



HAL
open science

Détection de mutations létales chez les ovins. Mise en évidence de segments chromosomiques homozygotes à partir d'informations moléculaires haute densité (60 000 SNPs)

Maxime Ben Braiek

► To cite this version:

Maxime Ben Braiek. Détection de mutations létales chez les ovins. Mise en évidence de segments chromosomiques homozygotes à partir d'informations moléculaires haute densité (60 000 SNPs). *Génétique animale*. 2019. hal-02941680

HAL Id: hal-02941680

<https://hal.inrae.fr/hal-02941680v1>

Submitted on 17 Sep 2020

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

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

MINISTERE DE L'AGRICULTURE, DE L'AGROALIMENTAIRE ET DE LA FORET

ECOLE NATIONALE SUPERIEURE des SCIENCES AGRONOMIQUES de BORDEAUX
AQUITAINE

1, cours du Général de Gaulle - CS 40201 – 33175 GRADIGNAN cedex

MEMOIRE de fin d'études

pour l'obtention du titre

d'Ingénieur de Bordeaux Sciences Agro

Détection de mutations létales chez les ovins. Mise en évidence de segments chromosomiques homozygotes à partir d'informations moléculaires haute densité (60 000 SNPs).

Detection of lethal mutations in sheep. Research for homozygous chromosome segments with high density molecular information (60 000 SNPs)

BEN BRAIEK, Maxime

Spécialisation : Filières Animales Durables (FAD)

Étude réalisée à : Équipes GESPR (Génétique et Sélection des Petits Ruminants) & GenROC (Génomique des Ruminants Ovins et Caprins) - Institut National de la Recherche Agronomique, UMR GenPhySE, 24, chemin de Borde-Rouge – CS 52627 31320 Castanet-Tolosan

MINISTÈRE DE L'AGRICULTURE, DE L'AGROALIMENTAIRE ET DE LA FORÊT

ÉCOLE NATIONALE SUPÉRIEURE des SCIENCES AGRONOMIQUES de BORDEAUX
AQUITAINE

1, cours du Général de Gaulle - CS 40201 – 33175 GRADIGNAN cedex

MEMOIRE de fin d'études

pour l'obtention du titre

d'Ingénieur de Bordeaux Sciences Agro

Détection de mutations létales chez les ovins. Mise en évidence de segments chromosomiques homozygotes à partir d'informations moléculaires haute densité (60 000 SNP).

Detection of lethal mutations in sheep. Research for homozygous chromosome segments with high density molecular information (60 000 SNPs)

BEN BRAIEK, Maxime

Spécialisation : Filières Animales Durables (FAD)

Étude réalisée à : Équipes GESPR (Génétique et Sélection des Petits Ruminants) & GenROC (Génomique des Ruminants Ovins et Caprins) - Institut National de la Recherche Agronomique, UMR GenPhySE, 24, chemin de Borde-Rouge – CS 52627 31320 Castanet-Tolosan

Tuteur école : Alain Ducos

Maîtres de stage : Carole Moreno (équipe GESPR) et Stéphane Fabre (équipe GenROC)

Remerciements

Tout d'abord, je tiens tout particulièrement à remercier mes deux maîtres de stage, Carole Moreno-Romieux et Stéphane Fabre, pour leur encadrement et l'ensemble des nouvelles connaissances qu'ils ont pu me transmettre. Merci de me donner ma chance de poursuivre en thèse sur ce sujet si passionnant.

Merci à Chris Hoze et Sébastien Fritz pour votre accueil à l'INRA de Jouy-en-Josas, merci de m'avoir transmis vos connaissances sur le sujet.

Un grand merci à toi Jean-Michel Astruc, qui m'a fourni l'ensemble de mes données. Merci à l'ensemble des membres des équipes GESPR (notamment Hélène Larroque, Rachel Rupp, Isabelle Palhière) et GenROC, qui m'ont à un moment ou un autre, apporté des conseils très utiles pour avancer dans mes recherches.

Je tiens à remercier Alain Ducos, mon tuteur école, pour m'avoir suivi tout au long de ce stage, ainsi que l'ensemble de l'équipe pédagogique de la spécialisation Filières Animales Durables de Bordeaux Sciences Agro.

Merci à l'ensemble des thésards (Sophie, Audrey, Vanille, Tiphaine, Claire, Estelle, Louise), Harmonie, Océane, Chris et mon super compère Amaury, ainsi que l'ensemble des autres stagiaires pour ces moments inoubliables !!

Je tiens également à remercier la région Occitanie et APIS-GENE pour le financement de ma thèse.

Remerciements

Sommaire

Liste des illustrations

Liste des abréviations

Glossaire

Introduction..... 1

I. Synthèse bibliographique : les anomalies génétiques dans les populations d'élevage : de leur émergence à leur identification..... 2

A. Comprendre l'émergence des anomalies génétiques dans les populations ovines. 2

1. Données générales sur la filière ovine française : de l'organisation de la filière à l'utilisation de la sélection génomique 2

a. Description des cheptels ovins laitiers 2

b. Organisation de la sélection 2

c. De la sélection classique à la sélection génomique 4

i. Mise en place de la sélection génomique en ovins laitiers : construction d'une population de référence dans chaque race 4

ii. Stratégie de génotypage pour le choix des béliers reproducteurs 4

2. La structure des populations ovines pourrait conditionner l'émergence d'anomalies génétiques 5

a. Des populations avec des effectifs génétiques limités 5

b. Conséquence : augmentation de la consanguinité 5

c. Origine des anomalies génétiques : les mutations de l'ADN 6

d. Les mutations sont soumises à des mécanismes évolutifs 6

B. Le développement d'outils moléculaires facilite la détection des mutations responsables d'anomalies génétiques 7

1. Les techniques de génotypages 7

a. Les différents marqueurs utilisés en sélection 7

b. Utilisation des puces de génotypages haut-débit 7

2. Le séquençage des génomes, vers une meilleure connaissance moléculaire 8

C. Comment détecter les anomalies génétiques ? 8

1. Approche « *top-down* » : du phénotype à la mutation causale 9

a. Les observatoires d'anomalies génétiques, un système de surveillance des anomalies 9

b. Du phénotype à la mutation causale : la cartographie par homozygotie 9

2. Approche « *bottom-up* » : raisonnements de génétique inverse 10

a. Détection d'haplotypes en déficit d'homozygotie 10

b. Détection de variants délétères dans les données de séquences complètes 12

3. Prise en compte des anomalies génétiques dans les schémas de sélection 12

D.	Situation du sujet	13
II.	Matériel et Méthodes	14
A.	Description des animaux génotypés disponibles pour l'étude	14
1.	Description des animaux étudiés	14
2.	Contrôle qualité et phasage des marqueurs SNP	14
3.	Imputation des génotypages 15K vers la 50K	14
B.	Recherche d'haplotypes en déficit d'homozygotie.....	15
1.	Principe de la recherche d'haplotype en déficit d'animaux homozygotes	15
2.	Assignation des génotypes aux marqueurs contenus dans les haplotypes pour l'ensemble des individus de la population totale génotypée.....	16
3.	Calcul du déséquilibre de liaison entre haplotypes en déficit présents sur un même chromosome.....	16
C.	Recherche de mutations causales et de gènes candidats dans les régions du génome présentant un déficit en homozygotie.....	16
1.	Recherche de variants candidats présents dans les régions des haplotypes en déficit identifiés.....	16
2.	Recherche de gènes candidats fonctionnels et positionnels pour les haplotypes n'ayant pas d'animaux porteurs séquencés.....	17
D.	Effet des haplotypes sur les caractères sous sélection et phénotypes enregistrés en routine	17
1.	Analyses statistiques	17
a.	Statistiques descriptives.....	17
b.	Analyse de variance multifactorielle par l'utilisation de modèle linéaire	18
2.	Effet des haplotypes sur les caractères de production	18
3.	Effet des haplotypes sur la fertilité	19
4.	Effet des haplotypes sur la qualité de semences	19
5.	Ajustement des seuils de significativité par une correction de Bonferroni	19
III.	Résultats.....	20
A.	Recherche de déficit en homozygotie en races Lacaune et Manech à Tête Rousse	20
1.	Identification des haplotypes en déficit d'homozygotie	20
2.	Étude du lien entre les haplotypes présents sur un même chromosome	21
3.	Détermination des génotypes aux différents loci identifiés pour l'ensemble des individus de la population génotypée	22
B.	Recherche de mutations causales dans les loci identifiés	22
1.	Exploration des données de séquences tout génome en race Lacaune	22
a.	Identification de mutations candidates associées aux haplotypes présentant un déficit total d'individus homozygotes.....	23
b.	Identification de mutations candidates associées aux haplotypes présentant un déficit partiel d'individus homozygotes.....	24

c. Cas particuliers : identification de deux gènes sous sélection dans les régions LL_OAR3_2 et LL_OAR13_1	24
i. Lien entre LL_OAR3_2a et le gène de résistance aux mammites <i>SOCS2</i> en race Lacaune	24
ii. Lien entre LL_OAR13_1a et le gène de résistance/sensibilité à la tremblante <i>PRNP</i> en race Lacaune.....	25
2. Identification de gènes candidats dans les régions ne disposant pas d'animaux séquencés porteurs des haplotypes en déficit.....	25
C. Effet des haplotypes identifiés sur les performances	26
1. Étude d'un potentiel avantage des individus hétérozygotes sur les caractères en sélection	26
2. Évaluation de l'impact des haplotypes sur l'efficacité reproductive mâle	27
a. Étude de la fertilité.....	27
b. Qualités de semences	27
IV. Discussion et perspectives.....	27
A. Mise en évidence d'haplotypes présentant des déficits en homozygotie	28
1. Identification de gènes et de polymorphismes candidats pour la létalité embryonnaire	28
2. Identification de mutations affectant le jeune animal	30
3. Identification d'une mutation en contre-sélection : cas du gène de résistance aux mammites <i>SOCS2</i>	30
B. Réflexion sur la méthode de recherche de déficit employée	31
1. Recherche de déficits en homozygotie	31
2. Recherche de mutations causales.....	32
C. Perspectives	33
Conclusion	34
Références bibliographiques	36
Annexes	44
Liste des annexes	45

Liste des illustrations

FIGURES

FIGURE 1. BELIER LACAUNE.....	1
FIGURE 2. BELIER MANECH A TETE ROUSSE.....	1
FIGURE 3. GESTION PYRAMIDALE DU SCHEMA DE SELECTION EN OVIN LAITIER (SELECTION GENOMIQUE).....	3
FIGURE 4. ÉVOLUTION DES CRITERES DE SELECTION EN RACES LACAUNE ET MANECH A TETE ROUSSE	3
FIGURE 5. MARQUEURS SNP PRESENTS SUR LA PUCE BASSE DENSITE.	4
FIGURE 6. STRATEGIE DE GENOTYPAGE ET CAUSES D'ECART A LA SELECTION EN OVIN LAITIER	4
FIGURE 7. ÉVOLUTION DE LA CONSANGUINITE MOYENNE EN RACES OVINES LAITIERES DES BELIERS DU NOYAU DE SELECTION NES ENTRE 1970 ET 2013.	5
FIGURE 8. LES MUTATIONS PONCTUELLES DE L'ADN AFFECTANT LES REGIONS CODANTES	6
FIGURE 9. TRANSMISSION MENDELIENNE RECESSIVE D'UNE ANOMALIE GENETIQUE..	6
FIGURE 10. ÉVOLUTION DE LA FREQUENCE DE L'ALLELE BL A L'ORIGINE DE L'ANOMALIE BLAD EN RACE HOLSTEIN FRANÇAISE CHEZ LES VEAUX, LES TAUREAUX ET LES VACHES.	7
FIGURE 11. DIFFERENTS TYPES DE VARIATIONS NUCLEOTIDIQUES	7
FIGURE 12. FONCTIONNEMENT D'UNE PUCE SNP	8
FIGURE 13. PRINCIPE DU SEQUENÇAGE HAUT-DEBIT PAR LA TECHNOLOGIE ILLUMINA (NEW GENERATION SEQUENCING).....	9
FIGURE 14. NOMBRE DE MUTATIONS CAUSALES LIEES A UNE ANOMALIE GENETIQUE AVEC UN DETERMINISME MENDELIEN DANS LES DIFFERENTES ESPECES D'ELEVAGE.	10
FIGURE 15. PRINCIPE DE LA CARTOGRAPHIE PAR HOMOZYGOTIE	10
FIGURE 16. CARTOGRAPHIE PAR HOMOZYGOTIE PERMETTANT DE LOCALISER L'HAPLOTYPE HOMOZYGOTE ASSOCIE LA MUTATION DANS LE GENE FANCI.	11
FIGURE 17. CAS D'UN AGNEAU AFFECTE PAR L'EPIDERMOLYSE BULLEUSE EN RACE OVINE CHURRA	11
FIGURE 18. METHODOLOGIE POUR IDENTIFIER DES MUTATIONS RECESSIVES DELETERES PAR UN RAISONNEMENT DE GENETIQUE INVERSE EN UTILISANT LES DONNEES DE SEQUENCES DE GENOMES COMPLETS,	13
FIGURE 19. A. PRINCIPE DE L'IMPUTATION, B. PRINCIPE DU PHASAGE..	14
FIGURE 20. PRINCIPE DU BALAYAGE DU GENOME PAR UNE FENETRE GLISSANTE	15
FIGURE 21. ESTIMATION DU NOMBRE D'ANIMAUX HOMOZYGOTES ATTENDUS A L'HAPLOTYPE K.....	15
FIGURE 22. DESCRIPTION DES DIFFERENTS POSTES DE POINTAGES DE MORPHOLOGIE MAMMAIRE MESURES EN RACE LACAUNE	18
FIGURE 23. METHODE POUR DETERMINER LES HAPLOTYPES CONCATENES EN DEFICIT EN HOMOZYGOTIE A PARTIR DES HAPLOTYPES DE 10 ET 20 MARQUEURS.	20
FIGURE 24. REPRESENTATION DES REGIONS DE LOCALISATION PROBABLE DE MUTATIONS CAUSALES OU L'ON DETECTE DES DEFICITS EN HOMOZYGOTIE EN RACES LACAUNE ET MANECH A TETE ROUSSE.....	20
FIGURE 25. IDENTIFICATION DE DEUX MUTATIONS CANDIDATES DANS LES GENES PREB (PROLACTIN REGULATORY ELEMENT BINDING) ET CCDC65 (COILED-COIL DOMAIN CONTAINING 65) (OAR3) EN RACE LACAUNE. ...	24
FIGURE 26. GENOTYPES AU VARIANT (C>T ENTRAINANT LA SUBSTITUTION NON CONSERVATIVE R96C) DU GENE SOCS2 (OAR3) DES 6 ANIMAUX SEQUENCES TOUT GENOME, PORTEURS DE L'ALLELE T (ISSUS DES SEQUENCES TOUT GENOME DES 68 ANIMAUX DISPONIBLES AU LABORATOIRE)	30

TABLEAUX

TABLEAU 1. PRESENTATION DES PRINCIPALES RACES OVINES LAITIÈRES FRANÇAISES.....	2
TABLEAU 2. COMPARAISON DE LA TAILLE EFFICACE NE (GENOMIQUE ET GENEALOGIQUE) D'UNE RACE LAITIÈRE BOVINE ET DE 4 RACES LAITIÈRES OVINES	5
TABLEAU 3. PRINCIPALES MUTATIONS CAUSALES IDENTIFIÉES PAR DES RAISONNEMENTS DE GENÉTIQUE INVERSE	12
TABLEAU 4. DESCRIPTION DES INDIVIDUS GENOTYPES DE L'ÉTUDE EN RACES LACAUNE ET MANECH A TÊTE ROUSSE.....	14
TABLEAU 5. DESCRIPTION DES ANIMAUX UTILES POUR LA RECHERCHE DE RÉGIONS EN DÉFICIT EN HOMOZYGOTIE EN RACES LACAUNE ET MANECH A TÊTE ROUSSE.....	16
TABLEAU 6. TABLE DE CALCUL DU DÉSÉQUILIBRE DE LIAISON ENTRE DEUX LOCI 1 ET 2	16
TABLEAU 7. DESCRIPTION DES JEUX DE DONNÉES DES DYD (« DAUGHTER YIELD DEVIATION ») POUR LES BELIERS GENOTYPES EN SÉLECTION EN RACES LACAUNE (A) ET MANECH A TÊTE ROUSSE (B).....	18
TABLEAU 8. DONNÉES DE FERTILITÉ A L'IA ET FERTILITÉ GLOBALE POUR LES BELIERS GENOTYPES EN SÉLECTION EN RACES LACAUNE ET MANECH A TÊTE ROUSSE.....	19
TABLEAU 9. DONNÉES DE SEMENCES DU PREMIER PRÉLEVEMENT EN CIA OBTENUES EN RACES LACAUNE ET MANECH A TÊTE ROUSSE POUR LES BELIERS GENOTYPES DE L'ÉTUDE.....	19
TABLEAU 10. DONNÉES DE SEMENCES MOYENNES DES TROIS PREMIERS PRÉLEVEMENTS EN CIA OBTENUES EN RACES LACAUNE ET MANECH A TÊTE ROUSSE POUR LES BELIERS GENOTYPES DE L'ÉTUDE.	19
TABLEAU 11. DESCRIPTION DES HAPLOTYPES CONCATENÉS EN DÉFICIT D'ANIMAUX HOMOZYGOTES OBTENUS EN RACES LACAUNE (LL) ET MANECH A TÊTE ROUSSE (MTR).	21
TABLEAU 12. CALCUL DU DÉSÉQUILIBRE DE LIAISON ENTRE LES DIFFÉRENTS HAPLOTYPES EN DÉFICIT PRÉSENTS SUR UN MÊME CHROMOSOME EN RACES LACAUNE (A. OAR3 ; B. OAR18) ET MANECH A TÊTE ROUSSE (C. OAR1).....	22
TABLEAU 13. ASSIGNATION DES GENOTYPES AUX MARQUEURS CONSTITUTIFS DES DIFFÉRENTS HAPLOTYPES PRÉSENTS DANS LES LOCI IDENTIFIÉS DANS L'ENSEMBLE DE LA POPULATION GENOTYPÉE.....	22
TABLEAU 14. DESCRIPTION DES RESULTATS DE LA RECHERCHE DE VARIANTS DANS LES 8 RÉGIONS DONT L'ON DISPOSE D'ANIMAUX PORTEURS SÉQUENCES EN RACE LACAUNE.	23
TABLEAU 15. LISTE DES MUTATIONS CANDIDATES IDENTIFIÉES DANS LES RÉGIONS DE LOCALISATION DES HAPLOTYPES EN DÉFICIT EN RACE LACAUNE.	23
TABLEAU 16. LIEN ENTRE L'HAPLOTYPE LL_OAR3_2A ET L'ALLELE MUTE T DU GÈNE SOCS2 (OAR3) EN RACE LACAUNE.....	25
TABLEAU 17. LISTE DES HAPLOTYPES ASSOCIÉS A LA MUTATION SOCS2 (OAR3) PRÉSENTS DANS LA RÉGION DE LL_OAR3_2 EN RACE LACAUNE	25
TABLEAU 18. IDENTIFICATION DE 3 HAPLOTYPES LIÉS A L'ALLELE MUTE T DU GÈNE SOCS2 (OAR3) EN RACE LACAUNE.....	25
TABLEAU 19. LIEN ENTRE L'HAPLOTYPE LL_OAR3_2_T ET L'ALLELE MUTE T DU SOCS2 (OAR3) EN RACE LACAUNE.....	25
TABLEAU 20. LIEN ENTRE L'HAPLOTYPE LL_OAR13_1A ET L'ALLELE Q DE SENSIBILITÉ A LA TREMBLANTE (EXON 171) DU GÈNE PRNP (OAR13) EN RACE LACAUNE.....	26

