau de la séquence du génome complet.

VI.2. Matériel et méthodes

VI.2.1. Animaux et phénotypes

Dans cette étude, 1059 animaux de la race Charolaise (CHA), 1209 Limousins (LIM) et 935 Blonde Aquitaine (BLA) issus du programme Qualvigène (c.f. Partie III.3.1) ont été utilisés. Les animaux sont phénotypés pour 22 ou 24 caractères liés au muscle, les aptitudes bouchères, la qualité de la viande ainsi que les performances mesurées en routine pendant le testage sur descendance (Allais *et al.*, 2010).

Les caractères liés aux aptitudes bouchères ont été mesurés sur les animaux vivants et après abattage. Les animaux sont pointés durant les 4 semaines avant l'abattage (à 570 jours en moyenne en CHA, et à 650 kg en moyenne en LIM et BLA ce qui correspond à un âge situé entre 15 et 24 mois) et une note de morphologie a été donnée pour le développement musculaire (dmvif) et le développement squelettique (dsvif). Les caractères condition de naissance (condnais), gain moyen quotidien (gmq) et poids de naissance (pd_naiss) sont également enregistrés. Ensuite, le poids (poidab) et l'âge à l'abattage (ageab) ont été enregistrés. Le poids de carcasse a été enregistré à chaud sur la chaîne d'abattage et le rendement en carcasse (rendca) a été déduit (% de viande prêt à être commercialisée). Le poids de gras internes (poidgi) (gras de bassin et de rognon) a également été pesé. Une note a été donnée pour la conformation EUROP (confca) au moment du ressuage (étape importante de refroidissement des carcasses à 7 °C dans moins de 24 heures). La longueur *Jarret-symphyse* (longjs) et l'épaisseur de cuisse (epaicu) ont été enregistrées pour chaque demi-carcasse. La surface du muscle *Longissimus dorsi* (LD) (surfld) a été mesurée 24 heures après l'abattage, et le pourcentage de gras dans la 6ème côte a également été mesuré (dap6c) (rapport du poids de dépôts adipeux sur la somme du poids du muscle LD, du poids des autres muscles et des dépôts adipeux de cette côte), ainsi que des mesures de gras interne par ultrason (VOS : *Velocity Of Sound*) ont aussi été réalisées.

Pour les caractères liés au muscle, les mesures ont été réalisées sur un échantillon de la 7ème côte du muscle LD. La surface moyenne des fibres musculaires a été calculée par analyse d'images (mtafibr), ensuite le nombre moyen de fibres du muscle (nbfibr) a été déduit en utilisant la surface du muscle LD et la surface moyenne des fibres musculaires. La teneur en collagène total a été mesurée par la méthode décrite dans les normes NF V 04-415, basée sur la méthode de BERGMAN et LOXLEY (1963) qui utilise la concentration en acide hydroxyproline ($C_{collagène} = 7,5 * C_{hydroxyproline}$) (AFNOR, 2002). La teneur en collagène insoluble (mtxcoli) a été mesurée en triplicat par dosage après un traitement de l'échantillon libérant la fraction thermolabile du collagène soluble (BONNET & KOPP, 1986). La solubilité du collagène (solcol) a ensuite été calculée. La teneur en lipides intramusculaires (mtxlipi) a été mesurée par la méthode d'extraction à l'éther de pétrole sur des échantillons séchés (cité par Sophie Allais, 2011a).

Concernant les caractères de la qualité de la viande, la couleur (micoull) a été mesurée 5 fois à l'aide d'un spectrophotomètre (cité par Sophie Allais, 2011a). La force de cisaillement (mcocisa) (10 répétitions) et les qualités sensorielles ont été mesurées sur des steaks ayant 14 jours de maturation et grillés à 55 °C. Les qualités sensorielles sont des notes de tendreté (mnotend), de jutosité (mnojuto) et de flaveur (mnoflav) de la viande qui ont été obtenues par un jury de dégustation (composée de 12 dégustateurs) spécifique de la race (et souvent identique pour les 3 années du programme), en moyennant les notes des membres du jury.

VI.2.2. Imputation pan-génomique

Les animaux ont été génotypés à l'aide de la puce BovineSNP50 (50 k, Illumina Inc., San Diego). Les filtres utilisés pour le contrôle qualité des données sont ceux du système national d'évaluation français (D. Boichard et al., 2012) : taux de génotypage par individu supérieur à 95 %, taux de génotypage par SNP supérieur à 90 %, fréquence des allèles mineurs (MAF) supérieure à 1 % dans au moins une population bovine, fréquences génotypiques en équilibre de Hardy-Weinberg avec $P > 10^{-4}$. Les animaux qui étaient incompatibles avec leurs parents (lorsque tous les génotypes étaient disponibles) ont été exclus (Sanchez et al., 2020). Ensuite, les SNPs génotypés sur la puce 50 k ont été imputés au niveau de la séquence du génome complet. Une approche en deux étapes a été appliquée afin d'améliorer la précision des génotypes imputés (Van Binsbergen et al., 2014). Tout d'abord, la population GEMBAL (une population multiraciale) a été utilisée pour imputer les génotypes de la puce 50 k à la puce HD (777 k) en utilisant le logiciel FImpute (Sargolzaei, Chesnais, & Schenkel, 2014). Ensuite, le passage de la densité HD imputée à la séquence du génome complet a été effectué en utilisant les données 1000 génomes (une population composée de 2712 animaux multiraciales, y compris 147 Charolais, 104 Limousins et 41 Blonde d'Aquitaine) à l'aide du logiciel Minimac (Leidy, 2011). La sortie du logiciel minimac comprend une statistique R2 qui indique la qualité d'imputation (plus la statistique est proche de 1, plus l'imputation est de bonne qualité).

VI.2.3. Analyse d'association séquence imputée génome complet

Nous avons effectué une analyse d'association à un seul caractère entre les 14119747 marqueurs (polymorphes) et les 24 caractères liés au développement musculaire et à la qualité de la viande. Les marqueurs monomorphes (entre 571499 et 575764 en LIM, 635036 et 658656 en CHA et entre 639351 et 642992 en BLA) sont exclus de l'analyse. Toutes les analyses d'association ont été réalisées en utilisant l'option *mlma* du logiciel GCTA (version 1.26.0) (Yang, Lee, Goddard, & Visscher, 2011) qui applique un modèle linéaire mixte incluant le variant à tester :

$$y = \mathbf{1}\mu + \mathbf{x}b + \mathbf{u} + e$$

Avec y est le vecteur des phénotypes préajustés, en moyenne par individu, μ est la moyenne générale, b est l'effet fixe additif du variant à tester pour l'association, x est le vecteur des dosages d'allèles prédits (varie entre 0 et 2). $u \sim N(0, G\sigma_u^2)$, est le vecteur des effets polygéniques aléatoires, avec G la matrice de relation génomique (*Genomic relationship matrix: GRM*) et σ_u^2 est la variance polygénique, estimée sur la base du modèle nul ($y = \mathbf{1}\mu + \mathbf{u} + e$) et ensuite fixée en testant l'association entre chaque variant et le caractère d'intérêt. $e \sim N(0, I\sigma_e^2)$ est le vecteur des effets résiduels aléatoires, avec I la matrice d'identité et σ_e^2 la variance résiduelle. Les variants significatifs sont sélectionnés sur la base de leur qualité d'imputation ($R2 \geq 0,4$) et leur p-valeur ($\leq 10^{-7}$).

VI.2.4. Annotation des variants et identification de mutations candidates

Les variants significatifs issus de l'analyse d'association sont annotés à l'aide de l'outil Ensembl Variant Effect Predictor (VEP version ensembl 103) (McLaren et al., 2016). Cet outil produit une annotation génomique des SNPs en fonction de leur localisation sur les gènes. Pour déterminer les phénotypes liés à nos mutations candidates (variants situés au niveau des gènes), la base de donnée *Mouse Genome Informatics* (MGI) a été interrogée (Bult et al., 2019). Le but est de voir si ces gènes sont impliqués dans le caractère étudié dans le modèle souris.

L'expression des gènes candidats (issus de l'annotation des variants significatifs) dans le muscle bovin a été vérifiée sur la base de donnée *Gene Expression Atlas* (Fang et al., 2020). La capacité des variants génétiques à modifier les sites de fixation des facteurs de transcription (SFFT) a été prédite à l'aide d'un script personnalisé (développé par les équipes G2B et SIGENAE à l'INRAE de Jouy-en-Josas) qui utilise des modèles de SFFT provenant des bases de données JASPAR (collection JASPAR CORE 2018) (Sandelin, Alkema, Engström, Wasserman, & Lenhard, 2004), HOCOMOCO (version v10) (Khamis et al., 2018) et TRANSFAC (version v3.2 publique) (KNÜPPEL, DIETZE, LEHNBERG, FRECH, & WINGENDER, 1994). Ces bases de données contiennent des ensembles de modèles de liaison de facteurs de transcription représentés sous forme de matrices de poids de position (*Position Weight Matrices : PWM*), qui sont dérivés de collections publiées de SFFT eucaryotes définies expérimentalement. Seuls les PWM de vertébrés ont été téléchargés pour être utilisés dans notre étude.

La base de données NONCODE (v6.0) (L. Zhao et al., 2021) a été utilisée afin d'identifier les long non-coding RNA (lncRNA) chevauchant avec nos variants significatifs. La liste des lncRNA inter-géniques exprimés dans le muscle chez la Limousine identifiée par Billerey *et al* (2014) a été également utilisée dans cette étude (Billerey et al., 2014).

Enfin, la base de données AnimalQTLdb (Hu, Park, & Reecy, 2019) a été utilisée pour déterminer les QTL liés au développement musculaire et la qualité des viandes chevauchants avec nos variants GWAS.

VI.3. Résultats

VI.3.1. Nombre de variants significatifs par race et par caractère

L'analyse d'association pan-génomique est réalisée avec le logiciel GCTA sur 14 119 747, 13 257 513 et 13 532 474 SNP chez le CHA, LIM et BLA, respectivement. Cette analyse a permis d'identifier 8 893, 1 182 et 5 495 variants significatifs au seuil p-valeur 10^{-7} et 6 349, 504, 4 509 variants significatifs au seuil p-valeur 10^{-8} , associés aux caractères étudiés dans la race CHA, LIM et BLA, respectivement (Tableau 8). Dans ce chapitre, nous focaliserons sur les variants associés aux phénotypes liés aux qualités des viandes et les aptitudes bouchères.

Nous avons détecté 7 variants significatifs au seuil p-valeur 10^{-7} (un seul au seuil 10^{-8}) pour la note de tendreté (mnotend) dans le CHA situés en position 43-46 Mb sur le chromosome 29, alors qu'aucun variant n'a été détecté pour ce caractère dans les deux autres races (Figure 16 & Tableau 8 & Annexe 1). Pour la note de jutosité (mnojuto), nous avons détecté 5 variants significatifs au seuil 10^{-7} sur le chromosome 10 (position 32,09-32,21 Mb) dans la race LIM, et aucun variant significatif n'a été détecté dans les deux autres races.

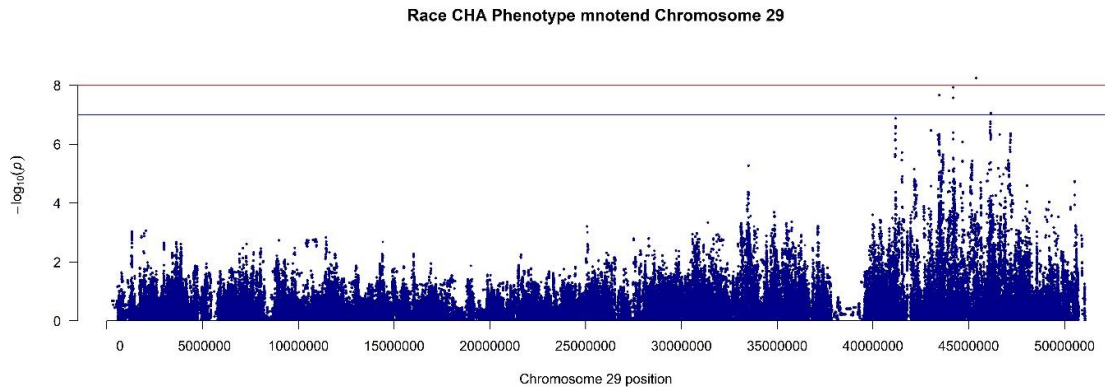


Figure 16: Distribution des variants associés au caractère note de tendreté sur le BTA 29 du CHA

Nous avons également détecté 106 variants significatifs au seuil 10^{-7} (y compris 4 au seuil 10^{-8} en position 27 Mb du chromosome 3) associés au caractère force de cisaillement dans la race LIM, 6 dans le CHA significatifs au seuil 10^{-8} (en position 28,2 Mb du chromosome 23) et un seul variant dans la BLA (au seuil 10^{-7} en position 45,1 Mb du chromosome 22) pour le même caractère.

Enfin, nous avons identifié 21 variants au seuil 10^{-7} (dont 7 au seuil 10^{-8} situés sur le chromosome 2 en position ~6,2 - 7,3 Mb) associés au taux de lipides intramusculaires (mtxlipi) dans le CHA et aucun variant significatif n'a été détecté dans la BLA et la LIM pour ce caractère. De plus, un seul variant significatif au seuil 10^{-7} associé au taux de collagène insoluble (mtxcoli) a été identifié dans le CHA (en position 7 341 014 sur le chromosome 2) (Figure 17).

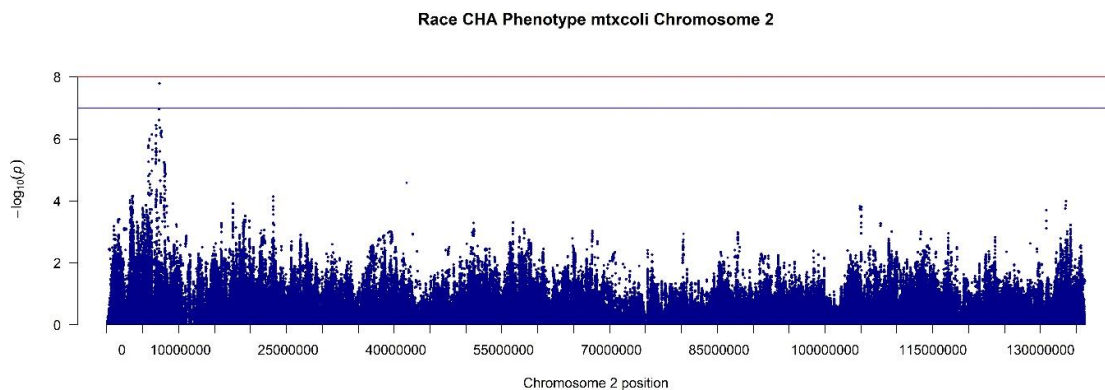


Figure 17: Distribution des variants associés au caractère taux de collagène insoluble sur le BTA2 du CHA

Une centaine de variants significatifs associés aux caractères du muscle (développement musculaire, rendement et conformation de carcasse) ont été identifiés sur le chromosome 2 dans les 3 races. Ces variants sont proches (~1 Mb) du gène de la myostatine (en position 2 :678 630), un gène très connu pour son rôle dans le développement musculaire chez le bovin (Grobet *et al.*, 1997 ; Doyle *et al.*, 2020).

Tableau 8: Nombre de variants significatifs au seuil p-valeur 10^{-7} (p-valeur 10^{-8}) par race et par caractère.

caractère	Nombre de variants significatifs au seuil 10^{-7} (10^{-8})		
	Charolaise	Limousine	Blonde Aquitaine
cond_naiss	0 (0)	22 (1)	NA
confca	1514 (1237)	14 (0)	408 (264)
dap6c	1182 (973)	0 (0)	956 (897)
dmvif	334 (68)	0 (0)	361 (287)
dsvif	1 (0)	129 (61)	912 (774)
epaicu	1145 (733)	81 (43)	0 (0)
gmq	0 (0)	154 (0)	75 (0)
longjs	0 (0)	379 (149)	923 (730)
mcocisa	6 (4)	106 (4)	1 (0)
micoull	42 (10)	2 (0)	0 (0)
mnoflav	0 (0)	0 (0)	0 (0)
mnotend	7 (1)	0 (0)	0 (0)
mnojuto	0 (0)	5 (0)	0 (0)
mtafibr	3 (1)	0 (0)	0 (0)
mtxcoli	1 (0)	0 (0)	0 (0)
mtxlipi	20 (7)	0 (0)	0 (0)
nbfibr	33 (11)	0 (0)	0 (0)
pds_naiss	0 (0)	0 (0)	NA
poidab	NA	119 (78)	443 (314)
poidgi	10 (4)	0 (0)	307 (199)
rendca	3717 (2971)	171 (168)	1109 (1044)
solcol	0 (0)	0 (0)	0 (0)
surfld	821 (307)	0 (0)	0 (0)
vos	57 (22)	0 (0)	0 (0)
ageab	0 (0)	NA	NA
Total	8893 (6349)	1182 (504)	5495 (4509)

ageab: âge à l'abattage, **cond_naiss**: condition de naissance, **confca**: conformation de carcasse, **dap6c**: pourcentage de gras dans la 6ème côte, **dmvif**: développement musculaire, **dsvif**: développement squelettique, **epaicu**: épaisseur de la cuisse, **gmq**: gain

moyen quotidien, **longjs**: longueur jarret-symphyse, **mcocisa**: force de cisaillement, **micoull**: mesure de la couleur (clarté), **mnoflav**: note de flaveur, **mnotend**: note de tendreté, **mnojuto**: note de jutosité, **mtafibr**: surface moyenne des fibres musculaires, **mtxcoli**: teneur en collagène insoluble, **mtxlipi**: teneur en lipides intramusculaires, **nbfibr**: nombre moyen de fibres du muscle LD, **pds_naiss**: poids de naissance, **poiadab**: poids à l'abattage, **poidgi**: poids des gras internes, **rendca**: rendement en carcasse, **solcol**: solubilité du collagène, **surfld**: surface *longissimus dorsi*, **vos**: velocity of sound (gras interne).

Seuls, quelques variants en commun pour le même phénotype entre deux races ont été observés dans cette étude et aucun variant commun entre les 3 races n'a été identifié pour l'ensemble des polymorphismes étudiés au seuil 10^{-7} (Figure 18).

Les 42 variants en commun entre CHA et BLA (Figure 18 A) pour le caractère rendcar sont situés sur le chromosome 2 en position 5,6 – 7,5 Mb. Ces variants sont localisés dans les gènes *INPP1*, *MSTN*, *COL5A2*, *COL3A1*, *NAB1*, *HIBCH* et un micro ARN (*bta-mir-2917*), alors que les 6 variants en commun entre ces 2 races pour le caractère confca sont situés sur le chromosome 2 en position 5,9 – 6,2 Mb et localisés sur les gènes *MSTN* et *INPP1*. De plus, 2 et 17 variants en commun pour le caractère dmvif (localisés sur le gène *INPP1*) et dap6c (localisés sur le gène *INPP1*, *MSTN*, *COL5A2*, *bta-mir-2917* et *COL3A1*) (Figure 18 K - L), respectivement, ont été identifiés dans les mêmes races. Seulement 5 variants en commun entre la race BLA et la race LIM ont été identifiés. Ces variants intergéniques sont associés au caractère longueur jarret-symphyse (longjs) (Figure 18 F).

Plusieurs variants pléiotropiques (variants associés à des caractères différents) ont été identifiés dans cette étude (1750 dans la BLA, 1666 dans le CHA et 154 dans la LIM). Certains variants sont associés jusqu'à 11 caractères différents (ex. Les variants 2:6861781 et rs799943285 dans le CHA (Figure 19).

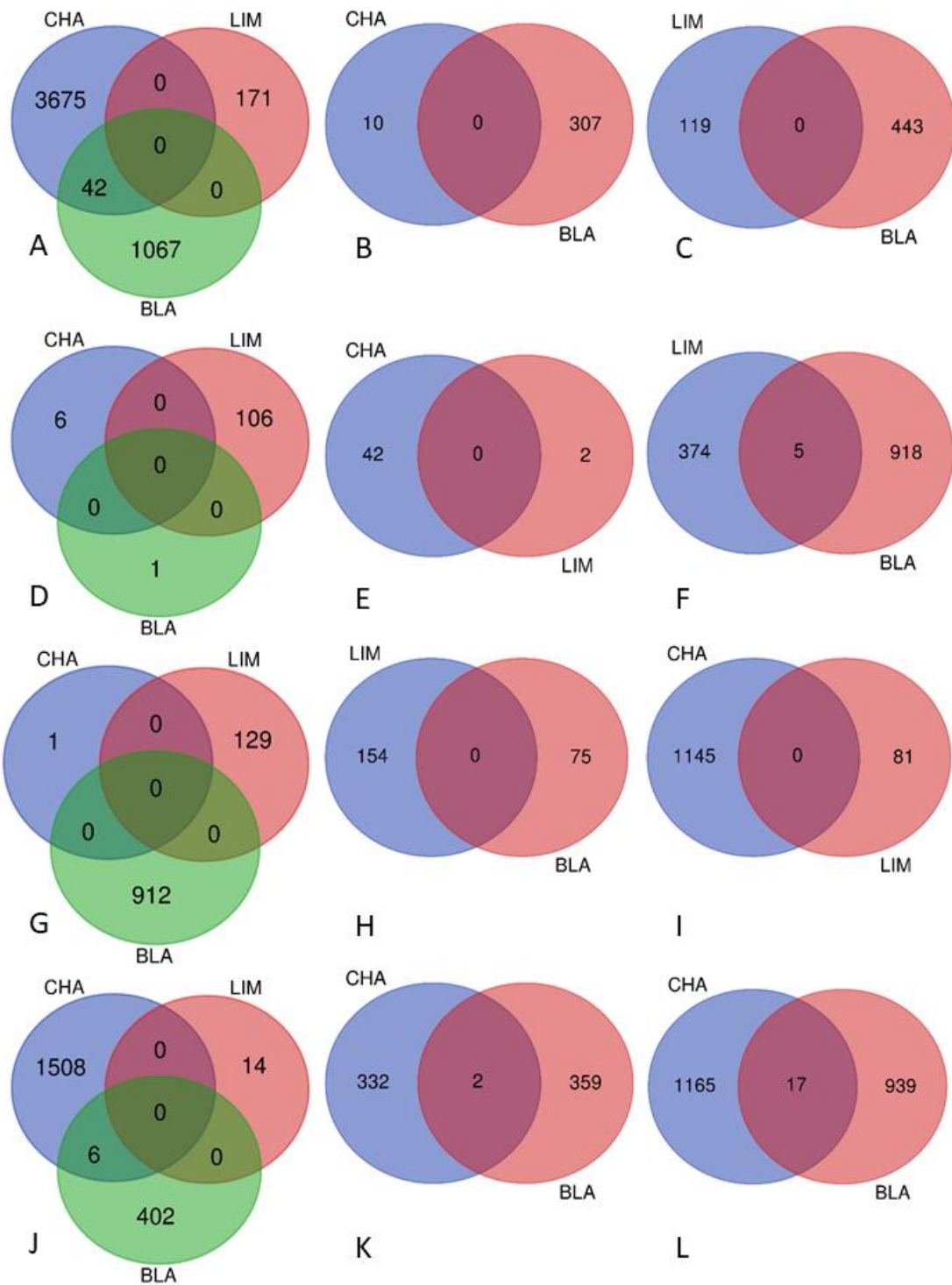


Figure 18: Diagramme de Venn des variants en commun entre races au seuil 10^{-7} . A : rendcar, B: poidgi, C: poidab, D : mcoicisa, E: micoull, F: longjs, G : dsvif, H : gmq, I : epaicu, J : confca, K : dmvif, L : dap6c

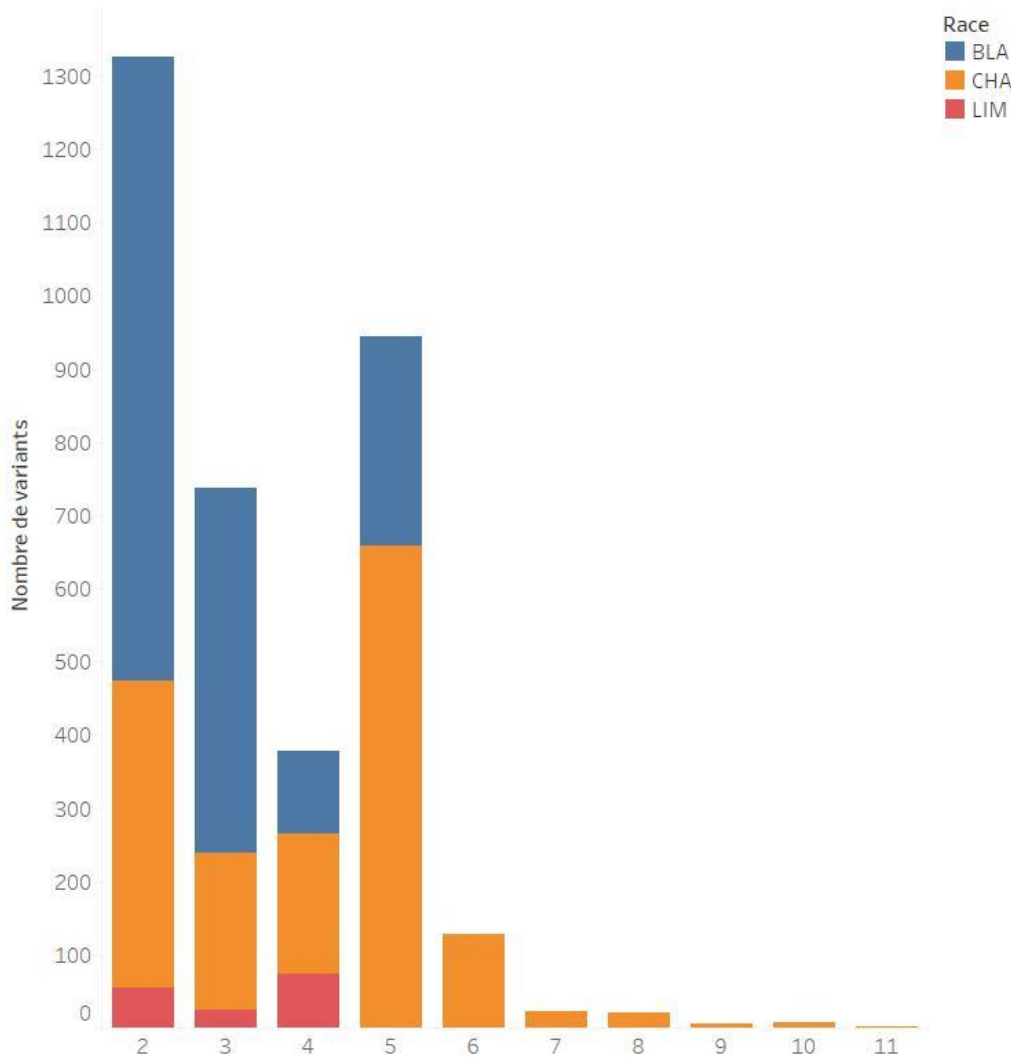


Figure 19: Nombre de variants par pléiotropie dans les différentes races CHA, LIM et BLA

VI.3.2. Annotation des variants et identification de mutations candidates

Les variants significatifs au seuil p-valeur 10^{-7} ont été annotés à l'aide de l'outil VEP. Cet outil permet une annotation fonctionnelle des variants selon leur position sur le génome et identifie l'effet de ces variants sur les gènes, les transcrits et la séquence des protéines, ainsi que sur les régions régulatrices. La majorité des variants annotés dans les 3 races est située dans des régions introniques (39,8 %) ou intergéniques (50,37 %), et un nombre important de variants sont localisés dans des régions 5' (*upstream variants*) (3,92 %) ou 3' des gènes (*downstream variant*) (4,55 %) (Tableau 9).

Les mutations faux-sens (*missense variant*) sont des mutations ponctuelles induisant un changement d'acide aminé de la protéine. Ce type de mutations peut altérer la fonction de la protéine produite. Les 5 variants faux-sens annotés chez le CHA sont situés sur le chromosome 1 (un seul variant rs79827398 dans l'exon 1 du gène *TMSB4*) et le chromosome 2 (4 variants dans les gènes *TUBGCP5*, *AMER3*, *ARHGEF4* et *UGGT1*). Le variant rs798273598 situé dans l'exon 1 du gène *TMSB4* (*thymosin beta-4*) est un variant délétère (SIFT 0,01), tandis que les autres variants faux-sens sont tolérés (SIFT entre 0,07 et 0,44).

Les 4 variants faux-sens chez la LIM sont situés sur le chromosome 6 en position 35,9 – 37,4 Mb et sont localisés dans les gènes *LCORL*, *FAM13A*, *SPP1* et *FAM184B*. Parmi ces mutations faux-sens, le variant rs110329232 dans l'exon 1 du gène *SPP1* (*secreted phosphoprotein 1*) est considéré par VEP comme délétère avec peu de confiance (SIFT = 0,02). Chez la BLA, 7 variants faux-sens situés sur le chromosome 2 en position 5,8 – 6,5 Mb sont dans les gènes *MFSD6*, *MSTN*, *PMS1* et *OSGEPL1*, alors que les autres variants sont situés sur le chromosome 6 en position 36 – 37,2 Mb et sont dans les gènes *FAM13A*, *MEPE* et *FAM184B*. Le variant pléiotropique rs110065568 correspond à la mutation F94L connue chez la LIM (S Allais et al., 2010). Elle est localisée dans l'exon 1 du gène de la *MSTN* est délétère (SIFT=0). Les variants rs208066431 (SIFT = 0,01), rs134577412 (SIFT= 0,03), rs210785796 (SIFT = 0,03) localisés sur les gènes *FAM13A*, *MEPE* et *FAM184B*, respectivement, sont également délétères.

Tableau 9: Nombre de variants annotés par conséquence dans les 3 races.

Conséquence	Race			Nombre total de variants uniques (nombre de gènes différents)
	CHA	LIM	BLA	
Variant intronique	1440	453	804	2658 (64)
Variant downstream	179	12	148	304 (55)
Variant 5' UTR	2	2	3	7 (6)
Variant intergenique	1943	347	1111	3361 (1)
Variant upstream	140	21	112	262 (49)
Mutation faux sens	5	4	7	16 (14)
Mutation synonyme	13	3	13	28 (17)
Variant 3' UTR	9	3	9	21 (11)
Variant non-sens (stop gained)	1	0	0	1 (1)
Variant d'épissage	4	3	6	13 (7)
Mutation cadre de lecture (Frame shift)	0	0	1	1 (1)

Le seul variant (rs110344317) induisant un codon stop (*stop gained*) chez le CHA est la mutation Q204X du gène de la myostatine responsable du phénotype culard (S Allais et al., 2010). Nous avons trouvé que ce variant pléiotropique est associé à plusieurs caractères mesurés (conformation de carcasse, dépôt adipeux, développement musculaire, épaisseur de la cuisse, couleur, teneur en lipides intramusculaire, nombre de fibres, rendement de carcasse, surface LD et vos).

Les variants situés dans des régions d'épissage peuvent conduire à des isoformes différents, et par conséquent peuvent impacter la structure du gène. Chez le CHA, ces variants sont dans le gène *UGGT1* (*UDP-glucose glycoprotein glucosyltransferase 1*). Chez la LIM, ces variants sont localisés dans les gènes *ABCG2*, *FAM13A* et *ATP1A1* alors que chez la BLA, ils sont localisés dans les gènes *OSGEPL1*, *WDR75* et *MMRN1* (Tableau 10).

Tableau 10: Liste des variants situés dans des régions d'épissage.

Race	Variant ID	Position	Ensembl ID	Symbole
CHA	rs209648063	2:4366951	ENSBTAG00000005444	<i>UGGT1</i>
CHA	rs378432500	2:4366952	ENSBTAG00000005444	<i>UGGT1</i>
CHA	rs380082764	2:4392067	ENSBTAG00000005444	<i>UGGT1</i>
CHA	rs207983400	2:4396210	ENSBTAG00000005444	<i>UGGT1</i>
LIM	rs110362902	6:36567616	ENSBTAG00000017704	<i>ABCG2</i>
LIM	rs29026552	6:35975140	ENSBTAG00000011187	<i>FAM13A</i>
LIM	rs797830828	3:26917858	ENSBTAG00000001246	<i>ATP1A</i>
BLA	rs210504134	2:6601375	ENSBTAG00000012557	<i>OSGEPL1</i>
BLA	rs133570721	2:6946998	ENSBTAG00000003002	<i>WDR75</i>
BLA	rs134865524	2:6947004	ENSBTAG00000003002	<i>WDR75</i>
BLA	rs109772311	2:6954063	ENSBTAG00000003002	<i>WDR75</i>
BLA	rs135449420	2:6975818	ENSBTAG00000003002	<i>WDR75</i>
BLA	rs135790137	6:34695942	ENSBTAG00000010285	<i>MMRN1</i>

De plus, 140, 21 et 112 variants « upstream » ont été identifiés dans le CHA, LIM et BLA, respectivement. À noter que tous les variants « upstream » dans le CHA sont situés sur le chromosome 2. Certains de ces variants sont localisés dans des gènes connus pour leur rôle dans le développement musculaire, comme le gène de la *MSTN* (Allais *et al.*, 2010), *NAB1* et le gène *MYO7B* (*myosin VII B*) (Doyle *et al.*, 2020).

Parmi les variants « upstream » chez le CHA, LIM et BLA, respectivement, 34, 6 et 59 modifient potentiellement 117, 20 et 223 sites de fixation à un facteur de transcription (SFFT) exprimés dans le muscle chez les bovins. Dans la race LIM, nous avons identifié 2 variants (rs384338078 et rs383473904) annotés dans la région « upstream » du gène de la myostatine (gène situé sur le chromosome 2 et exprimé dans le muscle). Ces variants sont associés aux caractères conformation de carcasse et rendement de carcasse. Un de ces variants (rs384338078) modifie probablement deux sites de fixation à un facteur de transcription exprimés dans le muscle. En « upstream » de ce gène, on a également identifié 19 variants pléiotropiques associés aux caractères conformation de carcasse, épaisseur de la cuisse, rendement en carcasse, surface LD, dépôt adipeux, développement musculaire et teneur en lipides intramusculaire dans la race Charolaise. De plus, chez la BLA, 4 variants annotés dans la région « upstream » du gène de la *MSTN* ont été identifiés. Ces variants sont associés aux caractères conformation de carcasse, dépôt adipeux et rendement en carcasse.

Sur la totalité des variants introniques chez le CHA, 5 variants situés sur le BTA29 en position (44,19 – 46,15 Mb) sont associés à la tendreté de la viande. Ces variants sont localisés dans les gènes *PACS1*, *PITPNM1* et *TESMIN*. De plus, 3 variants sur le BTA2 en position 7,2 – 7,3 Mb sont associés à la teneur en lipides intramusculaires. Ces variants sont liés aux gènes *COL5A2* et *COL3A1*. Chez la BLA, le variant intronique (rs209337704) en position 45 117 502 sur le BTA22 est associé à la force de cisaillement, ce variant est présent dans l'intron 14 du

gène *ERC2* (*ELKS/RAB6-interacting/CAST family member 2*). Enfin, 19 variants introniques associés au caractère force de cisaillement ont été identifiés chez la LIM. Ces variants sont situés en position 26,9 – 38,2 Mb du BTA3 et sont localisés dans les gènes *ATP1A1*, *MAB21L3* et un lncRNA (ENSBTAG00000048546). L'ensemble des variants introniques dans les 3 races sont positionnés sur 65 gènes différents, et parmi ces gènes 51 (78,46 %) sont exprimés dans le muscle et 53 (81,53 %) sont exprimés dans le tissu adipeux chez le bovin.

La liste des lncRNA intergéniques exprimés dans le muscle dans la race Limousine identifiés par Billerey et ses collaborateurs (Billerey et al., 2014) a été utilisée dans cette étude afin d'identifier nos variants présents dans ces lncRNA. La base de données NONCODE a été également utilisée.

Un seul variant (rs109580735) situé sur un lncRNA intergénique (XLOC_022489) exprimé dans le muscle bovin a été identifié en comparant nos variants avec les données de (Billerey et al., 2014). Ce variant est situé sur le BTA24 chez la BLA en position 7 001 206 et associé au caractère rendement de carcasse.

En comparant avec la base de données NONCODE, 4 variants chevauchant avec 3 lncRNA ont été identifiés sur le BTA2 chez le CHA. Les 4 variants sont associés aux caractères dépôt adipeux (*dap6c*), épaisseur de la cuisse (*epaicu*) et rendement de carcasse (*rendca*). Chez la LIM, les 3 variants associés aux caractères développement squelettique (*dsvif*), épaisseur de la cuisse (*epaicu*), longueur jarret-symphyse (*longjs*) et poids d'abatage (*poidab*) se chevauchent avec 3 lncRNA de la base de données, et pour la BLA, les 3 variants associés aux caractères *confca*, *dap6c*, développement musculaire vif (*dmvif*), poids de gras interne (*poidgi*) et *rendca* sont situés dans 2 lncRNA (Annexe 2).

Nous avons utilisé la base de données MGI pour identifier les variants localisés dans des gènes ayant lorsque muter, des phénotypes similaires à nos caractères dans le modèle souris. Parmi la totalité des variants annotés dans les 3 races, 3 630, 818 et 2 062 sont localisés dans 55, 27 et 39 gènes différents dans le CHA, LIM et BLA, respectivement et sur l'ensemble de ces gènes, seulement 7 en CHA, 6 en LIM et 7 en BLA sont liés à 38 phénotypes équivalant à nos caractères chez la souris.

VI.3.3. Comparaison avec des données bibliographiques

Le projet Qualvigène a déjà fait l'objet d'une détection de QTL par la méthode LDLA (Allais et al., 2014) à l'aide de la puce Bovine SNP50 en utilisant l'assemblage Btau 4.0 (Liu et al., 2009). Nous n'avons détecté aucun variant en commun avec les résultats de Allais et ses collaborateurs, pour l'ensemble des caractères étudiés dans les 3 races. Par contre, nous avons identifié 9 variants proches des QTLs identifiés par Allais et ses collaborateurs (distance moyenne de 68,5 kb, min = 6,5 kb et max = 235,8 kb) (Tableau 11).

Cinq variants (2 en upstream et 3 intergéniques) en position 5,60 – 5,70 Mb sur le BTA2 du CHA sont associés au caractère taux de lipides intramusculaire. À proximité de ces variants dans la même race, Allais *et al.*, ont identifié un QTL associé au même caractère (S. Allais et al., 2014) (Tableau 11). Les deux variants (en position 5,78 Mb du BTA2) sont annotés en « upstream » du gène *NEMP2* exprimé dans le muscle et ils sont localisés à une distance de ~78 kb du QTL identifié par Allais et collègues. Sur le chromosome 29 dans la même race (en position 44,19 Mb), nous avons identifié 2 variants introniques annotés dans le gène *PACS1* et associés au caractère note de tendreté. Allais et ses collègues ont identifié un QTL associé au même caractère à proximité de ces deux variants (en position 44,9 Mb) dans la race CHA et localisé dans le même gène.

En position 46,11 - 48,68 Mb du chromosome 2 de la race LIM, nous avons identifié 2 variants significatifs associés à la couleur de la viande. L'un de ces variants (en position 48,68 Mb) est situé en « downstream » du gène *GPR17*. À proximité de ces variants, Allais et al ont identifié un QTL pour le même caractère et dans la même race (Tableau 11) mais dans la région 3'UTR du gène *POLR2D* (annotation nouvel assemblage ARS-UCD1.2).

Tableau 11: Comparaison de nos variants avec les résultats de Allais et al, 2014

BTA	Race	Nos variants			Variants Allais et al 2014			Gène	Distance
		Position	Car	gène	IlluminaMarker	Position	Pvaleur		
2	CHA	5603129	mtxlipi	-	Hapmap44381-BTA-47399	5640288	8.5 x 10 ⁻⁷	-	104597
2	CHA	5699321	mtxlipi	-	Hapmap44381-BTA-47399	5640288	8.5 x 10 ⁻⁷	-	8616
2	CHA	5701317	mtxlipi	-	Hapmap44381-BTA-47399	5640288	8.5 x 10 ⁻⁷	-	6586
2	CHA	5785502	mtxlipi	<i>NEM P2</i>	Hapmap44381-BTA-47399	5640288	8.5 x 10 ⁻⁷	-	78096
2	CHA	5785838	mtxlipi	<i>NEM P2</i>	Hapmap44381-BTA-47399	5640288	8.5 x 10 ⁻⁷	-	78432
29	CHA	4419054 1	mnotend	<i>PACS 1</i>	BTA-66033-nors	44900940	4.1 x 10 ⁻¹⁰	<i>PACS1</i>	44730
29	CHA	4419884 5	mnotend	<i>PACS 1</i>	BTA-66033-nors	44900940	4.1 x 10 ⁻¹⁰	<i>PACS1</i>	36425
2	LIM	4611909	micoull	-	Hapmap53000-ss46526222	4565625	1.4 x 10 ⁻⁴	<i>POLR2 D</i>	22829
2	LIM	4868575	micoull	<i>GPR1 7</i>	Hapmap53000-ss46526222	4565625	1.4 x 10 ⁻⁴	<i>POLR2 D</i>	235764

Nous avons également comparé nos résultats avec les analyses de Gutiérrez et collaborateurs effectuées dans la race Angus sur une population de 2285 animaux (Leal-Gutiérrez et al., 2020a). Nous n'avons identifié aucun variant en commun en comparant les mêmes positions sur le génome. En revanche, nous avons identifié 43 variants significatifs (9 en CHA et 34 en LIM) situés à une distance inférieure à 500kb de 81 QTL identifiés dans l'étude de Gutiérrez et al.

Chez le CHA, 2 variants significatifs sur le BTA29 (rs42192427 et rs42193634) associés à la note de tendreté et localisés dans l'intron du gène *PACS1* sont situés à une distance inférieure à 60 kb de 4 QTL différents (rs110425960, rs42193726, rs136852361, rs209751522) identifiés par Gutiérrez *et al* dans le même gène (Tableau 12). Le variant intergénique rs42194138 (en position 29:43466339) associé à la note de tendreté est situé à une distance moyenne de 140 kb de 19 QTL identifiés par (Leal-Gutiérrez et al., 2020a) dans le gène de la calpain 1 (*CAPN1*) (Tableau 12), un gène associé à la qualité de la viande chez les bovins (S. Allais et al., 2011a ; H. J. Lee et al., 2019b ; S. H. Lee et al., 2014 ; Page et al., 2002 ; Sun et al., 2018). Parmi ces variants identifiés par Gutiérrez et al, une mutation faux sens (rs17872000) sur le gène de la *CAPN1* est situé à une distance de 60,6 kb de notre variant significatif rs42194138.

Une comparaison avec les travaux de Wang et Bedhane (Bedhane et al., 2019 ; Yining Wang et al., 2020) a été aussi effectuée, mais aucun variant en commun n'a été identifié pour les caractères équivalents à ceux de notre étude et aucun gène en commun.

Enfin, nous avons comparé nos résultats avec la base de données AnimalQTLdb. Seuls les QTL associés aux caractères de la qualité de la viande, carcasse et production ont été pris en compte. Un total de 119 variants associés aux caractères surface *Longissimus dorsi* (surfld), couleur de la viande, note de tendreté et gain moyen quotidien (dmq) situés à proximité de 15 variants pour les mêmes phénotypes et mêmes races ont été identifiés dans la base de donnée (Annexe 3). Le variant en position 6,28 Mb du BTA2 associé au caractère surfld est situé à proximité de 12 régions QTL de la base de données (QTL ID : 192847, 192856, 164702, 192586, 192652, 192617, 192144, 192613, 192752, 192744, 192793 et 192007) associées aux mêmes caractères (Annexe 3). Ces QTL sont situés au niveau du gène de la *MSTN*.

Tableau 12: Liste des variants significatifs associés à la note de tendreté chez le CHA et proches des QTL identifiés par (Leal-Gutiérrez et al., 2020a) sur le BTA29.

Nos variants				Variants de (Leal-Gutiérrez et al., 2020a)			
BTA	Position	rsID	Race	Gène	rsID	Position	Distance
29	44190541	rs42192427	CHA	<i>PACS1</i>	rs110425960	29:44270964	60655
					rs42193726	29:44270185	59876
					rs136852361	29:44272881	62572
					rs209751522	29:44199591	9051
	44198845	rs42193634	CHA	<i>PACS1</i>	rs110425960	29:44270964	52350
					rs42193726	29:44270185	51571
					rs136852361	29:44272881	54267
					rs209751522	29:44199591	746
	43466339	rs42194138	CHA	<i>CAPN1</i>	rs17871986	29 : 43404781	61784
					rs478744282	29:43406285	60280
					rs135499561	29:43406229	60336
					rs17872006	29:43405989	60575
					rs17872000	29:43405875	60689
					rs17872001	29:43405854	60710
					rs17872002	29:43405846	60718
					rs17872003	29:43405624	60940
					rs382718589	29:43412894	53672
					rs380363716	29:43412838	53728
					rs42192102	29:43407404	59162
rs110434759					29:43412643	53923	
rs207535103					29:43399278	67344	
rs378526749					29:43407056	59509	
rs132863588					29:43406807	59758	
rs42192091	29:43395655	70966					
rs42192094	29:43395808	70813					
rs42192093	29:43395676	70945					
rs42192090	29:43395557	71064					

VI.4. Mutations candidates

On considère comme gène candidat, les gènes contenant des variants significatifs associés aux caractères étudiés, mais qui sont aussi révélés dans d'autres études par différentes méthodes (bases de données QTL, MGI, rSNP, fonction connue, études d'association, etc.). En combinant les différents résultats obtenus dans cette étude et ceux de la littérature, on a identifié une liste de gènes potentiellement, candidats associés à la qualité de la viande dans la race CHA (Tableau 13).

Tableau 13: Liste de mutations candidates pour les caractères d'aptitudes bouchères et qualité des viandes

Race	variant	caractère	BTA	Position	Remarques	Références	Annot	gene
CHA	rs42192427 rs42193634	mnotend	29	44190541	QTLdb ID 36480	(Allais et al., 2014) & (Leal- Gutiérrez et al., 2020a)	Intronique intronique	<i>PACS1</i>
CHA	rs109485444	mtxcoli	2	7341014	-	(Jacinto et al., 2020)	intronique	<i>COL5A2</i>
CHA	rs42194138	mnotend	29	43466339	Situé à 60,6 kb de la mutation faux sens rs17872000	(Allais et al., 2011a; Casas et al., 2006; Sun et al., 2018)	intergénique	<i>CAPN1</i>
BLA	rs110065568	confca, dap6c, dmvif, poidgi, rendcar	2	6279278	Mutation F94L chez la LIM	(Lines, Pitchford, Kruk, & Bottema, 2009)	Mutation faux sens	<i>MSTN</i>
CHA	rs110344317	confca, dap6c, dmvif, epaicu, micoull, mtxlipi, nbfibr, rendca, surfld, vos	2	6281434	Utation Q204X	(Allais et al., 2014)	Stop	<i>MSTN</i>

VI.5. Discussion

Dans cette étude, nous avons détecté des milliers de variants significatifs associés aux caractères de la qualité des viandes, au muscle et aux aptitudes bouchères (Tableau 8). La plupart des variants significatifs identifiés dans les 3 races est intergénique. Cependant, en l'état actuel de nos connaissances, la plupart des études se focalisent sur les gènes dans la partie codante, bien que la majorité des polymorphismes associés à des caractères étudiés se situe sur des régions intergéniques. Cela est dû à des difficultés d'interprétation fonctionnelle des SNP intergéniques (Kern et al., 2021).

La majorité de ces variants intergéniques sont associés aux aptitudes bouchères (conformation carcasse, rendement carcasse et développement musculaire), des caractères d'une héritabilité plutôt élevée et en amélioration depuis plusieurs années. Chez le Charolais, 7 variants significatifs associés à la tendreté de la viande ont été identifiés sur le chromosome 29 en position 43,4 – 46,15 Mb. Deux variants (rs42192427 et rs42193634) parmi les 7 sont situés respectivement, à une distance de 44,7 et 36,4 kb d'un QTL (ID : 36480) identifié sur la base de données Cattle QTLdb. Ce QTL, associé à la tendreté de la viande dans le CHA a été identifié par Allais et collaborateurs dans une analyse d'association par la méthode LDLA sur la même population Qualvigène (S. Allais et al., 2014). Ces deux variants introniques sont situés au niveau du gène *PACS1*. Ce gène est exprimé de manière ubiquitaire (Fang et al., 2020) et, à ce jour, son rôle dans la tendreté de la viande est inconnu. Le SNP rs42192427 dans l'intron 1 du gène *PACS1* est un cis-eQTL⁶ SNP pour *CTSF* (Cathepsine F) dans l'utérus mais pas dans le muscle. Or, ce gène est associé à la tendreté (WBSF) chez le porc (Davoli et al., 2017). Chez le porc Gonzalez-Prendes *et al.* ont trouvé un cis eQTL SNP (dans les tissus muscle et foie) dans la région orthologue (González-Prendes et al., 2019). Il est donc possible que le SNP intronique du gène *PACS1* exerce un effet sur l'expression de *CTSF* qui lui impacte la tendreté. Ce SNP et le gène *CTSF* sont séparés d'environ 410 kb. D'autres cathepsines, telles que les cathepsines B, H et L affectent la tendreté et les processus de maturation post-mortem de la viande (Balatsky et al., 2016 ; Kaur, Hui, & Boland, 2020 ; Kaur et al., 2021 ; Warner, Dunshea, Ponnampalam, & Cottrell, 2005). Ces informations pourraient expliquer l'effet de cette région sur la tendreté de la viande dans la race CHA.

De plus, le seul variant significatif (rs109485444) associé au taux de collagène insoluble chez le CHA est annoté dans la région intronique (au milieu de l'intron 43) du gène *COL5A2* (*Collagen Type V Alpha 2 Chain*). Ce gène exprimé dans le muscle et le tissu adipeux chez le bovin, code pour la chaîne alpha des collagènes fibrillaires. Une mutation au niveau de ce gène chez le CHA a été identifiée récemment comme responsable d'une maladie ressemblant au syndrome d'Ehlers-Danlos, un trouble héréditaire du tissu conjonctif caractérisé par des degrés variables d'hyper-extensibilité et de fragilité de la peau, de cicatrices atrophiques et d'hypermobilité articulaire généralisée (Jacinto et al., 2020). Plusieurs mutations au niveau du gène *COL5A2* ont été identifiées comme responsable du syndrome chez l'homme (Ma et al., 2021 ; Watanabe et al., 2016).

Une mutation faux-sens (rs110329232) en position 36 700 131 sur le BTA6 associée au caractère longueur Jarret-symphyse a été identifiée chez la LIM dans l'intron du gène *SPP1*. Ce gène, également connu sous le nom d'ostéopontine (OPN), est un gène candidat pour des caractères de la carcasse chez le bovin. Il possède 8 exons couvrant 57 376 pb sur BTA6 et code une cytokine multifonctionnelle ayant diverses fonctions, comme le remodelage osseux, la cicatrisation des plaies et l'apoptose (Singh, Gill, Kaur, Amhmed, & Jakhu, 2018). Il a été démontré qu'un InDel dans la région promotrice (T9/T10) de ce gène influence les caractères de croissance (Allan et al., 2007). Un SNP dans l'intron 4 de ce gène a également un effet sur le poids corporel (Pareek et al., 2008).

Très peu de variants en commun entre deux races ont été identifiés pour l'ensemble des caractères étudiés et aucun variant en commun pour les 3 races. Cela est dû probablement à la différence d'architecture génétique et les bases moléculaires entre les races. Par exemple, Allais et collaborateurs ont identifié des variants associés à la tendreté de la viande dans les gènes *CAST* et *CAPN1*, mais les mutations ne sont pas conservées d'une race à l'autre (S.

⁶ D'après la base de données Cattle GTEEx (<https://cgtex.roslin.ed.ac.uk/search/>).

Allais et al., 2011b). Un autre exemple de 2 mutations différentes (Q204X en CHA et F94L en LIM) du gène de la myostatine associées au caractère développement musculaire ont été révélées par Allais et al., 2010.

Nous avons également identifié un nombre très élevé de variants significatifs à proximité du gène de la *MSTN* pour des caractères associés au muscle dans les 3 races. Ce gène responsable du phénotype culard chez les bovins a été cité dans plusieurs études (Akanno et al., 2018 ; Yining Wang et al., 2020). De plus, le seul variant (rs110344317) induisant un codon stop chez le CHA est la mutation Q204X du gène de la myostatine. Ce marqueur pléiotropique est associé à plusieurs caractères liés au muscle et le gras intramusculaire. Ce variant a été identifié par Allais et ses collaborateurs chez la BLA pour son association au taux de gras intramusculaire (Allais et al., 2014). Dans la race Charolaise, les jeunes taureaux porteurs de la mutation Q204X présentaient une carcasse avec moins de graisse, des teneurs en lipides intramusculaires et en collagène moindre et une viande plus claire et plus tendre que celles de bovins homozygotes-normaux. La viande de ces animaux avait également un peu moins de saveur (Allais et al., 2010).

Il est à noter qu'aucun variant en commun (même position, et même race) n'a été identifié en comparant nos résultats à ceux d'Allais et ses collègues faite sur la même population Qualvigène, et seulement 9 variants parmi les 135 annotés sur le nouvel assemblage identifié par Allais et collègues sont relativement proches (distance 22,8 kb et 235,7 kb) de nos marqueurs significatifs pour les mêmes caractères et les mêmes races. De même, en comparant nos résultats avec d'autres travaux (Bedhane et al., 2019; Leal-Gutiérrez et al., 2020b; Yining Wang et al., 2020) effectués sur des populations de bovins allaitantes différentes en utilisant des données imputées du génome complet, aucun variant en commun n'a été identifié. Cependant, des variants situés sur le même gène *PACS1* ont été identifiés en comparant avec les résultats de Leal-Gutiérrez et al. (2020a). Le variant rs42192427 dans un intron du gène *PACS1* associé à la note de tendreté a été identifié par Leal-Gutiérrez et al. (2020a) pour le même caractère. Ce gène est également associé à la tendreté de la viande chez le porc (Davoli et al., 2017). Le variant intergénique rs42194138 sur le BTA29 du CHA est situé à une distance de 60,6 kb de la mutation faux-sens rs17872000 dans le gène *CAPN1*. Plusieurs mutations dans le gène de la *CAPN1* ont été identifiées dans la littérature dans des races bovines différentes. Par exemple, Lee et al ont identifié une mutation *CAPN1*:c.1589G>A ayant un effet significatif sur la tendreté et la jutosité de la viande chez des races Hanwoo coréens, ce qui est fortement soutenu par la modélisation moléculaire de ce variant qui inhibe la liaison de la protéine CAST plus fortement que la protéine sauvage, ce qui peut expliquer son effet sur la tendreté de la viande (Lee et al., 2014). La mutation *CAPN1*:c.948G>C (rs17872000) située à 60,6 kb de notre variant significatif (rs42194138) a été validée avec 6 autres mutations du gène *CAPN1* dans une large population de bovins Hanwoo coréens (Lee et al., 2019a).

Plusieurs hypothèses peuvent expliquer l'absence de variants en commun entre nos résultats et ceux de Allais et al. (2014) effectués sur la même population Qualvigène. D'un côté, le seuil de significativité que nous avons fixé est très strict (p-valeur 10^{-7}), par conséquent les marqueurs en dessous du seuil sont exclus. Par exemple, la mutation rs17872000 associée à la tendreté de la viande dans différentes races sort à une p-valeur de 0,009 ($R^2 = 0,53$) dans le CHA, donc en appliquant un seuil p-valeur de 10^{-7} , elle sera exclue. D'un autre côté, il semble qu'il y ait des erreurs d'imputation des génotypes au niveau de la séquence. Par exemple, la mutation F94L connue chez la LIM sort à une p-valeur de $7,19 \times 10^{-10}$ dans notre étude pour la même race, en revanche sa qualité d'imputation est très faible ($R^2 = 0,008$),

donc elle est automatiquement exclue de l'analyse. La même mutation sort en BLA avec une p-valeur de $6,52 \times 10^{-11}$ et un R2 de 0,97. Pour améliorer cette qualité, il faut avoir plus d'animaux allaitants dans le jeu de données 1000 génomes utilisé dans l'imputation qui sont majoritairement de race laitière. Une 3eme hypothèse pourrait être liée à la méthodologie utilisée dans les deux études. En effet, l'étude de Allais et al. (2014) est réalisée par la méthode LDLA qui repose sur le DL et l'analyse d'association et qui est différente de notre approche GCTA qui repose uniquement sur le DL entre marqueurs. Cela pourrait expliquer les différences de résultats obtenues.

Concernant les variants d'autres études (Bedhane et al., 2019 ; Leal-Gutiérrez et al., 2020b ; Yining Wang et al., 2020), l'absence de marqueurs significatifs en commun avec nos travaux, peut être due essentiellement à l'effectif observé dans nos 3 races Qualvigène. Un tel effectif n'est pas suffisamment important pour détecter facilement les effets faibles des polymorphismes étudiés. En effet, l'étude de Bedhane est réalisée sur une population de 2110 bovins de la race Hanwoo, Leal-Gutiérrez a utilisé une population de 2285 bovins de la race Angus et Wang a utilisé une population multiraciale de 3 354 à 3 984 individus. De plus, les races utilisées dans notre étude sont différentes de celles citées dans les études précédentes.

Parmi les SNPs qui sont fortement associés aux caractères étudiés, 39 SNPs sont localisés dans 39 lncRNA, dont 2 lncRNA différents, exprimés dans le muscle (Billerey et al., 2014). Les lncRNA sont bien connus pour leur implication dans la régulation transcriptionnelle et épigénétique au niveau de la chromatine par le biais d'interactions avec les régulateurs de la chromatine (Caretti et al., 2006; Korostowski, Sedlak, & Engel, 2012). Et donc, il est possible que ces variants dans des lncRNA impactent la variabilité de nos caractères étudiés.

VI.6. Conclusion

Dans cette étude, nous avons détecté plusieurs variants significatifs associés aux caractères de la qualité des viandes, liés au muscle et aux aptitudes bouchères. Nous avons également identifié certains gènes candidats pour les caractères étudiés. Les résultats d'association pangénomiques doivent être répliqués avec d'autres animaux. Des tests fonctionnels permettraient de valider certaines des mutations candidates identifiées lors de ces travaux.

Discussion Générale

Discussion générale

Dans cette thèse, les travaux ont compris 2 parties : une analyse du déséquilibre de liaison à longue distance (LRLD) dans les populations allaitantes françaises à l'aide de puces de haute densité de marqueurs et des analyses d'association pangénomiques pour détecter les variants significatifs associés à la qualité de la viande et aux aptitudes bouchères dans la population Qualvigène à l'aide des données de séquence du génome complet.

Dans la première partie de la thèse, nous avons étudié dans le détail pour la première fois le LRLD dans les populations allaitantes françaises et nous avons proposé des hypothèses qui peuvent être à l'origine de ce LRLD, comme les goulots d'étranglement, le mélange des populations et la sélection. Nous avons notamment identifié plusieurs interactions potentielles fonctionnelles entre des gènes dans des paires de blocs haplotypiques en déséquilibre de liaison longue distance, ce qui suggère une sélection épistasique. Par exemple, dans le CHA, on a identifié des marqueurs en LRLD dans les gènes *SAP130* et *UGGT1* qui sont impliqués dans l'acétylation de l'histone H3 et dans le métabolisme des protéines (El Hou et al, 2021). Un autre exemple des gènes *MYL6B* et *MYL6* dans le CHA ont été identifiés dans la race Brown Hanwoo (bovins coréens) dans une région soumise à une sélection positive récente (Lim et al., 2016b). Cette région contient un QTL associé au phénotype muscle squelettique (Stella et al., 2011b). L'épistasie est responsable d'une partie de ce que l'on appelle « l'héritabilité manquante » qui n'a pas encore été intensément explorée dans le cadre des GWAS (Ritchie, 2015). En effet, une partie du challenge de l'analyse de l'épistasie dans les GWAS est le nombre exponentiel de combinaisons possibles et les temps de calcul qui en résultent. Par exemple, un jeu de données incluant 1 million de SNP nécessiterait de tester 5×10^{11} combinaisons dans un modèle à 2 SNP, $1,7 \times 10^{17}$ dans un modèle à 3 SNP, $4,2 \times 10^{22}$ dans un modèle à 4 SNP etc... La connaissance a priori des régions en LRLD et donc potentiellement en épistasie permettrait de réduire la complexité et ainsi de pouvoir intégrer les interactions épistasiques dans les modèles de GWAS.

Les analyses GWAS avec les données de séquence imputée génome complet sur la population Qualvigène ont révélé 8 893, 1 182 et 5 495 variants significatifs associés à nos caractères d'intérêt en CHA, LIM et BLA, respectivement. Certains de ces variants (3 083) sont localisés dans 91 gènes différents. On a également mis en évidence 6 mutations candidates qui peuvent avoir un effet fonctionnel direct sur 4 gènes différents. Comme par exemple les mutations rs42192427 et rs42193634 sur le gène *PACS1* associés à la tendreté de la viande chez le CHA. En revanche, le gène *PACS1* n'est pas loin du gène *CAPN1* (~427,5 kb) très connu pour son association avec la tendreté de la viande, donc il est possible que l'association ait dû aux variants dans le gène *CAPN1* et pas ceux du gène *PACS1*. Pour vérifier cette hypothèse, il serait intéressant de faire des GWAS conditionnelles en fixant les génotypes pour les mutations connues de *CAPN1* (Yang et al., 2012). Un autre exemple de mutation candidate associée à la tendreté de la viande chez le CHA est le variant rs42194138 dans le gène de la *CAPN1*, situé à proximité (60,6 kb) de la mutation rs17872000 du gène *CAPN1* déjà connue dans la littérature pour son association avec la tendreté de la viande (ex. (Lee et al., 2019a) et (Allais et al., 2011a ; Casas et al., 2006 ; Sun et al., 2018)). De plus, la mutation pléiotropique rs110065568 dans le gène *MSTN* connue sous le nom de F94L chez la LIM (Lines et al., 2009) a été identifiée chez la BLA.