TABLEAU 21. CALCUL DE LA FREQUENCE DE L'ALLELE Q (CODON 171 DU GENE PRNP) EN UTILISANT LE CALCUL D'EQUILIBRE DE HARDY-WEINBERG EN RACE LACAUNE	26
TABLEAU 22. IDENTIFICATION DES PRINCIPAUX GENES CANDIDATS FONCTIONNELS LOCALISES DANS LES REGIONS PROBABLEMENT ASSOCIEES A DES MUTATIONS LETALES NE DISPOSANT PAS D'ANIMAUX SEQUENCES (DEFICIT 100%) EN RACES LACAUNE ET MANECH A TETE ROUSSE.	26
TABLEAU 23. ÉTUDE D'UN POTENTIEL AVANTAGE SELECTIF ENTRE INDIVIDUS PORTEURS HETEROZYGOTES ET NON PORTEURS DES HAPLOTYPES IDENTIFIES EN RACES LACAUNE ET MANECH A TETE ROUSSE.....	27
TABLEAU 24. ÉTUDE D'UN POTENTIEL AVANTAGE SELECTIF ENTRE INDIVIDUS PORTEURS ET NON PORTEURS DES HAPLOTYPES ASSOCIES AU VARIANT SOCS2 (OAR3).	28
TABLEAU 25. DESEQUILIBRE DE LIAISON ENTRE L'HAPLOTYPE LL_OAR3_4A ET LES HAPLOTYPES ASSOCIES A LA MUTATION SOCS2 (OAR3).....	28

Liste des abréviations

50K	: Puce moyenne densité de 54 241 SNP
ADN	: Acide D éoxyribo N ucléique
CCS	: Comptage Cellules Somatiques
CDEO	: Centre D épartemental de l'Élevage O vin
CE	: Centre d'Élevage des jeunes mâles
Chr	: C hromosome
CIA	: Centre d'Insémination Animale
CIIRPO	: Centre I nterrégional d' I nformation et de R echerche en P roduction O vine
CLO	: Contrôle Laitier O fficiel
CNBL	: Comité National des B rebis L aitières
CORAM	: Collectif des R aces locales des M assifs
ddNTP	: d idésoxyribonucléotide
DYD	: <i>Daughter Yield Deviation</i>
ES	: Entreprise de Sélection
FANCI	: <i>FA Complementation Group I</i>
FGE	: France G énétique É levage
GART	: <i>Phosphoribosylglycinamide Formyltransferase</i>
Hap	: H aplotype
IA	: Insémination Animale
IDELE	: Institut d e l'Élevage
InDels	: I nsertion- D élétion
INRA	: Institut National de la R echerche A gronomique
ISOL	: Index de Synthèse O vin L ait
ITGB4	: <i>Integrin Subunit Beta 4</i>
kb	: kilobases
LD/15K	: <i>Low Density</i> , puce de 15000 SNP
LL	: Lacaune L ait
MAF	: <i>Minimum Allele Frequency</i>
Mb	: M égabases
MN	: Monte N aturelle
MTR	: Manech à T ête R ousse
Ne	: Taille efficace
NGS	: <i>Next Generation Sequencing</i>
NOAI	: <i>Nitric Oxide Associated 1</i>

NGS : *Next Generation Sequencing*

NOAI : *Nitric Oxide Associated 1*

OAR : Chromosome (*Ovis aries*)

OMIA : **O**nline **M**endelian **I**nheritance in **A**nimals

OMIM : **O**nline **M**endelian **I**nheritance in **M**an

ONAB : **O**bservatoire National des **A**nomalies **B**ovines

OS : **O**rganisme de **S**élection

PAB : **P**ère à **B**éliers

PAF : **P**ères à **F**illes

pb : **p**aires de **b**ases

PCR : *Polymerase Chain Reaction*

pG : **P**ression de **S**élection **g**énomique

PNAGRT : **P**rogramme National d'**A**mélioration **G**énétique pour la **R**ésistance à la Tremblante

QMG : **Q**uantité de **M**atière **G**rasse

QMP : **Q**uantité de **M**atière **P**rotéique

QTL : **Q**uantitative **T**rait **L**oci

ROLP : **R**aces **O**vines **L**aitières des **P**yrénées

SLC37A2 : *Solute Carrier Family 37 Member 2*

SMC2 : *Structural Maintenance Of Chromosome 2*

SNP : *Single Nucleotide Polymorphism*

Spz : Spermatozoïdes

TB : **T**aux **B**utyreux

TP : **T**aux **P**rotéique

UTR : **U**ntranslated **R**egions

VCF : *Variant Calling Format*

1,000 bull genomes : Projet international mis en place en bovins afin d'acquérir de nombreux séquençages tout génome de taureaux issus de races différentes.

Anomalie génétique : « phénotype déviant par rapport à la population » (Boichard et al., 2016).

BLAD (*Bovine Leucocyte Adhesion Deficiency*) : Anomalie détectée en race bovine Holstein liée à une mutation faux-sens récessive dans le gène *ITGB2*. Elle est à l'origine d'une déficience du système immunitaire, causant la mort de l'animal durant sa première année de vie (onab.fr).

Braschyspina : Anomalie embryonnaire létale détectée en race bovine Holstein (déterminisme récessif). Quelques animaux atteints peuvent être viables mais ont une colonne vertébrale et des membres réduits (onab.fr).

Capacitation : Étape de développement du spermatozoïde.

Clonage positionnel : Méthode qui consiste à identifier des gènes responsables de phénotypes particuliers en utilisant les informations de position sur le génome (Vignal, 2011).

CVM (*Complex Vertebral Malformation*) : Anomalie létale découverte en race Holstein (déterminisme récessif), associée à des déformations articulaires. Aucun individu est viable (mortalité en cours de gestation) (onab.fr).

Déséquilibre de liaison : Correspond à l'association non aléatoire entre allèles à des locus différents (Barbezant et al., 2010).

Fertilité : Capacité d'un individu à assurer une descendance viable.

Génétique inverse : Méthode de recherche consistant à exploiter les données de séquences et/ou génotypes afin de repérer des mutations causales sans *a priori* sur le génotype associé (Georges et al., 2019).

Haplotype : Combinaison d'allèles phasés (Georges et al., 2019).

Imputation : Processus statistique permettant de prédire le génotype des SNP non typés à partir des nombreux SNP typés dans la population de référence. Cette méthode repose sur le déséquilibre de liaison entre SNP typés et non typés (Georges et al., 2019).

Lissencephalie : Malformation du cerveau (période néonatale) liée à un problème de migration neuronale lors du développement embryonnaire (orpha.net ORPHA :48471)

Millésime : Année de naissance de l'animal.

Mortalité embryonnaire : Mortalité intervenant lors de la fécondation jusqu'à l'organogenèse.

Mortalité fœtale : Mortalité au stade fœtal (après la phase embryonnaire).

Mortalité néonatale : Mortalité de l'animal au cours de la première semaine de sa naissance.

Mortalité périnatale : Mortalité de l'animal au cours du premier mois de naissance.

Mortalité post-néonatale : l'individu meurt entre 1 mois et 1 an.

Mortinatalité : avortement ou fausse couche.

Mutation létale : mutation conduisant à la mort de l'animal.

Phasage : Méthode qui consiste à déterminer l'origine de l'allèle paternel et l'allèle maternel du SNP (Georges et al., 2019).

Polymorphe : ce dit d'un SNP qui présente au moins deux allèles ayant des fréquences >1%.

Population de référence : population d'individus de la même race phénotypés et génotypés (Georges et al., 2019).

Prolificité : Nombre d'agneaux par portée.

QTL (*Quantitative Trait Loci*) : région du génome associée à un caractère quantitatif d'intérêt.

RFLP (*Restriction Fragment Length Polymorphism*) : Le typage par RFLP repose sur une digestion enzymatique. Certains SNP créent un site reconnu et digéré par une enzyme de restriction. Cette technique permet donc par simple lecture de donner le génotype de chaque individu pour le polymorphisme étudié.

Sélection génomique : utilisation de l'information moléculaire des marqueurs SNP pour prédire la valeur génétique d'un animal (Barbezant et al., 2010).

SNP (*Single Nucleotide Polymorphism*) : substitution d'une base par une autre (Vignal, 2011).

Star system : pratique d'élevage consistant à choisir un nombre restreint de taureaux pour la pratique de l'insémination animale.

Taux de conception : correspond au succès reproducteur (nombre de femelles ayant mis bas sur le nombre de femelles ayant été mises à la reproduction).

Tremblante ovine : maladie à prion (Encéphalopathie Spongiforme Subaigües Transmissibles, ESST), affectant le système nerveux central conduisant à la mort de l'animal (Elsen et al., 1997).



Figure 1. Bélier Lacaune (Crédit photo : Diane Buisson)



Figure 2. Bélier Manech à Tête Rousse (Crédit photo : Diane Buisson)

Introduction

Contexte

Au sein des populations d'élevage, des anomalies génétiques**, correspondant à des « phénotypes déviant par rapport à la population », peuvent émerger sous l'effet de la consanguinité et être maintenues dans ces populations (Boichard et al., 2016). Elles peuvent apparaître à des stades plus ou moins tardifs (Ducos et al., 2002), entraînant des conséquences économiques non négligeables impactant la santé et le bien-être animal des individus atteints et l'efficacité reproductive de ces animaux (Georges et al., 2019). En effet, la reproduction et la fertilité** des animaux sont au cœur des systèmes de production en élevage. Au sein des populations, il existe des mutations naturelles qui à l'état homozygote ont des conséquences néfastes sur le développement embryonnaire et/ou sur le très jeune animal conduisant à sa mort précoce (mutations létales**). Cela impacte notamment les résultats enregistrés de fertilité des femelles. Repérer et gérer ces mutations (en évitant les croisements à risque) permettrait d'améliorer la fertilité globale d'une population, la santé et le bien-être animal. En bovins, des travaux portant sur les anomalies génétiques en repérant des déficits en animaux homozygotes pour certaines régions du génome (Fritz et al., 2013), à partir de données haut-débit issues de la sélection génomique**, a permis l'identification de nombreuses mutations causales entraînant une mortalité embryonnaire** ou du jeune animal (Georges et al., 2019).

Le déploiement de la sélection génomique est en cours également chez les ovins pour deux races laitières, la Lacaune (Figure 1) et la Manech Tête Rousse (Figure 2). En dehors du fait que cette approche de sélection génomique permet la prédiction de la valeur génétique de l'animal dès la naissance et accélère le progrès génétique, elle met à disposition pour la recherche les données de milliers de génotypes individuels sur la base des « puces » à 50 000 marqueurs SNP** (*Single Nucleotide Polymorphism*). Cela offre donc la possibilité de repérer des mutations conduisant à des anomalies dans les populations ovines à l'image des travaux réalisés en bovins.

Objectifs de l'étude

Sur la base des données de génotypes dans les deux populations, Lacaune et Manech à Tête Rousse, il s'agira dans un premier temps de repérer (approche statistique de génétique quantitative) des zones du génome en déficit d'homozygotie (dérogation à la loi d'équilibre des génotypes d'Hardy-Weinberg). Dans un second temps, il s'agira par des approches de génétique moléculaire d'identifier et de valider la présence de mutations létales dans les gènes présents dans ces zones du génome.

* renvoie à la liste des abréviations, ** renvoie au glossaire

Tableau 1. Présentation des principales races ovines laitières françaises, d'après Astruc et al. (2016) et IDELE, CNBL (2018). Les productions sont données pour les animaux soumis au contrôle laitier officiel (CLO*).

Bassin	Roquefort	Pyrénées-Atlantiques			Corse
Race	Lacaune	Manech à Tête Rousse	Manech à Tête Noire	Basco-Béarnaise	Corse
Effectif total des brebis	890 000 *	274 000	80 000	80 000	83 000
Production moyenne (durée lactation)	329,9 litres (171 j)	239,2 litres (162 j)	171,0 litres (149 j)	208,9 litres (148 j)	159,9 litres (188 j)

* Total des deux Entreprises de sélection en race Lacaune (Lacaune Ovitest et Service Élevage de la Confédération de Roquefort).

I. Synthèse bibliographique : les anomalies génétiques dans les populations d'élevage : de leur émergence à leur identification

A. Comprendre l'émergence des anomalies génétiques dans les populations ovines

Connaitre l'organisation de la filière ovine permet de comprendre comment des anomalies génétiques peuvent émerger dans des populations en sélection.

1. Données générales sur la filière ovine française : de l'organisation de la filière à l'utilisation de la sélection génomique

Actuellement, il existe 57 races ovines reconnues (Prache et al., 2018). On distingue les races laitières, allaitantes (races bouchères et rustiques) et en conservation (Races de France, 2019).

a. Description des cheptels ovins laitiers

A la fin de l'année 2018, la France comptait 7,2 millions d'ovins. Ce cheptel est constitué de 2/3 d'animaux allaitants et de 1/3 d'animaux laitiers (Agreste, 2019). Le cheptel ovin allaitant est essentiellement localisé en Nouvelle-Aquitaine et en Occitanie (45% du cheptel). Le cheptel ovin laitier est principalement localisé en Occitanie, Nouvelle-Aquitaine et Corse (95% du cheptel). En France, les sélectionneurs ont choisi de travailler en races pures. Parmi les races laitières, on distingue cinq races majeures réparties dans trois grands bassins de production (Tableau 1) (IDELE*, CNBL*, 2018 ; CORAM*, 2019 ; FGE*, 2019) :

- **Rayon de Roquefort** : on retrouve la Lacaune, race la plus productive et importante par les effectifs (60% de l'effectif total des brebis laitières) en France. Elle est originaire du massif central. Son lait est utilisé pour la production de fromage AOP *Roquefort*.
- **Bassin des Pyrénées-Atlantiques** : Les races ovines laitières des Pyrénées (ROLP*) sont constituées de la Manech à Tête Rousse, la Manech à Tête Noire et de la race Basco-Béarnaise. On retrouve ces races dans les zones de montagnes. Elles sont connues pour leur rusticité et leur facilité de traite. Elles sont utilisées pour la production de fromage AOP *Ossau-Iraty*. La Manech Tête Rousse est la plus productive des ROLP.
- **Corse** : Cette race est rustique puisqu'elle est adaptée à des conditions climatiques difficiles (sécheresse ...). Son lait est utilisé pour la production du fromage AOP *Brocciu*.

Par la suite, une attention particulière sera donnée aux races Lacaune lait et Manech à Tête Rousse, sujets de notre étude.

b. Organisation de la sélection

Les cinq races ovines laitières citées précédemment font l'objet d'un programme de sélection (contre 20/52 seulement pour les races allaitantes) (Prache et al., 2018). La (ou les) performance(s) « P » d'un individu pour un caractère d'intérêt donné est le reflet du potentiel

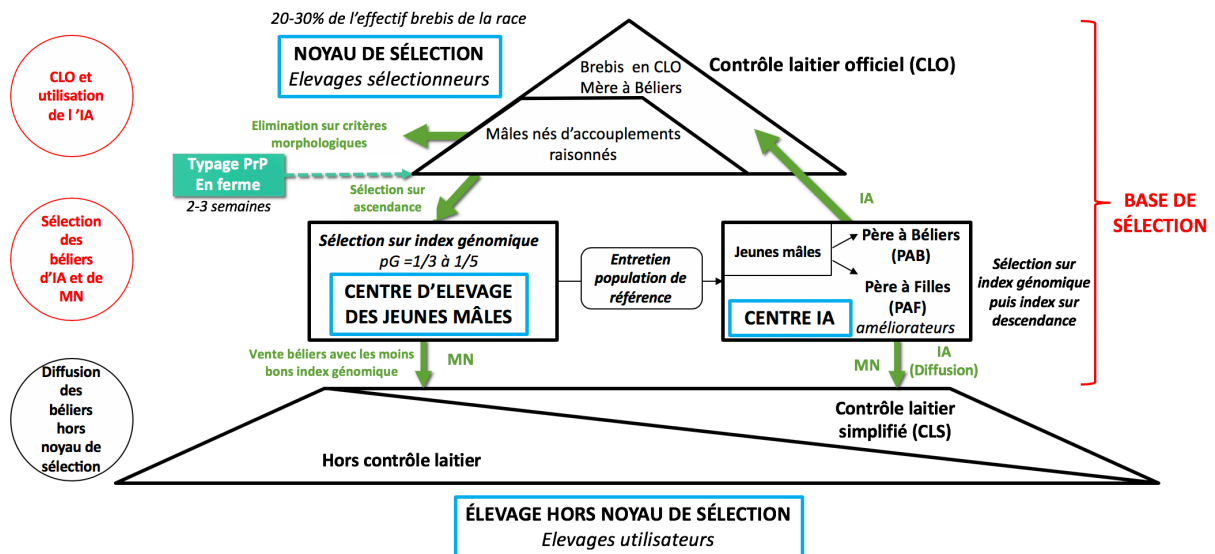


Figure 3. Gestion pyramidale du schéma de sélection en ovin laitier (sélection génomique), d'après Barillet et al. (2016). L'ensemble des béliers sont typés en ferme à l'âge de 3 semaines pour la tremblante** (Annexe i), seuls les béliers de génotype homozygote ARR/ARR sont gardés. Les animaux avec les meilleures valeurs génétiques sur ascendance et correspondant aux critères morphologiques intègrent le centre d'élevage (CE*). Une sélection des meilleurs mâles sur index génomique est réalisée pour pouvoir ensuite diffuser le progrès génétique dans les élevages utilisateurs (béliers d'Insémination Animale (IA*) et monte naturelle (MN*)). Les béliers qui n'ont pas été sélectionnés pour le centre d'IA (CIA*) sont classés en 3 catégories : +/+/+/+ suivant leur index génomique pour servir de reproducteurs en monte naturelle. Les béliers retenus pour le CIA sont qualifiés soit de Père à Béliers (PAB*, béliers élites), qui participent aux accouplements raisonnés, soit Pères à Filles (PAF*, permet de diffuser le progrès génétique en élevage). Lorsque les béliers ont une descendance connue (2,5 ans ou +), la sélection devient classique (index sur descendance). **Abréviation :** pG *: pression de sélection génomique.

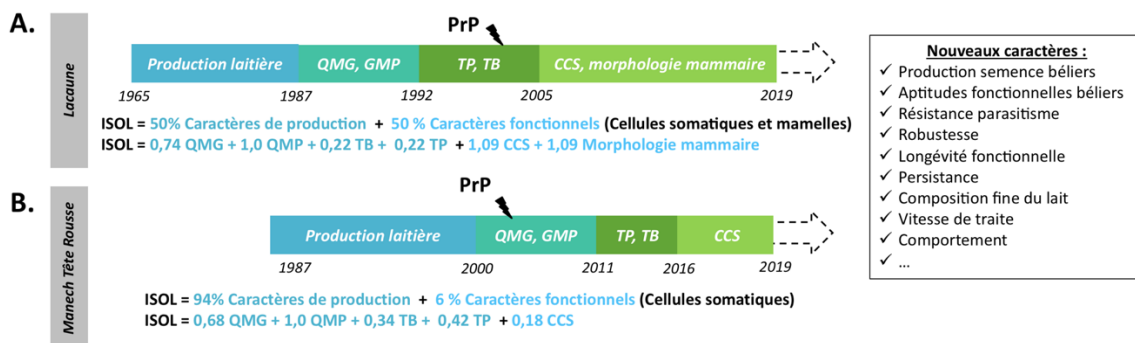


Figure 4. Évolution des critères de sélection en races Lacaune et Manech à Tête Rousse, d'après Barillet et al. (2016) et Astruc, Buisson (2018b). **Caractères laitiers :** LAIT: quantité de lait à la traite (en décilitre) ; QMG*: Quantité de Matière Grasse (en hg) ; QMP*: Quantité de Matière Protéique (en hg) ; TP*: Taux Protéique (en g/L) ; TB*: Taux Butyreux (en g/L) ; **Caractères fonctionnels :** CCS*: Comptage des cellules somatiques. **A.** Évolution des critères de sélection pris en compte dans le calcul de l'ISOL* (Index de Synthèse Ovin Lait) en race Lacaune. **B.** Évolution des critères de sélection pris en compte dans le calcul de l'ISOL (Index de Synthèse Ovin Lait) en race Manech à Tête Rousse. Dans les deux races, l'ISOL prend en compte des caractères laitiers et des caractères fonctionnels. A noter qu'en 2004, un nouvel objectif de sélection a dû être pris en compte pour que les animaux soient résistants à la tremblante (Annexe i). De nos jours, de nouveaux caractères sont étudiés pour être intégrés dans les objectifs de sélection.

génétique de l'animal (G, combinant l'effet additif des gènes sur le caractère, noté A, et des effets non additifs, de dominance et d'épistasie, notés respectivement D et I) et de l'effet de très nombreux facteurs de l'environnement (E) (Minvielle, 2010). L'objectif en sélection génétique est de choisir comme reproducteurs les animaux disposant des meilleures valeurs génétiques additives. Cette sélection classique passe, chez les ruminants, par une étape de testage sur descendance (évaluation de la valeur génétique d'un individu basée sur l'enregistrement de performances chez les descendants de l'animal à évaluer).

Ce sont les organismes de sélection (OS*) qui sont en charge des différents schémas de sélection. En race Lacaune, la sélection est orchestrée par l'OS UPRA Lacaune. Deux entreprises de sélection (ES*) contribuent à sa mise en œuvre : Lacaune Ovitest et Service Élevage de la Confédération de Roquefort. En race ROLP, l'OS ROLP repose sur l'entreprise de sélection CDEO*. L'ensemble des ES sont accréditées pour réaliser le contrôle de performances (Astruc, Buisson, 2018b).

L'organisation de la sélection est **pyramidale** (Figure 3) : on distingue les éleveurs du noyau de sélection et les éleveurs utilisateurs. Les animaux du noyau de sélection sont soumis à un **contrôle de performances** (Contrôle Laitier Officiel, CLO*) et de nombreux **testages sur descendance de béliers** d'Insémination Animale (IA*) sont effectués permettant ainsi de sélectionner les mâles avec les meilleures valeurs génétiques (Astruc et al., 2016). Le progrès génétique se fait principalement sur la **voie mâle**. Pour les populations laitières, la création et la diffusion du progrès génétique se fait principalement par IA. En effet, le recours à l'IA est de 85% en race Lacaune et de 60% en race Manech à Tête Rousse dans les noyaux de sélection (Astruc et al., 2016). Du fait de la **saisonnalité** et de l'utilisation de **semence fraîche**, le nombre de mâles présents en centres d'IA (CIA*) est important (1400 béliers Lacaune et 600 béliers Manech en CIA (Buisson et al., 2018)). Cette situation est assez différente de celle qui prévaut dans l'espèce bovine pour laquelle l'utilisation de semence congelée est possible. Ainsi, un taureau produit plusieurs (dizaines de) milliers de dose/an, contre 1200 au maximum pour un bélier (Buisson et al., 2018).

Dans l'ensemble des races laitières, la sélection a été orientée dans un premier temps sur le caractère de **production laitière** (Figure 4) puis a évolué vers des **caractères fromagers**. Depuis 2005, des **aptitudes fonctionnelles** (résistance aux mammites et morphologie mammaire) entrent dans les objectifs de sélection en race Lacaune. C'est seulement en 2016 que les races ROLP intègrent la résistance aux mammites (Barillet et al., 2016). L'ensemble de ces critères sont pris en compte dans un Index de Synthèse Ovin Lait (ISOL*) servant à classer les animaux selon leur valeur génétique mais sans tenir compte de l'information de génome avant la mise en place de la sélection génomique.

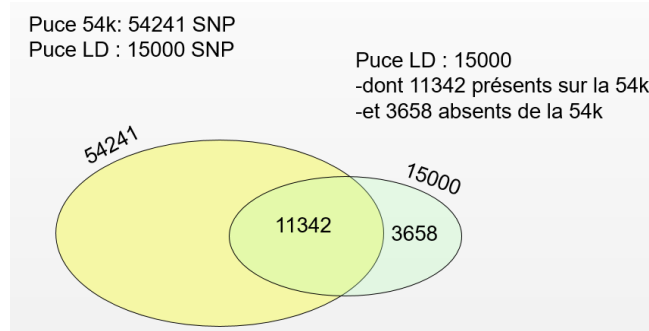


Figure 5. Marqueurs SNP présents sur la puce basse densité (Astruc, Buisson, 2018a). La puce basse densité (LD* : *Low density*) contient 15 000 marqueurs SNP dont 11 342 sont communs aux 54 241 marqueurs de la puce 50K. En sélection, la puce basse densité est utilisée pour génotyper les candidats à la sélection présents en centre d'élevage.

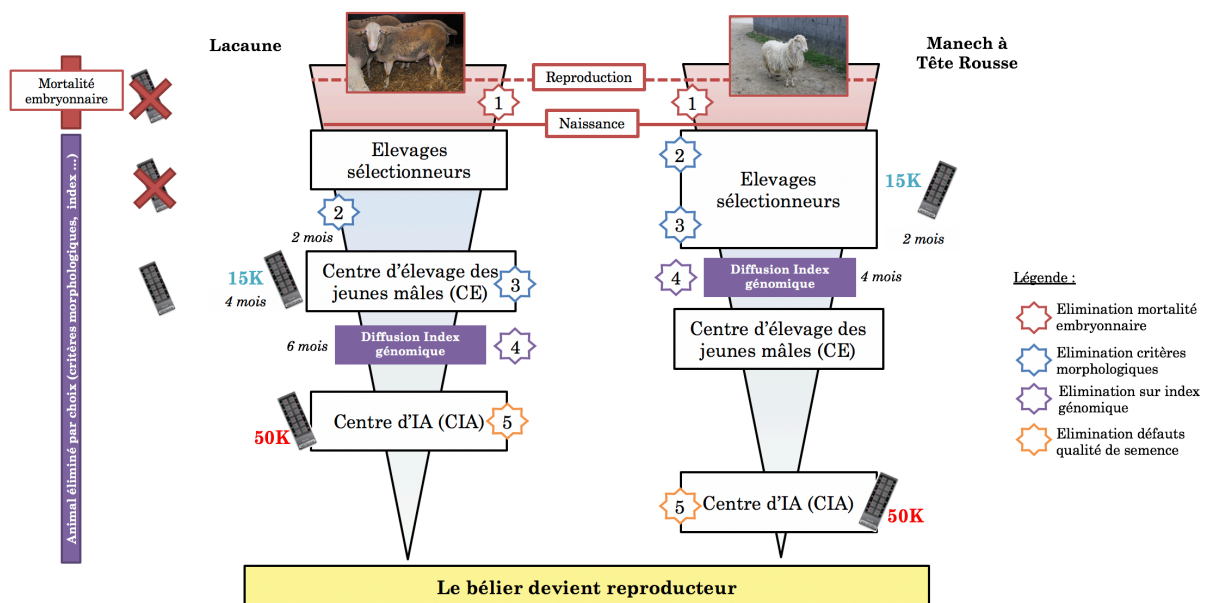


Figure 6. Stratégie de génotypage et causes d'écart à la sélection en ovin laitier, d'après Astruc, Buisson (2018a) (Crédit photo : Diane Buisson). Au sein des élevages sélectionneurs, les accouplements sont raisonnés entre les meilleurs reproducteurs mâles et femelles. Dans certains cas, la gestation ne va pas être menée à terme (mortalité embryonnaire liée à l'environnement ou à la génétique), le produit de l'accouplement n'a donc aucune chance d'être candidat à la sélection (Figure 6-1). Seuls les meilleurs agneaux issus de ces accouplements vont être sélectionnés parmi lesquels une faible proportion de béliers va devenir reproducteur. Une partie de ces agneaux sont éliminés sur critères morphologiques avant de rentrer en centre d'élevage des jeunes mâles (CE*) (Figure 6-2). En race Lacaune, les béliers candidats vont être génotypés sur la puce basse densité (15K) à l'âge de 4 mois en CE pour une diffusion des index génomiques vers l'âge de 6 mois. En Manech, le CE ne permet pas d'accueillir l'ensemble des agneaux. De ce fait, les éleveurs sélectionneurs gardent les agneaux en ferme où ils sont génotypés à l'âge de 2 mois (index génomiques disponibles à 4 mois). En attendant la diffusion des index, la moitié des individus Lacaune (et 15% des candidats Manech) sont éliminés sur critères morphologiques et fonctionnels (Figure 6-3). Au moment de la diffusion des index, une forte pression de sélection est exercée (Figure 6-4) : un tiers seulement des meilleurs béliers (sur index génomique ISOL) intègrent les CIA. Ils vont alors être à nouveau génotypés mais, cette fois-ci, sur la puce moyenne densité OvineSNP50® (50K). Enfin, un dernier tri s'effectue lors de l'évaluation de la qualité de semence (motilité, volume et concentration) des béliers en centre d'IA (Figure 6-5).

c. De la sélection classique à la sélection génomique

La sélection génomique permet de prédire la valeur génétique additive d'un candidat reproducteur en analysant un grand nombre marqueurs SNP répartis sur l'ensemble de son génome. Cette prédiction utilise les données génomiques et phénotypiques des individus de la même race présents dans la **population de référence**** (Georges et al., 2019). Ainsi la prédiction de la valeur génétique d'un animal peut se faire dès le plus jeune âge sur la base d'un simple génotypage. En 2009, le consortium international ovin a permis le développement d'un outil facilitant la sélection génomique, il s'agit de la puce ovine « *OvineSNP50®* » commercialisée par l'entreprise Illumina® et est constituée de 54 241 marqueurs SNP (Rupp et al., 2016). Trois autres puces sont également disponibles (Georges et al., 2019) : une puce très haute densité (*Ovine SNP600K Bead Chip array*), une puce basse densité (*Sheep LD, 15 000 SNP* (Figure 5)) et une puce très basse densité (*AgResearch OvineLD Imputation tool, 5 000 SNP*).

i. Mise en place de la sélection génomique en ovins laitiers : construction d'une population de référence dans chaque race

L'efficacité d'un schéma de sélection génomique repose sur un nombre suffisant d'individus génotypés et phénotypés dans la population de référence et l'enregistrement de l'apparentement entre animaux. En ovins laitiers, ces individus sont exclusivement des béliers testés sur descendance pour les différents caractères en sélection (Astruc et al., 2016). Fin 2017, les populations de référence en Lacaune et Manech à Tête Rousse étaient composées respectivement de 5 368 et 2 238 béliers génotypés et testés (Astruc, Buisson, 2018c).

ii. Stratégie de génotypage pour le choix des béliers reproducteurs

La stratégie de génotypage est similaire dans les races Lacaune et Manech à Tête Rousse (Astruc, Buisson, 2018a). Avant la mise en place de la sélection génomique, les béliers génotypés sont des béliers testés sur descendance (génotypage sur la puce 50K* à un âge plus tardif >1an). Depuis la mise en place de la sélection génomique (Figure 6), facilitée par le développement de la puce basse densité (15K), il est possible d'avoir l'information moléculaire de nombreux individus. Cependant, tous les agneaux candidats à la sélection ne peuvent prétendre à être génotypés. En effet, seuls les meilleurs mâles (index sur ascendance) et conformes aux standards de la race, intègrent le Centre d'élevage (CE). Ils vont alors être génotypés sur la puce basse densité. Une partie des candidats vont être éliminés sur critères morphologiques et standards de la race (50% en Lacaune, 15% en Manech), sur index génomique (calcul réalisé par imputation** de la puce basse densité 15K vers la densité de la

Tableau 2. Comparaison de la taille efficace N_e (génomique et généalogique) d'une race laitière bovine et de 4 races laitières ovines (IDELE, 2015). L'estimation de la taille efficace généalogique prend en compte le pedigree des ascendants. L'estimation de la taille efficace génomique repose sur l'estimation du déséquilibre de liaison** (correspondant à l'association non aléatoire entre allèles à des locus différents) obtenue par les données moléculaires de la puce SNP 50K. Les deux méthodes donnent des estimations similaires.

Espèce et filière	Race	N_e génomique	N_e généalogique
Bovins laitiers	Montbéliarde	216	231
	Lacaune	303	312
Ovins laitiers	Manech à Tête Rousse	145	148
	Manech à Tête Noire	92	82
	Basco-Béarnaise	98	91

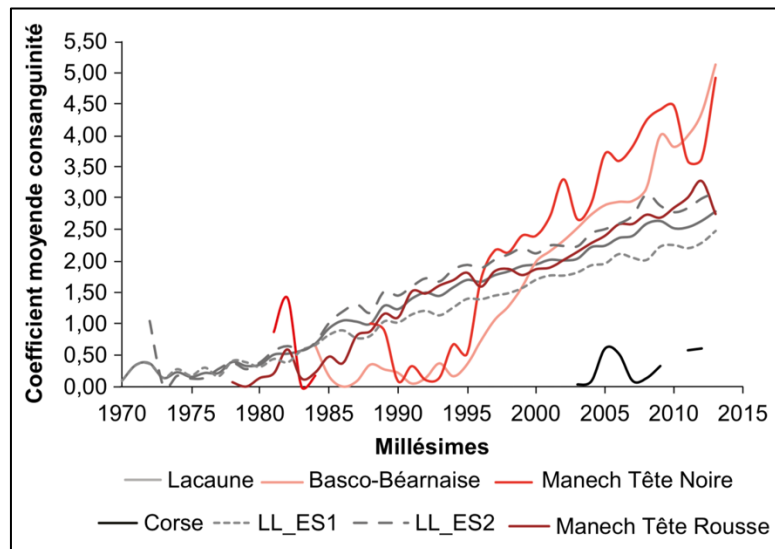


Figure 7. Évolution de la consanguinité moyenne en races ovines laitières des béliers du noyau de sélection nés entre 1970 et 2013 (Buisson et al., 2018). Pour la race Lacaune, l'accroissement de la consanguinité est similaire pour les béliers issus des deux entreprises de sélection (LL_ES1 et LL_ES2). Les races Basco-Béarnaise et Manech à Tête Noire ont les coefficients de consanguinité (~5%) les plus élevés par rapport aux autres races laitières.

puce 50K, pression de sélection de 1/3). Seuls les meilleurs candidats intègrent le CIA où ils vont alors être régénotypés sur la puce 50K. Un dernier tri s'effectue lors de l'évaluation de la qualité de semence (motilité, volume et concentration) des béliers en CIA. Ainsi, la sélection génomique repose sur deux puces : la puce basse densité (choix des candidats) et la puce moyenne densité (génotypage des reproducteurs).

2. La structure des populations ovines pourrait conditionner l'émergence d'anomalies génétiques

La **variabilité génétique** d'une population est définie par la « diversité des allèles rencontrés » (IDELE, 2015).

a. Des populations avec des effectifs génétiques limités

Au sein des populations en sélection, le nombre de reproducteurs est limité. Gérer la variabilité génétique des populations d'élevage est un enjeu majeur pour garantir la pérennité de la race. Le principal indicateur de la diversité génétique des populations en sélection est la **taille efficace** (N_e^*) qui correspond au nombre d'individus d'une population idéale qui « subirait le même taux de perte de variabilité génétique que la population étudiée » (IDELE, 2015). Les races Lacaune et Manech à Tête Rousse ont une taille efficace > 100 (Tableau 2), témoignant d'une bonne gestion de la diversité. Cependant, la sélection pratiquée dans l'ensemble de ces races a conduit à choisir un faible nombre d'individus reproducteurs à l'origine de l'apparition de goulots d'étranglements notamment liés à la sélection sur le gène de résistance à la tremblante ovine (Buisson et al., 2018) (Annexe i). Dans les races laitières, l'IA contribue à diminuer le nombre de reproducteurs. Une prise de conscience a été faite par les gestionnaires des races pour favoriser un plus grand nombre de reproducteurs et limiter le *star system***. Un compromis entre variabilité génétique et pression de sélection est désormais possible par la sélection génomique afin d'assurer des schémas de sélection efficaces à court, moyen et long terme.

b. Conséquence : augmentation de la consanguinité

La **consanguinité** (notée F) est le résultat de l'accouplement entre reproducteurs apparentés. Une des conséquences directes de cette consanguinité est l'augmentation du taux d'homozygotie dans la population (Buisson et al., 2018). Les races Lacaune et Manech à Tête Rousse (Figure 7) présentent une augmentation annuelle de consanguinité relativement limitée (respectivement + 0,08% et +0,11% entre 2000 et 2009). Les coefficients de consanguinité moyens dans ces deux races sont similaires ($F \sim 2,5-3\%$). Outre l'augmentation de la consanguinité, ce qui caractérise les populations de faible effectif c'est la fluctuation aléatoire des fréquences alléliques à chaque génération (phénomène de dérive, pouvant expliquer l'augmentation parfois brutale de la fréquence de certaines anomalies) (Charlesworth, Willis,

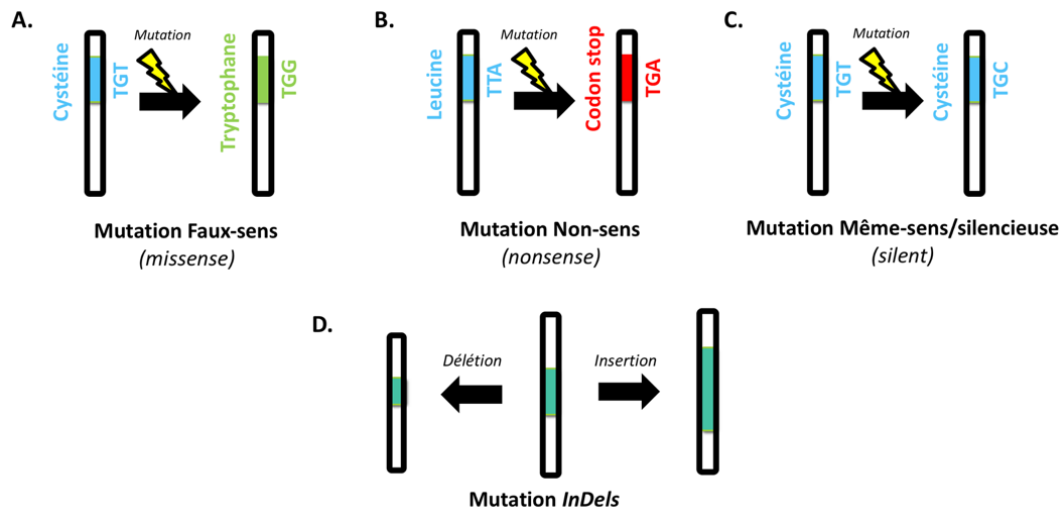


Figure 8. Les mutations ponctuelles de l'ADN affectant les régions codantes, d'après Dupont (2001). On distingue la (A.) **mutation faux-sens** (remplacement d'un nucléotide par un autre, l'acide aminé sera différent), (B.) la **mutation non-sens** (le changement de nucléotide conduit à l'apparition d'un codon stop), (C.) la **mutation silencieuse** (le changement de nucléotide n'a pas d'effet sur la protéine, différents codons codent pour le même acide aminé (redondance du code génétique)). Des **insertions ou délétions (D.)** de nucléotides peuvent également se produire, cela entraîne un décalage du cadre de lecture durant la traduction.