L'absence de mutations en commun avec les travaux d'Allais et al sur la même population Qualvigène est probablement lié au problème de la qualité d'imputation et aux seuils de

significativité fixés. À l'inverse des imputations multiraciales actuelles, et afin d'améliorer la qualité d'imputation dans les races allaitantes, il sera important de faire des imputations avec des séquences que d'animaux intra-races et aussi avec séquence de référence race-spécifique (pangénome) pour inclure l'ensemble des séquences et variants intra-races.

Parmi la totalité des variants significatifs annotés sur le nouvel assemblage (ARS-UCD1.2), la majorité sont introniques ou intergéniques. En l'état actuel de nos connaissances sur le génome, il est encore difficile d'interpréter fonctionnellement ces SNP intergéniques (Kern et al., 2021). Les avancées technologiques et l'intégration de nouvelles données, nous permettrons d'obtenir d'autres informations sur ces variants dans le futur.

Par conséquent, il est nécessaire d'approfondir les résultats de ces analyses GWAS par des analyses de cartographie fine des QTL et par des analyses d'enrichissement, d'analyse des voies métaboliques, des tests de validation sur d'autres animaux et par des analyses moléculaires afin de valider un effet fonctionnel des mutations candidates identifiées pour les caractères de la qualité des viandes et les aptitudes bouchères. Toutefois, même en présence de données de séquence génome complet, une taille de population relativement faible et l'absence de mesures en routine des caractères de la qualité des viandes, rend la SAM ou la sélection génomique difficile, d'où la nécessité d'identifier et valider des mutations candidates pour ces caractères d'intérêt.

Il serait très intéressant de combiner les résultats de l'analyse GWAS avec les résultats de l'analyse LRLD afin d'identifier des variants significatifs situés sur des LRLD, ainsi que des variants à effet épistatiques identifiés par GCTA sur des blocs en LRLD. Par exemple, vérifier si parmi les variants significatifs d'un même caractère, certains sont dans des régions LRLD ce qui laisse penser à une interaction épistatique entre ces régions.

Ce travail de thèse est un point de départ pour des analyses de validation afin d'identifier des mutations causales pour la qualité de la viande et les aptitudes bouchères. En termes d'application, ce travail pourra aider les chercheurs à prendre en considération l'information LRLD lors des études d'association GWAS et la recherche de relations d'épistasie. Ce travail est aussi un point de départ pour des analyses fonctionnelles plus approfondies qui vont sûrement révéler des gènes et des mutations causales associées à la qualité de la viande chez les bovins, et par conséquent aider la filière à atteindre l'objectif principal de satisfaire le consommateur et produire une viande de qualité.

Conclusion Générale et perspectives

Conclusion générale et perspectives

Dans ce manuscrit, nous avons fait dans un premier temps, un point sur l'élevage des bovins allaitants en France ainsi que la production et la consommation de la viande bovine. Nous avons également abordé la question de l'amélioration génétique des races à viande en France en passant par les méthodes de sélection et leurs limites, l'organisation de la filière à l'échelle nationale et européenne et les étapes d'un programme de sélection. Nous avons abordé dans la 3^{ème} partie de cette thèse, les méthodes de détection des QTL chez les bovins ainsi qu'un état des lieux sur les gènes et marqueurs associés à qualité des viandes chez les bovins. La dernière partie bibliographique de cette thèse est consacrée au déséquilibre de liaison dans les populations animales, son importance et ses méthodes de calcul, ainsi qu'un état des lieux sur les études du DL chez les bovins.

Nous avons analysé en détails pour la première fois le déséquilibre de liaison à longue distance dans les 3 principales races bovines allaitantes françaises. Des hypothèses pour expliquer ce LRLD ont été proposées, tel que les goulots d'étranglement, le mélange des populations et la sélection, d'autant que plusieurs phénomènes d'interaction génique ont été identifiés dans ces paires de blocs en LRLD. Les résultats de cette étude et l'approche suivie devraient être pris en compte pour les futures études d'association à l'échelle du génome.

En réalisant des GWAS à l'aide de données de séquence imputées génome complet de plusieurs milliers de bovins issus des 3 principales races à viande (Charolaise, Limousine et Blonde d'Aquitaine), nous avons identifié des milliers de variants significativement associés à la qualité de la viande, aux caractères liés au muscle et aux aptitudes bouchères. Nous avons également identifié une liste de gènes, candidats pour ces caractères étudiés. Ces résultats préliminaires nécessitent d'autres analyses complémentaires, génétiques, moléculaires et *in silico*, telles que des analyses d'association conditionnelles, l'analyse de réseaux d'interaction biologique, l'analyse d'enrichissement et l'analyse des voies métaboliques afin d'identifier des gènes candidats ou des marqueurs utilisables en sélection animale. Ces résultats devront également être réinterprétés, en particulier les variants intergéniques, lorsque de nouvelles informations d'annotation fonctionnelles du génome seront disponibles.

Références Bibliographiques

Références bibliographiques

- Abecasis, G. R., & Cookson, W. O. C. (2000). GOLD - Graphical Overview of Linkage Disequilibrium. *Bioinformatics*, 16(2), 182–183. <https://doi.org/10.1093/bioinformatics/16.2.182>
- Abiola, O., Angel, J. M., Avner, P., Bachmanov, A. A., Belknap, J. K., Bennett, B., ... Consortium, C. T. (2003). The nature and identification of quantitative trait loci: a community's view. *Nature Reviews. Genetics*, 4(11), 911–916. <https://doi.org/10.1038/nrg1206>
- AFNOR. (2002). viande, produits à base de viande et produits de la pêche Détermination de la teneur en L (-) hyd-roxyproline et calcul de la teneur en collagène AFNORV44C.
- Akanno, E. C., Chen, L., Abo-Ismaïl, M. K., Crowley, J. J., Wang, Z., Li, C., ... Plastow, G. S. (2018). Genome-wide association scan for heterotic quantitative trait loci in multi-breed and crossbred beef cattle. *Genetics Selection Evolution*, 50(1), 48. <https://doi.org/10.1186/s12711-018-0405-y>
- Allais, S. (2011a). Détection et validation de marqueurs génétiques impliqués dans la qualité de la viande bovine.
- Allais, S. (2011b). *Détection et validation de marqueurs génétiques impliqués dans la qualité de la viande bovine*. AgroParisTech.
- Allais, S., Journaux, L., Levéziel, H., Payet-Duprat, N., Raynaud, P., Hocquette, J. F., ... Renand, G. (2011a). Effects of polymorphisms in the calpastatin and μ -calpain genes on meat tenderness in 3 French beef breeds. *Journal of Animal Science*, 89(1), 1–11. <https://doi.org/10.2527/jas.2010-3063>
- Allais, S., Journaux, L., Levéziel, H., Payet-Duprat, N., Raynaud, P., Hocquette, J. F., ... Renand, G. (2011b). Effects of polymorphisms in the calpastatin and μ -calpain genes on meat tenderness in 3 French beef breeds1. *Journal of Animal Science*, 89(1), 1–11. <https://doi.org/10.2527/jas.2010-3063>
- Allais, S., Levéziel, H., Hocquette, J. F., Rousset, S., Denoyelle, C., Journaux, L., & Renand, G. (2014). Fine mapping of quantitative trait loci underlying sensory meat quality traits in three French beef cattle breeds. *Journal of Animal Science*, 92(10), 4329–4341. <https://doi.org/10.2527/jas.2014-7868>
- Allais, S., Levéziel, H., Payet-Duprat, N., Hocquette, J. F., Lepetit, J., Rousset, S., ... Renand, G. (2010). The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds. *Journal of Animal Science*, 88(2), 446–454. <https://doi.org/10.2527/jas.2009-2385>
- Allais, S., Levéziel, H., Payet-Duprat, N., Hocquette, J. F., Lepetit, J., Rousset, S., ... Renand, G. (2010). The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds1. *Journal of Animal Science*, 88(2), 446–454. <https://doi.org/10.2527/jas.2009-2385>
- Allan, M. F., Thallman, R. M., Cushman, R. A., Echtenkamp, S. E., White, S. N., Kuehn, L. A., ... Smith, T. P. L. (2007). Association of a single nucleotide polymorphism in SPP1 with growth traits and twinning in a cattle population selected for twinning rate1,2. *Journal of Animal Science*, 85(2), 341–347. <https://doi.org/10.2527/jas.2006-460>
- Anton, I., Kovács, K., Holló, G., Farkas, V., Lehel, L., Hajda, Z., & Zsolnai, A. (2011). Effect of leptin, DGAT1 and TG gene polymorphisms on the intramuscular fat of Angus cattle in Hungary. *Livestock Science*, 135(2–3), 300–303. <https://doi.org/10.1016/j.livsci.2010.07.012>
- Ardlie, K. G., Kruglyak, L., & Seielstad, M. (2002). Patterns of linkage disequilibrium in the human genome. *Nature Reviews Genetics*, 3, 299. Retrieved from <https://doi.org/10.1038/nrg777>
- Azzu, V., & Brand, M. D. (2010). The on-off switches of the mitochondrial uncoupling proteins. *Trends in Biochemical Sciences*, 35(5), 298–307. <https://doi.org/10.1016/j.tibs.2009.11.001>
- Balatsky, V., Bankovska, I., Pena, R. N., Saienko, A., Buslyk, T., Korinnyi, S., & Doran, O.

- (2016). Polymorphisms of the porcine cathepsins, growth hormone-releasing hormone and leptin receptor genes and their association with meat quality traits in Ukrainian Large White breed. *Molecular Biology Reports*, 43(6), 517–526. <https://doi.org/10.1007/s11033-016-3977-z>
- Barendse, W., Bunch, R. J., Harrison, B. E., & Thomas, M. B. (2006). The growth hormone 1 GH1:c.457C>G mutation is associated with intramuscular and rump fat distribution in a large sample of Australian feedlot cattle. *Animal Genetics*, 37(3), 211–214. <https://doi.org/10.1111/j.1365-2052.2006.01432.x>
- Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2), 263–265. <https://doi.org/10.1093/bioinformatics/bth457>
- Bedhane, M., van der Werf, J., Gondro, C., Duijvesteijn, N., Lim, D., Park, B., ... Clark, S. (2019). Genome-Wide Association Study of Meat Quality Traits in Hanwoo Beef Cattle Using Imputed Whole-Genome Sequence Data. *Frontiers in Genetics*, 10(November), 1–13. <https://doi.org/10.3389/fgene.2019.01235>
- Beghain, J., Boitard, S., Weiss, B., Boussaha, M., Gut, I., & Rocha, D. (2013). Genome-wide linkage disequilibrium in the Blonde d'Aquitaine cattle breed. *Journal of Animal Breeding and Genetics*, 130(4), 294–302. <https://doi.org/10.1111/j.1439-0388.2012.01020.x>
- Bhuiyan, M. S. A., Yu, S. L., Jeon, J. T., Yoon, D., Cho, Y. M., Park, E. W., ... Lee, J. H. (2009). DNA polymorphisms in SREBF1 and FASN genes affect fatty acid composition in Korean cattle (Hanwoo). *Asian-Australasian Journal of Animal Sciences*, 22(6), 765–773. <https://doi.org/10.5713/ajas.2009.80573>
- Billerey, C., Boussaha, M., Esquerré, D., Rebours, E., Djari, A., Meersseman, C., ... Rocha, D. (2014). Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing. *BMC Genomics*, 15(1), 499. <https://doi.org/10.1186/1471-2164-15-499>
- Bohmanova, J., Sargolzaei, M., & Schenkel, F. S. (2010). Characteristics of linkage disequilibrium in North American Holsteins. *BMC Genomics*, 11(1), 421. <https://doi.org/10.1186/1471-2164-11-421>
- Boichard, D., Fritz, S., Rossignol, M., Yvonne, M., Malafosse, A., Colleau, J. J., ... Malafosse, A. (2002). Implementation of marker-assisted selection in french dairy cattle. In *7th World Congress on Genetics Applied to Livestock Production, August 19-23, Montpellier, France*.
- Boichard, D., Guillaume, F., Baur, A., Croiseau, P., Rossignol, M. N., Boscher, M. Y., ... Fritz, S. (2012). Genomic selection in French dairy cattle. *Animal Production Science*, 52(2–3), 115–120. <https://doi.org/10.1071/AN11119>
- Boichard, D., Maignel, L., Verrier, E., Boichard, D., Maignel, L., & Verrier, E. (1996). Analyse généalogique des races bovines laitières françaises. *INRA Prod. Anim*, 9, 323–335.
- Bollongino, R., Burger, J., Powell, A., Mashkour, M., Vigne, J. D., & Thomas, M. G. (2012). Modern taurine cattle descended from small number of near-eastern founders. *Molecular Biology and Evolution*, 29(9), 2101–2104. <https://doi.org/10.1093/molbev/mss092>
- BONNET, M., & KOPP, J. (1986). Dosage du collagène dans les tissus conjonctifs de la viande et des produits carnés. *Viandes et Produits Carnés*, 7(6), 263–266.
- Bonny, S. P. F., Gardner, G. E., Pethick, D. W., Allen, P., Legrand, I., Wierzbicki, J., ... Hocquette, J. F. (2017). Untrained consumer assessment of the eating quality of European beef: 2. Demographic factors have only minor effects on consumer scores and willingness to pay. *Animal*, 11(8), 1399–1411. <https://doi.org/10.1017/S1751731117000076>
- Boulesteix, P. (2018). Méthodes et résultats de l'évaluation IBOVAL 2018 pour les races bovines à viande. Retrieved from <http://www4.jouy.inra.fr/gabi>
- Bouquet, A., Renand, G., & Phocas, F. (2009). Evolution de la diversité génétique des populations françaises de bovins allaitants spécialisés de 1979 à 2008. *Productions Animales*, 22(4), 317–330.
- Bult, C. J., Blake, J. A., Smith, C. L., Kadin, J. A., Richardson, J. E., Anagnostopoulos, A., ...

- Zhu, Y. (2019). Mouse Genome Database (MGD) 2019. *Nucleic Acids Research*, 47(D1), D801–D806. <https://doi.org/10.1093/nar/gky1056>
- Calder, P. C., & Yaqoob, P. (2009). Omega-3 polyunsaturated fatty acids and human health outcomes. *BioFactors*, 35(3), 266–272. <https://doi.org/10.1002/biof.42>
- Carbon, S., Douglass, E., Dunn, N., Good, B., Harris, N. L., Lewis, S. E., ... Westerfield, M. (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1), D330–D338. <https://doi.org/10.1093/nar/gky1055>
- Caretti, G., Schiltz, R. L., Dilworth, F. J., Di Padova, M., Zhao, P., Ogryzko, V., ... Sartorelli, V. (2006). The RNA Helicases p68/p72 and the Noncoding RNA SRA Are Coregulators of MyoD and Skeletal Muscle Differentiation. *Developmental Cell*, 11(4), 547–560. <https://doi.org/10.1016/j.devcel.2006.08.003>
- Casas, E., White, S. N., Wheeler, T. L., Shackelford, S. D., Koohmaraie, M., Riley, D. G., ... Smith, T. P. L. (2006). Effects of calpastatin and μ -calpain markers in beef cattle on tenderness traits. *Journal of Animal Science*, 84(3), 520–525. <https://doi.org/10.2527/2006.843520x>
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1). <https://doi.org/10.1186/s13742-015-0047-8>
- Cho, S., Tae, S. P., Yoon, D. H., Hyun, S. C., Namgoong, S., Byung, L. P., ... Hyoung, D. S. (2008). Identification of genetic polymorphisms in FABP3 and FABP4 and putative association with back fat thickness in Korean native cattle. *Journal of Biochemistry and Molecular Biology*, 41(1), 29–34. <https://doi.org/10.5483/bmbrep.2008.41.1.029>
- Choy, M.-K., Javierre, B. M., Williams, S. G., Baross, S. L., Liu, Y., Wingett, S. W., ... Keavney, B. D. (2018). Promoter interactome of human embryonic stem cell-derived cardiomyocytes connects GWAS regions to cardiac gene networks. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-04931-0>
- CIV, & Inrae. (2009). *Valeurs nutritionnelles des viandes (Nutritional values of meats)*. Retrieved from http://www.lessentiellesdesviandes-pro.org/pdf/PDF-tous_morceaux.pdf
- Claudia, S., Josef, H., Walther, V., & Günter, A. (2005). Genetic variability in a genomic region with long-range linkage disequilibrium reveals traces of a bottleneck in the history of the European population. *Human Genetics*, 118, 276–286. <https://doi.org/10.1007/s00439-005-0056-2>
- Craig, D. W., & Stephan, D. A. (2005). Applications of whole-genome high-density SNP genotyping. *Expert Review of Molecular Diagnostics*, 5(2), 159–170. <https://doi.org/10.1586/14737159.5.2.159>
- CREDOC, Tavoularis, S. (2018). Consommation et mode de vie : Les nouvelles générations transforment la consommation de viande. *CREDOC Centre de Recherche Pour l'Étude et l'Observation Des Conditions de Vie*, (300), 1–4. Retrieved from <https://www.credoc.fr/publications/les-nouvelles-generations-transforment-la-consommation-de-viande>
- Daix, M., Pirotte, C., Bister, J. L., Wergifosse, F., Cuvelier, C., Cabaraux, J. F., ... Paquay, R. (2008). Relationship between leptin content, metabolic hormones and fat deposition in three beef cattle breeds. *Veterinary Journal*, 177(2), 273–278. <https://doi.org/10.1016/j.tvjl.2007.04.004>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Davoli, R., Schivazappa, C., Zambonelli, P., Braglia, S., Rossi, A., & Virgili, R. (2017). Association study between single nucleotide polymorphisms in porcine genes and pork quality traits for fresh consumption and processing into Italian dry-cured ham. *Meat Science*, 126, 73–81. <https://doi.org/10.1016/j.meatsci.2016.11.018>
- Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science*, 295(5558), 1306–1311. <https://doi.org/10.1126/science.1067799>
- Di Stasio, L., Destefanis, G., Brugiapaglia, A., Albera, A., & Rolando, A. (2005). Polymorphism of the GHR gene in cattle and relationships with meat production and

- quality. *Animal Genetics*, 36(2), 138–140. <https://doi.org/10.1111/j.1365-2052.2005.01244.x>
- Donadelli, M., Dando, I., Fiorini, C., & Palmieri, M. (2014). UCP2, a mitochondrial protein regulated at multiple levels. *Cellular and Molecular Life Sciences*, 71(7), 1171–1190. <https://doi.org/10.1007/s00018-013-1407-0>
- Doyle, J. L., Berry, D. P., Veerkamp, R. F., Carthy, T. R., Evans, R. D., Walsh, S. W., & Purfield, D. C. (2020). Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds. *Genetics Selection Evolution*, 52(1). <https://doi.org/10.1186/s12711-020-0523-1>
- Drinkwater, R. D., Li, Y., Lenane, I., Davis, G. P., Shorthose, R., Harrison, B. E., ... Barendse, W. (2006). Detecting quantitative trait loci affecting beef tenderness on bovine chromosome 7 near calpastatin and lysyl oxidase. *Australian Journal of Experimental Agriculture*, 46(2), 159–164. Retrieved from <https://doi.org/10.1071/EA05185>
- Dudouet, C. (2010). *La production des bovins allaitants: Conduite, Qualité, Gestion*. (E. F. Agricole, Ed.) (3e édition).
- Dujková, R., Ranganathan, Y., Dufek, A., Macák, J., & Bezdíček, J. (2015). Polymorphic effects of FABP4 and SCD genes on intramuscular fatty acid profiles in longissimus muscle from two cattle breeds. *Acta Veterinaria Brno*, 84(4), 327–336. <https://doi.org/10.2754/avb201584040327>
- El Hou, A., Rocha, D., Venot, E., Blanquet, V., & Philippe, R. (2021). Long-range linkage disequilibrium in French beef cattle breeds. *Genetics Selection Evolution*, 53(1), 63. <https://doi.org/10.1186/s12711-021-00657-8>
- Elango, R., Ball, R. O., & Pencharz, P. B. (2009). Amino acid requirements in humans: With a special emphasis on the metabolic availability of amino acids. *Amino Acids*, 37(1), 19–27. <https://doi.org/10.1007/s00726-009-0234-y>
- Elsen, J. M. (2011). Sélection sur gènes et sélection assistée par marqueurs. *INRA Prod. Anim*, 24(4), 357–362.
- Enriquez-Valencia, C. E., Pereira, G. L., Malheiros, J. M., de Vasconcelos Silva, J. A. I. I., Albuquerque, L. G., de Oliveira, H. N., ... Curi, R. A. (2017). Effect of the g.98535683A>G SNP in the CAST gene on meat traits of Nellore beef cattle (*Bos indicus*) and their crosses with *Bos taurus*. *Meat Science*, 123, 64–66. <https://doi.org/10.1016/j.meatsci.2016.09.003>
- Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, 1, 117693430500100. <https://doi.org/10.1177/117693430500100003>
- Fang, L., Cai, W., Liu, S., Canela-Xandri, O., Gao, Y., Jiang, J., ... Liu, G. E. (2020). Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Research*, 30(5), 790–801. <https://doi.org/10.1101/gr.250704.119>
- Farnir, F., Coppieiers, W., Arranz, J. J., Berzi, P., Cambisano, N., Grisart, B., ... Georges, M. (2000). Extensive genome-wide linkage disequilibrium in cattle. *Genome Research*, 10(2), 220–227. <https://doi.org/10.1101/gr.10.2.220>
- France AgriMer. (2019). Consommation des produits carnés en 2018, 35. Retrieved from <http://www.franceagrimer.fr/content/download/40104/372599/file/STA-VIA-CONSO-2014-aout2015.pdf>
- Gagaoua, M., & Picard, B. (2020). Current Advances in Meat Nutritional, Sensory and Physical Quality Improvement. *Foods*, 9(3), 321. <https://doi.org/10.3390/foods9030321>
- GEB, I. de L. (2019). *Les chiffres clés du GEB, Bovins 2019. Production lait et viande*.
- Gerald Wiener, R. R. (2009). *L'amélioration génétique animale* (Quae).
- Girardot, M., Guibert, S., Laforet, M. P., Gallard, Y., Larroque, H., & Oulmouden, A. (2006). The insertion of a full-length *Bos taurus* LINE element is responsible for a transcriptional deregulation of the Normande Agouti gene. *Pigment Cell Research*, 19(4), 346–355. <https://doi.org/10.1111/j.1600-0749.2006.00312.x>
- González-Prendes, R., Mármol-Sánchez, E., Quintanilla, R., Castelló, A., Zidi, A., Ramayo-

- Caldas, Y., ... Amills, M. (2019). About the existence of common determinants of gene expression in the porcine liver and skeletal muscle. *BMC Genomics*, 20(1), 518. <https://doi.org/10.1186/s12864-019-5889-5>
- Gorlov, I. F., Fedunin, A. A., Randelin, D. A., & Sulimova, G. E. (2014). Polymorphisms of bGH, RORC, and DGAT1 Genes, 50(12), 1302–1307. <https://doi.org/10.1134/S1022795414120035>
- Griffon, L., Boulesteix, P., Delpeuch, A., Govignon-Gion, A., Guerrier, J., Leudet, O., ... Tribout, T. (2017). La sélection génétique des races bovines allaitantes en France: Un dispositif et des outils innovants au service des filières viande. *Productions Animales*, 30(2), 107–124. <https://doi.org/10.20870/productions-animales.2017.30.2.2237>
- Grobet, L., Martin, L. J., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., ... Georges, M. (1997). A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nature Genetics*, 17(1), 71–74. <https://doi.org/10.1038/ng0997-71>
- Gudmundsson, J., Sulem, P., Gudbjartsson, D. F., Blondal, T., Gylfason, A., Agnarsson, B. A., ... Stefansson, K. (2009). Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nature Genetics*, 41(10), 1122–1126. <https://doi.org/10.1038/ng.448>
- Guillaume, F., Boichard, D., Ducrocq, V., Fritz, S., Guillaume, F., Boichard, D., ... De, U. (2011). Utilisation de la sélection génomique chez les bovins laitiers To cite this version : Utilisation de la sélection génomique chez les bovins laitiers. *Innovative Food Science & Emerging Technologies*, 4(24), 363–368.
- Guillaume, F., Fritz, S., Boichard, D., Druet, T., Guillaume, F., Fritz, S., ... Druet, T. (2008). Estimation by simulation of the efficiency of the French marker-assisted selection program in dairy cattle. *Genetics Selection Evolution*, 40, 91–102.
- Gunderson, K. L., Steemers, F. J., Lee, G., Mendoza, L. G., & Chee, M. S. (2005). A genome-wide scalable SNP genotyping assay using microarray technology. *Nature Genetics*, 37(5), 549–554. <https://doi.org/10.1038/ng1547>
- Haley, C. S., Knott, S. A., & Elsen, J. M. (1994). Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics*, 136(3), 1195–1207.
- Hedrick, P., & Thomson, G. (1986). A two-locus neutrality test: application to humans, *E. coli* and Lodgepole pine. *Genetics*, 112, 135–156.
- Henderson, C. R. (1963). Selection Index and Expected Genetic Advance. *Statistical Genetics and Plant Breeding*, 982(NAS-NRC).
- Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics*, 31(2), 423–447.
- Higgs, J. D. (2000). The changing nature of red meat: 20 years of improving nutritional quality. *Trends in Food Science and Technology*, 11(3), 85–95. [https://doi.org/10.1016/S0924-2244\(00\)00055-8](https://doi.org/10.1016/S0924-2244(00)00055-8)
- Hill, W., & Roberston, A. (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, 38, 226–231.
- Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, 6(2), 95–108. <https://doi.org/10.1038/nrg1521>
- Hocquette, J., Lehnert, S., Barendse, W., & Picard, B. (2007). Recent advances in cattle functional genomics and their application to beef quality, 159–173. <https://doi.org/10.1017/S1751731107658042>
- Hou, G.-Y., Yuan, Z.-R., Zhou, H.-L., Zhang, L.-P., Li, J.-Y., Gao, X., ... Xu, S.-Z. (2011). Association of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. *Molecular Biology Reports*, 38(7), 4705–4708. <https://doi.org/10.1007/s11033-010-0605-1>
- Hu, Z. L., Park, C. A., & Reecy, J. M. (2019). Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Research*, 47(D1), D701–D710. <https://doi.org/10.1093/nar/gky1084>
- Institut de L'élevage. (2018). *Indicateurs des flux de reproducteurs Bovins Allaitants* -

compagne de naissance 2018.

- Jacinto, J. G. P., Häfliger, I. M., Veiga, I. M. B., Letko, A., Benazzi, C., Bolcato, M., & Drögemüller, C. (2020). A Heterozygous Missense Variant in the COL5A2 in Holstein Cattle Resembling the Classical Ehlers–Danlos Syndrome. *Animals*, 10(11), 2002. <https://doi.org/10.3390/ani10112002>
- Jennings, H. (1917). The numerical result of diverse systems of breedings with respect to two pairs of characters. *Genetics*, 2, 97–106.
- Kaplanová, K., Dufek, A., Dračková, E., Simeonovová, J., Šubrt, J., Vrtková, I., & Dvořák, J. (2013). The association of CAPN1, CAST, SCD, and FASN polymorphisms with beef quality traits in commercial crossbred cattle in the czech republic. *Czech Journal of Animal Science*, 58(11), 489–496. <https://doi.org/10.17221/7044-cjas>
- Kaur, L., Hui, S. X., & Boland, M. (2020). Changes in Cathepsin Activity during Low-Temperature Storage and Sous Vide Processing of Beef Brisket. *Food Science of Animal Resources*, 40(3), 415–425. <https://doi.org/10.5851/kosfa.2020.e21>
- Kaur, L., Hui, S. X., Morton, J. D., Kaur, R., Chian, F. M., & Boland, M. (2021). Endogenous Proteolytic Systems and Meat Tenderness: Influence of Post-Mortem Storage and Processing. *Food Science of Animal Resources*, 41(4), 589–607. <https://doi.org/10.5851/kosfa.2021.e27>
- Kern, C., Wang, Y., Xu, X., Pan, Z., Halstead, M., Chanthavixay, G., ... Zhou, H. (2021). Functional annotations of three domestic animal genomes provide vital resources for comparative and agricultural research. *Nature Communications*, 12(1), 1821. <https://doi.org/10.1038/s41467-021-22100-8>
- Kerry, J. P., & Ledward, D. (2009). *Improving the sensory and nutritional quality of fresh meat*. (J. Kerry, Ed.) (1st Editio).
- Khamis, A. M., Motwalli, O., Oliva, R., Jankovic, B. R., Medvedeva, Y. A., Ashoor, H., ... Bajic, V. B. (2018). A novel method for improved accuracy of transcription factor binding site prediction. *Nucleic Acids Research*, 46(12). <https://doi.org/10.1093/nar/gky237>
- Khatkar, M. S., Collins, A., Cavanagh, J. A. L., Hawken, R. J., Hobbs, M., Zenger, K. R., ... Raadsma, H. W. (2006). A first-generation metric linkage disequilibrium map of bovine chromosome 6. *Genetics*, 174(1), 79–85. <https://doi.org/10.1534/genetics.106.060418>
- Khatkar, M. S., Nicholas, F. W., Collins, A. R., Zenger, K. R., Cavanagh, J. A. L., Barris, W., ... Raadsma, H. W. (2008). Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics*, 9. <https://doi.org/10.1186/1471-2164-9-187>
- Knott, S. A., Elsen, J. M., & Haley, C. S. (1996). Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theoretical and Applied Genetics*, 93(1–2), 71–80. <https://doi.org/10.1007/BF00225729>
- KNÜPPEL, R., DIETZE, P., LEHNBERG, W., FRECH, K., & WINGENDER, E. (1994). TRANSFAC Retrieval Program: A Network Model Database of Eukaryotic Transcription Regulating Sequences and Proteins. *Journal of Computational Biology*, 1(3), 191–198. <https://doi.org/10.1089/cmb.1994.1.191>
- Koch, E., Ristroph, M., & Kirkpatrick, M. (2013). Long Range Linkage Disequilibrium across the Human Genome. *PLOS ONE*, 8(12), 1–10. <https://doi.org/10.1371/journal.pone.0080754>
- Koohmaraie, M. (1996). Biochemical factors regulating the toughening and tenderization processes of meat. *Meat Science*, 43S1, 193–201. [https://doi.org/10.1016/0309-1740\(96\)00065-4](https://doi.org/10.1016/0309-1740(96)00065-4)
- Korostowski, L., Sedlak, N., & Engel, N. (2012). The Kcnq1ot1 Long Non-Coding RNA Affects Chromatin Conformation and Expression of Kcnq1, but Does Not Regulate Its Imprinting in the Developing Heart. *PLoS Genetics*, 8(9), e1002956. <https://doi.org/10.1371/journal.pgen.1002956>
- Krauss, S., Zhang, C. Y., & Lowell, B. B. (2005). The mitochondrial uncoupling-protein homologues. *Nature Reviews Molecular Cell Biology*, 6(3), 248–261. <https://doi.org/10.1038/nrm1592>
- Le Roy, P., & Elsen, J. M. (2000). Principes de l'utilisation des marqueurs génétiques pour la