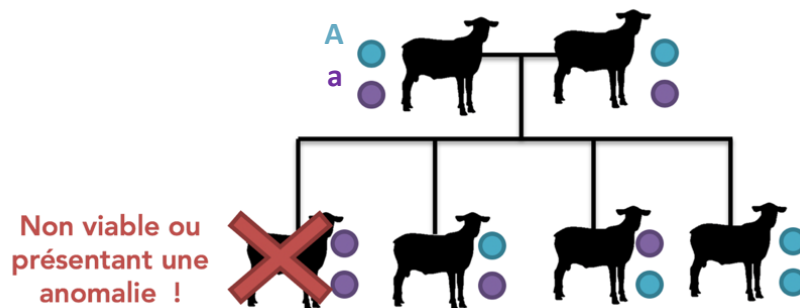


Figure 9. Transmission mendélienne récessive d'une anomalie génétique. La mutation apparaît chez un individu (allèle muté : a ; allèle sauvage A). Si deux reproducteurs ont un ancêtre commun porteur de cette mutation, ils peuvent être porteurs hétérozygotes à la mutation (ainsi que de son haplotype qui l'entoure). Ainsi, des descendants peuvent être homozygotes à la mutation entraînant la mort de l'individu ou présentant une anomalie congénitale.

2009). Chez les bovins laitiers, l'utilisation massive de taureaux reproducteurs (utilisation de semence congelée) a contribué à augmenter la consanguinité (F~ 4-5% pour les principales races laitières (IDELE, 2018)) et donc à l'émergence d'anomalies génétiques. La sélection en races pures et la reproduction en semence fraîche des ovins a permis de limiter la consanguinité (Buisson et al., 2018). Bien qu'elle soit plus faible que chez les bovins, elle reste quand même non négligeable, une bonne gestion de la diversité doit être réalisée.

c. Origine des anomalies génétiques : les mutations de l'ADN

Les gènes, support de l'information génétique, sont portés par les chromosomes. Le génome ovin (*Ovis aries*) est composé de deux chromosomes sexuels (XX pour la femelle et XY pour le mâle) et 26 paires d'autosomes ($2n = 54$). Les mutations peuvent avoir des origines environnementales qui induisent de modifications biochimiques des bases nucléotidiques (rayonnements, carences) ou génétique (Ducos et al., 2002). Lors de la réplication de l'ADN*, des erreurs peuvent se produire : cela conduit à l'apparition de néo-mutations qui peuvent être à l'origine d'anomalies génétiques héréditaires si elles affectent l'ADN des cellules germinales des reproducteurs (Boichard et al., 2016). Pour la plupart des anomalies, les mutations associées sont localisées dans les séquences codantes (Figure 8) ou régulatrices. Les mécanismes expliquant la transmission héréditaire des anomalies sont très variables. Les plus étudiées sont les anomalies à transmission mendélienne, et, parmi celles-ci, celles dont le déterminisme génétique est autosomique récessif (Figure 9). On parle de **mutation létale** quand la mutation conduit à la mort de l'animal. Dans l'espèce humaine, on estime le taux de mutation de 10^{-4} à 10^{-8} par base nucléotidique et par génération (Ducos et al., 2002). Ainsi, un individu serait porteur hétérozygote de 1 à 5 mutations récessives hautement délétères (Georges et al., 2019), c'est le fardeau génétique.

d. Les mutations sont soumises à des mécanismes évolutifs

Une fois ces mutations apparues, il est important de comprendre comment leurs fréquences évoluent au cours du temps. Deux mécanismes évolutifs contribuent à leur diffusion : la **dérive génétique et la sélection** (avantage de l'hétérozygote) (Boichard et al., 2016).

La **dérive génétique** correspond à la variation aléatoire de la fréquence des allèles. Plus la taille efficace de la population est faible plus l'effet de la dérive sera important. Citons les exemples des anomalies BLAD** (*Bovine Leucocyte Adhesion Deficiency* (Shuster et al., 1992)) et CVM** (*Complex Vertebral Malformation* (Thomsen et al., 2006)) apparues en race bovine Holstein. Leur émergence s'explique par l'utilisation massive des taureaux d'insémination *Osborndale Invanhoe* (pour BLAD) et *Bell* (pour CVM) qui ont largement



Figure 10. Évolution de la fréquence de l'allèle BL à l'origine de l'anomalie BLAD en race Holstein Française chez les veaux, les taureaux et les vaches (Boichard et al., 2016). La fréquence de l'allèle BL responsable de l'anomalie BLAD a atteint les 10% dans la population française Holstein en 1991.



Figure 11. Différents types de variations nucléotidiques, d'après Vignal (2011). Afin d'obtenir une séquence de référence du génome ovin commune à toutes les communautés scientifiques, la race Texel a été choisie (A). Cette séquence sert de référence dans les différentes bases de données internationales (Ensembl, NCBI, ...) et est annotée (emplacement des gènes, des transcrits, etc ...). On distingue des polymorphismes de type SNP (B) et des polymorphismes de longueur (C, D). Seul un des deux brins d'ADN est représenté ici.

contribué à transmettre leur allèle délétère à leur descendance (**Figure 10**) (Boichard et al., 2016).

Il se peut que la mutation à l'**état hétérozygote** donne un avantage phénotypique, expliquant son maintien dans la population. Par exemple, Kadri et al. (2014) en recherchant des QTL* de fertilité chez la vache ont mis en évidence une délétion de 600 kb* qui est létale à l'état homozygote (entraîne une diminution de fertilité pour les mères porteuses de ces embryons) mais procure un avantage aux hétérozygotes dans les populations laitières scandinaves. En effet, les individus hétérozygotes ont une production laitière plus importante et un lait plus riche en protéines et matières grasses.

La sélection de l'allèle responsable de l'anomalie peut également être indirect. Il s'agit de l'effet « **auto-stop** », c'est-à-dire que l'allèle muté est en déséquilibre de liaison avec un locus d'intérêt en sélection (Boichard et al., 2016).

A retenir : Les populations en sélection ont un nombre de reproducteurs réduit, à l'origine de l'émergence d'anomalies génétiques. Sous l'effet de la consanguinité, des mutations délétères (voir létales) peuvent alors apparaître à l'état homozygote. Leur fréquence peut augmenter sous l'effet de mécanismes évolutifs (dérive génétique, avantage à l'hétérozygote).

B. Le développement d'outils moléculaires facilite la détection des mutations responsables d'anomalies génétiques

Le développement de nouvelles technologies haut-débit (puces SNP et séquençage de génomes complets) permet d'identifier et de caractériser plus rapidement les anomalies génétiques. Au fil des années, les méthodes et outils technologiques ont évolué. On est passé d'une sélection basée sur une évaluation génétique obtenue à partir des performances des individus, à une sélection génomique reposant sur l'utilisation de marqueurs SNP balisant l'intégralité du génome (Boichard et al., 2015).

1. Les techniques de génotypages

a. Les différents marqueurs utilisés en sélection

La séquence du génome est légèrement variable entre individus appartenant à une même espèce. Ces variations de l'ADN sont appelées **polymorphismes** (Tagu et al., 2018c). Ces derniers peuvent être des **SNP** (*Single Nucleotide Polymorphism*, **Figure 11-B**) correspondant au remplacement d'une base par une autre, des **délétions ou insertions** (**Figure 11-C**) ou des **microsatellites** correspondant à des répétitions de motifs de 2 à 4 bases (**Figure 11-D**).

b. Utilisation des puces de génotypages haut-débit

Les puces à SNP permettent d'obtenir le génotype (dès le plus jeune âge) d'un grand nombre de marqueurs pour chaque individu dans différentes espèces d'élevage. Les marqueurs SNP

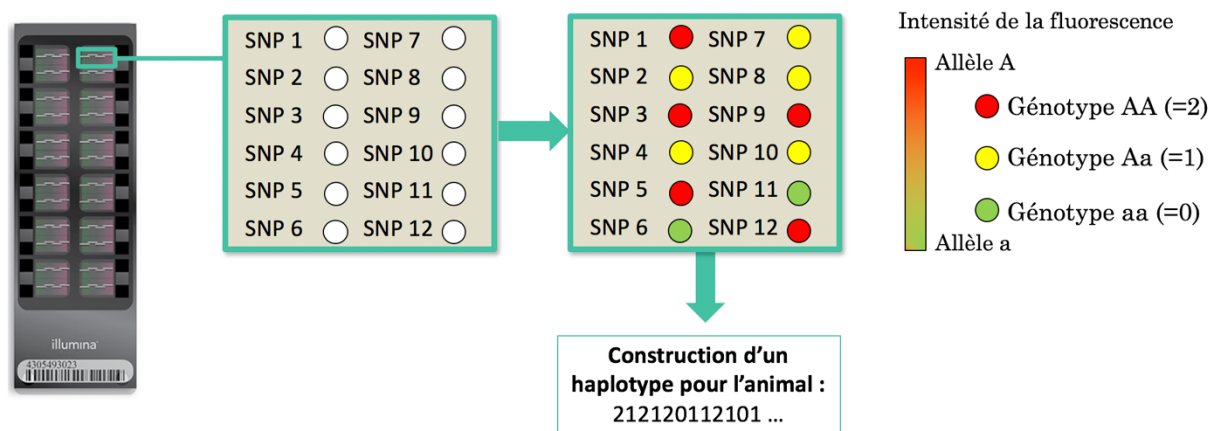


Figure 12. Fonctionnement d'une puce SNP, d'après [Illumina \(2015\)](#) et [Tagu et al. \(2018c\)](#). A partir d'un échantillon biologique (sang, cartilage ...) les puces SNP permettent d'obtenir le génotype de nombreux marqueurs SNP pour un grand nombre d'individus simultanément (24 individus sur la puce *OvineSNP50*[®]). Le génotypage d'un SNP repose sur l'hybridation d'une amorce sur l'ADN génomique. Il s'agit d'une automatisation d'un génotypage simple marqueur. Les SNP ajoutés lors de l'élongation sont marqués par des fluophores. La lecture de la fluorescence permet d'attribuer le génotype pour chacun des SNP. Sur le schéma, on observe le résultat de génotypage de 12 SNP. Cela permet d'aboutir à la construction d'un haplotype codifié.

Encadré 1. Principe de l'équilibre de Hardy-Weinberg ([Minvielle, 2010](#))

- Ce principe a été énoncé en 1908.
- Soit un SNP bi-allélique A et B, de fréquences alléliques $f(A)$ et $f(B)$,
- Dans une population de grande taille, sans migration (régime fermé), non soumise à la sélection, sans mutation alors **les fréquences alléliques [$p=f(A)$ et $q=f(B)$] restent constantes d'une génération à l'autre.**
- De plus, s'il y a **panmixie**, on est alors capable de déduire à partir des fréquences alléliques, les fréquences génotypiques (et donc le nombre d'individus attendus pour chacun des génotypes) :

	AA	AB	BB
Nombre d'animaux attendus (théorique)	$N p^2$	$N 2pq$	$N q^2$
Nombre d'animaux observés	N_{AA}	N_{AB}	N_{BB}

Si le nombre d'individus observés correspond au nombre d'individus attendus pour chacun des génotypes, on dit que le marqueur SNP est en équilibre. Si ce n'est pas le cas, le marqueur n'est plus à l'équilibre (écart à la loi de Hardy-Weinberg).

présents sur la puce ont été choisis de telle sorte *i)* qu'ils couvrent l'ensemble du génome *ii)* et sont polymorphes** dans plusieurs races d'une même espèce. Le principe du génotypage repose, après une étape d'amplification par PCR*, sur l'hybridation de sondes d'ADN fluorescentes complémentaires aux deux allèles du SNP (Illumina, 2015). Cette fluorescence est lue par un scanner à haute résolution (Figure 12). Cela permet d'attribuer le génotype de l'animal pour chaque marqueur SNP présent sur la puce suivant l'intensité de la fluorescence. *A posteriori*, un contrôle qualité est réalisé pour éliminer les erreurs de typages, dont notamment ceux qui ne respectent pas l'équilibre de Hardy-Weinberg (Encadré 1). Le génotypage par puces présente de nombreux avantages : coût de génotypage réduit, typages simultanés d'un nombre important de SNP, précision des typages (Illumina, 2015).

2. Le séquençage des génomes, vers une meilleure connaissance moléculaire

Séquencer un génome c'est connaître le « nombre, la nature et l'ordre des nucléotides » qui composent les séquences (Tagu et al., 2018a). Cela permet de connaître tous les variants de chaque animal séquencé par rapport au génome de référence, ce qui est très utile pour identifier des mutations causales.

De nouvelles technologies haut-débit, les NGS* (*New Generation Sequencing*), ont été développées à partir de la fin des années 2000. La technique de séquençage la plus utilisée actuellement pour les espèces agronomiques est celle du séquençage de l'ADN par la technologie HiSeq d'Illumina® (Tagu et al., 2018b). Elle repose sur le séquençage simultané de nombreux fragments (200 à 500pb*) sur une lame de verre (Figure 13). Grâce à l'ajout de ddNTP* lors de chaque cycle, la lecture de la fluorescence va permettre de déterminer chacun des nucléotides des fragments pour chaque cluster. Ainsi, cette technologie s'est avérée compétitive puisqu'aujourd'hui elle permet de séquencer un génome pour un coût raisonnable (<1000€) dans un délai court (quelques jours), avec un taux de précision >99% (Servin et al., 2014).

A retenir : Le développement de la sélection génomique passe par l'utilisation des **marqueurs SNP**. Cela permet pour un coût raisonnable (moins de 100€/animal) de favoriser l'accès à l'information moléculaire de nombreux individus pour prédire leurs performances. **L'essor des NGS** permet de mieux étudier le génome (annotations fonctionnelles, recherche de variants...). Ces technologies peuvent s'avérer utiles pour chercher et caractériser des anomalies génétiques.

C. Comment détecter les anomalies génétiques ?

Dans un contexte de sélection, la mise à la reproduction des meilleurs individus (mâles en particuliers) contribue à la transmission de mutations récessives délétères. Différentes stratégies

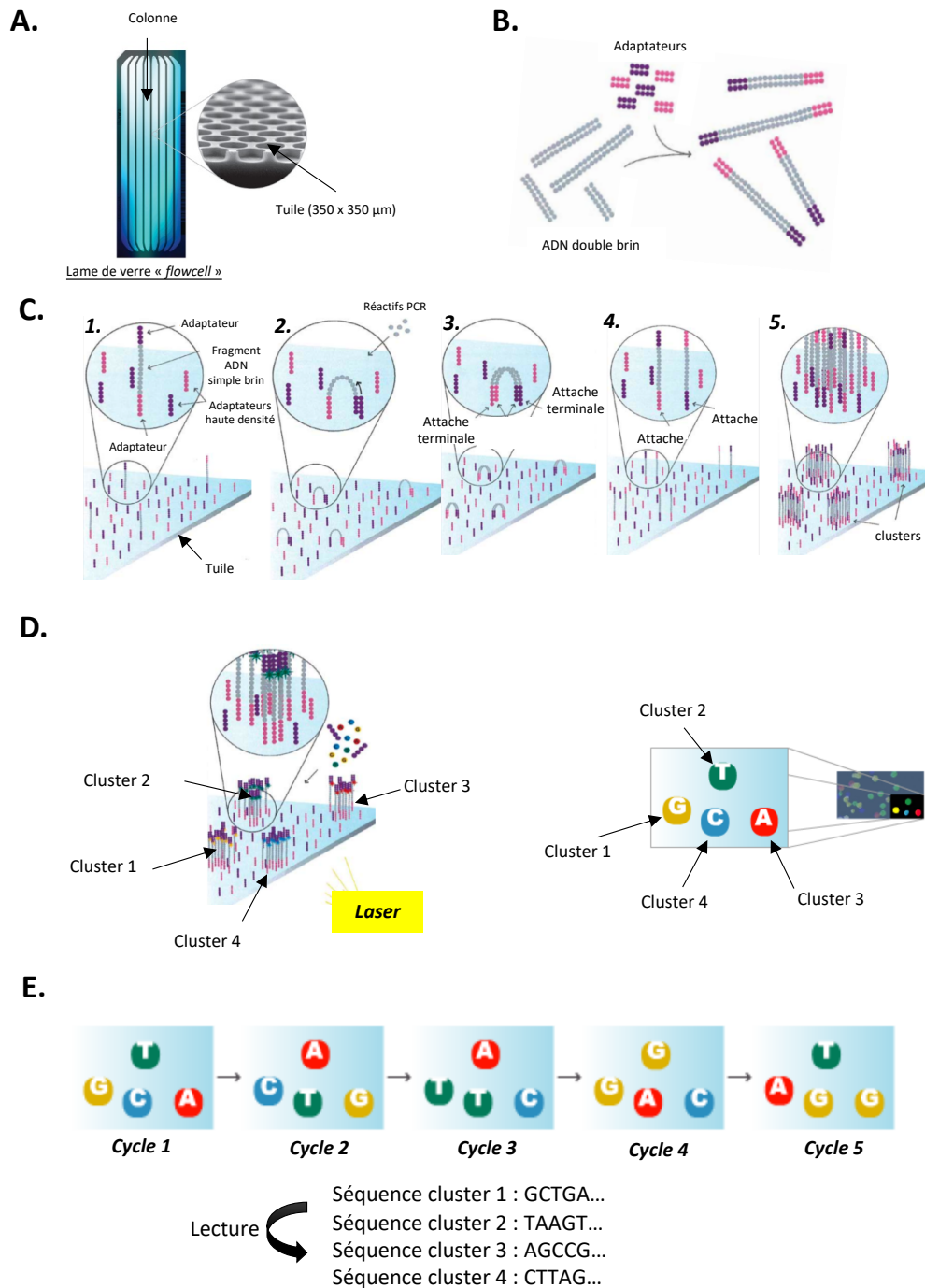


Figure 13. Principe du séquençage haut-débit par la technologie Illumina (New Generation Sequencing), d'après Tagu et al. (2018b) et Illumina (2010) **A.** Le séquençage se fait sur une lame de verre appelée «flowcell», constituée de 8 canaux. **B.** L'ADN est morcelé en fragments de petite taille où l'on ajoute des adaptateurs. **C.** Chaque fragment va se fixer à l'adaptateur haut débit qui lui est complémentaire (1). Les fragments vont alors établir un pont (2) pour se fixer sur les adaptateurs libres à proximité, ils sont ensuite amplifiés par PCR (3) et les deux fragments vont se dissocier (4). L'ensemble de ces étapes se répète (5). **D.** Cela aboutit aux regroupements des séquences identiques en cluster. Ici, 4 clusters sont représentés. **E.** A chaque cycle, la fluorescence est lue pour chaque cluster.

ont été développées pour identifier ces anomalies génétiques à déterminisme récessif. La stratégie « *top-down* » consiste à collecter et étudier les phénotypes d'animaux atteints puis d'identifier les mutations causales responsables de ces anomalies. Si l'information phénotypique n'est pas disponible, en partant des données haut-débit de génotypage ou de séquençage NGS, on peut réaliser un raisonnement de génétique inverse « *bottom-up* », sans *a priori* sur le phénotype pour identifier des mutations récessives délétères.

1. Approche « *top-down* » : du phénotype à la mutation causale

Les structures alliant les différents partenaires des filières d'élevage (de l'éleveur à l'institut de recherche) organisées en observatoires sont des outils essentiels pour faire remonter les cas d'anomalies repérés en élevage. Une des principales méthodes utilisées pour caractériser ces anomalies est la cartographie par homozygotie.

a. Les observatoires d'anomalies génétiques, un système de surveillance des anomalies

L'ONAB* pour les bovins (Observatoire des Anomalies Bovines, <https://www.onab.fr/>), ou l'observatoire des anomalies porcines se montrent efficaces pour identifier de nouvelles anomalies (Grohs et al., 2016). Le rôle des observatoires est multiple : *i*) identifier l'émergence de nouvelles anomalies (fiche descriptive remplie par les éleveurs, les techniciens d'élevage, les vétérinaires), *ii*) comprendre l'anomalie et *iii*) gérer son élimination progressive. A ce jour aucun observatoire permettant de recenser les anomalies n'est opérationnel en petits ruminants. L'ensemble des anomalies des espèces animales sont décrites dans la base de données OMIA* (Online Inheritance in Animals, <https://omia.org/home/>) (Figure 14). Cette base est complémentaire à la base OMIM* (Online Inheritance in Man, <https://www.omim.org/>) disponible pour l'espèce humaine.

b. Du phénotype à la mutation causale : la cartographie par homozygotie

Grâce au développement des techniques de **clonage positionnel**** permis par l'utilisation des marqueurs SNP en déséquilibre de liaison avec la mutation causale, il est désormais possible *i*) d'identifier l'haplotype en association avec la mutation par une approche de « cartographie par homozygotie » et *ii*) d'identifier la mutation causale à partir des données de séquences des individus atteints de l'anomalie (Georges et al., 2019). L'utilisation de cette méthode repose sur l'enregistrement précis des phénotypes et le prélèvement d'échantillons sanguins des individus porteurs des anomalies.

Le **principe de cartographie par homozygotie** (Figure 15) est le suivant : lors d'un accouplement consanguin, deux reproducteurs porteurs d'une mutation venant d'un ancêtre commun peuvent donner naissance à des individus homozygotes à cette mutation. L'acquisition

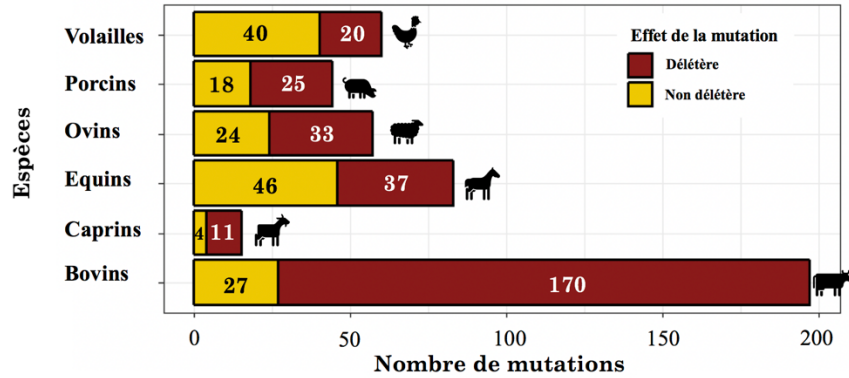


Figure 14. Nombre de mutations causales liées à une anomalie génétique avec un déterminisme mendélien dans les différentes espèces d'élevage (inventaire réalisé le 24 mai 2019, d'après OMIA <https://omia.org/home/>). Pour chaque espèce, le nombre de mutations comprend les mutations délétères (entraînant la mort de l'animal) et des mutations non délétères. On compte 57 entrées en ovins.

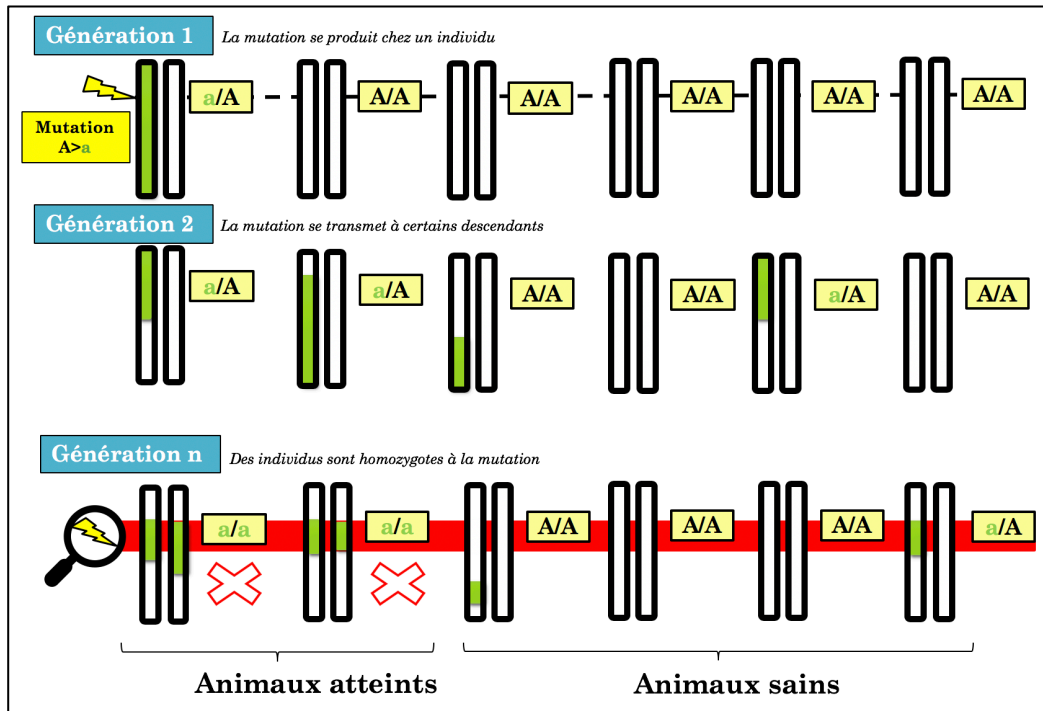


Figure 15. Principe de la cartographie par homozygotie d'après Duchesne et al. (2016). Une mutation « a » apparaît (éclair jaune) chez un individu. Si cet individu a des descendants, il va transmettre la mutation à certains de ses descendants. Au cours des générations suivantes, si deux reproducteurs ont un ancêtre commun porteur de cette mutation, des descendants homozygotes à cette mutation peuvent apparaître (cas cliniques visibles). Ainsi, la mutation peut être observée en cherchant un haplotype homozygote (en association avec la mutation) chez les individus atteints mais jamais observé à l'état homozygote chez les individus sains.

des génotypages d'individus atteints et témoins permet, par un calcul statistique (Charlier et al., 2008), d'identifier un segment chromosomique présent à l'état homozygote chez l'ensemble des individus atteints de l'anomalie (quelques cas sont suffisants, mais plus le nombre de témoins est important plus le « filtrage » est efficace, et plus le nombre de cas est important plus on réduit la zone contenant la mutation). Ce segment n'est jamais observé à l'état homozygote chez les individus témoins (Duchesne et al., 2016). La seconde étape consiste à explorer les fichiers de variants issus des séquençages tout génome pour identifier des variants candidats. Un génotypage à grande échelle et des tests fonctionnels permettent de valider ou non la mutation candidate identifiée.

Par cette méthode, Charlier et al. (2012) ont identifié une mutation létale en Holstein (délétion de 3,3 kb), associée à l'anomalie *Brachyspina*** , dans le gène *FANCI** (Figure 16). Cette mutation conduit généralement à un avortement durant les 40 premiers jours de gestation. Les quelques individus observés à la naissance ont une colonne vertébrale réduite, des pattes plus longues et un corps plus court. L'utilisation massive de deux taureaux porteurs de la mutation a contribué à son émergence dans la population Holstein. Même s'il n'existe pas encore d'observatoire d'anomalies en petits ruminants, des mutations délétères peuvent être identifiées en race ovine. En effet, Suárez-Vega et al. (2015) ont identifié une délétion dans le gène *ITGB4* à l'origine de l'*Epidermolyse bulleuse* dans la race ovine Churra (Figure 17).

2. Approche « bottom-up » : raisonnements de génétique inverse

Pour pouvoir réaliser l'approche « top-down », il est nécessaire d'avoir accès au phénotype, mais ceci n'est pas toujours évident. En effet, lorsque l'anomalie conduit à une mortalité embryonnaire, elle est confondu avec le phénotype de **fertilité** de la mère par exemple (VanRaden et al., 2011 ; Sonstegard et al., 2013 ; McClure et al., 2013). En sélection, un mâle reproducteur porteur hétérozygote d'une mutation délétère peut largement diffuser l'allèle dans la population entraînant des conséquences économiques (Georges et al., 2019). Grâce aux différentes technologies utilisées en élevage, de nombreux génotypages et/ou données de séquences tout génome sont accessibles pour identifier de telles mutations.

a. Détection d'haplotypes en déficit d'homozygotie

L'essor de la **sélection génomique** a permis de disposer de nombreux typages d'animaux en sélection. VanRaden et al. (2011) et Fritz et al. (2013) ont développé des méthodes pour *i)* repérer des déficits en homozygotie afin d'identifier des mutations létales sans avoir l'information phénotypique. En effet, si **on fait l'hypothèse qu'au sein d'une population ségrége un haplotype avec une fréquence relativement importante mais qu'aucun de ces haplotypes n'est observé à l'état homozygote, cela signifie que les individus porteurs homozygotes à cet haplotype n'ont pas été génotypés car ils n'ont jamais été candidats à**

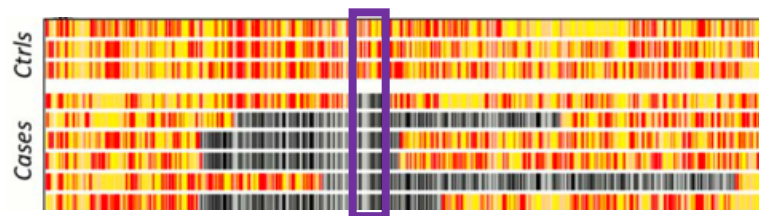
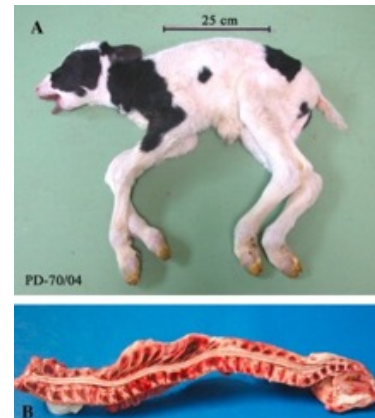
A.**B.**

Figure 16. Cartographie par homozygotie permettant de localiser l'haplotype homozygote associé la mutation dans le gène *FANCI* (Testoni et al., 2008 ; Charlier et al., 2012). **A.** La cartographie par homozygotie a permis d'identifier un segment de 2,51Mb (encadré violet) chez les 6 individus atteints et absent à l'état homozygote chez les 15 animaux sains. Les génotypes homozygotes apparaissent en jaune ou blanc et les génotypes hétérozygotes en rouge. Les segments homozygotes sont représentés en noir et blanc. **B.** Description clinique des animaux atteints de l'anomalie. Ils ont un corps plus court, une colonne vertébrale réduite, des pattes plus longues.



Figure 17. Cas d'un agneau affecté par l'*Epidermolyse bulleuse* en race ovine Churra (Suárez-Vega et al., 2015). Les animaux homozygotes à la mutation présentent une peau très fragile au niveau des extrémités des pattes, des oreilles et des voies nasales entraînant une mortalité périnatale. Une approche gènes candidats pour la maladie permet d'arriver à la mutation causale (facilité par la connaissance de l'*Epidermolyse bulleuse* existante dans de nombreuses espèces).

la sélection parce qu'ils ne sont jamais nés. Une fois de tels haplotypes identifiés, *ii*) le statut (porteur/non porteur) des mâles reproducteurs est déterminé pour chaque haplotype afin de confirmer un éventuel effet de la mutation sur la fertilité. Pour cela, deux caractères sont évalués : le taux de conception** et la mortinatalité** des accouplements entre mâles porteurs x filles de mâles porteurs *versus* les autres accouplements. L'étape suivante consiste à *iii*) typer à grande échelle (ajout de la mutation sur la puce SNP) la mutation candidate identifiée à partir des données de séquences afin de confirmer qu'aucun individu homozygote est observé.

Avec cette approche, initiée par [VanRaden et al. \(2011\)](#) en recherchant des haplotypes en déficits en homozygotie à partir de génotypages de 58 453 animaux Holstein, 5 288 animaux Jersiais et 1 991 animaux Brown Swiss, **11 régions candidates**, dont 5 nouvelles mutations délétères récessives associées à une baisse fertilité, ont été identifiées : 3 en Holstein (HH1, HH2 et HH3), 1 en Jersiaise (JH1) et 1 en Brown Swiss (BH1). Cette recherche a également permis de montrer des déficits pour des anomalies déjà connues telles que *CVM*, *Braschyspina* et *BLAD*.

Avec cette même approche, [Fritz et al. \(2013\)](#) ont détecté **33 régions** présentant des déficits en homozygotie. L'analyse a porté sur trois races bovines laitières françaises : 47 878 animaux Holstein, 16 833 animaux Montbéliard et 11 466 animaux Normand typés sur la puce BovineSNP50®. Ainsi, 19 régions ont été détectées en Holstein, 11 en Montbéliarde et 7 en Normande. Les longueurs de ces régions variaient entre 1 et 4,8 Mb*, avec des fréquences des haplotypes en déficit d'homozygotie comprises entre 1,7 et 9%. Une dégradation de la fertilité a été observée pour 10 de ces haplotypes. De plus, 7 haplotypes avaient des effets significatifs sur la mortinatalité. La disponibilité des nombreuses séquences bovines issues du consortium *1,000 bull genomes*** a permis d'étudier les polymorphismes dans les régions d'intérêt des séquences de taureaux porteurs et non porteurs (n=45). Ainsi, en 2013, quatre nouvelles mutations causales (validées à grande échelle) ont pu être découvertes dans les gènes *GART**, *SMC2**, *SLC37A2** et *NOAI**. Par exemple, *GART* est un gène qui code une enzyme impliquée dans la synthèse des purines (l'adénine et la guanine) ou le gène *SMC2* codant une protéine impliquée la condensation des chromosomes. Depuis ces travaux, de nombreux haplotypes et mutations causales ont été découverts (**Tableau 3**). Les mutations découvertes peuvent avoir des conséquences sur la fertilité mais aucune d'entre elles n'est directement impliquée dans la modification de gènes liés à reproduction. Les mutations interviennent dans des gènes assurant le bon développement embryonnaire ou intervenant dans la structure de l'ADN. Par conséquent, une mutation dans un de ces gènes explique aisément la non viabilité du descendant.

Tableau 3. Principales mutations causales identifiées par des raisonnements de génétique inverse : *i*) recherche de déficit en homozygotie puis recherche de la mutation causale dans les données de séquences (symbolisé par ☼) et/ou *ii*) recherche de variants délétères dans les données de séquences de génomes complets (symbolisé par ☆). Les symboles représentés à côté du nom du gène indiquent la méthode employée permettant d'arriver à la mutation causale. **Abréviations :** Hap* : Haplotype ; Chr* : Chromosome, BBB* : Blanc Bleu Belge, NZ* : New-Zealand.