- détection des gènes influençant les caractères quantitatifs. *Productions Animales, October*(SPEC.ISS.), 211–215. <https://doi.org/10.20870/productions-animales.2000.13.hs.3840>
- Leal-Gutiérrez, J. D., Rezende, F. M., Reecy, J. M., Kramer, L. M., Peñagaricano, F., & Mateescu, R. G. (2020a). Whole Genome Sequence Data Provides Novel Insights Into the Genetic Architecture of Meat Quality Traits in Beef. *Frontiers in Genetics, 11*(September), 1–18. <https://doi.org/10.3389/fgene.2020.538640>
- Leal-Gutiérrez, J. D., Rezende, F. M., Reecy, J. M., Kramer, L. M., Peñagaricano, F., & Mateescu, R. G. (2020b). Whole Genome Sequence Data Provides Novel Insights Into the Genetic Architecture of Meat Quality Traits in Beef. *Frontiers in Genetics, 11*. <https://doi.org/10.3389/fgene.2020.538640>
- Lee, H. J., Jin, S., Kim, H.-J., Bhuiyan, M. S. A., Lee, D. H., Lee, S. H., ... Lee, S. H. (2019a). Validation Study of SNPs in CAPN1-CAST Genes on the Tenderness of Muscles (Longissimus thoracis and Semimembranosus) in Hanwoo (Korean Cattle). *Animals, 9*(9), 691. <https://doi.org/10.3390/ani9090691>
- Lee, H. J., Jin, S., Kim, H. J., Bhuiyan, M. S. A., Lee, D. H., Lee, S. H., ... Lee, S. H. (2019b). Validation study of SNPs in CAPN1-CAST genes on the tenderness of muscles (Longissimus thoracis and semimembranosus) in Hanwoo (Korean cattle). *Animals, 9*(9), 1–8. <https://doi.org/10.3390/ani9090691>
- Lee, S.-H., Kim, S.-C., Chai, H.-H., Cho, S.-H., Kim, H.-C., Lim, D., ... Hong, S.-K. (2014). Mutations in calpastatin and μ -calpain are associated with meat tenderness, flavor and juiciness in Hanwoo (Korean cattle): Molecular modeling of the effects of substitutions in the calpastatin/ μ -calpain complex. *Meat Science, 96*(4), 1501–1508. <https://doi.org/10.1016/j.meatsci.2013.11.026>
- Lee, S. H., Kim, S. C., Chai, H. H., Cho, S. H., Kim, H. C., Lim, D., ... Hong, S. K. (2014). Mutations in calpastatin and μ -calpain are associated with meat tenderness, flavor and juiciness in Hanwoo (Korean cattle): Molecular modeling of the effects of substitutions in the calpastatin/ μ -calpain complex. *Meat Science, 96*(4), 1501–1508. <https://doi.org/10.1016/j.meatsci.2013.11.026>
- Leidy. (2011). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Bone, 23*(1), 1–7. <https://doi.org/10.1038/ng.2354>
- Lettice, L. A., Heaney, S. J. H., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., ... de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human Molecular Genetics, 12*(14), 1725–1735. <https://doi.org/10.1093/hmg/ddg180>
- Lewontin, R. C. (1964). The interaction of selection and linkage. I. General considerations: heterotic models. *Genetics, 49*(1), 49–67. Retrieved from <https://www.genetics.org/content/49/1/49>
- Li, X., Ekerljung, M., Lundström, K., & Lundén, A. (2013). Association of polymorphisms at DGAT1, leptin, SCD1, CAPN1 and CAST genes with color, marbling and water holding capacity in meat from beef cattle populations in Sweden. *Meat Science, 94*(2), 153–158. <https://doi.org/10.1016/j.meatsci.2013.01.010>
- Li, Y. X., Jin, H. G., Yan, C. G., Seo, K. S., Zhang, L. C., Ren, C. Y., & Jin, X. (2013). Association of CAST gene polymorphisms with carcass and meat quality traits in Yanbian cattle of China. *Molecular Biology Reports, 40*(2), 1875–1881. <https://doi.org/10.1007/s11033-012-2243-2>
- Lim, D., Strucken, E. M., Choi, B. H., Chai, H. H., Cho, Y. M., Jang, G. W., ... Lee, S. H. (2016a). Genomic Footprints in Selected and Unselected Beef Cattle Breeds in Korea. *PloS One, 11*(3), e0151324. <https://doi.org/10.1371/journal.pone.0151324>
- Lim, D., Strucken, E. M., Choi, B. H., Chai, H. H., Cho, Y. M., Jang, G. W., ... Lee, S. H. (2016b). Genomic Footprints in Selected and Unselected Beef Cattle Breeds in Korea. *PLOS ONE, 11*(3), e0151324. <https://doi.org/10.1371/journal.pone.0151324>
- Lines, D. S., Pitchford, W. S., Kruk, Z. A., & Bottema, C. D. K. (2009). Limousin myostatin F94L variant affects semitendinosus tenderness. *Meat Science, 81*(1), 126–131. <https://doi.org/10.1016/j.meatsci.2008.07.007>

- Liu, Y., Qin, X., Song, X. Z. H., Jiang, H., Shen, Y., Durbin, K. J., ... Gibbs, R. A. (2009). Bos taurus genome assembly. *BMC Genomics*, *10*, 1–11. <https://doi.org/10.1186/1471-2164-10-180>
- Livak, K. J. (1999). Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genetic Analysis - Biomolecular Engineering*, *14*(5–6), 143–149. [https://doi.org/10.1016/S1050-3862\(98\)00019-9](https://doi.org/10.1016/S1050-3862(98)00019-9)
- Lu, D., Sargolzaei, M., Kelly, M., Li, C., Voort, G. Vander, Wang, Z., ... Miller, S. P. (2012). Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Frontiers in Genetics*, *3*(AUG), 1–10. <https://doi.org/10.3389/fgene.2012.00152>
- Luo, Z. W., & Kearsley, M. J. (1989). Maximum likelihood estimation of linkage between a marker gene and a quantitative locus. *Heredity*, *63*(3), 401–408. <https://doi.org/10.1038/hdy.1989.114>
- Ma, N., Zhu, Z., Liu, J., Peng, Y., Zhao, X., Tang, W., ... Du, J. (2021). Clinical and genetic analysis of classical Ehlers-Danlos syndrome patient caused by synonymous mutation in COL5A2. *Molecular Genetics & Genomic Medicine*, *9*(5). <https://doi.org/10.1002/mgg3.1632>
- Malheiros, J. M., Enríquez-Valencia, C. E., da Silva Duran, B. O., de Paula, T. G., Curi, R. A., de Vasconcelos Silva, J. A. I. I., ... Chardulo, L. A. L. (2018). Association of CAST2, HSP90AA1, DNAJA1 and HSPB1 genes with meat tenderness in Nellore cattle. *Meat Science*, *138*(January), 49–52. <https://doi.org/10.1016/j.meatsci.2018.01.003>
- Matsuhashi, T., Maruyama, S., Uemoto, Y., Kobayashi, N., Mannen, H., Abe, T., ... Kobayashi, E. (2011a). Effects of bovine fatty acid synthase, stearyl-coenzyme A desaturase, sterol regulatory element-binding protein 1, and growth hormone gene polymorphisms on fatty acid composition and carcass traits in Japanese Black cattle. *Journal of Animal Science*, *89*(1), 12–22. <https://doi.org/10.2527/jas.2010.3121>
- Matsuhashi, T., Maruyama, S., Uemoto, Y., Kobayashi, N., Mannen, H., Abe, T., ... Kobayashi, E. (2011b). Effects of bovine fatty acid synthase, stearyl-coenzyme A desaturase, sterol regulatory element-binding protein 1, and growth hormone gene polymorphisms on fatty acid composition and carcass traits in Japanese Black cattle. *Journal of Animal Science*, *89*(1), 12–22. <https://doi.org/10.2527/jas.2010.3121>
- Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., ... Van Tassell, C. P. (2009). Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *PLOS ONE*, *4*(4), 1–13. <https://doi.org/10.1371/journal.pone.0005350>
- McKay, S. D., Schnabel, R. D., Murdoch, B. M., Matukumalli, L. K., Aerts, J., Coppieters, W., ... Moore, S. S. (2007). Whole genome linkage disequilibrium maps in cattle. *BMC Genetics*, *8*, 1–12. <https://doi.org/10.1186/1471-2156-8-74>
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., ... Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 1–14. <https://doi.org/10.1186/s13059-016-0974-4>
- Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). [Meuwissen_et al2001Genetics1819.pdf](https://doi.org/10.1186/1471-2156-2-24). <https://doi.org/10.1186/1471-2156-2-24>
- Michal, J. J., Zhang, Z. W., Gaskins, C. T., & Jiang, Z. (2006). The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. *Animal Genetics*, *37*(4), 400–402. <https://doi.org/10.1111/j.1365-2052.2006.01464.x>
- Mokry, F., Buzanskas, M., de Alvarenga, M., do Amaral, G., Higa, R., Ventura, R., ... de Almeida Regitano, L. (2014). Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics*, *15 Suppl 7*(Suppl 7), S6--S6. <https://doi.org/10.1186/1471-2164-15-S7-S6>
- Monin, G. (1991). Facteurs biologiques des qualités de la viande bovine. *Productions Animales* *2* (4), 151-160. (1991), *4*(2), 151–160.
- Muir, P. D., Deaker, J. M., & Bown, M. D. (1998). Effects of forage- and grain-based feeding systems on beef quality: A review. *New Zealand Journal of Agricultural Research*, *41*(4), 623–635. <https://doi.org/10.1080/00288233.1998.9513346>

- Mwangi, F. W., Charmley, E., Gardiner, C. P., Malau-Aduli, B. S., Kinobe, R. T., & Malau-Aduli, A. E. O. (2019). Diet and genetics influence beef cattle performance and meat quality characteristics. *Foods*, 8(12), 1–24. <https://doi.org/10.3390/foods8120648>
- Nei, M., & Li, W.-H. (1973). LINKAGE DISEQUILIBRIUM IN SUBDIVIDED POPULATIONS. *Genetics*, 75(1), 213–219. Retrieved from <https://www.genetics.org/content/75/1/213>
- Newton, K. G., & Gill, C. O. (1981). The microbiology of DFD fresh meats: A review. *Meat Science*, 5(3), 223–232. [https://doi.org/10.1016/0309-1740\(81\)90005-X](https://doi.org/10.1016/0309-1740(81)90005-X)
- Page, B. T., Casas, E., Heaton, M. P., Cullen, N. G., Hyndman, D. L., Morris, C. A., ... Smith, T. P. L. (2002). Evaluation of single-nucleotide polymorphisms in CAPN1 for association with meat tenderness in cattle. *Journal of Animal Science*, 80(12), 3077–3085. <https://doi.org/10.2527/2002.80123077x>
- Pareek, C. S., Czarnik, U., Pierzchała, M., & Zwierzchowski, L. (2008). An association between the C> T single nucleotide polymorphism within intron IV of osteopontin encoding gene (SPP1) and body weight of growing Polish Holstein-Friesian cattle. *Animal Science Papers and Reports*, 26(4), 251–257.
- Park, L. (2019). Population-specific long-range linkage disequilibrium in the human genome and its influence on identifying common disease variants. *Scientific Reports*, 9(1), 1–13. <https://doi.org/10.1038/s41598-019-47832-y>
- Payet N., Malafosse A., Renand G., Menissier F., H. J. (2006). Le projet Qualvigène : mise en évidence du polymorphisme des gènes impliqués dans les qualités de la viande en races Charolaise , Limousine et Blonde d ' Aquitaine. *Renc. Rech. Ruminants*, (1), 264.
- Pereira, P. M. de C. C., & Vicente, A. F. dos R. B. (2013). Meat nutritional composition and nutritive role in the human diet. *Meat Science*, 93(3), 586–592. <https://doi.org/10.1016/j.meatsci.2012.09.018>
- Phocas, F., Bloch, C., Chapelle, P., Bécherel, F., Renand, G., & Ménissier, F. (1998). Developing a breeding objective for a French purebred beef cattle selection programme. *Livestock Production Science*, 57(1), 49–65. [https://doi.org/10.1016/S0301-6226\(98\)00157-2](https://doi.org/10.1016/S0301-6226(98)00157-2)
- Pierre Sellier, Didier Boichard, E. V. (2019). “La génétique animale à l’INRA”, Histoire de la recherche contemporaine, Tome VIII-n°1 (pp. 86–97).
- Ramayo-Caldas, Y., Renand, G., Ballester, M., Saintilan, R., & Rocha, D. (2016). Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. *Genetics Selection Evolution*, 48(1), 1–9. <https://doi.org/10.1186/s12711-016-0216-y>
- Randhawa, I. A. S., Khatkar, M. S., Thomson, P. C., & Raadsma, H. W. (2015). Composite Selection Signals for Complex Traits Exemplified Through Bovine Stature Using Multibreed Cohorts of European and African Bos taurus. *G3: Genes, Genomes, Genetics*, 5(7), 1391–1401. <https://doi.org/10.1534/g3.115.017772>
- Rebai, A., Goffinet, B., & Mangin, B. (1995). Comparing Power of Different Methods for QTL Detection. *Biometrics*, 51(1), 87–99.
- Rempel, L. A., Casas, E., Shackelford, S. D., & Wheeler, T. L. (2012). Relationship of polymorphisms within metabolic genes and carcass traits in crossbred beef cattle. *Journal of Animal Science*, 90(4), 1311–1316. <https://doi.org/10.2527/jas.2011-4302>
- Ritchie, M. D. (2015). Finding the Epistasis Needles in the Genome-Wide Haystack (pp. 19–33). https://doi.org/10.1007/978-1-4939-2155-3_2
- Robbins, R. (1918). Some applications of mathematics to breeding problems III. *Genetics*, 3, 375–389.
- Ron, G., Globerson, Y., Moran, D., & Kaplan, T. (2017). Promoter-enhancer interactions identified from Hi-C data using probabilistic models and hierarchical topological domains. *Nature Communications*, 8(1). <https://doi.org/10.1038/s41467-017-02386-3>
- Roy, R., Taourit, S., Zaragoza, P., Eggen, A., & Rodellar, C. (2005). Genomic structure and alternative transcript of bovine fatty acid synthase gene (FASN): Comparative analysis of the FASNgene between monogastric and ruminant species. *Cytogenetic and Genome Research*, 111(1), 65–73. <https://doi.org/10.1159/000085672>
- Ruiz-Larrañaga, O., Langa, J., Rendo, F., Manzano, C., Iriando, M., & Estonba, A. (2018).

- Genomic selection signatures in sheep from the Western Pyrenees. *Genetics Selection Evolution*, 50(1). <https://doi.org/10.1186/s12711-018-0378-x>
- Russo, G. L. (2009). Dietary n - 6 and n - 3 polyunsaturated fatty acids: From biochemistry to clinical implications in cardiovascular prevention. *Biochemical Pharmacology*, 77(6), 937–946. <https://doi.org/10.1016/j.bcp.2008.10.020>
- Ryu, J., & Lee, C. (2016). Genetic association of marbling score with intragenic nucleotide variants at selection signals of the bovine genome. *Animal*, 10(4), 566–570. <https://doi.org/10.1017/S1751731115002633>
- Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., ... Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mammalian Genome*, 15(1), 23–34. <https://doi.org/10.1007/s00335-033-2317-5>
- Sanchez, M. P., Guatteo, R., Davergne, A., Saout, J., Grohs, C., Deloche, M. C., ... Boichard, D. (2020). Identification of the ABCC4, IER3, and CBFA2T2 candidate genes for resistance to paratuberculosis from sequence-based GWAS in Holstein and Normande dairy cattle. *Genetics Selection Evolution*, 52(1), 1–17. <https://doi.org/10.1186/s12711-020-00535-9>
- Sanchez, M. P., Ramayo-Caldas, Y., Wolf, V., Laithier, C., El Jabri, M., Michenet, A., ... Boichard, D. (2019). Sequence-based GWAS, network and pathway analyses reveal genes co-associated with milk cheese-making properties and milk composition in Montbéliarde cows. *Genetics Selection Evolution*, 51(1), 1–19. <https://doi.org/10.1186/s12711-019-0473-7>
- Sandelin, A., Alkema, W., Engström, P., Wasserman, W. W., & Lenhard, B. (2004). JASPAR: An open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Research*, 32(DATABASE ISS.). <https://doi.org/10.1093/nar/gkh012>
- Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, 15(1). <https://doi.org/10.1186/1471-2164-15-478>
- Sati, S., & Cavalli, G. (2017). Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma*, 126(1), 33–44. <https://doi.org/10.1007/s00412-016-0593-6>
- Schenkel, F. S., Miller, S. P., Jiang, Z., Mandell, I. B., Ye, X., Li, H., & Wilton, J. W. (2006). Association of a single nucleotide polymorphism in the calpastatin gene with carcass and meat quality traits of beef cattle. *Journal of Animal Science*, 84(2), 291–299. <https://doi.org/10.2527/2006.842291x>
- Schenkel, F. S., Miller, S. P., Ye, X., Moore, S. S., Nkrumah, J. D., Li, C., ... Williams, J. L. (2005). Association of single nucleotide polymorphisms in the leptin gene with carcass and meat quality traits of beef cattle. *Journal of Animal Science*, 83(9), 2009–2020. <https://doi.org/10.2527/2005.8392009x>
- Serão, N. V. L., González-Peña, D., Beever, J. E., Bollero, G. A., Southey, B. R., Faulkner, D. B., & Rodriguez-Zas, S. L. (2013). Bivariate genome-wide association analysis of the growth and intake components of feed efficiency. *PLoS One*, 8(10), e78530. <https://doi.org/10.1371/journal.pone.0078530>
- Sevane, N., Armstrong, E., Wiener, P., Pong Wong, R., & Dunner, S. (2014). Polymorphisms in twelve candidate genes are associated with growth, muscle lipid profile and meat quality traits in eleven European cattle breeds. *Molecular Biology Reports*, 41(7), 4721–4731. <https://doi.org/10.1007/s11033-014-3343-y>
- Shin, S. C., & Chung, E. R. (2007). Association of SNP marker in the leptin gene with carcass and meat quality traits in Korean cattle. *Asian-Australasian Journal of Animal Sciences*, 20(1), 1–6. <https://doi.org/10.5713/ajas.2007.1>
- Singh, A., Gill, G., Kaur, H., Amhmed, M., & Jakhu, H. (2018). Role of osteopontin in bone remodeling and orthodontic tooth movement: a review. *Progress in Orthodontics*, 19(1), 18. <https://doi.org/10.1186/s40510-018-0216-2>
- Smith, S. B., Gill, C. A., Lunt, D. K., & Brooks, M. A. (2009). Regulation of fat and fatty acid composition in beef cattle. *Asian-Australasian Journal of Animal Sciences*, 22(9), 1225–

1233. <https://doi.org/10.5713/ajas.2009.r.10>
- Soller, M., Brody, T., & Genizi, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theoretical and Applied Genetics*, 47(1), 35–39. <https://doi.org/10.1007/BF00277402>
- Soller, M., & Genizi, A. (1978). The Efficiency of Experimental Designs for the Detection of Linkage between a Marker Locus and a Locus Affecting a Quantitative Trait in Segregating Populations. *Biometrics*, 34(1), 47. <https://doi.org/10.2307/2529587>
- Song, Y., Xu, L., Chen, Y., Zhang, L., Gao, H., Zhu, B., ... Li, J. (2016). Genome-wide association study reveals the PLAG1 gene for Knuckle, Biceps and Shank weight in Simmental beef cattle. *PLoS ONE*, 11(12), 1–10. <https://doi.org/10.1371/journal.pone.0168316>
- Špehar, M., Vincek, D., & Žgur, S. (2008). Beef Quality: Factors Affecting Tenderness and Marbling. *Stočarstvo : Časopis Za Unapređenje Stočarstva*, 62(6), 463–478.
- Stella, R., Biancotto, G., Krogh, M., Angeletti, R., Pozza, G., Sorgato, M. C., ... Andrighetto, I. (2011a). Protein Expression Changes in Skeletal Muscle in Response to Growth Promoter Abuse in Beef Cattle. *Journal of Proteome Research*, 10(6), 2744–2757. <https://doi.org/10.1021/pr101255c>
- Stella, R., Biancotto, G., Krogh, M., Angeletti, R., Pozza, G., Sorgato, M. C., ... Andrighetto, I. (2011b). Protein Expression Changes in Skeletal Muscle in Response to Growth Promoter Abuse in Beef Cattle. *Journal of Proteome Research*, 10(6), 2744–2757. <https://doi.org/10.1021/pr101255c>
- Sun, X., Wu, X., Fan, Y., Mao, Y., Ji, D., Huang, B., & Yang, Z. (2018). Effects of polymorphisms in CAPN1 and CAST genes on meat tenderness of Chinese Simmental cattle. *Archives Animal Breeding*, 61(4), 433–439. <https://doi.org/10.5194/aab-61-433-2018>
- Taniguchi, M., Utsugi, T., Oyama, K., Mannen, H., Kobayashi, M., Tanabe, Y., ... Tsuji, S. (2004). Genotype of stearoyl-CoA desaturase is associated with fatty acid composition in Japanese Black cattle. *Mammalian Genome*, 15(2), 142–148. <https://doi.org/10.1007/s00335-003-2286-8>
- Tian, J., Zhao, Z., Zhang, L., Zhang, Q., Yu, Z., Li, J., & Yang, R. (2013). Association of the leptin gene E2-169T>C and E3-299T>A mutations with carcass and meat quality traits of the Chinese Simmental-cross steers. *Gene*, 518(2), 443–448. <https://doi.org/10.1016/j.gene.2012.11.071>
- Uto, Y. (2016). Recent progress in the discovery and development of stearoyl CoA desaturase inhibitors. *Chemistry and Physics of Lipids*, 197, 3–12. <https://doi.org/10.1016/j.chemphyslip.2015.08.018>
- Vallejo, R. L., Li, Y. L., Rogers, G. W., & Ashwell, M. S. (2003). Genetic diversity and background linkage disequilibrium in the North American Holstein cattle population. *Journal of Dairy Science*, 86(12), 4137–4147. [https://doi.org/10.3168/jds.S0022-0302\(03\)74028-4](https://doi.org/10.3168/jds.S0022-0302(03)74028-4)
- Vallejo, R. L., Silva, R. M. O., Evenhuis, J. P., Gao, G., Liu, S., Parsons, J. E., ... Palti, Y. (2018). Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels: Evidence that long-range LD is a major contributing factor. *Journal of Animal Breeding and Genetics*, 135(4), 263–274. <https://doi.org/10.1111/jbg.12335>
- Valsta, L. M., Tapanainen, H., & Männistö, S. (2005). Meat fats in nutrition. *Meat Science*, 70(3 SPEC. ISS.), 525–530. <https://doi.org/10.1016/j.meatsci.2004.12.016>
- Van Binsbergen, R., Bink, M. C. A. M., Calus, M. P. L., Van Eeuwijk, F. A., Hayes, B. J., Hulsege, I., & Veerkamp, R. F. (2014). Accuracy of imputation to whole-genome sequence data in Holstein Friesian cattle. *Genetics Selection Evolution*, 46(1), 1–13. <https://doi.org/10.1186/1297-9686-46-41>
- Vermunt, M. W., Zhang, D., & Blobel, G. A. (2019). The interdependence of gene-regulatory elements and the 3D genome. *Journal of Cell Biology*, 218(1), 12–26. <https://doi.org/10.1083/jcb.201809040>

- Wais, M. K. (2012). *Méthodes statistiques pour la détection de QTL : nouveaux développements et applications chez le canard mulard*. AgroParis- Tech,.
- Wang, L., Zhang, L., Yan, H., Liu, X., Li, N., Liang, J., ... Wang, L. (2014). Genome-wide association studies identify the loci for 5 exterior traits in a large white x Minzhu pig population. *PLoS ONE*, *9*(8). <https://doi.org/10.1371/journal.pone.0103766>
- Wang, Y., Yang, W., Gui, L., Wang, H., & Zan, L. (2016). Association and expression analyses of the Ucp2 and Ucp3 gene polymorphisms with body measurement and meat quality traits in Qinchuan cattle. *Journal of Genetics*, *95*(4), 939–946. <https://doi.org/10.1007/s12041-016-0720-4>
- Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., ... Li, C. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics*, *21*(1), 38. <https://doi.org/10.1186/s12864-019-6273-1>
- Warner, R. D., Dunshea, F. R., Ponnampalam, E. N., & Cottrell, J. J. (2005). Effects of nitric oxide and oxidation in vivo and postmortem on meat tenderness. *Meat Science*, *71*(1), 205–217. <https://doi.org/10.1016/j.meatsci.2005.04.008>
- Watanabe, M., Nakagawa, R., Naruto, T., Kohmoto, T., Suga, K., Goji, A., ... Imoto, I. (2016). A novel missense mutation of COL5A2 in a patient with Ehlers–Danlos syndrome. *Human Genome Variation*, *3*(1), 16030. <https://doi.org/10.1038/hgv.2016.30>
- Westerling, D. B., & Hedrick, H. B. (1979). Fatty Acid Composition of Bovine Lipids as Influenced by Diet, Sex and Anatomical Location and Relationship to Sensory Characteristics. *Journal of Animal Science*, *48*(6), 1343–1348. <https://doi.org/10.2527/jas1979.4861343x>
- Williams, P. (2007). Nutritional composition of red meat. *Nutrition and Dietetics*, *64*(SUPPL. 4), 5–7. <https://doi.org/10.1111/j.1747-0080.2007.00197.x>
- Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Richardson, R. I., & Sheard, P. R. (1999). 1999_Manipulating meat quality&composition_Wood et al. *READ*, *44*(July 1998), 363–370.
- Wu, G. (2009). Amino acids: Metabolism, functions, and nutrition. *Amino Acids*, *37*(1), 1–17. <https://doi.org/10.1007/s00726-009-0269-0>
- Wu, X. X., Yang, Z. P., Shi, X. K., Li, J. Y., Ji, D. J., Mao, Y. J., ... Gao, H. J. (2012). Association of SCD1 and DGAT1 SNPs with the intramuscular fat traits in Chinese Simmental cattle and their distribution in eight Chinese cattle breeds. *Molecular Biology Reports*, *39*(2), 1065–1071. <https://doi.org/10.1007/s11033-011-0832-0>
- Wyness, L. (2016). The role of red meat in the diet: Nutrition and health benefits. *Proceedings of the Nutrition Society*, *75*(3), 227–232. <https://doi.org/10.1017/S0029665115004267>
- Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Madden, P. A. F., Heath, A. C., ... Visscher, P. M. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics*, *44*(4), 369–375. <https://doi.org/10.1038/ng.2213>
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., ... Li, C. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: feed efficiency and component traits. *BMC Genomics*, *21*(1), 1–22. <https://doi.org/10.1186/s12864-019-6362-1>
- Zhang, L., Liu, J., Zhao, F., Ren, H., Xu, L., Lu, J., ... Du, L. (2013). Genome-Wide Association Studies for Growth and Meat Production Traits in Sheep. *PLoS ONE*, *8*(6). <https://doi.org/10.1371/journal.pone.0066569>
- Zhang, M., Hill, M. C., Kadow, Z. A., Suh, J. H., Tucker, N. R., Hall, A. W., ... Martin, J. F. (2019). Long-range Pitx2c enhancer promoter interactions prevent predisposition to atrial fibrillation. *Proceedings of the National Academy of Sciences*, *116*(45), 22692–22698. <https://doi.org/10.1073/pnas.1907418116>

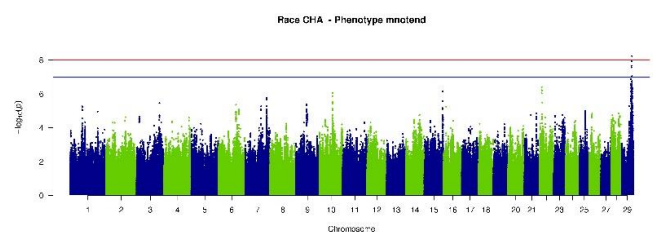
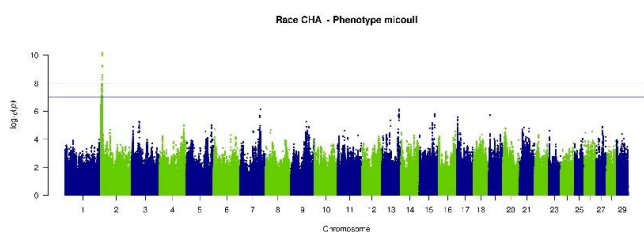
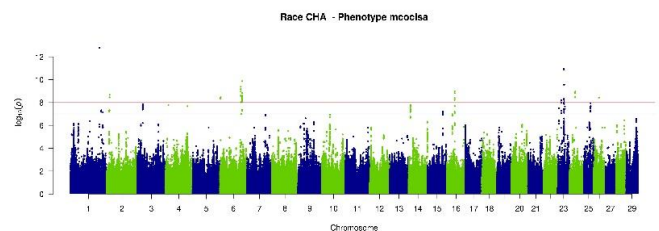
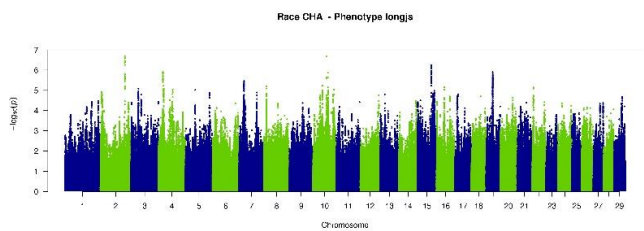
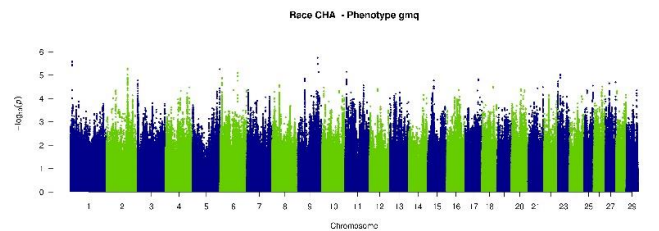
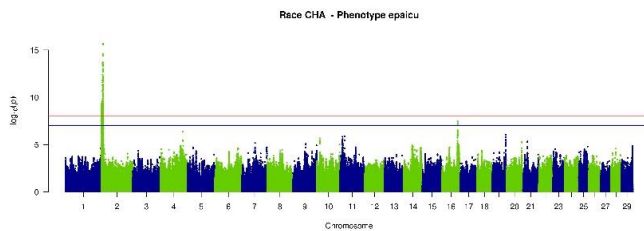
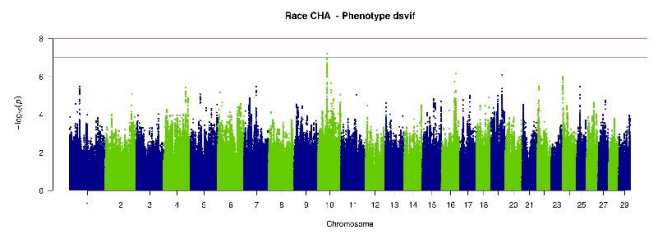
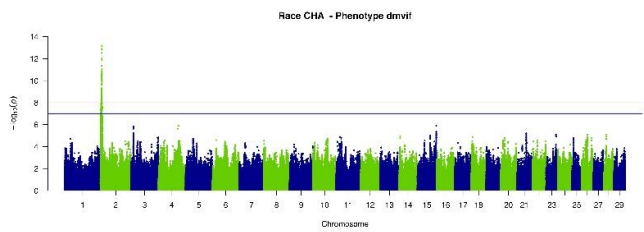
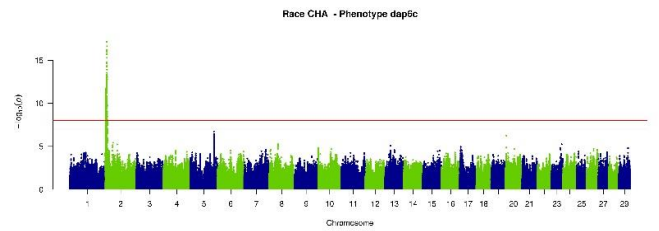
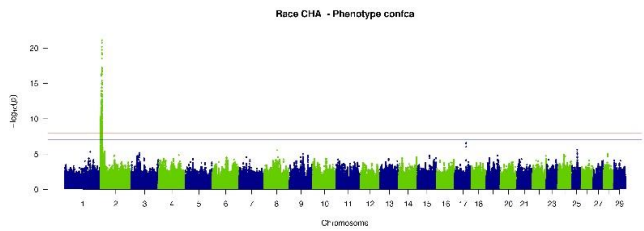
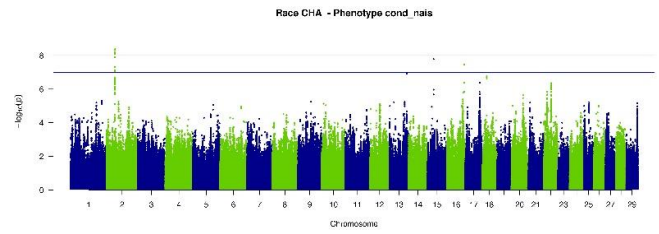
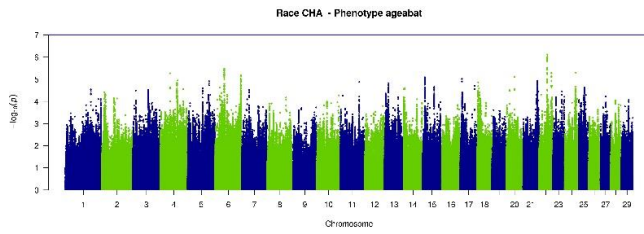
- Zhang, S., Knight, T. J., Reecy, J. M., & Beitz, D. C. (2008). DNA polymorphisms in bovine fatty acid synthase are associated with beef fatty acid composition. *Animal Genetics*, 39(1), 62–70. <https://doi.org/10.1111/j.1365-2052.2007.01681.x>
- Zhao, H., Nettleton, D., & Dekkers, J. C. M. (2007). Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms. *Genetical Research*, 89(1), 1–6. <https://doi.org/10.1017/S0016672307008634>
- Zhao, L., Wang, J., Li, Y., Song, T., Wu, Y., Fang, S., ... He, S. (2021). NONCODEV6: An updated database dedicated to long non-coding RNA annotation in both animals and plants. *Nucleic Acids Research*, 49(D1), D165–D171. <https://doi.org/10.1093/nar/gkaa1046>
- Zhu, B., Niu, H., Zhang, W., Wang, Z., Liang, Y., Guan, L., ... Li, J. (2017). Genome wide association study and genomic prediction for fatty acid composition in Chinese Simmental beef cattle using high density SNP array. *BMC Genomics*, 18(1), 1–15. <https://doi.org/10.1186/s12864-017-3847-7>