Espèce	Race	Hap	Chr	Gène	Description	Référence
Bovin	Ayshire	AH1	17	☼ <i>PIRM/UBE3B</i>	Syndrome PIRM (retard de croissance, mortalité) lié à la mutation du gène <i>UBE3B</i> codant pour l'enzyme de conjugaison de l'ubiquitine E2 (modification post-transcriptionnelle)	Venhoranta et al. (2014)
Bovin	BBB*	-	10	☆ <i>MYH6</i>	Formation et contraction des myofibrilles et développement cardiaque	Charlier et al. (2016)
Bovin	BBB	-	11	☆ <i>SNAPC4</i>	Permet la transcription des ARN polymérase II et III à partir de petits ARN nucléaires	Charlier et al. (2016)
Bovin	BBB	-	11	☆ <i>RPIA</i>	Conversion entre le ribose-5-phosphate et le ribulose-5-phosphate dans la voie des Pentose-phosphate	Charlier et al. (2016)
Bovin	BBB	-	14	☆ <i>EXOSC4</i>	Participe à la synthèse et la dégradation de l'ARN	Charlier et al. (2016)
Bovin	BBB	-	11	☆ <i>MED22</i>	Intervient dans la régulation de la transcription	Charlier et al. (2016)
Bovin	Brune	BH2	19	☼ <i>TUBD1</i>	Intervient dans la polymérisation des microtubules	Schwarzenbacher et al. (2016)
Bovin	Holstein	HHY	21	☼ <i>FANCI</i>	Réparation de la réticulation de l'ADN (associée à <i>Brachyospina</i>)	Charlier et al. (2012) ; Fritz et al. (2013)
Bovin	Holstein	HCD	11	☼ <i>APOB</i>	Structure de l'apolipoprotéine B - Cholestérol LDL (associée à <i>Cholesterol deficiency</i>)	Schütz et al. (2016)
Bovin	Holstein	HH1	5	☼ <i>APAF1</i>	Protéine codant un domaine WD40 important pour les interactions entre protéines	VanRaden et al. (2011) ; Fritz et al. (2013) ; Adams et al. (2016)
Bovin	Holstein	HH3	8	☼☆ <i>SMC2</i>	Protéine essentielle à la condensation des chromosomes	VanRaden et al. (2011) ; Daetwyler et al., (2014) ; McClure et al. (2014)
Bovin	Holstein	HH4	1	☼ <i>GART</i>	Enzyme impliquée dans les synthèses des purines	Fritz et al. (2013)
Bovin	Holstein	HH5	9	☼ <i>TFB1M</i>	Initiation de la traduction des protéines dans les mitochondries	Fritz et al. (2013) ; Schütz et al. (2016)
Bovin	Holstein	HH6	16	☼ <i>SDE2</i>	Maintenance des télomères	Fritz et al. (2018)
Bovin	Holstein	HHC	3	☼ <i>SLC35A3</i>	Transporteur UDP-N-acetylglucosamine, lié à <i>CVM</i>	Thomsen et al. (2006) ; VanRaden et al. (2011) ; Fritz et al. (2013)
Bovin	Holstein NZ*	-	11	☆ <i>TTF1</i>	Régulation de la transcription ribosomale	Charlier et al. (2016)
Bovin	Holstein NZ	-	3	☆ <i>RABGGTB</i>	Ajout post-transcriptionnel d'un groupe geranylgeranyl au Rab GTPases	Charlier et al. (2016)
Bovin	Holstein NZ	-	8	☆ <i>RNF20</i>	Régulation de la structure des chromosomes (histone H2B)	Charlier et al. (2016)
Bovin	Jersey	JH1	15	☼ <i>CWC15</i>	Protéine associée au spliceosome	VanRaden et al. (2011) ; Sonstegard et al. (2013)

b. Détection de variants délétères dans les données de séquences complètes

Grâce aux différents consortiums internationaux, l'acquisition de nombreuses séquences d'animaux en sélection (facilitée par l'évolution technologique des NGS) est un outil efficace pour détecter des variants délétères entraînant une perte de fonction de la protéine à l'état homozygote (Michot, 2017). Cette méthode de génétique inverse repose sur plusieurs étapes (Figure 18). La première étape consiste à exploiter les données de génomes complets afin d'identifier de nombreux variants candidats sur la base de leur **annotation fonctionnelle**. Des outils bio-informatiques permettent de prédire l'effet d'une mutation sur l'expression et/ou la fonction du gène. Dans un second temps, il est nécessaire de **valider cette mutation à grande échelle**. Les mutations candidates peuvent être rajoutées sur les puces SNP afin d'identifier les animaux de la sélection génomique porteurs de ces variants délétères. Cependant, cette validation ne reste efficace que si la fréquence de la mutation reste élevée dans la population en sélection. Des croisements dit « à risque » entre deux individus porteurs peuvent également être réalisés pour étudier la mutation.

En bovin, le projet *1,000 bull genomes* a permis d'identifier de tels variants (Tableau 3). Par exemple, Daetwyler et al. (2014) ont mis en évidence une mutation létale embryonnaire dans le gène *SMC2* en Holstein (associée à l'haplotype HH3 mis en évidence par Fritz et al. (2013)). En utilisant ce raisonnement de génétique inverse, Charlier et al. (2016) ont identifié **9 nouvelles mutations** (5 en Blanc Bleu Belge, 3 en Holstein-Friesian et 1 en Jersiaise de Nouvelle-Zélande) à l'origine de mortalité embryonnaire ségrégeant à des fréquences comprises entre 1,2 et 6,6%. Cette étude montre que **15% des variants entraînant une perte de fonction et 6% des variants non synonymes délétères sont à l'origine de mutations létales embryonnaires**. Michot et al. (2015, 2017) ont étudié les données de séquences du projet *1,000 bull genomes* pour rechercher des variants à l'origine d'anomalies génétiques dans 15 races européennes. Cette recherche a conduit à identifier **2 489 variants** délétères candidats (fréquence > 5% dans la population). Ces travaux semblent très prometteurs pour découvrir de nouvelles anomalies génétiques encore non publiées.

3. Prise en compte des anomalies génétiques dans les schémas de sélection

En raison de la détection d'un nombre croissant de mutations à l'origine d'anomalies génétiques, il est nécessaire de réfléchir à l'intégration de cette information dans les schémas de sélection. L'élimination systématique des individus porteurs conduirait à une perte de diversité génétique importante et rapide (Georges et al., 2019). Deux méthodes alternatives sont envisageables pour gérer ces anomalies génétiques : *i*) intégrer les anomalies dans les index de

Bovin	Jersey NZ	-	26	★ <i>OBFC1</i>	Initiation de la réplication de l'ADN et protection des télomères	Charlier et al. (2016)
Bovin	Montbéliarde	MH1	19	⊗ <i>PFAS</i>	Protéine intervenant dans la biosynthèse des purines	Fritz et al. (2013) ; Michot et al. (2017)
Bovin	Montbéliarde	MH2	29	⊗ <i>SLC37A2</i>	Antiport du glucose-6-Phosphate	Fritz et al. (2013)
Bovin	Normande	NH2	11	⊗ <i>CAD</i>	Protéine essentielle pour la biosynthèse des acides nucléiques	Mesbah-Uddin et al. (2019)
Bovin	Simmental	FH2	1	⊗ <i>SLC2A2</i>	Code pour le transporteur de glucose 2	Pausch et al. (2015)
Bovin	Simmental	FH4	12	⊗ <i>SUGT1</i>	Intervient dans le contrôle de la fonction des kinétochores pendant la division cellulaire lors de la métaphase	Pausch et al. (2015)
Porc	Duroc	DU1	-	⊗ <i>TADA2A</i>	Intervient dans la régulation de la transcription	Derks et al. (2019)
Porc	Landrace	LA1	-	⊗ <i>POLR1B</i>	Intervient dans la régulation de la transcription	Derks et al. (2019)
Porc	Landrace	LA2	-	⊗ <i>URB1</i>	Intervient dans la translation	Derks et al. (2019)
Porc	Landrace	LA3	-	⊗ <i>PNKP</i>	Réparation de l'ADN	Derks et al. (2019)
Porc	Large White	LW	-	⊗ <i>BMPER</i>	Code pour protéine intervenant dans la structure osseuse	Derks et al. (2018)
Porc	Large White	LW	18	⊗ <i>BBS9</i>	Code pour protéine intervenant dans la structure osseuse	Derks et al. (2018)

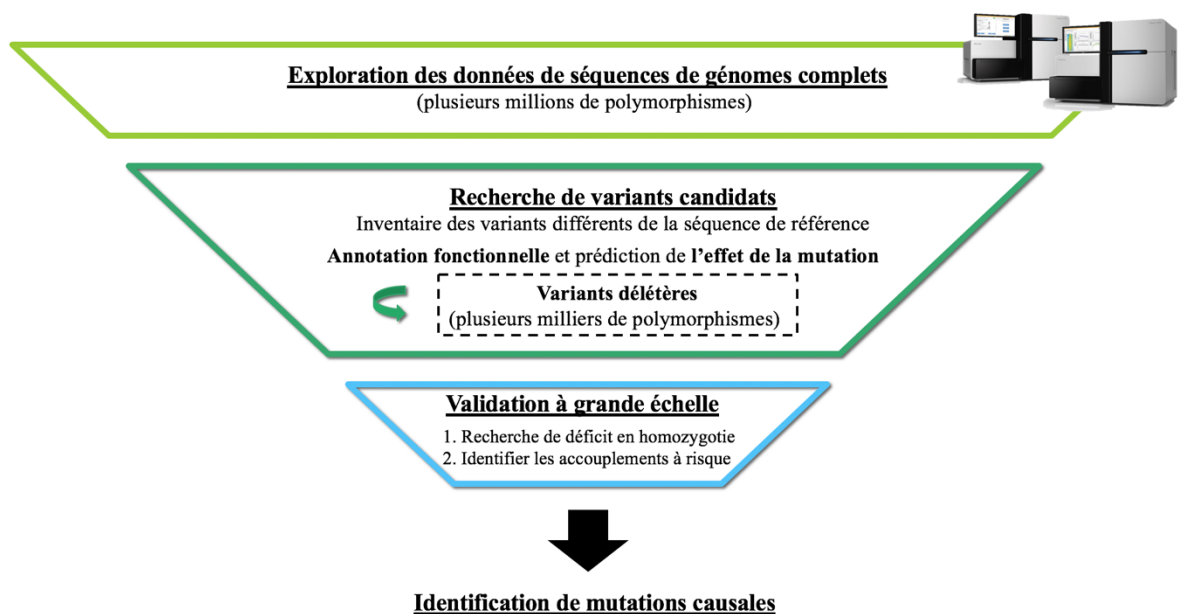


Figure 18. Méthodologie pour identifier des mutations récessives délétères par un raisonnement de génétique inverse en utilisant les données de séquences de génomes complets, d'après Michot (2017). L'identification de mutations causales à l'origine d'anomalie délétère passe par trois étapes : *i*) un recensement des différents polymorphismes identifiés, *ii*) les polymorphismes sont filtrés suite à l'annotation fonctionnelle et l'effet prédit de la mutation sur la fonction de la protéine et *iii*) une validation à grande échelle de la mutation causale.

sélection *ii*) génotyper la population femelle pour éviter les croisements à risque (limitée en ovin puisque le génotypage représente un certain coût pour l'éleveur).

A retenir : La méthode de **cartographie par homozygotie** repose sur la disponibilité de quelques cas cliniques permettant d'arriver rapidement à la **mutation causale (NGS)**. Cependant, la disponibilité des phénotypes n'est pas toujours évidente. D'autant plus lorsque l'anomalie conduit à une mortalité précoce de l'animal. Grâce à l'essor de la sélection génomique, l'utilisation de données moléculaires par un raisonnement de **génétique inverse** a permis de repérer des déficits en homozygotie et/ou d'explorer les données de séquences pour identifier des variants délétères à l'origine d'anomalies génétiques/mutations létales. La découverte des mutations létales impose une bonne gestion des schémas de sélection pour permettre de les intégrer et de les gérer.

D. Situation du sujet

Dans les populations d'élevage, l'utilisation de l'information génomique permise par le développement de différents outils technologiques facilite l'identification des mutations létales en ségrégation dans ces populations. Il est important de repérer ces mutations car elles ont des conséquences sur la fertilité. Cependant, elles peuvent être maintenues par dérive ou parce qu'elles procurent un avantage sélectif aux porteurs. Repérer ces mutations constitue un enjeu majeur pour améliorer la fertilité d'une population (coûts économiques engendrés importants). Dans certaines espèces, la présence d'un observatoire dédié permet de faire remonter les cas repérés en élevage. Cependant, si le phénotype n'est pas disponible (cas de mortalité embryonnaire), l'anomalie passe généralement inaperçue. En bovins, l'utilisation de raisonnements de génétique inverse a fait ses preuves (Fritz et al., 2013). Ces raisonnements commencent à être utilisés dans d'autres espèces : porc (Derks, Lopes, et al., 2018) et volailles (Derks, Megens, et al., 2018). En ovins, l'avènement de la sélection génomique dans les races Lacaune et Manech à Tête Rousse a permis de collecter de nombreux génotypages. Une étude préliminaire avait été réalisée par Petit (2014) montrant des régions du génome en déficit d'homozygotie, mais les effectifs étaient réduits par rapport aux données disponibles actuellement, limitant la puissance de l'analyse. Ainsi, nous allons utiliser ces données de génotypages pour identifier des déficits en homozygotie dans ces deux populations en utilisant la méthode développée par Fritz et al. (2013) pour détecter des mutations létales en recherchant des segments chromosomiques présentant des déficits en homozygotie.

On fait l'hypothèse que de telles mutations existent dans les populations ovines laitières à l'origine *i*) d'une baisse de fertilité des mères porteuses de ces embryons et *ii*) sont conservées dans les populations liées à un certain avantage à l'hétérozygote.

Tableau 4. Description des individus génotypés de l'étude en races Lacaune et Manech à Tête Rousse. ¹Mâle, ²Femelle, ³Génotypage réalisé sur la puce 50 000 SNP, ⁴Béliers génotypés sur la puce basse densité puis sur la puce 50 000 SNP, ils correspondent aux béliers choisis pour devenir reproducteurs depuis la mise en place de la puce basse densité en 2017, ⁵Béliers génotypés exclusivement sur la puce basse densité SNP 15K.

Race	Lacaune		Manech à Tête Rousse	
	♂ ¹	♀ ²	♂	♀
Sexe	17 585	1 517	4 213	687
Génotypages disponibles	<u>50K</u> ³	<u>15K+50K</u> ⁴	<u>50K</u>	<u>15K+50K</u>
	12 090	510	3 184	297
Nombre d'animaux séquencés et génotypés	24		0	
	15K ⁵		15K	
	6 502		1419	
Total	19 102		4900	

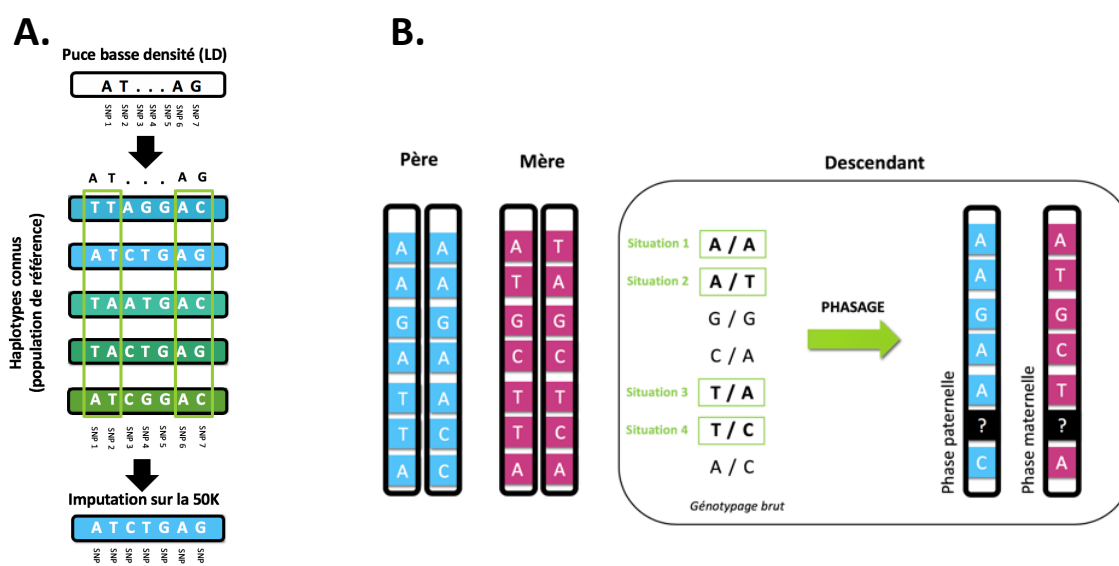


Figure 19. A. Principe de l'imputation, d'après Marchini et Howie (2010). Il s'agit d'une méthode statistique qui permet de prédire les génotypes des marqueurs non typés sur la puce basse densité présents sur la puce moyenne densité et de corriger les génotypes manquants des marqueurs (erreurs génotypage). Pour cela, on prend en compte les génotypes des animaux de la population de référence et la généalogie. Cette méthode repose sur le déséquilibre de liaison entre les haplotypes basse densité et moyenne densité d'une même région. **B. Principe du phasage.** A partir des génotypes des parents, il est possible de déduire l'origine des allèles du descendant génotypé suivant les marqueurs (Astruc, Buisson, 2018a). Le phasage se fait directement pour les situations 1, 2 et 3. Cependant, certaines situations nécessitent une imputation pour phaser le marqueur (situation 4). La reconstruction des phases est simple dans trois situations : *i*) Le descendant est homozygote au marqueur. Chaque parent a donné le même allèle (Situation 1) *ii*) Le descendant est hétérozygote au marqueur et le père homozygote, on en déduit l'allèle issu du père et donc de la mère (Situation 2) *iii*) Le descendant est hétérozygote au marqueur et la mère homozygote, on en déduit l'allèle issu de la mère et donc du père (Situation 3). Dans les autres situations, les typages sont **imputés** afin d'avoir des informations de génotypages complètes. Par exemple, lorsque le père et la mère sont hétérozygotes (Situation 4), il est impossible *a priori* d'attribuer l'origine des allèles du descendant. L'imputation permet par la prise en compte des liens de parenté et des génotypages de la population de référence de reconstruire les haplotypes phasés.

II. Matériel et Méthodes

A. Description des animaux génotypés disponibles pour l'étude

1. Description des animaux étudiés

La recherche de segments en déficit en homozygotie s'est réalisée avec 19 102 animaux Lacaune lait (millésimes** 1996 à 2019) et 4 900 animaux Manech à Tête Rousse (millésimes 1998 à 2019) génotypés issus des bases de données Valogène pour la sélection génomique (**Tableau 4**, voir **Annexe ii** pour la description plus détaillée des individus génotypés). Au laboratoire, les séquences complètes des génomes de 68 ovins de races différentes sont disponibles. Parmi ces animaux, 28 sont de race Lacaune lait (millésime 2001 à 2012) dont 24 sont génotypés sur puce SNP 50K, mais aucun individu Manech à Tête Rousse n'est séquencé. L'ensemble des séquences ont été alignées sur le génome de référence (*génomme Ovis aries v4.0*) et les variants ont été annotés à l'aide des outils SnpEFF et snpSIFT (**Cingolani et al., 2012**). Ces séquences et les fichiers de variants ont été mises à disposition par l'équipe bio-informatique SIGENAE de l'INRA de Toulouse.

2. Contrôle qualité et phasage des marqueurs SNP

L'ensemble des animaux de l'étude ont été génotypés sur la puce basse densité (15K, *sheepLD*) et/ou moyenne densité (50K, Illumina Ovine SNP50 BeadChip assay) (**Tableau 4**). Les génotypages pour la sélection génomique ont été réalisés par Labogéna (<http://www.labogena.fr>) et Aveyron Labo (<http://www.aveyron-labo.com>). Les données sont stockées sur la plateforme Valogène (hébergées au CTIG* de Jouy-en-Josas). Les 54 241 marqueurs SNP présents sur la puce 50K ont subi un contrôle qualité. Seuls les marqueurs présents sur les **26 autosomes** ont été gardés. Les marqueurs avec *i*) une fréquence de l'allèle minimal (**MAF*** : *Minimum Allele Frequency*) supérieure à 1%, *ii*) un pourcentage de typages > 97%, *iii*) et respectant l'équilibre de Hardy-Weinberg ($p > 10^{-5}$) ont été conservés (**Astruc et al., 2016**). En effet, on ne prend pas en compte ces marqueurs car dans la pratique, ils sont associés à des problèmes de génotypages. Il y a peu de chance que les marqueurs présents sur la puce soient associés individuellement à une mutation létale car ils ont été choisis sur le fait qu'ils soient polymorphes avec une MAF importante dans plusieurs populations. De ce fait, comme on recherche des mutations causales qui dérogent à l'équilibre de Hardy-Weinberg, on raisonne sur des haplotypes qui ont le même comportement que la mutation.

3. Imputation des génotypages 15K vers la 50K

Les génotypages basse densité (15K) ont été **imputés** vers la 50K (**Figure 19-A**) (processus statistique permettant de prédire le génotype des SNP non présents sur la puce 15K à partir des

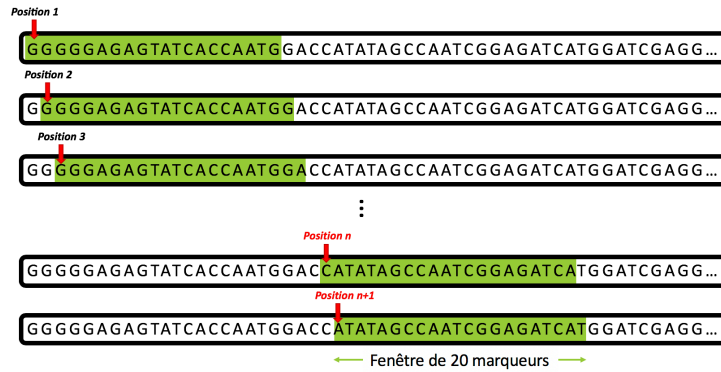


Figure 20. Principe du balayage du génome par une fenêtre glissante. Dans cet exemple la fenêtre est fixée à 20 marqueurs (encadré vert). La fenêtre glissante se trouve en position 1 : les fréquences de l'ensemble des haplotypes présents dans la fenêtre sont calculées sur les phases d'origine maternelle. Si cette fréquence est supérieure à 1%, des approches probabilistes permettent d'estimer le nombre d'animaux homozygotes attendus et de rendre compte de l'écart à loi de Hardy-Weinberg. Ensuite, la fenêtre se décale d'un marqueur vers la droite (passage de la fenêtre du marqueur n au marqueur $n+1$), l'ensemble des calculs sont de nouveau réalisés pour l'ensemble des haplotypes présents dans la fenêtre glissante.

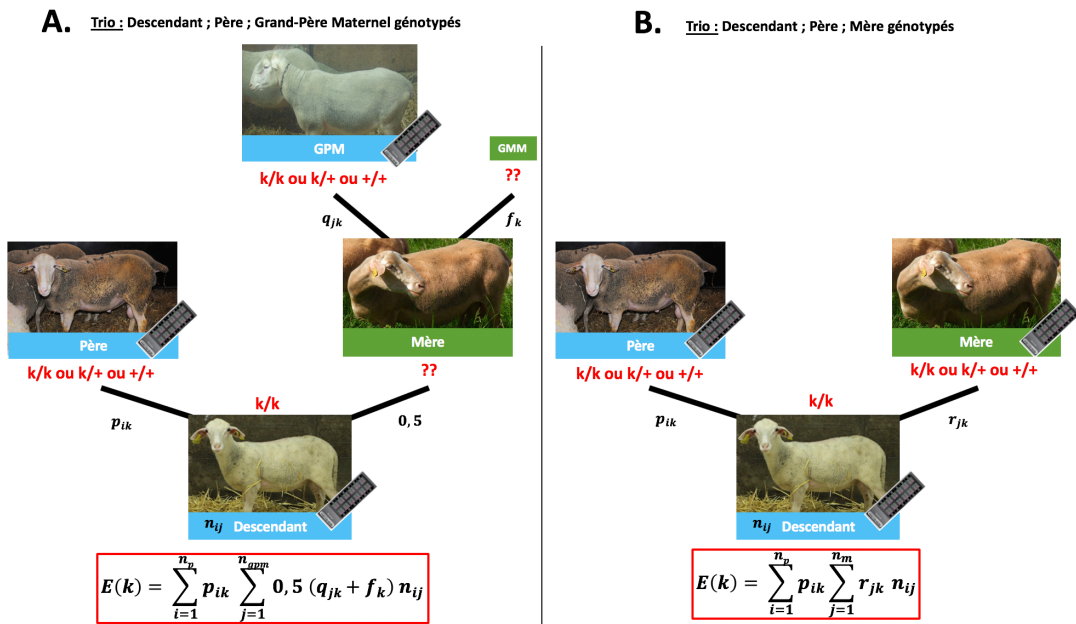


Figure 21. Estimation du nombre d'animaux homozygotes attendus à l'haplotype k , d'après Fritz et al. (2013) (Crédit photo : Diane Buisson). GPM : Grand-Père Maternel, GMM : Grand-Mère Maternel, + : indique un autre haplotype que k . **A. Calcul du nombre d'animaux attendus en prenant en compte le trio descendant, père et grand-père maternel :** $E(k)$: Nombre d'individus homozygotes attendus à l'haplotype k ; n_p : Nombre de pères ; n_{gpm} : Nombre de grand-pères maternels ; p_{ik} : Probabilité de transmettre l'haplotype k par le père i à ses descendants (0 si non porteur $+/+$ ou 0,5 si le père est hétérozygote $k/+$ ou 1 s'il est homozygote k/k) ; q_{jk} : Probabilité de transmettre l'haplotype k par le gpm j à ses descendants (0 si non porteur $+/+$ ou 0,5 si le gpm est hétérozygote $k/+$ ou 1 s'il est homozygote k/k) ; f_k : Fréquence de l'haplotype k dans la population ; n_{ij} : Nombre de descendants issus du père i et du gpm j . **B. Calcul du nombre d'animaux attendus en prenant en compte le trio descendant, père et mère :** n_m : Nombre de mères ; p_{ik} : Probabilité de transmettre l'haplotype k par le père i à ses descendants (0 si non porteur $+/+$ ou 0,5 si le père est hétérozygote $k/+$ ou 1 s'il est homozygote k/k) ; r_{jk} : Probabilité de transmettre l'haplotype k par la mère j à ses descendants (0 si non porteur $+/+$ ou 0,5 si la mère est hétérozygote $k/+$ ou 1 si elle est homozygote k/k) ; n_{ij} : Nombre de descendants issus du père i et de la mère j .

nombreux SNP typés de la puce 50K à partir de la population de référence). L'ensemble des typages ont été **phasés** (*ie* déterminer l'allèle issu de la mère et l'allèle issu du père pour chaque marqueur, (**Figure 19-B**)). Pour imputer et reconstruire les phases, le logiciel *FImpute* (**Sargolzaei et al., 2014**) a été utilisé. Au final **38 696 marqueurs SNP** ont été gardés pour la recherche de déficits en homozygotie. Les positions des marqueurs sont définies sur la carte ovine *version 4.0*.

B. Recherche d'haplotypes en déficit d'homozygotie

1. Principe de la recherche d'haplotype en déficit d'animaux homozygotes

En utilisant la méthode initiée par **VanRaden et al. (2011)** puis développée par **Fritz et al. (2013)**, on recherche des déficits en homozygotie sur les 26 autosomes ovins. Le programme de détection de déficit en homozygotie (langage awk) nous a été fourni par Sébastien Fritz et Chris Hoze de l'unité GABI (INRA Jouy-en-Josas) et a été adapté pendant mon stage à son utilisation en ovins.

La détection de déficit en homozygotie se fait **intra-race**. Elle repose sur l'utilisation moléculaire des marqueurs SNP qui balisent le génome. On parcourt le génome de marqueur en marqueur sur l'ensemble de chaque chromosome par une **fenêtre glissante de 10 ou 20 marqueurs** (**Figure 20**). Les fréquences des différents haplotypes observés dans la fenêtre glissante sont calculées sur les phases d'origine maternelle. Seuls les haplotypes présentant une fréquence supérieure à 1% sont retenus. Pour chaque haplotype (noté ***k***), on compare par un test du χ^2 le **nombre d'animaux attendus (noté $E(k)$) au nombre d'animaux homozygotes observés (noté $O(k)$)** à cet haplotype *k*.

Le calcul du nombre d'animaux attendus à un haplotype $E(k)$ repose sur des trios génotypés *i*) l'individu, son père et grand-père maternel (**Figure 21-A**) ou *ii*) l'individu, son père et sa mère (**Figure 21-B**). Si l'on ne dispose pas du génotype de la mère, son génotype est prédit en utilisant la fréquence de l'haplotype dans la population et le statut de son père (grand-père maternel, GPM). Au cours de l'analyse, les deux formules sont combinées. Tous les animaux génotypés disponibles ne sont pas inclus dans l'analyse (peu informatif car l'animal est non associé à un trio apparenté et génotypé). Au final, **16 346 animaux Lacaune et 3 661 animaux Manech à Tête Rousse** ont été utilisés pour détecter des haplotypes en déficit d'homozygotie (**Tableau 5**).

Un haplotype est significativement en déficit en homozygotie quand **$O(k) < E(k)$ par un test du χ^2** sous l'hypothèse de neutralité (ddl=1) avec un **seuil P-value $<10^{-3}$** (Étant donné le nombre de tests réalisés pour effectuer la recherche, il convient d'appliquer une correction de

Tableau 5. Description des animaux utiles pour la recherche de régions en déficit en homozygotie en races Lacaune et Manech à Tête Rousse. 15 568 descendants ont leur père et grand-père maternel ou père et mère de génotypés en race Lacaune. Respectivement, on a 3 290 descendants en race Manech à Tête Rousse. Un animal « utile » est un animal qui a été utilisé pour la recherche de déficit, appartenant à un trio.

Race	Nombre d'animaux génotypés disponibles	Nombre de descendants avec père et grand-père maternel de génotypés	Nombre de descendants avec père et mère de génotypés	Animaux utiles dans l'étude
Lacaune	19 102	15 530	38	16 346
Manech à Tête Rousse	4 900	3 259	31	3 661

Tableau 6. Table de calcul du déséquilibre de liaison entre deux loci 1 et 2 (Hill, Robertson, 1968). Si l'on considère deux locus 1 (A, a) et 2 (B, b) bi-alléliques présents sur le même chromosome, quatre haplotypes sont possibles : AB, Ab, aB, ab. On note les fréquences des allèles A, a, B, b : p_A, p_a, q_B et q_b et les fréquences haplotypiques : h_{AB}, h_{aB}, h_{Ab} et h_{ab} . On parle de déséquilibre de liaison quand le produit des fréquences alléliques est différent de la fréquence de l'haploype (ex : $h_{AB} \neq p_A * q_B$).

		Locus 1		
		A	a	
Locus 2	B	$h_{AB} \neq p_A \times q_B$	$h_{aB} \neq p_a \times q_B$	q_B
	b	$h_{Ab} \neq p_A \times q_b$	$h_{ab} \neq p_a \times q_b$	q_b
		p_A	p_a	1

Bonferroni en divisant le seuil global d'étude $\alpha=5\%$ par le nombre de tests indépendants réalisés, approché par le nombre de chromosome ($n=26$) et un **déficit supérieur ou égal à 75%**, fixé arbitrairement afin de repérer des haplotypes où l'on observe pas (100% déficit) ou peu (déficit compris entre 75,0 et 99,9%) d'animaux homozygotes (Défini par déficit (en %) = $\frac{E(k)-O(k)}{E(k)}$). Pour chaque fenêtre, plusieurs informations sont disponibles : numéro du chromosome, le numéro et la position (en Mb) du premier marqueur constituant l'haplotype, l'haplotype (fenêtre de marqueurs), le nombre d'animaux attendus et observés, la fréquence de l'haplotype, le pourcentage de déficit et la p-value du test du χ^2 . De plus, lorsque plusieurs haplotypes significativement en déficit sont chevauchants, on les regroupe pour obtenir un **haplotype concaténé**. Ainsi, chaque haplotype concaténé correspond à une région de localisation probable de mutation causale.

2. Assignation des génotypes aux marqueurs contenus dans les haplotypes pour l'ensemble des individus de la population totale génotypée

Tous les animaux génotypés disponibles n'ont pas été utilisés pour mettre en évidence des régions en déficit. A partir des haplotypes identifiés, on détermine le statut (homozygote muté « 2 », porteur hétérozygote « 1 » ou non porteur « 0 ») pour l'ensemble des animaux génotypés ($n=19\ 102$ en race Lacaune et $n=4\ 900$ en race Manech) disponibles dans l'étude. Un animal est défini comme non porteur lorsqu'il présente un haplotype différent de l'haplotype (sur les deux phases) en déficit identifié dans la région de localisation.

3. Calcul du déséquilibre de liaison entre haplotypes en déficit présents sur un même chromosome

Lorsque deux haplotypes sont présents sur un même chromosome, le calcul du déséquilibre de liaison permet de savoir si ces deux haplotypes sont liés (**Tableau 6**) et donc associés à la même mutation. A partir des fréquences haplotypiques et alléliques, on calcule le déséquilibre de liaison entre deux loci par $D = (h_{AB} * h_{ab}) - (h_{Ab} * h_{aB})$ (Hill, 1981), permettant ainsi de calculer un déséquilibre de liaison statistique (compris entre 0 et 1) : $r^2 = \frac{D^2}{p_A * p_a * q_B * q_b}$ (Hill, Robertson, 1968). Dans notre situation, on considère pour chaque locus bi-allélique que l'allèle a (respectivement b) correspond à l'haplotype en déficit d'homozygotie identifié et l'allèle A (respectivement B) correspond à l'ensemble des autres haplotypes que l'on observe dans la même région que l'haplotype en déficit.

C. Recherche de mutations causales et de gènes candidats dans les régions du génome présentant un déficit en homozygotie

1. Recherche de variants candidats présents dans les régions des haplotypes en déficit identifiés

Nous avons établi le statut porteur hétérozygote ou non porteur des haplotypes concaténés pour les 24 animaux Lacaune séquencés et génotypés avec la puce SNP 50K. Les fichiers VCF** contenant l'ensemble des variants (SNP et petites InDels) des 68 animaux séquencés présents dans les régions des haplotypes concaténés identifiés ont été générés grâce à l'instance bio-informatique Galaxy (<https://sigenae-worbench.toulouse.inra.fr>) et l'utilisation de l'outil *SnpSift Filter*. L'ensemble des polymorphismes peuvent être visualisés grâce au logiciel IGV (*Integrative Genome Viewer*, v.2.5.3, Thorvaldsdottir et al., 2013). Pour chacune des régions, les polymorphismes issus des fichiers de variants des 68 animaux (composés d'animaux Lacaune et d'animaux d'autres races) sont filtrés à partir des critères suivants : *i*) présents à l'état hétérozygote chez les individus porteurs, sur la base de l'haplotype, *ii*) absents chez les animaux non porteurs (homozygote référence), sur la base de l'haplotype, *iii*) non présents à l'état homozygote muté chez les animaux de la même race dont on ne connaît pas le génotype, sinon la mutation ne serait pas létale, *iv*) absents chez les individus d'autres races. Lorsqu'un variant candidat est identifié dans la séquence codante d'un gène et induit une substitution non-conservative dans la séquence protéique, l'effet de la mutation est prédit en utilisant l'outil *PolyPhen-2* (<http://genetics.bwh.harvard.edu/pph2/>) (Adzhubei et al., 2013).

2. Recherche de gènes candidats fonctionnels et positionnels pour les haplotypes n'ayant pas d'animaux porteurs séquencés

Quand l'information des séquences tout génome n'est pas disponible, une approche gènes candidats fonctionnels présents dans les régions des haplotypes d'intérêt a été réalisée. L'ensemble des gènes présents dans les zones d'intérêt ont été identifiés sous IGV puis choisis en fonction de l'annotation fonctionnelle du gène obtenue sur GeneCards (<https://www.genecards.org>) (Stelzer et al., 2016).

D. Effet des haplotypes sur les caractères sous sélection et phénotypes enregistrés en routine

Des analyses de variances ont été réalisées pour étudier l'effet du génotype (porteur hétérozygote vs non porteur) sur les caractères en sélection et sur des phénotypes enregistrés en routine (fertilité, qualité de semences) afin d'identifier un potentiel avantage à l'hétérozygote des individus porteurs de la mutation ou impactant les facultés reproductives de ces animaux.

1. Analyses statistiques

L'ensemble des analyses statistiques sur les performances des béliers ont été réalisées sur le logiciel SAS® (version 9.4, SAS Institute Inc., Cary, NC, USA).

a. Statistiques descriptives

La procédure UNIVARIATE de SAS® permet de représenter la distribution des différentes performances étudiées et de donner les éléments de statistiques descriptives. Nous avons pu

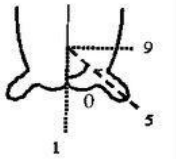
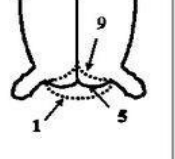
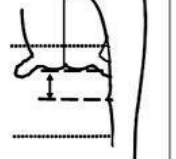
Angle des Trayons	Sillon	Distance Plancher-Jarret
		
1 = vertical 9 = horizontal	1 = absent 9 = bien marqué	1 = décroché 9 = ramassé

Figure 22. Description des différents postes de pointage de morphologie mammaire mesurés en race Lacaune (IDELE, 2017). Il s'agit de la grille de pointage utilisée dans le rayon de Roquefort. Une note de 1 à 9 est attribuée pour chaque poste. L'objectif étant d'avoir un angle de trayon compris entre 5 et 8, une profondeur de sillon supérieure à 5 et une distance plancher-jarret supérieure à 6.

Tableau 7. Description des jeux de données des DYD (« *Daughter Yield Deviation* ») pour les béliers génotypés en sélection en races Lacaune (A) et Manech à Tête Rousse (B). La production laitière (Lait) est exprimée en litres (L), les quantités (QMG : Quantité de Matière Grasse, QMP : Quantité de Matière Protéique) sont exprimées en hectogramme (hg), les taux (Butyreux (TB) et Protéique (TP)) sont exprimés en g/L et le comptage de cellules somatiques (CCS) et les caractères de morphologie mammaire (Angle, Sillon, Plancher-jarret (PLJ)) sont exprimés en écart-type génétique (σ_g).

A.

Caractère	Effectif (N)	Moyenne	Écart-type
Lait (L)	5478	32,2	16,1
QMG (hg)	5461	13,6	11,5
QMP (hg)	5462	10,7	8,4
TB (g/L)	5467	0,96	2,28
TP (g/L)	5460	0,74	1,33
CCS (σ_g)	5447	-0,03	0,51
Angle (σ_g)	4907	-0,25	0,39
Sillon (σ_g)	4855	-0,20	0,41
PLJ (σ_g)	5009	0,32	0,75

B.

Caractère	Effectif (N)	Moyenne	Écart-type
Lait (L)	1764	24,4	10,5
QMG (hg)	1714	4,3	7,5
QMP (hg)	1715	3,4	5,6
TB (g/L)	1718	0,13	2,34
TP (g/L)	1720	0,05	1,16
CCS (σ_g)	1571	0,04	0,54

ainsi vérifier la normalité de distribution de nos caractères, hypothèse à respecter pour réaliser une analyse de variance. D'autre part, les valeurs extrêmes supposées comme erronées (valeurs en dehors de l'intervalle de trois fois l'écart-type de la moyenne), sont éliminées de l'étude.

b. Analyse de variance multifactorielle par l'utilisation de modèle linéaire

L'objectif est d'étudier l'effet du génotype (deux modalités : non porteur ou porteur hétérozygote de l'haplotype) sur le phénotype. Les individus homozygotes à l'haplotype n'ont pas été pris en compte (effectif trop faible). Les conditions de réalisation de ce test ont été vérifiées : *i*) homoscedasticité (égalité de variances) des résidus (test de Bartlett avec la procédure HOVTEST de « GLM »), *ii*) normalité de la distribution des résidus (méthode graphique : QQPLOT) et *iii*) indépendance des distributions.