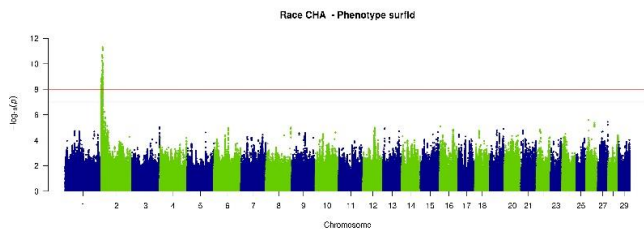
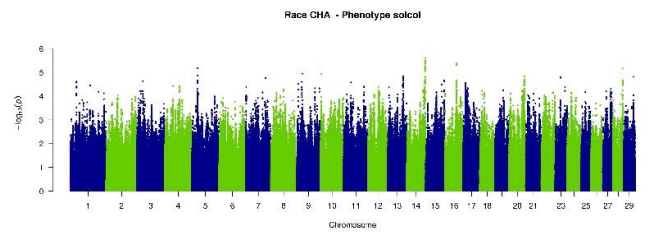
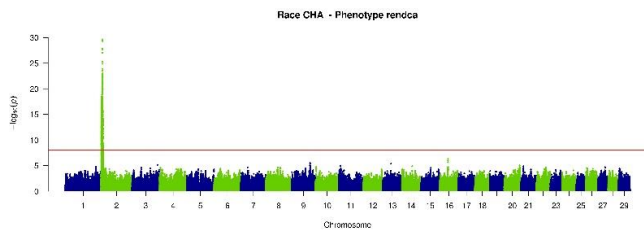
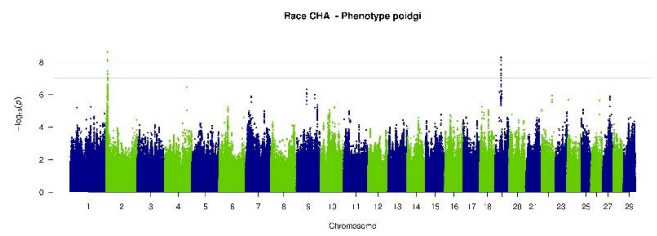
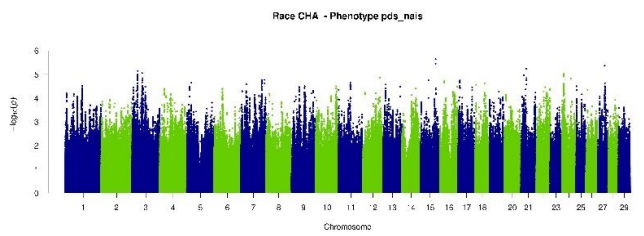
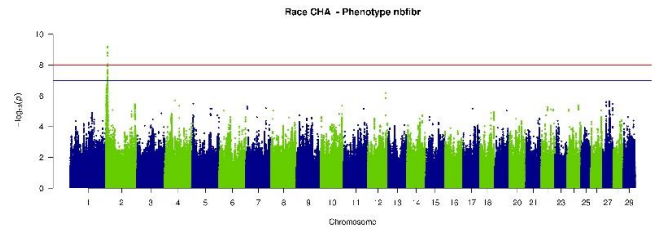
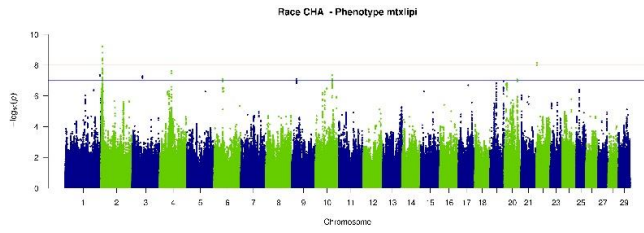
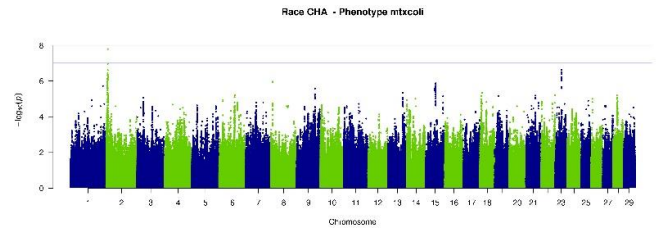
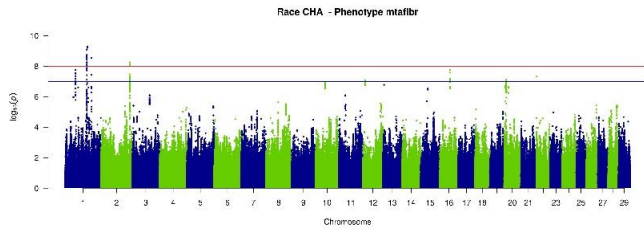
Annexes

Annexe 1. Manhattan plot des variants associés aux caractères étudiés dans chaque race	120
Annexe 2. Liste des variants chevauchants avec des lncRNA de la base de données NONCODE pour la race Charolaise (CHA), Limousine (LIM) et Blonde d'Aquitaine (BLA).	126
Annexe 3. Liste des variants chevauchants avec les QTL de la base de données CattleQTLdb.	127

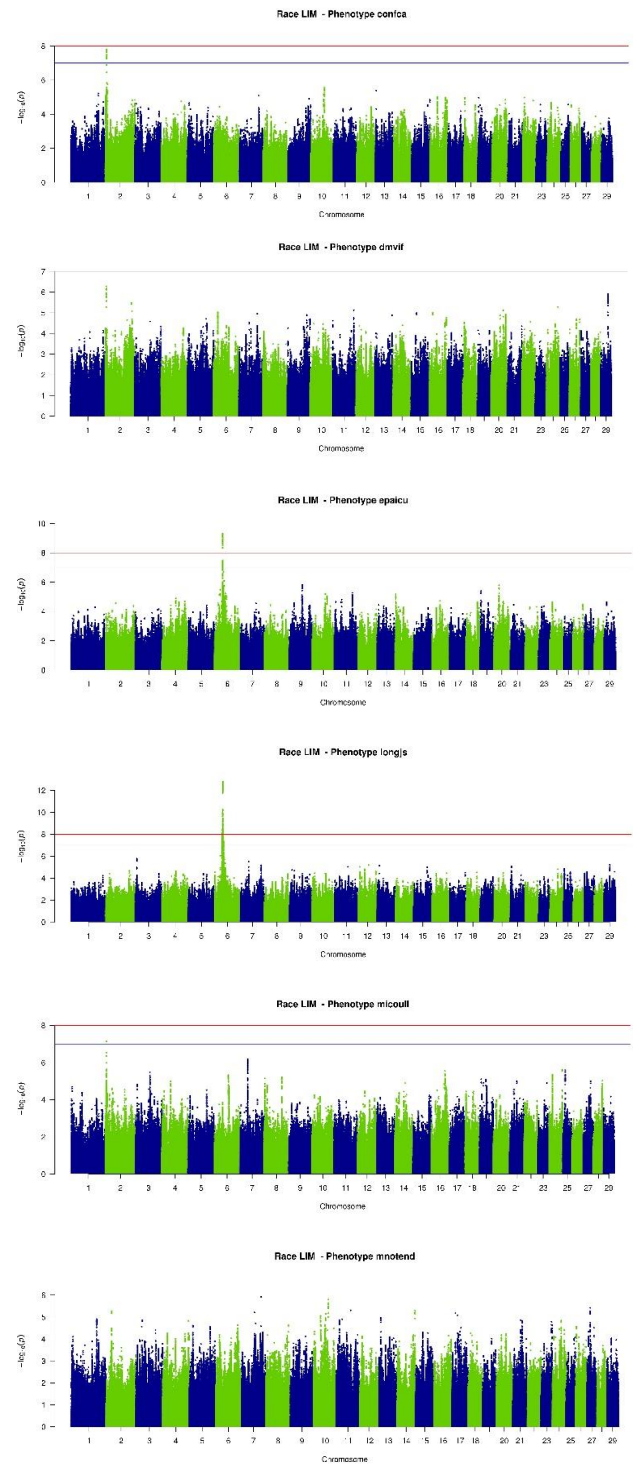
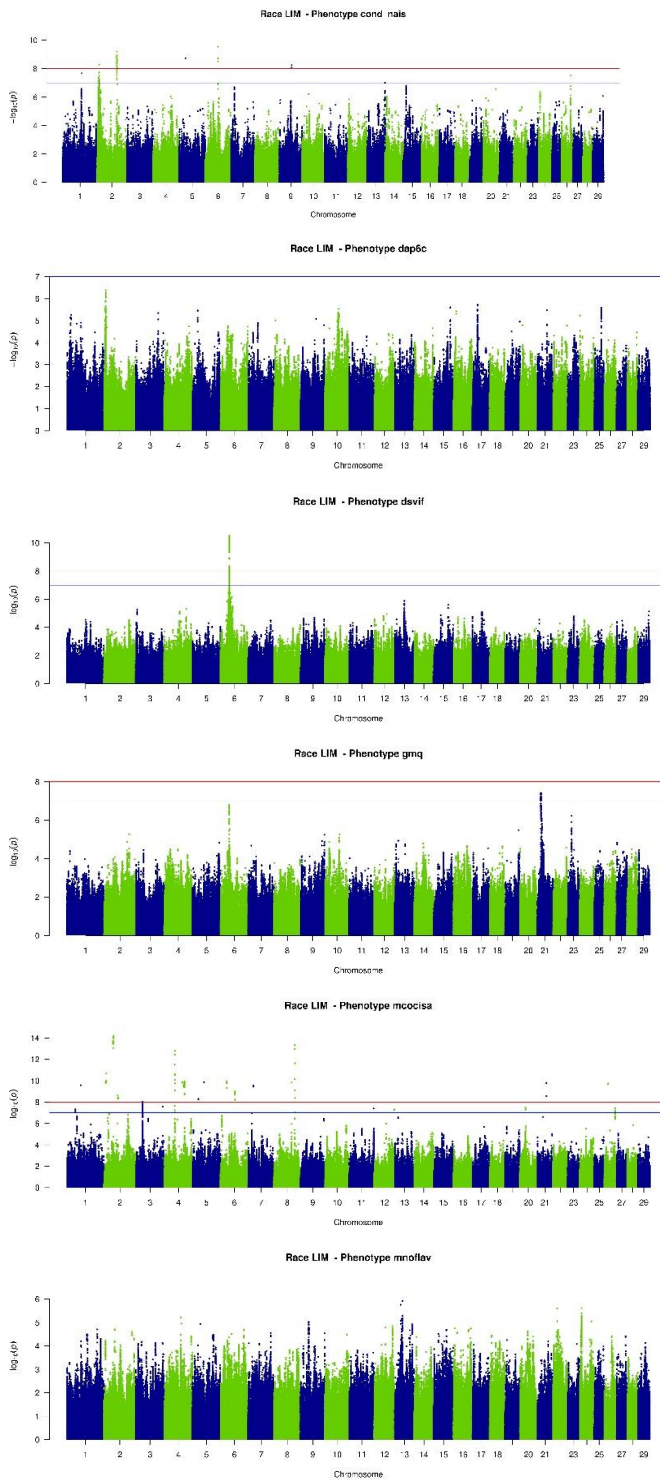
Annexe 1. Manhattan plot des variants associés aux caractères étudiés dans chaque race

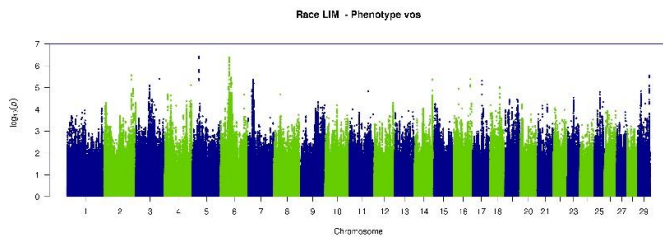
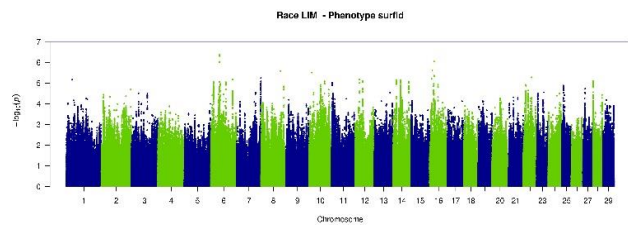
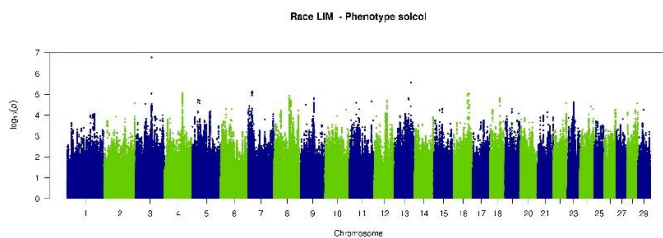
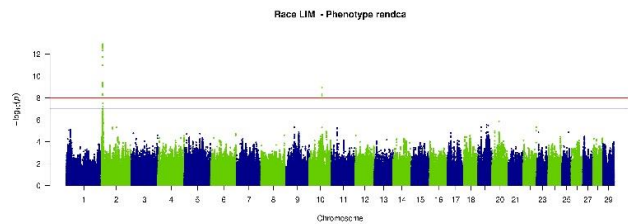
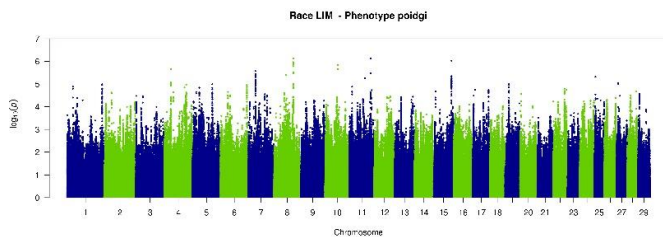
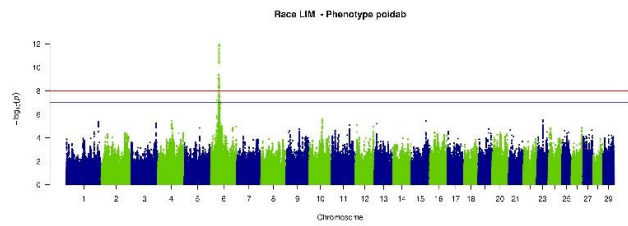
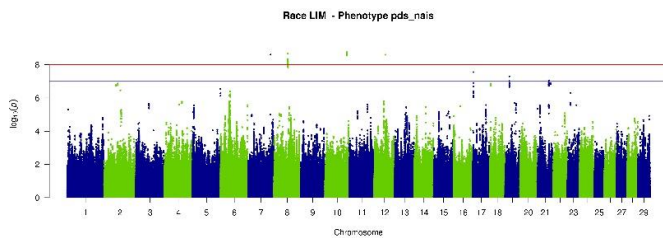
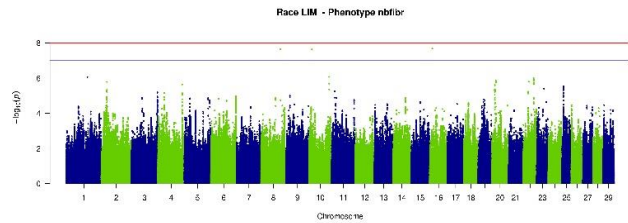
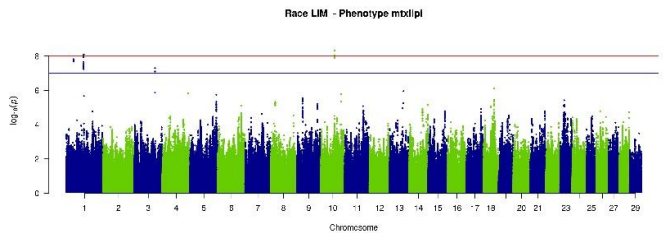
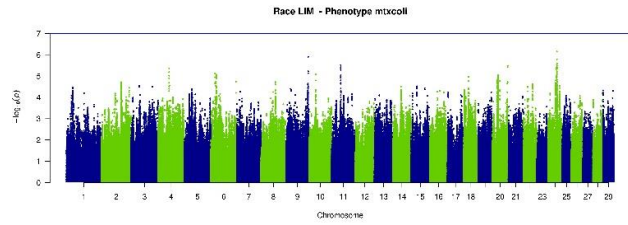
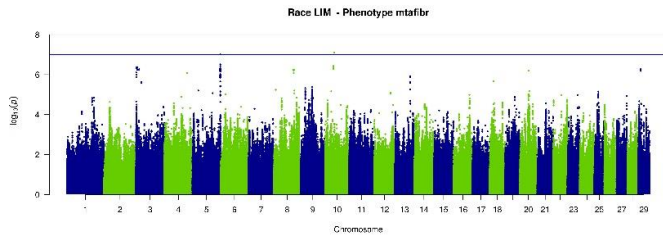
Race Charolaise :



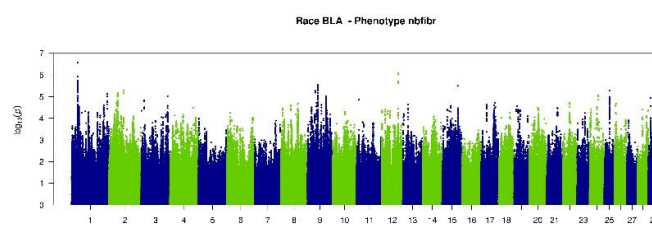
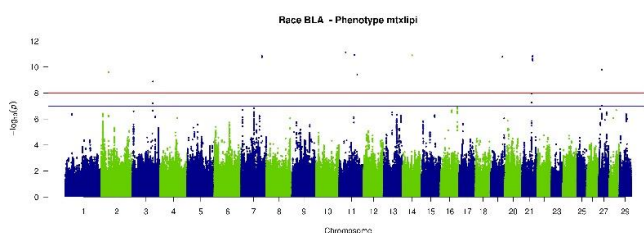
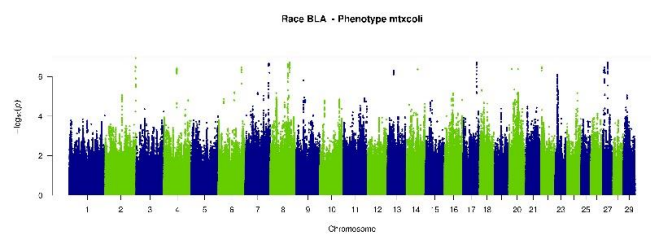
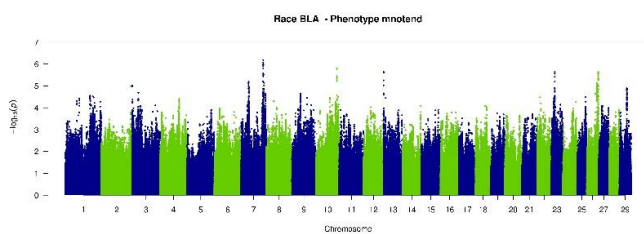
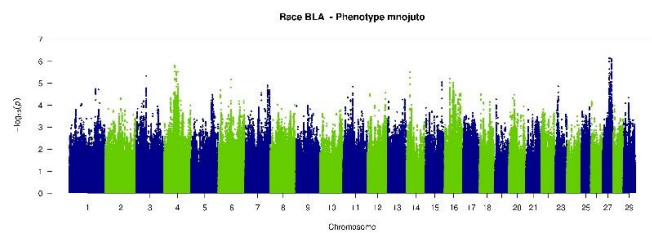
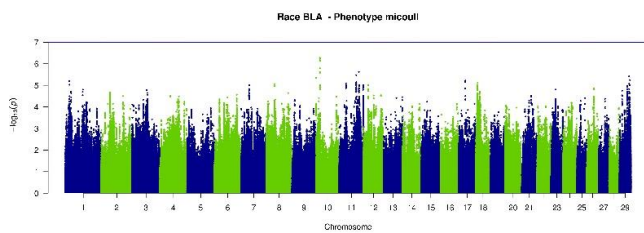
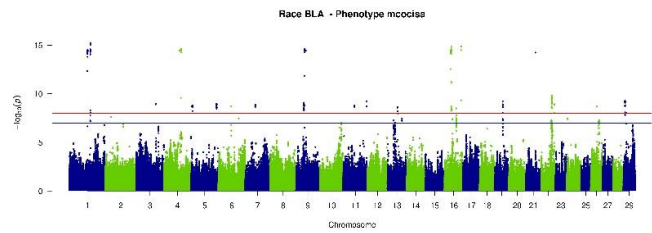
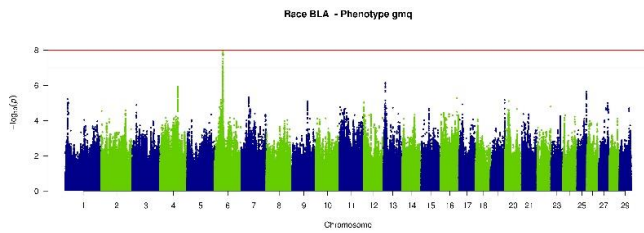
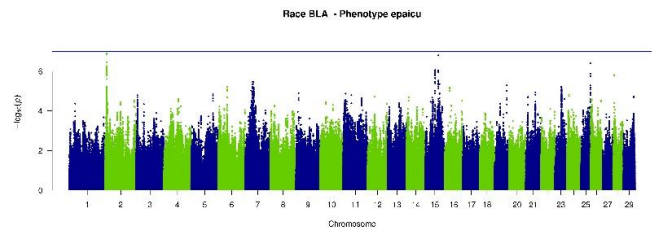
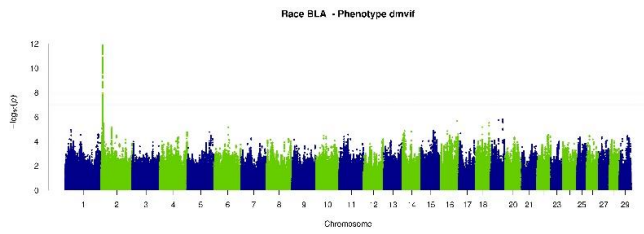
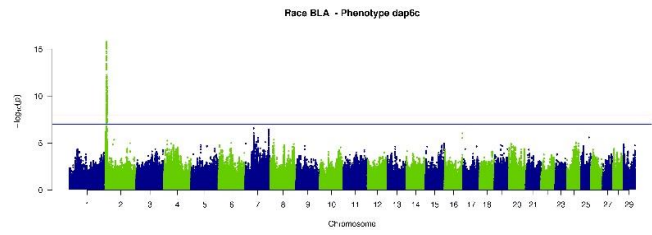
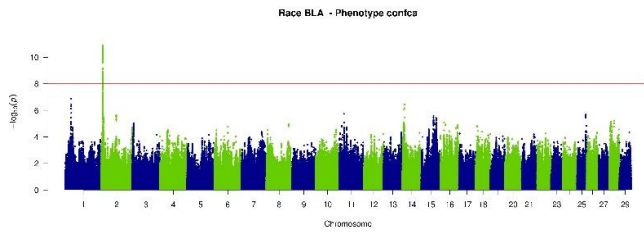


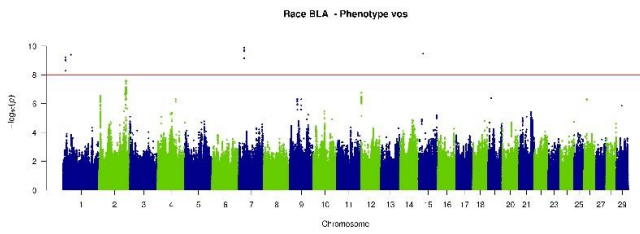
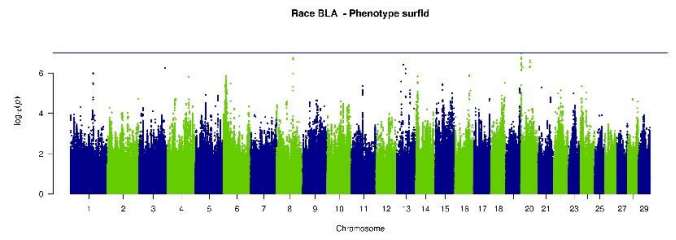
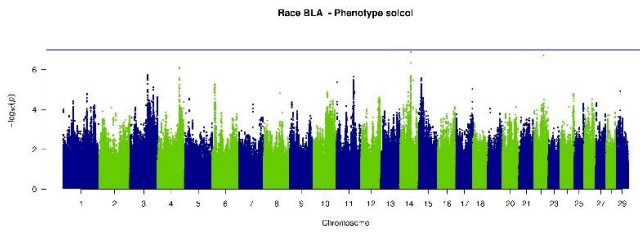
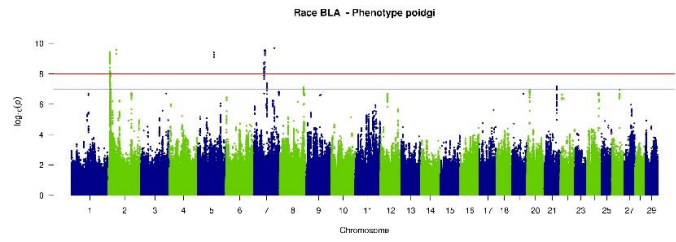
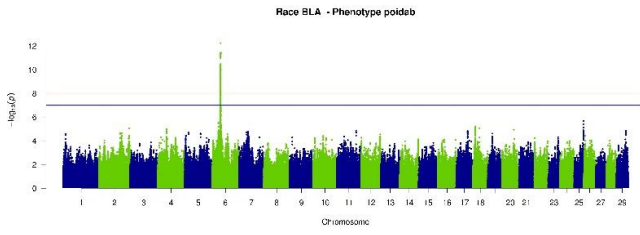
Race Limousine :





Race Blonde Aquitaine :





Annexe 2. Liste des variants chevauchants avec des lncRNA de la base de données NONCODE pour la race Charolaise (CHA), Limousine (LIM) et Blonde d'Aquitaine (BLA).

Race	Caractère	rsID	BTA	Position umd3	lncRNA	début	fin
BLA	dap6c	rs209356402	2	5563494	NONBTAG007064.2	5563321	5564080
BLA	rendca	rs209356402	2	5563494	NONBTAG007064.2	5563321	5564080
BLA	dap6c	rs210827020	2	5563785	NONBTAG007064.2	5563321	5564080
BLA	rendca	rs210827020	2	5563785	NONBTAG007064.2	5563321	5564080
BLA	confca	rs110802172	2	6037069	NONBTAG007069.2	6036020	6037511
BLA	dap6c	rs110802172	2	6037069	NONBTAG007069.2	6036020	6037511
BLA	dmvif	rs110802172	2	6037069	NONBTAG007069.2	6036020	6037511
BLA	poidgi	rs110802172	2	6037069	NONBTAG007069.2	6036020	6037511
BLA	rendca	rs110802172	2	6037069	NONBTAG007069.2	6036020	6037511
CHA	dap6c	rs381441534	2	1164109	NONBTAG007041.2	1163222	1164121
CHA	epaicu	rs381441534	2	1164109	NONBTAG007041.2	1163222	1164121
CHA	rendca	rs381441534	2	1164109	NONBTAG007041.2	1163222	1164121
CHA	rendca	rs384223252	2	4225175	NONBTAG007045.2	4224993	4225753
CHA	rendca	rs381471162	2	4333937	NONBTAG019933.1	4333917	4334560
CHA	rendca	rs383615000	2	4334158	NONBTAG019933.1	4333917	4334560
LIM	dsvif	rs378955293	6	37871185	NONBTAG021702.1	37871113	37872270
LIM	epaicu	rs378955293	6	37871185	NONBTAG021702.1	37871113	37872270
LIM	longjs	rs378955293	6	37871185	NONBTAG021702.1	37871113	37872270
LIM	poidab	rs378955293	6	37871185	NONBTAG021702.1	37871113	37872270
LIM	dsvif	rs383473904	6	37872592	NONBTAG021703.1	37872465	37873310
LIM	epaicu	rs383473904	6	37872592	NONBTAG021703.1	37872465	37873310
LIM	longjs	rs383473904	6	37872592	NONBTAG021703.1	37872465	37873310
LIM	poidab	rs383473904	6	37872592	NONBTAG021703.1	37872465	37873310
LIM	dsvif	rs380209716	6	37888602	NONBTAG014729.2	37888168	37888720
LIM	epaicu	rs380209716	6	37888602	NONBTAG014729.2	37888168	37888720
LIM	longjs	rs380209716	6	37888602	NONBTAG014729.2	37888168	37888720
LIM	poidab	rs380209716	6	37888602	NONBTAG014729.2	37888168	37888720

Dap6c : dépôt adipeux, **rendca** : rendement en carcasse, **confca** : conformation de carcasse, **dmvif** : développement musculaire, **poidgi** : poids de gras interne, **epaicu** : épaisseur de la cuisse, **dsvif** : développement squelettique, **longjs** : longueur jarret-symphise, **poidab** : poids à l'abattage

Annexe 3. Liste des variants chevauchants avec les QTL de la base de données CattleQTLdb.

race	carac	rsID	bta	pos_ars	Pval	distance	start	end	QTL_ID
CHA	surfld	rs382702657	2	6274134	9,3412E-09	473219	5800911	5800915	192847
CHA	surfld	rs382702657	2	6274134	9,3412E-09	475232	5798898	5798902	192856
CHA	surfld	rs382702657	2	6274134	9,3412E-09	25695	6248435	6248439	164702
CHA	surfld	rs382702657	2	6274134	9,3412E-09	15754	6258376	6258380	192586
CHA	surfld	rs382702657	2	6274134	9,3412E-09	21071	6253059	6253063	192652
CHA	surfld	rs382702657	2	6274134	9,3412E-09	4328	6269802	6269806	192617
CHA	surfld	rs382702657	2	6274134	9,3412E-09	1277	6275411	6275415	192144
CHA	surfld	rs382702657	2	6274134	9,3412E-09	443067	6717201	6717205	192613
CHA	surfld	rs382702657	2	6274134	9,3412E-09	443412	6717546	6717550	192752
CHA	surfld	rs382702657	2	6274134	9,3412E-09	458819	6732953	6732957	192744
CHA	surfld	rs382702657	2	6274134	9,3412E-09	458833	6732967	6732971	192793
CHA	surfld	rs382702657	2	6274134	9,3412E-09	458273	6732407	6732411	192007
CHA	surfld	rs378744284	2	6253061	1,594E-09	452146	5800911	5800915	192847
CHA	surfld	rs378744284	2	6253061	1,594E-09	454159	5798898	5798902	192856
CHA	surfld	rs378744284	2	6253061	1,594E-09	4622	6248435	6248439	164702
CHA	surfld	rs378744284	2	6253061	1,594E-09	5315	6258376	6258380	192586
CHA	surfld	rs378744284	2	6253061	1,594E-09	2	6253059	6253063	192652
CHA	surfld	rs378744284	2	6253061	1,594E-09	16741	6269802	6269806	192617
CHA	surfld	rs378744284	2	6253061	1,594E-09	22350	6275411	6275415	192144
CHA	surfld	rs378744284	2	6253061	1,594E-09	464140	6717201	6717205	192613
CHA	surfld	rs378744284	2	6253061	1,594E-09	464485	6717546	6717550	192752
CHA	surfld	rs378744284	2	6253061	1,594E-09	479892	6732953	6732957	192744
CHA	surfld	rs378744284	2	6253061	1,594E-09	479906	6732967	6732971	192793
CHA	surfld	rs378744284	2	6253061	1,594E-09	479346	6732407	6732411	192007
CHA	surfld	rs209833633	2	6264230	3,9131E-09	463315	5800911	5800915	192847
CHA	surfld	rs209833633	2	6264230	3,9131E-09	465328	5798898	5798902	192856
CHA	surfld	rs209833633	2	6264230	3,9131E-09	15791	6248435	6248439	164702
CHA	surfld	rs209833633	2	6264230	3,9131E-09	5850	6258376	6258380	192586
CHA	surfld	rs209833633	2	6264230	3,9131E-09	11167	6253059	6253063	192652
CHA	surfld	rs209833633	2	6264230	3,9131E-09	5572	6269802	6269806	192617
CHA	surfld	rs209833633	2	6264230	3,9131E-09	11181	6275411	6275415	192144
CHA	surfld	rs209833633	2	6264230	3,9131E-09	452971	6717201	6717205	192613
CHA	surfld	rs209833633	2	6264230	3,9131E-09	453316	6717546	6717550	192752
CHA	surfld	rs209833633	2	6264230	3,9131E-09	468723	6732953	6732957	192744
CHA	surfld	rs209833633	2	6264230	3,9131E-09	468737	6732967	6732971	192793
CHA	surfld	rs209833633	2	6264230	3,9131E-09	468177	6732407	6732411	192007
CHA	surfld	rs379161703	2	6264428	4,3066E-09	463513	5800911	5800915	192847
CHA	surfld	rs379161703	2	6264428	4,3066E-09	465526	5798898	5798902	192856
CHA	surfld	rs379161703	2	6264428	4,3066E-09	15989	6248435	6248439	164702
CHA	surfld	rs379161703	2	6264428	4,3066E-09	6048	6258376	6258380	192586
CHA	surfld	rs379161703	2	6264428	4,3066E-09	11365	6253059	6253063	192652
CHA	surfld	rs379161703	2	6264428	4,3066E-09	5374	6269802	6269806	192617

CHA	surflid	rs379161703	2	6264428	4,3066E-09	10983	6275411	6275415	192144
CHA	surflid	rs379161703	2	6264428	4,3066E-09	452773	6717201	6717205	192613
CHA	surflid	rs379161703	2	6264428	4,3066E-09	453118	6717546	6717550	192752
CHA	surflid	rs379161703	2	6264428	4,3066E-09	468525	6732953	6732957	192744
CHA	surflid	rs379161703	2	6264428	4,3066E-09	468539	6732967	6732971	192793
CHA	surflid	rs379161703	2	6264428	4,3066E-09	467979	6732407	6732411	192007
CHA	surflid	rs378569611	2	6265327	3,9078E-09	464412	5800911	5800915	192847
CHA	surflid	rs378569611	2	6265327	3,9078E-09	466425	5798898	5798902	192856
CHA	surflid	rs378569611	2	6265327	3,9078E-09	16888	6248435	6248439	164702
CHA	surflid	rs378569611	2	6265327	3,9078E-09	6947	6258376	6258380	192586
CHA	surflid	rs378569611	2	6265327	3,9078E-09	12264	6253059	6253063	192652
CHA	surflid	rs378569611	2	6265327	3,9078E-09	4475	6269802	6269806	192617
CHA	surflid	rs378569611	2	6265327	3,9078E-09	10084	6275411	6275415	192144
CHA	surflid	rs378569611	2	6265327	3,9078E-09	451874	6717201	6717205	192613
CHA	surflid	rs378569611	2	6265327	3,9078E-09	452219	6717546	6717550	192752
CHA	surflid	rs378569611	2	6265327	3,9078E-09	467626	6732953	6732957	192744
CHA	surflid	rs378569611	2	6265327	3,9078E-09	467640	6732967	6732971	192793
CHA	surflid	rs378569611	2	6265327	3,9078E-09	467080	6732407	6732411	192007
CHA	surflid	rs384344059	2	6265634	7,8082E-08	464719	5800911	5800915	192847
CHA	surflid	rs384344059	2	6265634	7,8082E-08	466732	5798898	5798902	192856
CHA	surflid	rs384344059	2	6265634	7,8082E-08	17195	6248435	6248439	164702
CHA	surflid	rs384344059	2	6265634	7,8082E-08	7254	6258376	6258380	192586
CHA	surflid	rs384344059	2	6265634	7,8082E-08	12571	6253059	6253063	192652
CHA	surflid	rs384344059	2	6265634	7,8082E-08	4168	6269802	6269806	192617
CHA	surflid	rs384344059	2	6265634	7,8082E-08	9777	6275411	6275415	192144
CHA	surflid	rs384344059	2	6265634	7,8082E-08	451567	6717201	6717205	192613
CHA	surflid	rs384344059	2	6265634	7,8082E-08	451912	6717546	6717550	192752
CHA	surflid	rs384344059	2	6265634	7,8082E-08	467319	6732953	6732957	192744
CHA	surflid	rs384344059	2	6265634	7,8082E-08	467333	6732967	6732971	192793
CHA	surflid	rs384344059	2	6265634	7,8082E-08	466773	6732407	6732411	192007
CHA	surflid	rs379694473	2	6266066	2,2292E-09	465151	5800911	5800915	192847
CHA	surflid	rs379694473	2	6266066	2,2292E-09	467164	5798898	5798902	192856
CHA	surflid	rs379694473	2	6266066	2,2292E-09	17627	6248435	6248439	164702
CHA	surflid	rs379694473	2	6266066	2,2292E-09	7686	6258376	6258380	192586
CHA	surflid	rs379694473	2	6266066	2,2292E-09	13003	6253059	6253063	192652
CHA	surflid	rs379694473	2	6266066	2,2292E-09	3736	6269802	6269806	192617
CHA	surflid	rs379694473	2	6266066	2,2292E-09	9345	6275411	6275415	192144
CHA	surflid	rs379694473	2	6266066	2,2292E-09	451135	6717201	6717205	192613
CHA	surflid	rs379694473	2	6266066	2,2292E-09	451480	6717546	6717550	192752
CHA	surflid	rs379694473	2	6266066	2,2292E-09	466887	6732953	6732957	192744
CHA	surflid	rs379694473	2	6266066	2,2292E-09	466901	6732967	6732971	192793
CHA	surflid	rs379694473	2	6266066	2,2292E-09	466341	6732407	6732411	192007
CHA	surflid	rs210991659	2	6269804	2,8061E-09	468889	5800911	5800915	192847
CHA	surflid	rs210991659	2	6269804	2,8061E-09	470902	5798898	5798902	192856
CHA	surflid	rs210991659	2	6269804	2,8061E-09	21365	6248435	6248439	164702
CHA	surflid	rs210991659	2	6269804	2,8061E-09	11424	6258376	6258380	192586

CHA	surflid	rs210991659	2	6269804	2,8061E-09	16741	6253059	6253063	192652
CHA	surflid	rs210991659	2	6269804	2,8061E-09	2	6269802	6269806	192617
CHA	surflid	rs210991659	2	6269804	2,8061E-09	5607	6275411	6275415	192144
CHA	surflid	rs210991659	2	6269804	2,8061E-09	447397	6717201	6717205	192613
CHA	surflid	rs210991659	2	6269804	2,8061E-09	447742	6717546	6717550	192752
CHA	surflid	rs210991659	2	6269804	2,8061E-09	463149	6732953	6732957	192744
CHA	surflid	rs210991659	2	6269804	2,8061E-09	463163	6732967	6732971	192793
CHA	surflid	rs210991659	2	6269804	2,8061E-09	462603	6732407	6732411	192007
CHA	surflid	rs110874471	2	6275413	1,1061E-09	474498	5800911	5800915	192847
CHA	surflid	rs110874471	2	6275413	1,1061E-09	476511	5798898	5798902	192856
CHA	surflid	rs110874471	2	6275413	1,1061E-09	26974	6248435	6248439	164702
CHA	surflid	rs110874471	2	6275413	1,1061E-09	17033	6258376	6258380	192586
CHA	surflid	rs110874471	2	6275413	1,1061E-09	22350	6253059	6253063	192652
CHA	surflid	rs110874471	2	6275413	1,1061E-09	5607	6269802	6269806	192617
CHA	surflid	rs110874471	2	6275413	1,1061E-09	2	6275411	6275415	192144
CHA	surflid	rs110874471	2	6275413	1,1061E-09	441788	6717201	6717205	192613
CHA	surflid	rs110874471	2	6275413	1,1061E-09	442133	6717546	6717550	192752
CHA	surflid	rs110874471	2	6275413	1,1061E-09	457540	6732953	6732957	192744
CHA	surflid	rs110874471	2	6275413	1,1061E-09	457554	6732967	6732971	192793
CHA	surflid	rs110874471	2	6275413	1,1061E-09	456994	6732407	6732411	192007
CHA	surflid	rs525967723	2	6275413	1,1061E-09	474498	5800911	5800915	192847
CHA	surflid	rs525967723	2	6275413	1,1061E-09	476511	5798898	5798902	192856
CHA	surflid	rs525967723	2	6275413	1,1061E-09	26974	6248435	6248439	164702
CHA	surflid	rs525967723	2	6275413	1,1061E-09	17033	6258376	6258380	192586
CHA	surflid	rs525967723	2	6275413	1,1061E-09	22350	6253059	6253063	192652
CHA	surflid	rs525967723	2	6275413	1,1061E-09	5607	6269802	6269806	192617
CHA	surflid	rs525967723	2	6275413	1,1061E-09	2	6275411	6275415	192144
CHA	surflid	rs525967723	2	6275413	1,1061E-09	441788	6717201	6717205	192613
CHA	surflid	rs525967723	2	6275413	1,1061E-09	442133	6717546	6717550	192752
CHA	surflid	rs525967723	2	6275413	1,1061E-09	457540	6732953	6732957	192744
CHA	surflid	rs525967723	2	6275413	1,1061E-09	457554	6732967	6732971	192793
CHA	surflid	rs525967723	2	6275413	1,1061E-09	456994	6732407	6732411	192007
CHA	surflid	rs110359757	2	6275477	2,3002E-09	474562	5800911	5800915	192847
CHA	surflid	rs110359757	2	6275477	2,3002E-09	476575	5798898	5798902	192856
CHA	surflid	rs110359757	2	6275477	2,3002E-09	27038	6248435	6248439	164702
CHA	surflid	rs110359757	2	6275477	2,3002E-09	17097	6258376	6258380	192586
CHA	surflid	rs110359757	2	6275477	2,3002E-09	22414	6253059	6253063	192652
CHA	surflid	rs110359757	2	6275477	2,3002E-09	5671	6269802	6269806	192617
CHA	surflid	rs110359757	2	6275477	2,3002E-09	62	6275411	6275415	192144
CHA	surflid	rs110359757	2	6275477	2,3002E-09	441724	6717201	6717205	192613
CHA	surflid	rs110359757	2	6275477	2,3002E-09	442069	6717546	6717550	192752
CHA	surflid	rs110359757	2	6275477	2,3002E-09	457476	6732953	6732957	192744
CHA	surflid	rs110359757	2	6275477	2,3002E-09	457490	6732967	6732971	192793
CHA	surflid	rs110359757	2	6275477	2,3002E-09	456930	6732407	6732411	192007
CHA	surflid	rs378726559	2	5538507	1,5972E-08	262404	5800911	5800915	192847
CHA	surflid	rs378726559	2	5538507	1,5972E-08	260391	5798898	5798902	192856

CHA	surflid	rs385090797	2	5541112	8,7893E-08	259799	5800911	5800915	192847
CHA	surflid	rs385090797	2	5541112	8,7893E-08	257786	5798898	5798902	192856
CHA	surflid	rs798522384	2	5603129	2,4841E-08	197782	5800911	5800915	192847
CHA	surflid	rs798522384	2	5603129	2,4841E-08	195769	5798898	5798902	192856
CHA	surflid	rs385792560	2	5665966	5,7027E-08	134945	5800911	5800915	192847
CHA	surflid	rs385792560	2	5665966	5,7027E-08	132932	5798898	5798902	192856
CHA	surflid	rs378836284	2	5699321	9,853E-08	101590	5800911	5800915	192847
CHA	surflid	rs378836284	2	5699321	9,853E-08	99577	5798898	5798902	192856
CHA	surflid	rs379506214	2	5701317	9,712E-08	99594	5800911	5800915	192847
CHA	surflid	rs379506214	2	5701317	9,712E-08	97581	5798898	5798902	192856
CHA	surflid	rs208682005	2	5785502	1,7486E-08	15409	5800911	5800915	192847
CHA	surflid	rs208682005	2	5785502	1,7486E-08	13396	5798898	5798902	192856
CHA	surflid	rs208682005	2	5785502	1,7486E-08	462933	6248435	6248439	164702
CHA	surflid	rs208682005	2	5785502	1,7486E-08	472874	6258376	6258380	192586
CHA	surflid	rs208682005	2	5785502	1,7486E-08	467557	6253059	6253063	192652
CHA	surflid	rs208682005	2	5785502	1,7486E-08	484300	6269802	6269806	192617
CHA	surflid	rs208682005	2	5785502	1,7486E-08	489909	6275411	6275415	192144
CHA	surflid	rs378426833	2	5785838	5,7515E-08	15073	5800911	5800915	192847
CHA	surflid	rs378426833	2	5785838	5,7515E-08	13060	5798898	5798902	192856
CHA	surflid	rs378426833	2	5785838	5,7515E-08	462597	6248435	6248439	164702
CHA	surflid	rs378426833	2	5785838	5,7515E-08	472538	6258376	6258380	192586
CHA	surflid	rs378426833	2	5785838	5,7515E-08	467221	6253059	6253063	192652
CHA	surflid	rs378426833	2	5785838	5,7515E-08	483964	6269802	6269806	192617
CHA	surflid	rs378426833	2	5785838	5,7515E-08	489573	6275411	6275415	192144
CHA	surflid	rs110029796	2	5792169	6,8814E-08	8742	5800911	5800915	192847
CHA	surflid	rs110029796	2	5792169	6,8814E-08	6729	5798898	5798902	192856
CHA	surflid	rs110029796	2	5792169	6,8814E-08	456266	6248435	6248439	164702
CHA	surflid	rs110029796	2	5792169	6,8814E-08	466207	6258376	6258380	192586
CHA	surflid	rs110029796	2	5792169	6,8814E-08	460890	6253059	6253063	192652
CHA	surflid	rs110029796	2	5792169	6,8814E-08	477633	6269802	6269806	192617
CHA	surflid	rs110029796	2	5792169	6,8814E-08	483242	6275411	6275415	192144
CHA	surflid	rs137683217	2	5807396	7,5609E-08	6481	5800911	5800915	192847
CHA	surflid	rs137683217	2	5807396	7,5609E-08	8494	5798898	5798902	192856
CHA	surflid	rs137683217	2	5807396	7,5609E-08	441039	6248435	6248439	164702
CHA	surflid	rs137683217	2	5807396	7,5609E-08	450980	6258376	6258380	192586
CHA	surflid	rs137683217	2	5807396	7,5609E-08	445663	6253059	6253063	192652
CHA	surflid	rs137683217	2	5807396	7,5609E-08	462406	6269802	6269806	192617
CHA	surflid	rs137683217	2	5807396	7,5609E-08	468015	6275411	6275415	192144
CHA	surflid	rs378005953	2	5963707	7,6481E-09	162792	5800911	5800915	192847
CHA	surflid	rs378005953	2	5963707	7,6481E-09	164805	5798898	5798902	192856
CHA	surflid	rs378005953	2	5963707	7,6481E-09	284728	6248435	6248439	164702
CHA	surflid	rs378005953	2	5963707	7,6481E-09	294669	6258376	6258380	192586
CHA	surflid	rs378005953	2	5963707	7,6481E-09	289352	6253059	6253063	192652
CHA	surflid	rs378005953	2	5963707	7,6481E-09	306095	6269802	6269806	192617
CHA	surflid	rs378005953	2	5963707	7,6481E-09	311704	6275411	6275415	192144
CHA	surflid	rs378659617	2	5968717	1,2724E-08	167802	5800911	5800915	192847

CHA	surflid	rs378659617	2	5968717	1,2724E-08	169815	5798898	5798902	192856
CHA	surflid	rs378659617	2	5968717	1,2724E-08	279718	6248435	6248439	164702
CHA	surflid	rs378659617	2	5968717	1,2724E-08	289659	6258376	6258380	192586
CHA	surflid	rs378659617	2	5968717	1,2724E-08	284342	6253059	6253063	192652
CHA	surflid	rs378659617	2	5968717	1,2724E-08	301085	6269802	6269806	192617
CHA	surflid	rs378659617	2	5968717	1,2724E-08	306694	6275411	6275415	192144
CHA	surflid	rs383343889	2	5975507	2,7661E-08	174592	5800911	5800915	192847
CHA	surflid	rs383343889	2	5975507	2,7661E-08	176605	5798898	5798902	192856
CHA	surflid	rs383343889	2	5975507	2,7661E-08	272928	6248435	6248439	164702
CHA	surflid	rs383343889	2	5975507	2,7661E-08	282869	6258376	6258380	192586
CHA	surflid	rs383343889	2	5975507	2,7661E-08	277552	6253059	6253063	192652
CHA	surflid	rs383343889	2	5975507	2,7661E-08	294295	6269802	6269806	192617
CHA	surflid	rs383343889	2	5975507	2,7661E-08	299904	6275411	6275415	192144
CHA	surflid	rs385420656	2	5978976	3,3634E-08	178061	5800911	5800915	192847
CHA	surflid	rs385420656	2	5978976	3,3634E-08	180074	5798898	5798902	192856
CHA	surflid	rs385420656	2	5978976	3,3634E-08	269459	6248435	6248439	164702
CHA	surflid	rs385420656	2	5978976	3,3634E-08	279400	6258376	6258380	192586
CHA	surflid	rs385420656	2	5978976	3,3634E-08	274083	6253059	6253063	192652
CHA	surflid	rs385420656	2	5978976	3,3634E-08	290826	6269802	6269806	192617
CHA	surflid	rs385420656	2	5978976	3,3634E-08	296435	6275411	6275415	192144
CHA	surflid	rs378472599	2	5982093	2,6042E-08	181178	5800911	5800915	192847
CHA	surflid	rs378472599	2	5982093	2,6042E-08	183191	5798898	5798902	192856
CHA	surflid	rs378472599	2	5982093	2,6042E-08	266342	6248435	6248439	164702
CHA	surflid	rs378472599	2	5982093	2,6042E-08	276283	6258376	6258380	192586
CHA	surflid	rs378472599	2	5982093	2,6042E-08	270966	6253059	6253063	192652
CHA	surflid	rs378472599	2	5982093	2,6042E-08	287709	6269802	6269806	192617
CHA	surflid	rs378472599	2	5982093	2,6042E-08	293318	6275411	6275415	192144
CHA	surflid	rs385222096	2	5984774	2,7453E-08	183859	5800911	5800915	192847
CHA	surflid	rs385222096	2	5984774	2,7453E-08	185872	5798898	5798902	192856
CHA	surflid	rs385222096	2	5984774	2,7453E-08	263661	6248435	6248439	164702
CHA	surflid	rs385222096	2	5984774	2,7453E-08	273602	6258376	6258380	192586
CHA	surflid	rs385222096	2	5984774	2,7453E-08	268285	6253059	6253063	192652
CHA	surflid	rs385222096	2	5984774	2,7453E-08	285028	6269802	6269806	192617
CHA	surflid	rs385222096	2	5984774	2,7453E-08	290637	6275411	6275415	192144
CHA	surflid	rs378072047	2	5985041	2,5E-08	184126	5800911	5800915	192847
CHA	surflid	rs378072047	2	5985041	2,5E-08	186139	5798898	5798902	192856
CHA	surflid	rs378072047	2	5985041	2,5E-08	263394	6248435	6248439	164702
CHA	surflid	rs378072047	2	5985041	2,5E-08	273335	6258376	6258380	192586
CHA	surflid	rs378072047	2	5985041	2,5E-08	268018	6253059	6253063	192652
CHA	surflid	rs378072047	2	5985041	2,5E-08	284761	6269802	6269806	192617
CHA	surflid	rs378072047	2	5985041	2,5E-08	290370	6275411	6275415	192144
CHA	surflid	rs1115676547	2	5986702	1,2906E-08	185787	5800911	5800915	192847
CHA	surflid	rs1115676547	2	5986702	1,2906E-08	187800	5798898	5798902	192856
CHA	surflid	rs1115676547	2	5986702	1,2906E-08	261733	6248435	6248439	164702
CHA	surflid	rs1115676547	2	5986702	1,2906E-08	271674	6258376	6258380	192586
CHA	surflid	rs1115676547	2	5986702	1,2906E-08	266357	6253059	6253063	192652

CHA	surflid	rs1115676547	2	5986702	1,2906E-08	283100	6269802	6269806	192617
CHA	surflid	rs1115676547	2	5986702	1,2906E-08	288709	6275411	6275415	192144
CHA	surflid	rs379378629	2	5989703	2,8227E-08	188788	5800911	5800915	192847
CHA	surflid	rs379378629	2	5989703	2,8227E-08	190801	5798898	5798902	192856
CHA	surflid	rs379378629	2	5989703	2,8227E-08	258732	6248435	6248439	164702
CHA	surflid	rs379378629	2	5989703	2,8227E-08	268673	6258376	6258380	192586
CHA	surflid	rs379378629	2	5989703	2,8227E-08	263356	6253059	6253063	192652
CHA	surflid	rs379378629	2	5989703	2,8227E-08	280099	6269802	6269806	192617
CHA	surflid	rs379378629	2	5989703	2,8227E-08	285708	6275411	6275415	192144
CHA	surflid	rs379957659	2	5990387	8,5413E-08	189472	5800911	5800915	192847
CHA	surflid	rs379957659	2	5990387	8,5413E-08	191485	5798898	5798902	192856
CHA	surflid	rs379957659	2	5990387	8,5413E-08	258048	6248435	6248439	164702
CHA	surflid	rs379957659	2	5990387	8,5413E-08	267989	6258376	6258380	192586
CHA	surflid	rs379957659	2	5990387	8,5413E-08	262672	6253059	6253063	192652
CHA	surflid	rs379957659	2	5990387	8,5413E-08	279415	6269802	6269806	192617
CHA	surflid	rs379957659	2	5990387	8,5413E-08	285024	6275411	6275415	192144
CHA	surflid	rs385074991	2	5991053	2,6016E-08	190138	5800911	5800915	192847
CHA	surflid	rs385074991	2	5991053	2,6016E-08	192151	5798898	5798902	192856
CHA	surflid	rs385074991	2	5991053	2,6016E-08	257382	6248435	6248439	164702
CHA	surflid	rs385074991	2	5991053	2,6016E-08	267323	6258376	6258380	192586
CHA	surflid	rs385074991	2	5991053	2,6016E-08	262006	6253059	6253063	192652
CHA	surflid	rs385074991	2	5991053	2,6016E-08	278749	6269802	6269806	192617
CHA	surflid	rs385074991	2	5991053	2,6016E-08	284358	6275411	6275415	192144
CHA	surflid	rs799243013	2	5991327	8,9512E-09	190412	5800911	5800915	192847
CHA	surflid	rs799243013	2	5991327	8,9512E-09	192425	5798898	5798902	192856
CHA	surflid	rs799243013	2	5991327	8,9512E-09	257108	6248435	6248439	164702
CHA	surflid	rs799243013	2	5991327	8,9512E-09	267049	6258376	6258380	192586
CHA	surflid	rs799243013	2	5991327	8,9512E-09	261732	6253059	6253063	192652
CHA	surflid	rs799243013	2	5991327	8,9512E-09	278475	6269802	6269806	192617
CHA	surflid	rs799243013	2	5991327	8,9512E-09	284084	6275411	6275415	192144
CHA	surflid	rs465822044	2	5991328	8,9035E-09	190413	5800911	5800915	192847
CHA	surflid	rs465822044	2	5991328	8,9035E-09	192426	5798898	5798902	192856
CHA	surflid	rs465822044	2	5991328	8,9035E-09	257107	6248435	6248439	164702
CHA	surflid	rs465822044	2	5991328	8,9035E-09	267048	6258376	6258380	192586
CHA	surflid	rs465822044	2	5991328	8,9035E-09	261731	6253059	6253063	192652
CHA	surflid	rs465822044	2	5991328	8,9035E-09	278474	6269802	6269806	192617
CHA	surflid	rs465822044	2	5991328	8,9035E-09	284083	6275411	6275415	192144
CHA	surflid	rs518174551	2	5994475	1,1871E-08	193560	5800911	5800915	192847
CHA	surflid	rs518174551	2	5994475	1,1871E-08	195573	5798898	5798902	192856
CHA	surflid	rs518174551	2	5994475	1,1871E-08	253960	6248435	6248439	164702
CHA	surflid	rs518174551	2	5994475	1,1871E-08	263901	6258376	6258380	192586
CHA	surflid	rs518174551	2	5994475	1,1871E-08	258584	6253059	6253063	192652
CHA	surflid	rs518174551	2	5994475	1,1871E-08	275327	6269802	6269806	192617
CHA	surflid	rs518174551	2	5994475	1,1871E-08	280936	6275411	6275415	192144
CHA	surflid	rs384298586	2	5994586	1,1992E-08	193671	5800911	5800915	192847
CHA	surflid	rs384298586	2	5994586	1,1992E-08	195684	5798898	5798902	192856

CHA	surflid	rs384298586	2	5994586	1,1992E-08	253849	6248435	6248439	164702
CHA	surflid	rs384298586	2	5994586	1,1992E-08	263790	6258376	6258380	192586
CHA	surflid	rs384298586	2	5994586	1,1992E-08	258473	6253059	6253063	192652
CHA	surflid	rs384298586	2	5994586	1,1992E-08	275216	6269802	6269806	192617
CHA	surflid	rs384298586	2	5994586	1,1992E-08	280825	6275411	6275415	192144
CHA	surflid	rs799772996	2	5994751	9,8386E-09	193836	5800911	5800915	192847
CHA	surflid	rs799772996	2	5994751	9,8386E-09	195849	5798898	5798902	192856
CHA	surflid	rs799772996	2	5994751	9,8386E-09	253684	6248435	6248439	164702
CHA	surflid	rs799772996	2	5994751	9,8386E-09	263625	6258376	6258380	192586
CHA	surflid	rs799772996	2	5994751	9,8386E-09	258308	6253059	6253063	192652
CHA	surflid	rs799772996	2	5994751	9,8386E-09	275051	6269802	6269806	192617
CHA	surflid	rs799772996	2	5994751	9,8386E-09	280660	6275411	6275415	192144
CHA	surflid	rs876438272	2	5994751	9,8386E-09	193836	5800911	5800915	192847
CHA	surflid	rs876438272	2	5994751	9,8386E-09	195849	5798898	5798902	192856
CHA	surflid	rs876438272	2	5994751	9,8386E-09	253684	6248435	6248439	164702
CHA	surflid	rs876438272	2	5994751	9,8386E-09	263625	6258376	6258380	192586
CHA	surflid	rs876438272	2	5994751	9,8386E-09	258308	6253059	6253063	192652
CHA	surflid	rs876438272	2	5994751	9,8386E-09	275051	6269802	6269806	192617
CHA	surflid	rs876438272	2	5994751	9,8386E-09	280660	6275411	6275415	192144
CHA	surflid	rs383512625	2	5999051	1,2258E-08	198136	5800911	5800915	192847
CHA	surflid	rs383512625	2	5999051	1,2258E-08	200149	5798898	5798902	192856
CHA	surflid	rs383512625	2	5999051	1,2258E-08	249384	6248435	6248439	164702
CHA	surflid	rs383512625	2	5999051	1,2258E-08	259325	6258376	6258380	192586
CHA	surflid	rs383512625	2	5999051	1,2258E-08	254008	6253059	6253063	192652
CHA	surflid	rs383512625	2	5999051	1,2258E-08	270751	6269802	6269806	192617
CHA	surflid	rs383512625	2	5999051	1,2258E-08	276360	6275411	6275415	192144
CHA	surflid	rs384641016	2	5999485	1,484E-08	198570	5800911	5800915	192847
CHA	surflid	rs384641016	2	5999485	1,484E-08	200583	5798898	5798902	192856
CHA	surflid	rs384641016	2	5999485	1,484E-08	248950	6248435	6248439	164702
CHA	surflid	rs384641016	2	5999485	1,484E-08	258891	6258376	6258380	192586
CHA	surflid	rs384641016	2	5999485	1,484E-08	253574	6253059	6253063	192652
CHA	surflid	rs384641016	2	5999485	1,484E-08	270317	6269802	6269806	192617
CHA	surflid	rs384641016	2	5999485	1,484E-08	275926	6275411	6275415	192144
CHA	surflid	rs384909931	2	6001495	7,3315E-09	200580	5800911	5800915	192847
CHA	surflid	rs384909931	2	6001495	7,3315E-09	202593	5798898	5798902	192856
CHA	surflid	rs384909931	2	6001495	7,3315E-09	246940	6248435	6248439	164702
CHA	surflid	rs384909931	2	6001495	7,3315E-09	256881	6258376	6258380	192586
CHA	surflid	rs384909931	2	6001495	7,3315E-09	251564	6253059	6253063	192652
CHA	surflid	rs384909931	2	6001495	7,3315E-09	268307	6269802	6269806	192617
CHA	surflid	rs384909931	2	6001495	7,3315E-09	273916	6275411	6275415	192144
CHA	surflid	rs380888646	2	6003295	1,0786E-08	202380	5800911	5800915	192847
CHA	surflid	rs380888646	2	6003295	1,0786E-08	204393	5798898	5798902	192856
CHA	surflid	rs380888646	2	6003295	1,0786E-08	245140	6248435	6248439	164702
CHA	surflid	rs380888646	2	6003295	1,0786E-08	255081	6258376	6258380	192586
CHA	surflid	rs380888646	2	6003295	1,0786E-08	249764	6253059	6253063	192652
CHA	surflid	rs380888646	2	6003295	1,0786E-08	266507	6269802	6269806	192617

CHA	surflid	rs380888646	2	6003295	1,0786E-08	272116	6275411	6275415	192144
CHA	surflid	rs379192175	2	6012867	8,9663E-09	211952	5800911	5800915	192847
CHA	surflid	rs379192175	2	6012867	8,9663E-09	213965	5798898	5798902	192856
CHA	surflid	rs379192175	2	6012867	8,9663E-09	235568	6248435	6248439	164702
CHA	surflid	rs379192175	2	6012867	8,9663E-09	245509	6258376	6258380	192586
CHA	surflid	rs379192175	2	6012867	8,9663E-09	240192	6253059	6253063	192652
CHA	surflid	rs379192175	2	6012867	8,9663E-09	256935	6269802	6269806	192617
CHA	surflid	rs379192175	2	6012867	8,9663E-09	262544	6275411	6275415	192144
CHA	surflid	rs378535388	2	6013759	8,6069E-09	212844	5800911	5800915	192847
CHA	surflid	rs378535388	2	6013759	8,6069E-09	214857	5798898	5798902	192856
CHA	surflid	rs378535388	2	6013759	8,6069E-09	234676	6248435	6248439	164702
CHA	surflid	rs378535388	2	6013759	8,6069E-09	244617	6258376	6258380	192586
CHA	surflid	rs378535388	2	6013759	8,6069E-09	239300	6253059	6253063	192652
CHA	surflid	rs378535388	2	6013759	8,6069E-09	256043	6269802	6269806	192617
CHA	surflid	rs378535388	2	6013759	8,6069E-09	261652	6275411	6275415	192144
CHA	surflid	rs379475871	2	6024310	5,7309E-09	223395	5800911	5800915	192847
CHA	surflid	rs379475871	2	6024310	5,7309E-09	225408	5798898	5798902	192856
CHA	surflid	rs379475871	2	6024310	5,7309E-09	224125	6248435	6248439	164702
CHA	surflid	rs379475871	2	6024310	5,7309E-09	234066	6258376	6258380	192586
CHA	surflid	rs379475871	2	6024310	5,7309E-09	228749	6253059	6253063	192652
CHA	surflid	rs379475871	2	6024310	5,7309E-09	245492	6269802	6269806	192617
CHA	surflid	rs379475871	2	6024310	5,7309E-09	251101	6275411	6275415	192144
CHA	surflid	rs384731408	2	6030773	6,5735E-09	229858	5800911	5800915	192847
CHA	surflid	rs384731408	2	6030773	6,5735E-09	231871	5798898	5798902	192856
CHA	surflid	rs384731408	2	6030773	6,5735E-09	217662	6248435	6248439	164702
CHA	surflid	rs384731408	2	6030773	6,5735E-09	227603	6258376	6258380	192586
CHA	surflid	rs384731408	2	6030773	6,5735E-09	222286	6253059	6253063	192652
CHA	surflid	rs384731408	2	6030773	6,5735E-09	239029	6269802	6269806	192617
CHA	surflid	rs384731408	2	6030773	6,5735E-09	244638	6275411	6275415	192144
CHA	surflid	rs382561696	2	6031814	4,5619E-10	230899	5800911	5800915	192847
CHA	surflid	rs382561696	2	6031814	4,5619E-10	232912	5798898	5798902	192856
CHA	surflid	rs382561696	2	6031814	4,5619E-10	216621	6248435	6248439	164702
CHA	surflid	rs382561696	2	6031814	4,5619E-10	226562	6258376	6258380	192586
CHA	surflid	rs382561696	2	6031814	4,5619E-10	221245	6253059	6253063	192652
CHA	surflid	rs382561696	2	6031814	4,5619E-10	237988	6269802	6269806	192617
CHA	surflid	rs382561696	2	6031814	4,5619E-10	243597	6275411	6275415	192144
CHA	surflid	rs456652025	2	6038100	4,3148E-09	237185	5800911	5800915	192847
CHA	surflid	rs456652025	2	6038100	4,3148E-09	239198	5798898	5798902	192856
CHA	surflid	rs456652025	2	6038100	4,3148E-09	210335	6248435	6248439	164702
CHA	surflid	rs456652025	2	6038100	4,3148E-09	220276	6258376	6258380	192586
CHA	surflid	rs456652025	2	6038100	4,3148E-09	214959	6253059	6253063	192652
CHA	surflid	rs456652025	2	6038100	4,3148E-09	231702	6269802	6269806	192617
CHA	surflid	rs456652025	2	6038100	4,3148E-09	237311	6275411	6275415	192144
CHA	surflid	rs385216156	2	6047619	5,0885E-09	246704	5800911	5800915	192847
CHA	surflid	rs385216156	2	6047619	5,0885E-09	248717	5798898	5798902	192856
CHA	surflid	rs385216156	2	6047619	5,0885E-09	200816	6248435	6248439	164702

CHA	surflid	rs385216156	2	6047619	5,0885E-09	210757	6258376	6258380	192586
CHA	surflid	rs385216156	2	6047619	5,0885E-09	205440	6253059	6253063	192652
CHA	surflid	rs385216156	2	6047619	5,0885E-09	222183	6269802	6269806	192617
CHA	surflid	rs385216156	2	6047619	5,0885E-09	227792	6275411	6275415	192144
CHA	surflid	rs384635768	2	6054968	1,1523E-09	254053	5800911	5800915	192847
CHA	surflid	rs384635768	2	6054968	1,1523E-09	256066	5798898	5798902	192856
CHA	surflid	rs384635768	2	6054968	1,1523E-09	193467	6248435	6248439	164702
CHA	surflid	rs384635768	2	6054968	1,1523E-09	203408	6258376	6258380	192586
CHA	surflid	rs384635768	2	6054968	1,1523E-09	198091	6253059	6253063	192652
CHA	surflid	rs384635768	2	6054968	1,1523E-09	214834	6269802	6269806	192617
CHA	surflid	rs384635768	2	6054968	1,1523E-09	220443	6275411	6275415	192144
CHA	surflid	rs384374992	2	6102781	4,415E-08	301866	5800911	5800915	192847
CHA	surflid	rs384374992	2	6102781	4,415E-08	303879	5798898	5798902	192856
CHA	surflid	rs384374992	2	6102781	4,415E-08	145654	6248435	6248439	164702
CHA	surflid	rs384374992	2	6102781	4,415E-08	155595	6258376	6258380	192586
CHA	surflid	rs384374992	2	6102781	4,415E-08	150278	6253059	6253063	192652
CHA	surflid	rs384374992	2	6102781	4,415E-08	167021	6269802	6269806	192617
CHA	surflid	rs384374992	2	6102781	4,415E-08	172630	6275411	6275415	192144
CHA	surflid	rs382032850	2	6104391	3,9771E-08	303476	5800911	5800915	192847
CHA	surflid	rs382032850	2	6104391	3,9771E-08	305489	5798898	5798902	192856
CHA	surflid	rs382032850	2	6104391	3,9771E-08	144044	6248435	6248439	164702
CHA	surflid	rs382032850	2	6104391	3,9771E-08	153985	6258376	6258380	192586
CHA	surflid	rs382032850	2	6104391	3,9771E-08	148668	6253059	6253063	192652
CHA	surflid	rs382032850	2	6104391	3,9771E-08	165411	6269802	6269806	192617
CHA	surflid	rs382032850	2	6104391	3,9771E-08	171020	6275411	6275415	192144
CHA	surflid	rs378897113	2	6104439	2,1041E-08	303524	5800911	5800915	192847
CHA	surflid	rs378897113	2	6104439	2,1041E-08	305537	5798898	5798902	192856
CHA	surflid	rs378897113	2	6104439	2,1041E-08	143996	6248435	6248439	164702
CHA	surflid	rs378897113	2	6104439	2,1041E-08	153937	6258376	6258380	192586
CHA	surflid	rs378897113	2	6104439	2,1041E-08	148620	6253059	6253063	192652
CHA	surflid	rs378897113	2	6104439	2,1041E-08	165363	6269802	6269806	192617
CHA	surflid	rs378897113	2	6104439	2,1041E-08	170972	6275411	6275415	192144
CHA	surflid	rs378762857	2	6133213	2,7071E-09	332298	5800911	5800915	192847
CHA	surflid	rs378762857	2	6133213	2,7071E-09	334311	5798898	5798902	192856
CHA	surflid	rs378762857	2	6133213	2,7071E-09	115222	6248435	6248439	164702
CHA	surflid	rs378762857	2	6133213	2,7071E-09	125163	6258376	6258380	192586
CHA	surflid	rs378762857	2	6133213	2,7071E-09	119846	6253059	6253063	192652
CHA	surflid	rs378762857	2	6133213	2,7071E-09	136589	6269802	6269806	192617
CHA	surflid	rs378762857	2	6133213	2,7071E-09	142198	6275411	6275415	192144
CHA	surflid	rs383391721	2	6205728	7,4362E-08	404813	5800911	5800915	192847
CHA	surflid	rs383391721	2	6205728	7,4362E-08	406826	5798898	5798902	192856
CHA	surflid	rs383391721	2	6205728	7,4362E-08	42707	6248435	6248439	164702
CHA	surflid	rs383391721	2	6205728	7,4362E-08	52648	6258376	6258380	192586
CHA	surflid	rs383391721	2	6205728	7,4362E-08	47331	6253059	6253063	192652
CHA	surflid	rs383391721	2	6205728	7,4362E-08	64074	6269802	6269806	192617
CHA	surflid	rs383391721	2	6205728	7,4362E-08	69683	6275411	6275415	192144

CHA	surflid	rs448262344	2	6213596	8,1845E-09	412681	5800911	5800915	192847
CHA	surflid	rs448262344	2	6213596	8,1845E-09	414694	5798898	5798902	192856
CHA	surflid	rs448262344	2	6213596	8,1845E-09	34839	6248435	6248439	164702
CHA	surflid	rs448262344	2	6213596	8,1845E-09	44780	6258376	6258380	192586
CHA	surflid	rs448262344	2	6213596	8,1845E-09	39463	6253059	6253063	192652
CHA	surflid	rs448262344	2	6213596	8,1845E-09	56206	6269802	6269806	192617
CHA	surflid	rs448262344	2	6213596	8,1845E-09	61815	6275411	6275415	192144
CHA	surflid	rs520792341	2	6213596	8,1845E-09	412681	5800911	5800915	192847
CHA	surflid	rs520792341	2	6213596	8,1845E-09	414694	5798898	5798902	192856
CHA	surflid	rs520792341	2	6213596	8,1845E-09	34839	6248435	6248439	164702
CHA	surflid	rs520792341	2	6213596	8,1845E-09	44780	6258376	6258380	192586
CHA	surflid	rs520792341	2	6213596	8,1845E-09	39463	6253059	6253063	192652
CHA	surflid	rs520792341	2	6213596	8,1845E-09	56206	6269802	6269806	192617
CHA	surflid	rs520792341	2	6213596	8,1845E-09	61815	6275411	6275415	192144
CHA	surflid	rs382943374	2	6240735	4,4915E-08	439820	5800911	5800915	192847
CHA	surflid	rs382943374	2	6240735	4,4915E-08	441833	5798898	5798902	192856
CHA	surflid	rs382943374	2	6240735	4,4915E-08	7700	6248435	6248439	164702
CHA	surflid	rs382943374	2	6240735	4,4915E-08	17641	6258376	6258380	192586
CHA	surflid	rs382943374	2	6240735	4,4915E-08	12324	6253059	6253063	192652
CHA	surflid	rs382943374	2	6240735	4,4915E-08	29067	6269802	6269806	192617
CHA	surflid	rs382943374	2	6240735	4,4915E-08	34676	6275411	6275415	192144
CHA	surflid	rs382943374	2	6240735	4,4915E-08	476466	6717201	6717205	192613
CHA	surflid	rs382943374	2	6240735	4,4915E-08	476811	6717546	6717550	192752
CHA	surflid	rs382943374	2	6240735	4,4915E-08	492218	6732953	6732957	192744
CHA	surflid	rs382943374	2	6240735	4,4915E-08	492232	6732967	6732971	192793
CHA	surflid	rs382943374	2	6240735	4,4915E-08	491672	6732407	6732411	192007
CHA	surflid	rs208300352	2	6248762	2,3066E-09	447847	5800911	5800915	192847
CHA	surflid	rs208300352	2	6248762	2,3066E-09	449860	5798898	5798902	192856
CHA	surflid	rs208300352	2	6248762	2,3066E-09	323	6248435	6248439	164702
CHA	surflid	rs208300352	2	6248762	2,3066E-09	9614	6258376	6258380	192586
CHA	surflid	rs208300352	2	6248762	2,3066E-09	4297	6253059	6253063	192652
CHA	surflid	rs208300352	2	6248762	2,3066E-09	21040	6269802	6269806	192617
CHA	surflid	rs208300352	2	6248762	2,3066E-09	26649	6275411	6275415	192144
CHA	surflid	rs208300352	2	6248762	2,3066E-09	468439	6717201	6717205	192613
CHA	surflid	rs208300352	2	6248762	2,3066E-09	468784	6717546	6717550	192752
CHA	surflid	rs208300352	2	6248762	2,3066E-09	484191	6732953	6732957	192744
CHA	surflid	rs208300352	2	6248762	2,3066E-09	484205	6732967	6732971	192793
CHA	surflid	rs208300352	2	6248762	2,3066E-09	483645	6732407	6732411	192007
CHA	surflid	rs466774071	2	6256648	4,7881E-12	455733	5800911	5800915	192847
CHA	surflid	rs466774071	2	6256648	4,7881E-12	457746	5798898	5798902	192856
CHA	surflid	rs466774071	2	6256648	4,7881E-12	8209	6248435	6248439	164702
CHA	surflid	rs466774071	2	6256648	4,7881E-12	1728	6258376	6258380	192586
CHA	surflid	rs466774071	2	6256648	4,7881E-12	3585	6253059	6253063	192652
CHA	surflid	rs466774071	2	6256648	4,7881E-12	13154	6269802	6269806	192617
CHA	surflid	rs466774071	2	6256648	4,7881E-12	18763	6275411	6275415	192144
CHA	surflid	rs466774071	2	6256648	4,7881E-12	460553	6717201	6717205	192613

CHA	surflid	rs466774071	2	6256648	4,7881E-12	460898	6717546	6717550	192752
CHA	surflid	rs466774071	2	6256648	4,7881E-12	476305	6732953	6732957	192744
CHA	surflid	rs466774071	2	6256648	4,7881E-12	476319	6732967	6732971	192793
CHA	surflid	rs466774071	2	6256648	4,7881E-12	475759	6732407	6732411	192007
CHA	surflid	rs209371527	2	6258378	2,3982E-09	457463	5800911	5800915	192847
CHA	surflid	rs209371527	2	6258378	2,3982E-09	459476	5798898	5798902	192856
CHA	surflid	rs209371527	2	6258378	2,3982E-09	9939	6248435	6248439	164702
CHA	surflid	rs209371527	2	6258378	2,3982E-09	2	6258376	6258380	192586
CHA	surflid	rs209371527	2	6258378	2,3982E-09	5315	6253059	6253063	192652
CHA	surflid	rs209371527	2	6258378	2,3982E-09	11424	6269802	6269806	192617
CHA	surflid	rs209371527	2	6258378	2,3982E-09	17033	6275411	6275415	192144
CHA	surflid	rs209371527	2	6258378	2,3982E-09	458823	6717201	6717205	192613
CHA	surflid	rs209371527	2	6258378	2,3982E-09	459168	6717546	6717550	192752
CHA	surflid	rs209371527	2	6258378	2,3982E-09	474575	6732953	6732957	192744
CHA	surflid	rs209371527	2	6258378	2,3982E-09	474589	6732967	6732971	192793
CHA	surflid	rs209371527	2	6258378	2,3982E-09	474029	6732407	6732411	192007
CHA	surflid	rs385605706	2	6260096	2,4274E-09	459181	5800911	5800915	192847
CHA	surflid	rs385605706	2	6260096	2,4274E-09	461194	5798898	5798902	192856
CHA	surflid	rs385605706	2	6260096	2,4274E-09	11657	6248435	6248439	164702
CHA	surflid	rs385605706	2	6260096	2,4274E-09	1716	6258376	6258380	192586
CHA	surflid	rs385605706	2	6260096	2,4274E-09	7033	6253059	6253063	192652
CHA	surflid	rs385605706	2	6260096	2,4274E-09	9706	6269802	6269806	192617
CHA	surflid	rs385605706	2	6260096	2,4274E-09	15315	6275411	6275415	192144
CHA	surflid	rs385605706	2	6260096	2,4274E-09	457105	6717201	6717205	192613
CHA	surflid	rs385605706	2	6260096	2,4274E-09	457450	6717546	6717550	192752
CHA	surflid	rs385605706	2	6260096	2,4274E-09	472857	6732953	6732957	192744
CHA	surflid	rs385605706	2	6260096	2,4274E-09	472871	6732967	6732971	192793
CHA	surflid	rs385605706	2	6260096	2,4274E-09	472311	6732407	6732411	192007
CHA	surflid	rs381762253	2	6263104	2,9493E-10	462189	5800911	5800915	192847
CHA	surflid	rs381762253	2	6263104	2,9493E-10	464202	5798898	5798902	192856
CHA	surflid	rs381762253	2	6263104	2,9493E-10	14665	6248435	6248439	164702
CHA	surflid	rs381762253	2	6263104	2,9493E-10	4724	6258376	6258380	192586
CHA	surflid	rs381762253	2	6263104	2,9493E-10	10041	6253059	6253063	192652
CHA	surflid	rs381762253	2	6263104	2,9493E-10	6698	6269802	6269806	192617
CHA	surflid	rs381762253	2	6263104	2,9493E-10	12307	6275411	6275415	192144
CHA	surflid	rs381762253	2	6263104	2,9493E-10	454097	6717201	6717205	192613
CHA	surflid	rs381762253	2	6263104	2,9493E-10	454442	6717546	6717550	192752
CHA	surflid	rs381762253	2	6263104	2,9493E-10	469849	6732953	6732957	192744
CHA	surflid	rs381762253	2	6263104	2,9493E-10	469863	6732967	6732971	192793
CHA	surflid	rs381762253	2	6263104	2,9493E-10	469303	6732407	6732411	192007
CHA	surflid	rs110344317	2	6281434	8,9026E-11	480519	5800911	5800915	192847
CHA	surflid	rs110344317	2	6281434	8,9026E-11	482532	5798898	5798902	192856
CHA	surflid	rs110344317	2	6281434	8,9026E-11	32995	6248435	6248439	164702
CHA	surflid	rs110344317	2	6281434	8,9026E-11	23054	6258376	6258380	192586
CHA	surflid	rs110344317	2	6281434	8,9026E-11	28371	6253059	6253063	192652
CHA	surflid	rs110344317	2	6281434	8,9026E-11	11628	6269802	6269806	192617

CHA	surflid	rs110344317	2	6281434	8,9026E-11	6019	6275411	6275415	192144
CHA	surflid	rs110344317	2	6281434	8,9026E-11	435767	6717201	6717205	192613
CHA	surflid	rs110344317	2	6281434	8,9026E-11	436112	6717546	6717550	192752
CHA	surflid	rs110344317	2	6281434	8,9026E-11	451519	6732953	6732957	192744
CHA	surflid	rs110344317	2	6281434	8,9026E-11	451533	6732967	6732971	192793
CHA	surflid	rs110344317	2	6281434	8,9026E-11	450973	6732407	6732411	192007
CHA	surflid	rs383490556	2	6331108	5,1647E-10	82669	6248435	6248439	164702
CHA	surflid	rs383490556	2	6331108	5,1647E-10	72728	6258376	6258380	192586
CHA	surflid	rs383490556	2	6331108	5,1647E-10	78045	6253059	6253063	192652
CHA	surflid	rs383490556	2	6331108	5,1647E-10	61302	6269802	6269806	192617
CHA	surflid	rs383490556	2	6331108	5,1647E-10	55693	6275411	6275415	192144
CHA	surflid	rs383490556	2	6331108	5,1647E-10	386093	6717201	6717205	192613
CHA	surflid	rs383490556	2	6331108	5,1647E-10	386438	6717546	6717550	192752
CHA	surflid	rs383490556	2	6331108	5,1647E-10	401845	6732953	6732957	192744
CHA	surflid	rs383490556	2	6331108	5,1647E-10	401859	6732967	6732971	192793
CHA	surflid	rs383490556	2	6331108	5,1647E-10	401299	6732407	6732411	192007
CHA	surflid	rs377873694	2	6331113	2,5251E-10	82674	6248435	6248439	164702
CHA	surflid	rs377873694	2	6331113	2,5251E-10	72733	6258376	6258380	192586
CHA	surflid	rs377873694	2	6331113	2,5251E-10	78050	6253059	6253063	192652
CHA	surflid	rs377873694	2	6331113	2,5251E-10	61307	6269802	6269806	192617
CHA	surflid	rs377873694	2	6331113	2,5251E-10	55698	6275411	6275415	192144
CHA	surflid	rs377873694	2	6331113	2,5251E-10	386088	6717201	6717205	192613
CHA	surflid	rs377873694	2	6331113	2,5251E-10	386433	6717546	6717550	192752
CHA	surflid	rs377873694	2	6331113	2,5251E-10	401840	6732953	6732957	192744
CHA	surflid	rs377873694	2	6331113	2,5251E-10	401854	6732967	6732971	192793
CHA	surflid	rs377873694	2	6331113	2,5251E-10	401294	6732407	6732411	192007
CHA	surflid	rs383080448	2	6811704	8,6885E-11	94499	6717201	6717205	192613
CHA	surflid	rs383080448	2	6811704	8,6885E-11	94154	6717546	6717550	192752
CHA	surflid	rs383080448	2	6811704	8,6885E-11	78747	6732953	6732957	192744
CHA	surflid	rs383080448	2	6811704	8,6885E-11	78733	6732967	6732971	192793
CHA	surflid	rs383080448	2	6811704	8,6885E-11	79293	6732407	6732411	192007
CHA	surflid	rs384022421	2	6829456	1,4753E-09	112251	6717201	6717205	192613
CHA	surflid	rs384022421	2	6829456	1,4753E-09	111906	6717546	6717550	192752
CHA	surflid	rs384022421	2	6829456	1,4753E-09	96499	6732953	6732957	192744
CHA	surflid	rs384022421	2	6829456	1,4753E-09	96485	6732967	6732971	192793
CHA	surflid	rs384022421	2	6829456	1,4753E-09	97045	6732407	6732411	192007
CHA	surflid	rs462025042	2	6844361	6,2924E-10	127156	6717201	6717205	192613
CHA	surflid	rs462025042	2	6844361	6,2924E-10	126811	6717546	6717550	192752
CHA	surflid	rs462025042	2	6844361	6,2924E-10	111404	6732953	6732957	192744
CHA	surflid	rs462025042	2	6844361	6,2924E-10	111390	6732967	6732971	192793
CHA	surflid	rs462025042	2	6844361	6,2924E-10	111950	6732407	6732411	192007
CHA	surflid	rs110571003	2	6856750	4,9393E-10	139545	6717201	6717205	192613
CHA	surflid	rs110571003	2	6856750	4,9393E-10	139200	6717546	6717550	192752
CHA	surflid	rs110571003	2	6856750	4,9393E-10	123793	6732953	6732957	192744
CHA	surflid	rs110571003	2	6856750	4,9393E-10	123779	6732967	6732971	192793
CHA	surflid	rs110571003	2	6856750	4,9393E-10	124339	6732407	6732411	192007

CHA	surflid	rs385811308	2	6859207	6,6569E-10	142002	6717201	6717205	192613
CHA	surflid	rs385811308	2	6859207	6,6569E-10	141657	6717546	6717550	192752
CHA	surflid	rs385811308	2	6859207	6,6569E-10	126250	6732953	6732957	192744
CHA	surflid	rs385811308	2	6859207	6,6569E-10	126236	6732967	6732971	192793
CHA	surflid	rs385811308	2	6859207	6,6569E-10	126796	6732407	6732411	192007
CHA	surflid	rs378397067	2	6860727	4,6117E-10	143522	6717201	6717205	192613
CHA	surflid	rs378397067	2	6860727	4,6117E-10	143177	6717546	6717550	192752
CHA	surflid	rs378397067	2	6860727	4,6117E-10	127770	6732953	6732957	192744
CHA	surflid	rs378397067	2	6860727	4,6117E-10	127756	6732967	6732971	192793
CHA	surflid	rs378397067	2	6860727	4,6117E-10	128316	6732407	6732411	192007
CHA	surflid	rs384224054	2	6861518	4,2956E-10	144313	6717201	6717205	192613
CHA	surflid	rs384224054	2	6861518	4,2956E-10	143968	6717546	6717550	192752
CHA	surflid	rs384224054	2	6861518	4,2956E-10	128561	6732953	6732957	192744
CHA	surflid	rs384224054	2	6861518	4,2956E-10	128547	6732967	6732971	192793
CHA	surflid	rs384224054	2	6861518	4,2956E-10	129107	6732407	6732411	192007
CHA	surflid	rs110129105	2	6867337	4,3927E-09	150132	6717201	6717205	192613
CHA	surflid	rs110129105	2	6867337	4,3927E-09	149787	6717546	6717550	192752
CHA	surflid	rs110129105	2	6867337	4,3927E-09	134380	6732953	6732957	192744
CHA	surflid	rs110129105	2	6867337	4,3927E-09	134366	6732967	6732971	192793
CHA	surflid	rs110129105	2	6867337	4,3927E-09	134926	6732407	6732411	192007
CHA	surflid	rs799943285	2	6867899	6,2795E-12	150694	6717201	6717205	192613
CHA	surflid	rs799943285	2	6867899	6,2795E-12	150349	6717546	6717550	192752
CHA	surflid	rs799943285	2	6867899	6,2795E-12	134942	6732953	6732957	192744
CHA	surflid	rs799943285	2	6867899	6,2795E-12	134928	6732967	6732971	192793
CHA	surflid	rs799943285	2	6867899	6,2795E-12	135488	6732407	6732411	192007
CHA	surflid	rs377937500	2	6867974	2,1253E-08	150769	6717201	6717205	192613
CHA	surflid	rs377937500	2	6867974	2,1253E-08	150424	6717546	6717550	192752
CHA	surflid	rs377937500	2	6867974	2,1253E-08	135017	6732953	6732957	192744
CHA	surflid	rs377937500	2	6867974	2,1253E-08	135003	6732967	6732971	192793
CHA	surflid	rs377937500	2	6867974	2,1253E-08	135563	6732407	6732411	192007
CHA	surflid	rs475819310	2	6868030	2,0055E-10	150825	6717201	6717205	192613
CHA	surflid	rs475819310	2	6868030	2,0055E-10	150480	6717546	6717550	192752
CHA	surflid	rs475819310	2	6868030	2,0055E-10	135073	6732953	6732957	192744
CHA	surflid	rs475819310	2	6868030	2,0055E-10	135059	6732967	6732971	192793
CHA	surflid	rs475819310	2	6868030	2,0055E-10	135619	6732407	6732411	192007
CHA	surflid	rs439944289	2	6868040	1,9764E-10	150835	6717201	6717205	192613
CHA	surflid	rs439944289	2	6868040	1,9764E-10	150490	6717546	6717550	192752
CHA	surflid	rs439944289	2	6868040	1,9764E-10	135083	6732953	6732957	192744
CHA	surflid	rs439944289	2	6868040	1,9764E-10	135069	6732967	6732971	192793
CHA	surflid	rs439944289	2	6868040	1,9764E-10	135629	6732407	6732411	192007
CHA	surflid	rs458460529	2	6868048	1,4717E-10	150843	6717201	6717205	192613
CHA	surflid	rs458460529	2	6868048	1,4717E-10	150498	6717546	6717550	192752
CHA	surflid	rs458460529	2	6868048	1,4717E-10	135091	6732953	6732957	192744
CHA	surflid	rs458460529	2	6868048	1,4717E-10	135077	6732967	6732971	192793
CHA	surflid	rs458460529	2	6868048	1,4717E-10	135637	6732407	6732411	192007
CHA	surflid	rs521447441	2	6868054	1,2564E-10	150849	6717201	6717205	192613

CHA	surflid	rs521447441	2	6868054	1,2564E-10	150504	6717546	6717550	192752
CHA	surflid	rs521447441	2	6868054	1,2564E-10	135097	6732953	6732957	192744
CHA	surflid	rs521447441	2	6868054	1,2564E-10	135083	6732967	6732971	192793
CHA	surflid	rs521447441	2	6868054	1,2564E-10	135643	6732407	6732411	192007
CHA	surflid	rs800245025	2	6868058	1,272E-10	150853	6717201	6717205	192613
CHA	surflid	rs800245025	2	6868058	1,272E-10	150508	6717546	6717550	192752
CHA	surflid	rs800245025	2	6868058	1,272E-10	135101	6732953	6732957	192744
CHA	surflid	rs800245025	2	6868058	1,272E-10	135087	6732967	6732971	192793
CHA	surflid	rs800245025	2	6868058	1,272E-10	135647	6732407	6732411	192007
CHA	surflid	rs480087053	2	6868061	1,2032E-10	150856	6717201	6717205	192613
CHA	surflid	rs480087053	2	6868061	1,2032E-10	150511	6717546	6717550	192752
CHA	surflid	rs480087053	2	6868061	1,2032E-10	135104	6732953	6732957	192744
CHA	surflid	rs480087053	2	6868061	1,2032E-10	135090	6732967	6732971	192793
CHA	surflid	rs480087053	2	6868061	1,2032E-10	135650	6732407	6732411	192007
CHA	surflid	rs447118544	2	6868077	1,0234E-10	150872	6717201	6717205	192613
CHA	surflid	rs447118544	2	6868077	1,0234E-10	150527	6717546	6717550	192752
CHA	surflid	rs447118544	2	6868077	1,0234E-10	135120	6732953	6732957	192744
CHA	surflid	rs447118544	2	6868077	1,0234E-10	135106	6732967	6732971	192793
CHA	surflid	rs447118544	2	6868077	1,0234E-10	135666	6732407	6732411	192007
CHA	surflid	rs209199169	2	6880817	8,6502E-12	163612	6717201	6717205	192613
CHA	surflid	rs209199169	2	6880817	8,6502E-12	163267	6717546	6717550	192752
CHA	surflid	rs209199169	2	6880817	8,6502E-12	147860	6732953	6732957	192744
CHA	surflid	rs209199169	2	6880817	8,6502E-12	147846	6732967	6732971	192793
CHA	surflid	rs209199169	2	6880817	8,6502E-12	148406	6732407	6732411	192007
CHA	surflid	rs482509687	2	6885704	8,589E-12	168499	6717201	6717205	192613
CHA	surflid	rs482509687	2	6885704	8,589E-12	168154	6717546	6717550	192752
CHA	surflid	rs482509687	2	6885704	8,589E-12	152747	6732953	6732957	192744
CHA	surflid	rs482509687	2	6885704	8,589E-12	152733	6732967	6732971	192793
CHA	surflid	rs482509687	2	6885704	8,589E-12	153293	6732407	6732411	192007
CHA	surflid	rs436214940	2	6896029	5,9539E-12	178824	6717201	6717205	192613
CHA	surflid	rs436214940	2	6896029	5,9539E-12	178479	6717546	6717550	192752
CHA	surflid	rs436214940	2	6896029	5,9539E-12	163072	6732953	6732957	192744
CHA	surflid	rs436214940	2	6896029	5,9539E-12	163058	6732967	6732971	192793
CHA	surflid	rs436214940	2	6896029	5,9539E-12	163618	6732407	6732411	192007
CHA	surflid	rs453302103	2	6897525	6,2365E-12	180320	6717201	6717205	192613
CHA	surflid	rs453302103	2	6897525	6,2365E-12	179975	6717546	6717550	192752
CHA	surflid	rs453302103	2	6897525	6,2365E-12	164568	6732953	6732957	192744
CHA	surflid	rs453302103	2	6897525	6,2365E-12	164554	6732967	6732971	192793
CHA	surflid	rs453302103	2	6897525	6,2365E-12	165114	6732407	6732411	192007
CHA	surflid	rs383530052	2	6949173	4,6884E-10	231968	6717201	6717205	192613
CHA	surflid	rs383530052	2	6949173	4,6884E-10	231623	6717546	6717550	192752
CHA	surflid	rs383530052	2	6949173	4,6884E-10	216216	6732953	6732957	192744
CHA	surflid	rs383530052	2	6949173	4,6884E-10	216202	6732967	6732971	192793
CHA	surflid	rs383530052	2	6949173	4,6884E-10	216762	6732407	6732411	192007
CHA	surflid	rs380481920	2	6954280	1,0132E-09	237075	6717201	6717205	192613
CHA	surflid	rs380481920	2	6954280	1,0132E-09	236730	6717546	6717550	192752

CHA	surflid	rs380481920	2	6954280	1,0132E-09	221323	6732953	6732957	192744
CHA	surflid	rs380481920	2	6954280	1,0132E-09	221309	6732967	6732971	192793
CHA	surflid	rs380481920	2	6954280	1,0132E-09	221869	6732407	6732411	192007
CHA	surflid	rs799093918	2	6961728	1,21E-09	244523	6717201	6717205	192613
CHA	surflid	rs799093918	2	6961728	1,21E-09	244178	6717546	6717550	192752
CHA	surflid	rs799093918	2	6961728	1,21E-09	228771	6732953	6732957	192744
CHA	surflid	rs799093918	2	6961728	1,21E-09	228757	6732967	6732971	192793
CHA	surflid	rs799093918	2	6961728	1,21E-09	229317	6732407	6732411	192007
CHA	surflid	rs384675900	2	6961995	1,7289E-09	244790	6717201	6717205	192613
CHA	surflid	rs384675900	2	6961995	1,7289E-09	244445	6717546	6717550	192752
CHA	surflid	rs384675900	2	6961995	1,7289E-09	229038	6732953	6732957	192744
CHA	surflid	rs384675900	2	6961995	1,7289E-09	229024	6732967	6732971	192793
CHA	surflid	rs384675900	2	6961995	1,7289E-09	229584	6732407	6732411	192007
CHA	surflid	rs384121826	2	6967389	3,8616E-10	250184	6717201	6717205	192613
CHA	surflid	rs384121826	2	6967389	3,8616E-10	249839	6717546	6717550	192752
CHA	surflid	rs384121826	2	6967389	3,8616E-10	234432	6732953	6732957	192744
CHA	surflid	rs384121826	2	6967389	3,8616E-10	234418	6732967	6732971	192793
CHA	surflid	rs384121826	2	6967389	3,8616E-10	234978	6732407	6732411	192007
CHA	surflid	rs380905932	2	6971775	1,0521E-09	254570	6717201	6717205	192613
CHA	surflid	rs380905932	2	6971775	1,0521E-09	254225	6717546	6717550	192752
CHA	surflid	rs380905932	2	6971775	1,0521E-09	238818	6732953	6732957	192744
CHA	surflid	rs380905932	2	6971775	1,0521E-09	238804	6732967	6732971	192793
CHA	surflid	rs380905932	2	6971775	1,0521E-09	239364	6732407	6732411	192007
CHA	surflid	rs378877606	2	6984509	7,7705E-08	267304	6717201	6717205	192613
CHA	surflid	rs378877606	2	6984509	7,7705E-08	266959	6717546	6717550	192752
CHA	surflid	rs378877606	2	6984509	7,7705E-08	251552	6732953	6732957	192744
CHA	surflid	rs378877606	2	6984509	7,7705E-08	251538	6732967	6732971	192793
CHA	surflid	rs378877606	2	6984509	7,7705E-08	252098	6732407	6732411	192007
CHA	surflid	rs477065135	2	6988026	8,0428E-08	270821	6717201	6717205	192613
CHA	surflid	rs477065135	2	6988026	8,0428E-08	270476	6717546	6717550	192752
CHA	surflid	rs477065135	2	6988026	8,0428E-08	255069	6732953	6732957	192744
CHA	surflid	rs477065135	2	6988026	8,0428E-08	255055	6732967	6732971	192793
CHA	surflid	rs477065135	2	6988026	8,0428E-08	255615	6732407	6732411	192007
CHA	surflid	rs383891876	2	6989886	8,9034E-08	272681	6717201	6717205	192613
CHA	surflid	rs383891876	2	6989886	8,9034E-08	272336	6717546	6717550	192752
CHA	surflid	rs383891876	2	6989886	8,9034E-08	256929	6732953	6732957	192744
CHA	surflid	rs383891876	2	6989886	8,9034E-08	256915	6732967	6732971	192793
CHA	surflid	rs383891876	2	6989886	8,9034E-08	257475	6732407	6732411	192007
CHA	surflid	rs385542897	2	6992981	6,317E-08	275776	6717201	6717205	192613
CHA	surflid	rs385542897	2	6992981	6,317E-08	275431	6717546	6717550	192752
CHA	surflid	rs385542897	2	6992981	6,317E-08	260024	6732953	6732957	192744
CHA	surflid	rs385542897	2	6992981	6,317E-08	260010	6732967	6732971	192793
CHA	surflid	rs385542897	2	6992981	6,317E-08	260570	6732407	6732411	192007
CHA	surflid	rs383633166	2	6997592	2,2574E-08	280387	6717201	6717205	192613
CHA	surflid	rs383633166	2	6997592	2,2574E-08	280042	6717546	6717550	192752
CHA	surflid	rs383633166	2	6997592	2,2574E-08	264635	6732953	6732957	192744

CHA	surflid	rs383633166	2	6997592	2,2574E-08	264621	6732967	6732971	192793
CHA	surflid	rs383633166	2	6997592	2,2574E-08	265181	6732407	6732411	192007
CHA	surflid	rs385706443	2	6243958	1,8241E-08	443043	5800911	5800915	192847
CHA	surflid	rs385706443	2	6243958	1,8241E-08	445056	5798898	5798902	192856
CHA	surflid	rs385706443	2	6243958	1,8241E-08	4477	6248435	6248439	164702
CHA	surflid	rs385706443	2	6243958	1,8241E-08	14418	6258376	6258380	192586
CHA	surflid	rs385706443	2	6243958	1,8241E-08	9101	6253059	6253063	192652
CHA	surflid	rs385706443	2	6243958	1,8241E-08	25844	6269802	6269806	192617
CHA	surflid	rs385706443	2	6243958	1,8241E-08	31453	6275411	6275415	192144
CHA	surflid	rs385706443	2	6243958	1,8241E-08	473243	6717201	6717205	192613
CHA	surflid	rs385706443	2	6243958	1,8241E-08	473588	6717546	6717550	192752
CHA	surflid	rs385706443	2	6243958	1,8241E-08	488995	6732953	6732957	192744
CHA	surflid	rs385706443	2	6243958	1,8241E-08	489009	6732967	6732971	192793
CHA	surflid	rs385706443	2	6243958	1,8241E-08	488449	6732407	6732411	192007
CHA	surflid	rs381363353	2	6244385	5,6306E-08	443470	5800911	5800915	192847
CHA	surflid	rs381363353	2	6244385	5,6306E-08	445483	5798898	5798902	192856
CHA	surflid	rs381363353	2	6244385	5,6306E-08	4050	6248435	6248439	164702
CHA	surflid	rs381363353	2	6244385	5,6306E-08	13991	6258376	6258380	192586
CHA	surflid	rs381363353	2	6244385	5,6306E-08	8674	6253059	6253063	192652
CHA	surflid	rs381363353	2	6244385	5,6306E-08	25417	6269802	6269806	192617
CHA	surflid	rs381363353	2	6244385	5,6306E-08	31026	6275411	6275415	192144
CHA	surflid	rs381363353	2	6244385	5,6306E-08	472816	6717201	6717205	192613
CHA	surflid	rs381363353	2	6244385	5,6306E-08	473161	6717546	6717550	192752
CHA	surflid	rs381363353	2	6244385	5,6306E-08	488568	6732953	6732957	192744
CHA	surflid	rs381363353	2	6244385	5,6306E-08	488582	6732967	6732971	192793
CHA	surflid	rs381363353	2	6244385	5,6306E-08	488022	6732407	6732411	192007
CHA	surflid	rs209416653	2	6247291	1,2069E-08	446376	5800911	5800915	192847
CHA	surflid	rs209416653	2	6247291	1,2069E-08	448389	5798898	5798902	192856
CHA	surflid	rs209416653	2	6247291	1,2069E-08	1144	6248435	6248439	164702
CHA	surflid	rs209416653	2	6247291	1,2069E-08	11085	6258376	6258380	192586
CHA	surflid	rs209416653	2	6247291	1,2069E-08	5768	6253059	6253063	192652
CHA	surflid	rs209416653	2	6247291	1,2069E-08	22511	6269802	6269806	192617
CHA	surflid	rs209416653	2	6247291	1,2069E-08	28120	6275411	6275415	192144
CHA	surflid	rs209416653	2	6247291	1,2069E-08	469910	6717201	6717205	192613
CHA	surflid	rs209416653	2	6247291	1,2069E-08	470255	6717546	6717550	192752
CHA	surflid	rs209416653	2	6247291	1,2069E-08	485662	6732953	6732957	192744
CHA	surflid	rs209416653	2	6247291	1,2069E-08	485676	6732967	6732971	192793
CHA	surflid	rs209416653	2	6247291	1,2069E-08	485116	6732407	6732411	192007
CHA	surflid	rs210222267	2	6247555	1,3174E-08	446640	5800911	5800915	192847
CHA	surflid	rs210222267	2	6247555	1,3174E-08	448653	5798898	5798902	192856
CHA	surflid	rs210222267	2	6247555	1,3174E-08	880	6248435	6248439	164702
CHA	surflid	rs210222267	2	6247555	1,3174E-08	10821	6258376	6258380	192586
CHA	surflid	rs210222267	2	6247555	1,3174E-08	5504	6253059	6253063	192652
CHA	surflid	rs210222267	2	6247555	1,3174E-08	22247	6269802	6269806	192617
CHA	surflid	rs210222267	2	6247555	1,3174E-08	27856	6275411	6275415	192144
CHA	surflid	rs210222267	2	6247555	1,3174E-08	469646	6717201	6717205	192613

CHA	surflid	rs210222267	2	6247555	1,3174E-08	469991	6717546	6717550	192752
CHA	surflid	rs210222267	2	6247555	1,3174E-08	485398	6732953	6732957	192744
CHA	surflid	rs210222267	2	6247555	1,3174E-08	485412	6732967	6732971	192793
CHA	surflid	rs210222267	2	6247555	1,3174E-08	484852	6732407	6732411	192007
CHA	surflid	rs208490483	2	6248085	6,7061E-09	447170	5800911	5800915	192847
CHA	surflid	rs208490483	2	6248085	6,7061E-09	449183	5798898	5798902	192856
CHA	surflid	rs208490483	2	6248085	6,7061E-09	350	6248435	6248439	164702
CHA	surflid	rs208490483	2	6248085	6,7061E-09	10291	6258376	6258380	192586
CHA	surflid	rs208490483	2	6248085	6,7061E-09	4974	6253059	6253063	192652
CHA	surflid	rs208490483	2	6248085	6,7061E-09	21717	6269802	6269806	192617
CHA	surflid	rs208490483	2	6248085	6,7061E-09	27326	6275411	6275415	192144
CHA	surflid	rs208490483	2	6248085	6,7061E-09	469116	6717201	6717205	192613
CHA	surflid	rs208490483	2	6248085	6,7061E-09	469461	6717546	6717550	192752
CHA	surflid	rs208490483	2	6248085	6,7061E-09	484868	6732953	6732957	192744
CHA	surflid	rs208490483	2	6248085	6,7061E-09	484882	6732967	6732971	192793
CHA	surflid	rs208490483	2	6248085	6,7061E-09	484322	6732407	6732411	192007
CHA	surflid	rs109558564	2	6248437	1,4732E-08	447522	5800911	5800915	192847
CHA	surflid	rs109558564	2	6248437	1,4732E-08	449535	5798898	5798902	192856
CHA	surflid	rs109558564	2	6248437	1,4732E-08	2	6248435	6248439	164702
CHA	surflid	rs109558564	2	6248437	1,4732E-08	9939	6258376	6258380	192586
CHA	surflid	rs109558564	2	6248437	1,4732E-08	4622	6253059	6253063	192652
CHA	surflid	rs109558564	2	6248437	1,4732E-08	21365	6269802	6269806	192617
CHA	surflid	rs109558564	2	6248437	1,4732E-08	26974	6275411	6275415	192144
CHA	surflid	rs109558564	2	6248437	1,4732E-08	468764	6717201	6717205	192613
CHA	surflid	rs109558564	2	6248437	1,4732E-08	469109	6717546	6717550	192752
CHA	surflid	rs109558564	2	6248437	1,4732E-08	484516	6732953	6732957	192744
CHA	surflid	rs109558564	2	6248437	1,4732E-08	484530	6732967	6732971	192793
CHA	surflid	rs109558564	2	6248437	1,4732E-08	483970	6732407	6732411	192007
CHA	surflid	rs132828579	2	6249305	1,2324E-08	448390	5800911	5800915	192847
CHA	surflid	rs132828579	2	6249305	1,2324E-08	450403	5798898	5798902	192856
CHA	surflid	rs132828579	2	6249305	1,2324E-08	866	6248435	6248439	164702
CHA	surflid	rs132828579	2	6249305	1,2324E-08	9071	6258376	6258380	192586
CHA	surflid	rs132828579	2	6249305	1,2324E-08	3754	6253059	6253063	192652
CHA	surflid	rs132828579	2	6249305	1,2324E-08	20497	6269802	6269806	192617
CHA	surflid	rs132828579	2	6249305	1,2324E-08	26106	6275411	6275415	192144
CHA	surflid	rs132828579	2	6249305	1,2324E-08	467896	6717201	6717205	192613
CHA	surflid	rs132828579	2	6249305	1,2324E-08	468241	6717546	6717550	192752
CHA	surflid	rs132828579	2	6249305	1,2324E-08	483648	6732953	6732957	192744
CHA	surflid	rs132828579	2	6249305	1,2324E-08	483662	6732967	6732971	192793
CHA	surflid	rs132828579	2	6249305	1,2324E-08	483102	6732407	6732411	192007
CHA	surflid	rs210856936	2	6249553	1,2053E-08	448638	5800911	5800915	192847
CHA	surflid	rs210856936	2	6249553	1,2053E-08	450651	5798898	5798902	192856
CHA	surflid	rs210856936	2	6249553	1,2053E-08	1114	6248435	6248439	164702
CHA	surflid	rs210856936	2	6249553	1,2053E-08	8823	6258376	6258380	192586
CHA	surflid	rs210856936	2	6249553	1,2053E-08	3506	6253059	6253063	192652
CHA	surflid	rs210856936	2	6249553	1,2053E-08	20249	6269802	6269806	192617

CHA	surflid	rs210856936	2	6249553	1,2053E-08	25858	6275411	6275415	192144
CHA	surflid	rs210856936	2	6249553	1,2053E-08	467648	6717201	6717205	192613
CHA	surflid	rs210856936	2	6249553	1,2053E-08	467993	6717546	6717550	192752
CHA	surflid	rs210856936	2	6249553	1,2053E-08	483400	6732953	6732957	192744
CHA	surflid	rs210856936	2	6249553	1,2053E-08	483414	6732967	6732971	192793
CHA	surflid	rs210856936	2	6249553	1,2053E-08	482854	6732407	6732411	192007
CHA	surflid	rs110018083	2	6249857	1,307E-08	448942	5800911	5800915	192847
CHA	surflid	rs110018083	2	6249857	1,307E-08	450955	5798898	5798902	192856
CHA	surflid	rs110018083	2	6249857	1,307E-08	1418	6248435	6248439	164702
CHA	surflid	rs110018083	2	6249857	1,307E-08	8519	6258376	6258380	192586
CHA	surflid	rs110018083	2	6249857	1,307E-08	3202	6253059	6253063	192652
CHA	surflid	rs110018083	2	6249857	1,307E-08	19945	6269802	6269806	192617
CHA	surflid	rs110018083	2	6249857	1,307E-08	25554	6275411	6275415	192144
CHA	surflid	rs110018083	2	6249857	1,307E-08	467344	6717201	6717205	192613
CHA	surflid	rs110018083	2	6249857	1,307E-08	467689	6717546	6717550	192752
CHA	surflid	rs110018083	2	6249857	1,307E-08	483096	6732953	6732957	192744
CHA	surflid	rs110018083	2	6249857	1,307E-08	483110	6732967	6732971	192793
CHA	surflid	rs110018083	2	6249857	1,307E-08	482550	6732407	6732411	192007
CHA	surflid	rs110204889	2	6251286	1,2992E-08	450371	5800911	5800915	192847
CHA	surflid	rs110204889	2	6251286	1,2992E-08	452384	5798898	5798902	192856
CHA	surflid	rs110204889	2	6251286	1,2992E-08	2847	6248435	6248439	164702
CHA	surflid	rs110204889	2	6251286	1,2992E-08	7090	6258376	6258380	192586
CHA	surflid	rs110204889	2	6251286	1,2992E-08	1773	6253059	6253063	192652
CHA	surflid	rs110204889	2	6251286	1,2992E-08	18516	6269802	6269806	192617
CHA	surflid	rs110204889	2	6251286	1,2992E-08	24125	6275411	6275415	192144
CHA	surflid	rs110204889	2	6251286	1,2992E-08	465915	6717201	6717205	192613
CHA	surflid	rs110204889	2	6251286	1,2992E-08	466260	6717546	6717550	192752
CHA	surflid	rs110204889	2	6251286	1,2992E-08	481667	6732953	6732957	192744
CHA	surflid	rs110204889	2	6251286	1,2992E-08	481681	6732967	6732971	192793
CHA	surflid	rs110204889	2	6251286	1,2992E-08	481121	6732407	6732411	192007
CHA	surflid	rs135709917	2	6252224	2,1543E-08	451309	5800911	5800915	192847
CHA	surflid	rs135709917	2	6252224	2,1543E-08	453322	5798898	5798902	192856
CHA	surflid	rs135709917	2	6252224	2,1543E-08	3785	6248435	6248439	164702
CHA	surflid	rs135709917	2	6252224	2,1543E-08	6152	6258376	6258380	192586
CHA	surflid	rs135709917	2	6252224	2,1543E-08	835	6253059	6253063	192652
CHA	surflid	rs135709917	2	6252224	2,1543E-08	17578	6269802	6269806	192617
CHA	surflid	rs135709917	2	6252224	2,1543E-08	23187	6275411	6275415	192144
CHA	surflid	rs135709917	2	6252224	2,1543E-08	464977	6717201	6717205	192613
CHA	surflid	rs135709917	2	6252224	2,1543E-08	465322	6717546	6717550	192752
CHA	surflid	rs135709917	2	6252224	2,1543E-08	480729	6732953	6732957	192744
CHA	surflid	rs135709917	2	6252224	2,1543E-08	480743	6732967	6732971	192793
CHA	surflid	rs135709917	2	6252224	2,1543E-08	480183	6732407	6732411	192007
CHA	surflid	rs381889129	2	6256818	4,3093E-09	455903	5800911	5800915	192847
CHA	surflid	rs381889129	2	6256818	4,3093E-09	457916	5798898	5798902	192856
CHA	surflid	rs381889129	2	6256818	4,3093E-09	8379	6248435	6248439	164702
CHA	surflid	rs381889129	2	6256818	4,3093E-09	1558	6258376	6258380	192586

CHA	surflid	rs381889129	2	6256818	4,3093E-09	3755	6253059	6253063	192652
CHA	surflid	rs381889129	2	6256818	4,3093E-09	12984	6269802	6269806	192617
CHA	surflid	rs381889129	2	6256818	4,3093E-09	18593	6275411	6275415	192144
CHA	surflid	rs381889129	2	6256818	4,3093E-09	460383	6717201	6717205	192613
CHA	surflid	rs381889129	2	6256818	4,3093E-09	460728	6717546	6717550	192752
CHA	surflid	rs381889129	2	6256818	4,3093E-09	476135	6732953	6732957	192744
CHA	surflid	rs381889129	2	6256818	4,3093E-09	476149	6732967	6732971	192793
CHA	surflid	rs381889129	2	6256818	4,3093E-09	475589	6732407	6732411	192007
CHA	surflid	rs210732995	2	6259967	1,5786E-08	459052	5800911	5800915	192847
CHA	surflid	rs210732995	2	6259967	1,5786E-08	461065	5798898	5798902	192856
CHA	surflid	rs210732995	2	6259967	1,5786E-08	11528	6248435	6248439	164702
CHA	surflid	rs210732995	2	6259967	1,5786E-08	1587	6258376	6258380	192586
CHA	surflid	rs210732995	2	6259967	1,5786E-08	6904	6253059	6253063	192652
CHA	surflid	rs210732995	2	6259967	1,5786E-08	9835	6269802	6269806	192617
CHA	surflid	rs210732995	2	6259967	1,5786E-08	15444	6275411	6275415	192144
CHA	surflid	rs210732995	2	6259967	1,5786E-08	457234	6717201	6717205	192613
CHA	surflid	rs210732995	2	6259967	1,5786E-08	457579	6717546	6717550	192752
CHA	surflid	rs210732995	2	6259967	1,5786E-08	472986	6732953	6732957	192744
CHA	surflid	rs210732995	2	6259967	1,5786E-08	473000	6732967	6732971	192793
CHA	surflid	rs210732995	2	6259967	1,5786E-08	472440	6732407	6732411	192007
CHA	surflid	rs211362058	2	6262506	7,7122E-09	461591	5800911	5800915	192847
CHA	surflid	rs211362058	2	6262506	7,7122E-09	463604	5798898	5798902	192856
CHA	surflid	rs211362058	2	6262506	7,7122E-09	14067	6248435	6248439	164702
CHA	surflid	rs211362058	2	6262506	7,7122E-09	4126	6258376	6258380	192586
CHA	surflid	rs211362058	2	6262506	7,7122E-09	9443	6253059	6253063	192652
CHA	surflid	rs211362058	2	6262506	7,7122E-09	7296	6269802	6269806	192617
CHA	surflid	rs211362058	2	6262506	7,7122E-09	12905	6275411	6275415	192144
CHA	surflid	rs211362058	2	6262506	7,7122E-09	454695	6717201	6717205	192613
CHA	surflid	rs211362058	2	6262506	7,7122E-09	455040	6717546	6717550	192752
CHA	surflid	rs211362058	2	6262506	7,7122E-09	470447	6732953	6732957	192744
CHA	surflid	rs211362058	2	6262506	7,7122E-09	470461	6732967	6732971	192793
CHA	surflid	rs211362058	2	6262506	7,7122E-09	469901	6732407	6732411	192007
CHA	surflid	rs110077073	2	6271317	1,0831E-08	470402	5800911	5800915	192847
CHA	surflid	rs110077073	2	6271317	1,0831E-08	472415	5798898	5798902	192856
CHA	surflid	rs110077073	2	6271317	1,0831E-08	22878	6248435	6248439	164702
CHA	surflid	rs110077073	2	6271317	1,0831E-08	12937	6258376	6258380	192586
CHA	surflid	rs110077073	2	6271317	1,0831E-08	18254	6253059	6253063	192652
CHA	surflid	rs110077073	2	6271317	1,0831E-08	1511	6269802	6269806	192617
CHA	surflid	rs110077073	2	6271317	1,0831E-08	4094	6275411	6275415	192144
CHA	surflid	rs110077073	2	6271317	1,0831E-08	445884	6717201	6717205	192613
CHA	surflid	rs110077073	2	6271317	1,0831E-08	446229	6717546	6717550	192752
CHA	surflid	rs110077073	2	6271317	1,0831E-08	461636	6732953	6732957	192744
CHA	surflid	rs110077073	2	6271317	1,0831E-08	461650	6732967	6732971	192793
CHA	surflid	rs110077073	2	6271317	1,0831E-08	461090	6732407	6732411	192007
CHA	surflid	rs137753575	2	6272604	2,3198E-08	471689	5800911	5800915	192847
CHA	surflid	rs137753575	2	6272604	2,3198E-08	473702	5798898	5798902	192856

CHA	surflid	rs137753575	2	6272604	2,3198E-08	24165	6248435	6248439	164702
CHA	surflid	rs137753575	2	6272604	2,3198E-08	14224	6258376	6258380	192586
CHA	surflid	rs137753575	2	6272604	2,3198E-08	19541	6253059	6253063	192652
CHA	surflid	rs137753575	2	6272604	2,3198E-08	2798	6269802	6269806	192617
CHA	surflid	rs137753575	2	6272604	2,3198E-08	2807	6275411	6275415	192144
CHA	surflid	rs137753575	2	6272604	2,3198E-08	444597	6717201	6717205	192613
CHA	surflid	rs137753575	2	6272604	2,3198E-08	444942	6717546	6717550	192752
CHA	surflid	rs137753575	2	6272604	2,3198E-08	460349	6732953	6732957	192744
CHA	surflid	rs137753575	2	6272604	2,3198E-08	460363	6732967	6732971	192793
CHA	surflid	rs137753575	2	6272604	2,3198E-08	459803	6732407	6732411	192007
CHA	surflid	rs384338078	2	6273651	2,8065E-09	472736	5800911	5800915	192847
CHA	surflid	rs384338078	2	6273651	2,8065E-09	474749	5798898	5798902	192856
CHA	surflid	rs384338078	2	6273651	2,8065E-09	25212	6248435	6248439	164702
CHA	surflid	rs384338078	2	6273651	2,8065E-09	15271	6258376	6258380	192586
CHA	surflid	rs384338078	2	6273651	2,8065E-09	20588	6253059	6253063	192652
CHA	surflid	rs384338078	2	6273651	2,8065E-09	3845	6269802	6269806	192617
CHA	surflid	rs384338078	2	6273651	2,8065E-09	1760	6275411	6275415	192144
CHA	surflid	rs384338078	2	6273651	2,8065E-09	443550	6717201	6717205	192613
CHA	surflid	rs384338078	2	6273651	2,8065E-09	443895	6717546	6717550	192752
CHA	surflid	rs384338078	2	6273651	2,8065E-09	459302	6732953	6732957	192744
CHA	surflid	rs384338078	2	6273651	2,8065E-09	459316	6732967	6732971	192793
CHA	surflid	rs384338078	2	6273651	2,8065E-09	458756	6732407	6732411	192007
LIM	micoull	rs384248403	2	4611909	7,4223E-08	23078	4634987	4634991	36370
LIM	micoull	rs378731556	2	4868575	7,224E-08	233584	4634987	4634991	36370
CHA	mnotend	rs42192427	29	44190541	2,678E-08	44725	44235266	44235270	36480
CHA	mnotend	rs42193634	29	44198845	1,196E-08	36421	44235266	44235270	36480
LIM	gmq	rs723213524	21	20580674	6,1855E-08	86929	20667603	20667607	69170

Résumé

Résumé

Les méthodes classiques de sélection animale ne permettent pas d'améliorer la qualité sensoriel et organoleptique de la viande vu la complexité de prise des mesures en routine de ces caractères. L'utilisation de la génomique pourra donc être l'alternative pour améliorer ces caractères en cherchant des marqueurs moléculaires qui sont en fort déséquilibre de liaison avec les mutations causales qui expliquent une part de la variabilité phénotypique. Mais tout d'abord, une analyse de la structure du déséquilibre de liaison à l'échelle du génome pourra être intéressante d'un point de vue fondamental afin de comprendre à quel point les allèles sont transmis ensemble de génération en génération.

Ma thèse consiste à réaliser une étude du déséquilibre de liaison à longue distance (LRLD) dans les principales races bovines allaitantes françaises (Charolaise, Limousine et Blonde d'Aquitaine). Le deuxième volet de ma thèse a consisté à rechercher des variants associés à la qualité de la viande et aux aptitudes bouchères chez les principales races bovines allaitantes françaises à l'aide des données de séquence du génome complet.

Mots-clés : Qualité de la viande, QTL, SNP, GWAS, DL, LRLD

Abstract

Classical methods of animal selection do not improve the sensory and organoleptic quality of meat due to the complexity of taking routine measurements of these qualities. The use of genomics could therefore be the alternative to improve these traits by looking for molecular markers that are in strong linkage disequilibrium with the causal mutation that explains part of the phenotypic variability. But first of all, an analysis of the structure of the linkage disequilibrium at the genome level could be interesting from a fundamental point of view in order to understand to which extent the alleles are transmitted together from generation to generation.

My PhD thesis consists of a study of long-range linkage disequilibrium (LRLD) in the main French beef cattle breeds (Charolaise, Limousine and Blonde d'Aquitaine), in the framework of the GEMBAL project. The second part of my thesis consisted in searching for variants associated with meat quality and meat aptitude in the main French beef cattle breeds using whole genome sequence data. in the main French beef cattle breeds, in the framework of the Qualvigène project

Keywords : Meat quality, QTL, SNP, GWAS, LD, LRLD