L'utilisation d'un modèle linéaire pour réaliser une analyse de variance multifactorielle permet de prendre en compte différents effets fixes tels que le millésime, le génotype. Pour chacun des effets fixes du modèle, on élimine les classes pour lesquels il y a peu d'individus (fréquence <1%). Pour chaque effet, on vérifie également qu'il y a au moins 5 animaux pour les différentes classes et groupes (porteurs/non porteurs). L'ensemble de ces calculs ont été réalisés par la procédure FREQ. Chacun des effets fixes est testé par un test exact de Fischer. L'ensemble des effets fixes qui ressortent significatifs peuvent être pris en compte dans le modèle. Le modèle linéaire est réalisé par la procédure GLM suivant le modèle suivant :

$$Y = \mu + X\beta + e$$

Où Y correspond au vecteur des performances, X : la matrice d'incidence des effets fixes, β : le vecteur des effets fixes, e : vecteur des effets aléatoires résiduels du modèle linéaire. S'il existe une différence significative pour la performance étudiée entre les différents génotypes (t-test), un test *post-hoc* est réalisé pour obtenir une estimation de la valeur de la performance corrigée par les effets fixes pour chacun des statuts, obtenue par le LSMEANS.

2. Effet des haplotypes sur les caractères de production

Pour les caractères de production, les DYD (« *Daughter Yield Deviation* ») des béliers ont été analysés. Ils correspondent à la « performance moyenne des filles de chaque bélier, corrigée pour les effets du milieu et la valeur génétique moyenne des mères » (Astruc et al., 2016). Ainsi, 6 caractères de production, disponibles en races Lacaune et Manech, sont mesurés en CLO : la production laitière (LAIT, calcul de Fleischmann corrigé pour la durée de traite), le taux butyreux (TB), le taux protéique (TP), QMG (=LAITxTB), QMP (=LAITxTP), comptage du nombre de cellules somatiques (CCS) (exprimé en écart-type génétique du caractère) (IDELE et al., 2016). De plus, en race Lacaune, trois postes de pointage de morphologie mammaire (Figure 22) sont enregistrés (angle des trayons, profondeur du sillon et distance plancher-

Tableau 8. Données de fertilité à l'IA et fertilité Globale pour les béliers génotypés en sélection en races Lacaune et Manech à Tête Rousse.

	Caractère	Effectif (N)	Moyenne	Écart-type
Lacaune	Fertilité à l'IA	5176	0,70	0,07
	Fertilité Globale	5031	0,96	0,02
Manech à Tête Rousse	Fertilité à l'IA	1961	0,61	0,07
	Fertilité Globale	1974	0,94	0,04

Tableau 9. Données de semences du premier prélèvement en CIA obtenues en races Lacaune et Manech à Tête Rousse pour les béliers génotypés de l'étude. Spz : spermatozoïdes.

	Caractère	Effectif (N)	Moyenne	Écart-type
Lacaune	Motilité (0-5)	4592	3,97	0,74
	Volume (mL)	4668	0,91	0,38
	Concentration (Milliard spz/mL)	3839	3,94	0,73
Manech à Tête Rousse	Motilité (0-5)	1878	4,69	0,15
	Volume (mL)	1894	1,18	0,58
	Concentration (Milliard spz/mL)	1906	2,88	0,64

Tableau 10. Données de semences moyennes des trois premiers prélèvements en CIA obtenues en races Lacaune et Manech à Tête Rousse pour les béliers génotypés de l'étude. Spz : spermatozoïdes.

	Caractère	Effectif (N)	Moyenne	Écart-type
Lacaune	Motilité (0-5)	13389	4,00	0,63
	Volume (mL)	13823	0,96	0,40
	Concentration (Milliard spz/mL)	11546	3,87	0,75
Manech à Tête Rousse	Motilité (0-5)	5563	4,69	0,15
	Volume (mL)	5612	1,26	0,62
	Concentration (Milliard spz/mL)	5642	2,93	0,63

jarret). Ils sont exprimés en écart-type génétique. Pour l'analyse, les DYD calculés pour un bélier repose sur au moins 20 lactations. Les millésimes pris en compte dans l'analyse sont compris entre 2000 et 2016 en race Lacaune et compris entre 2002 et 2016 en race Manech. Le **Tableau 7** décrit les jeux de données utilisés. En plus de l'effet haplotype, l'effet millésime a été utilisé dans le modèle.

3. Effet des haplotypes sur la fertilité

Les données de **fertilité à l'IA** et de **fertilité globale** sont disponibles pour étudier l'effet du statut sur la fertilité des béliers. La fertilité à l'IA correspond au rapport entre le nombre de mises bas par IA et le nombre d'IA total. La fertilité globale correspond au rapport entre le nombre de mises bas (mises bas sur IA ou sur retour, cela inclut toutes les mises bas même si le bélier est le bélier de retour) et le nombre de mises bas par IA. Les millésimes pris en compte dans l'analyse sont compris entre 2001 et 2016 en race Lacaune et compris entre 2001 et 2015 en race Manech (**Tableau 8**). Un modèle linéaire prenant en compte les effets millésime et haplotype de l'animal a été utilisé pour ce caractère également.

4. Effet des haplotypes sur la qualité de semences

Depuis, 2005, la qualité de la semence est mesurée dans les CIA selon trois principaux paramètres : la motilité massale (note de 0 à 5), le volume (en mL) et la concentration (en milliards spz/mL). Dans les deux races, on étudie les millésimes de 2005 à 2018 sur jeunes béliers (de 0-1-2 ans en race Lacaune et de 1 à 2 ans en race Manech). Deux modèles ont été mis en place prenant en compte *i*) le premier prélèvement de la carrière d'un jeune bélier et *ii*) moyenne des 3 premiers prélèvements afin d'étudier si les animaux porteurs des haplotypes en déficit sont écartés suite à des défauts de qualité de semences lors de leur entrée en CIA. Les jeux de données sont décrits dans les **Tableaux 9 et 10**. Le modèle d'analyse inclut les effets : âge au prélèvement, mois et année de collecte, centre de prélèvement et haplotype. Pour le caractère moyenné, l'effet du mois de collecte n'a pas été pris en compte.

5. Ajustement des seuils de significativité par une correction de Bonferroni

La correction de Bonferroni consiste à corriger le seuil de significativité α lors de la réalisation de tests multiples sur un même jeu de données. Ici, on divise le seuil $\alpha=5\%$ par le nombre de tests effectués (nombre haplotypes testés \times nombre de caractères étudiés). En race Lacaune, 12 haplotypes concaténés ont été testés sur 12 performances indépendantes (QMG et QMP étant directement liées à la production laitière et aux taux). En race Manech, 6 haplotypes concaténés ont été testés sur 9 performances indépendantes. Le seuil de détection ($\alpha=5\%$) est de $3,5 \times 10^{-4}$ en race Lacaune et $9,3 \times 10^{-4}$ en race Manech.

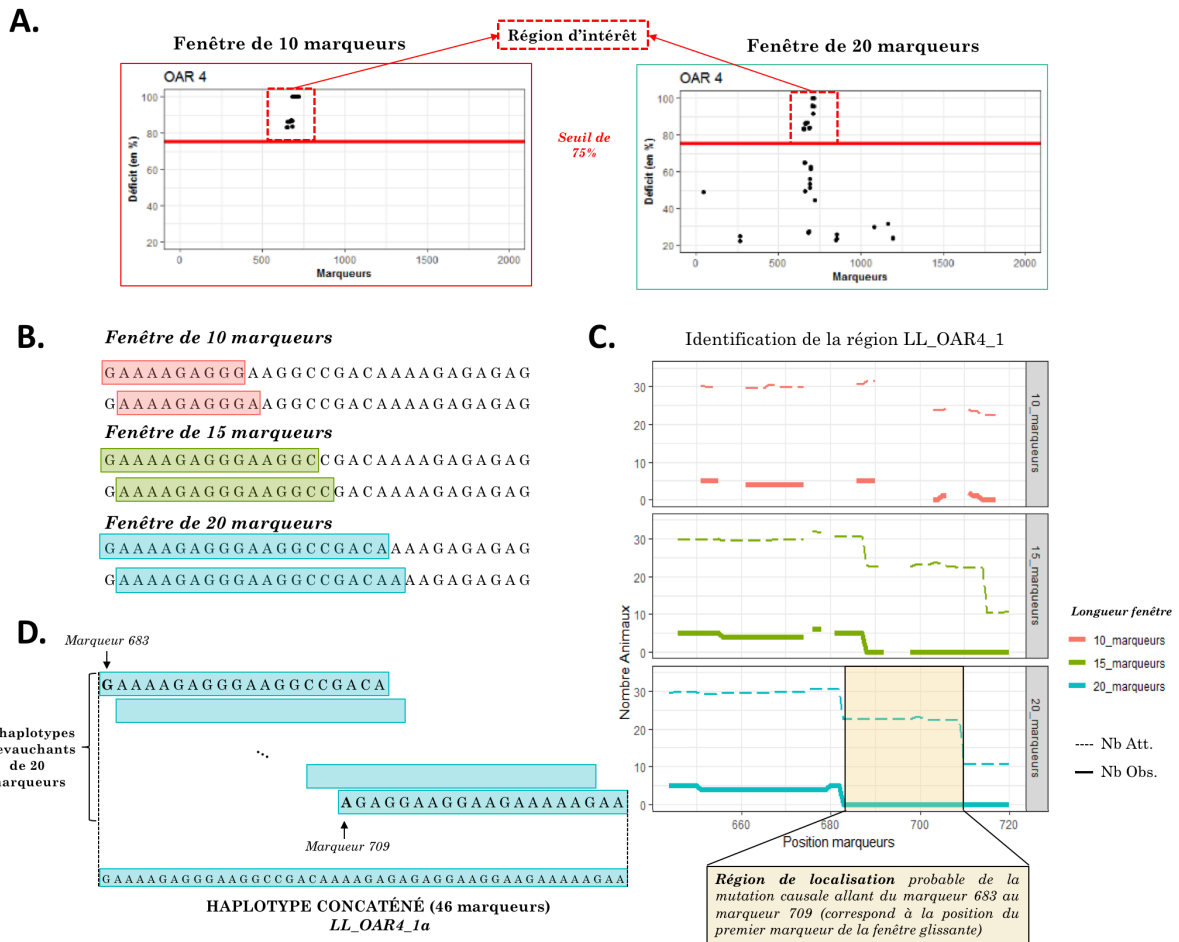


Figure 23. Méthode pour déterminer les haplotypes concaténés en déficit en homozygotie à partir des haplotypes de 10 et 20 marqueurs (exemple de l'haplotype LL_OAR4_1a). **A.** Exécution du programme de recherche de déficit pour des fenêtres glissantes de 10 et 20 marqueurs. Cela conduit à la mise en évidence d'une région d'intérêt. Chaque point représente le numéro du premier marqueur de l'haplotype de 10 ou 20 marqueurs. **B.** Exécution du programme dans la région d'intérêt pour des fenêtres glissantes de 10, 15 et 20 marqueurs. **C.** Représentation graphique du nombre d'animaux attendus (---) et du nombre d'animaux observés (—) pour chaque longueur d'haplotype constituant la fenêtre glissante en fonction du premier marqueur de chaque haplotype de la fenêtre glissante. Cela permet de mettre en évidence une région de localisation (encadré jaune). **D.** Obtention de l'haplotype concaténé à partir des haplotypes chevauchants présentant les mêmes estimations.

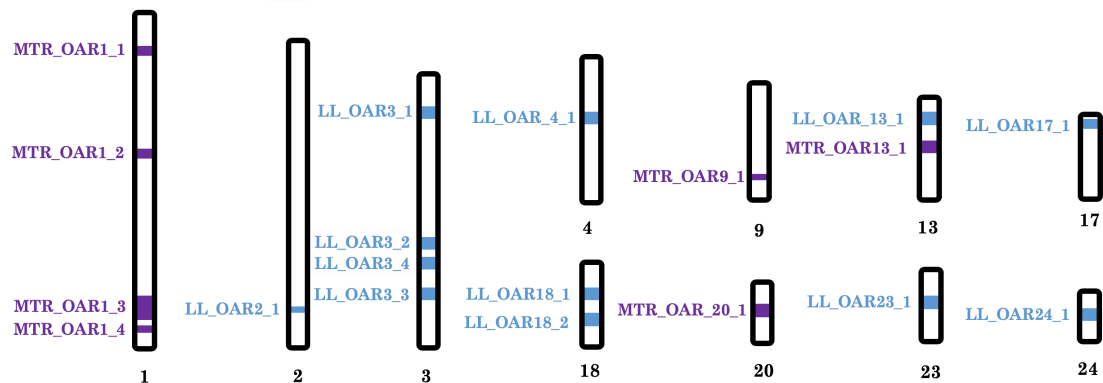


Figure 24. Représentation des régions de localisation probable de mutations causales où l'on détecte des déficits en homozygotie en races Lacaune et Manech à Tête Rousse. Les 12 régions identifiées en race Lacaune sont représentées en bleu et les 7 régions en race Manech sont représentées en violet. Chaque bâton représente un chromosome (OAR).

III. Résultats

A. Recherche de déficit en homozygotie en races Lacaune et Manech à Tête Rousse

1. Identification des haplotypes en déficit d'homozygotie

En balayant le génome avec des fenêtres glissantes de **10 et 20 marqueurs**, on recherche des déficits en homozygotie dans chacune des deux populations ovines à partir des génotypages phasés de 16 346 animaux Lacaune utiles et 3 661 animaux Manech à Tête Rousse utiles. Comme nous recherchons des mutations létales ou affectant le jeune animal, on émet l'hypothèse que ces mutations sont liées à des haplotypes où l'on n'observe aucun ou peu d'animaux homozygotes. Ainsi, seuls les haplotypes présentant un **déficit > 75%** sont conservés. Cette recherche met en évidence **131 haplotypes de 10 marqueurs et 202 haplotypes de 20 marqueurs** en race Lacaune ; **30 haplotypes de 10 marqueurs et 52 haplotypes de 20 marqueurs** en race Manech à Tête Rousse présentant un déficit d'homozygotie supérieur 75% au seuil de $\alpha=10^{-3}$.

Pour les différentes régions d'intérêt mises en évidence (**Figure 23-A**), on refait tourner le programme de recherche de déficit avec des fenêtres glissantes de 10, 15 et 20 marqueurs (**Figure 23-B**) afin d'identifier la taille de fenêtre la plus optimale permettant d'obtenir la plus grande région en déficit d'homozygotie (correspondant à des haplotypes chevauchants significativement en déficit). Une fois la taille de fenêtre choisie, on identifie la **région de localisation** définie par les différents haplotypes chevauchants présentant le même minimum d'animaux observés (**Figure 23-C**). Ces haplotypes sont regroupés pour définir un **haplotype concaténé** (**Figure 23-D**). On exécute de nouveau le programme avec une fenêtre glissante correspondant au nombre de marqueurs constituant l'haplotype concaténé afin de déterminer les différentes caractéristiques de cet haplotype. On émet l'hypothèse que les haplotypes concaténés définissent une région de localisation probable d'une mutation causale.

Ainsi, la recherche de segments en déficit en homozygotie conduit à identifier **12 haplotypes concaténés en race Lacaune et 7 haplotypes concaténés en race Manech à Tête Rousse** (**Figure 24, Tableau 11**). Par la suite, les régions de localisation sont nommées de la manière suivante : race (LL ou MTR) + numéro du chromosome (OAR) + numéro de la région chromosomique concernée par le déficit d'homozygotie (1,2,3 ...). Les différents haplotypes concaténés, que l'on retrouve dans chacune des régions de localisation, sont identifiés par le nom de la région suivi d'une lettre de a à z (*ex : LL_OAR4_1a*).

Au sein des deux races, on distingue deux cas de figures (**Tableau 11**). D'un côté on observe des haplotypes où l'on a aucun animal homozygote (100% de déficit), on peut faire

Tableau 11. Description des haplotypes concaténés en déficit d'animaux homozygotes obtenus en races Lacaune (LL) et Manech à Tête Rousse (MTR). ¹OAR : Chromosome, ²marqueur, ³fréquence de l'haplotype calculée sur les phases d'origine maternelle, ⁴P-value obtenue par un test du χ^2 (ddl=1). Les haplotypes sont classés par race et par pourcentage décroissant de déficit. Les haplotypes potentiellement associés à des mutations létales sont indiqués en orange (déficit de 100%). *pb* : paires de bases, *Mb* : Mégabases, *les positions sont données sur la version du génome ovin 4.0*.

Race	Nom Haplotype concaténé	OAR ¹	Nombre de marqueurs composant l'haplotype concaténé	Borne de la région de localisation			Nombre homozygotes				
				Position marq ² début (pb)	Position marq ² fin (pb)	Position marq ² fin (pb)	Taille (Mb)	Fréquence ³ (%)	Attendus	Observés	Déficit (%)
Lacaune	LL_OAR3_1a	3	31	32 341 158	34 040 951	1,70	2,4%	10	0	100%	1,6E-03
Lacaune	LL_OAR3_4a	3	17	132 151 018	133 151 531	1,00	2,2%	12	0	100%	4,8E-04
Lacaune	LL_OAR4_1a	4	46	43 354 129	46 320 680	2,97	3,3%	21	0	100%	3,8E-06
Lacaune	LL_OAR13_1a	13	22	44 880 292	46 565 715	1,69	2,9%	17	0	100%	2,9E-05
Lacaune	LL_OAR3_3a	3	27	135 970 984	137 458 257	1,49	5,4%	72	3	96%	4,7E-16
Lacaune	LL_OAR17_1a	17	23	140 420	1 337 222	1,20	1,9%	12	1	92%	1,5E-03
Lacaune	LL_OAR23_1a	23	12	29 299 212	30 256 424	0,96	1,4%	16	2	87%	5,2E-04
Lacaune	LL_OAR24_1a	24	15	16 573 708	17 855 812	1,28	3,7%	16	2	87%	4,8E-04
Lacaune	LL_OAR2_1a	2	10	162 361 305	162 800 670	0,44	2,9%	21	3	86%	9,8E-05
Lacaune	LL_OAR18_1a	18	23	25 730 034	27 530 707	1,80	3,1%	14	2	86%	1,5E-03
Lacaune	LL_OAR18_2a	18	26	32 943 056	34 598 799	1,66	2,7%	17	3	82%	8,8E-04
Lacaune	LL_OAR3_2a	3	18	129 113 785	130 473 652	1,36	3,5%	21	4	81%	1,6E-04
MTR	MTR_OAR1_2a	1	10	98 276 401	98 948 647	0,67	5,1%	10	0	100%	1,4E-03
MTR	MTR_OAR1_3a	1	34	229 877 877	232 208 143	2,33	4,5%	10	0	100%	1,6E-03
MTR	MTR_OAR1_4a	1	21	233 410 109	235 168 365	1,76	4,3%	11	0	100%	1,4E-03
MTR	MTR_OAR20_1a	20	31	21 296 641	23 043 970	1,75	5,2%	13	0	100%	3,7E-04
MTR	MTR_OAR1_1a	1	10	9 849 067	10 363 151	0,51	5,3%	15	2	86%	1,0E-03
MTR	MTR_OAR9_1a	9	10	71 751 035	72 383 768	0,63	4,3%	14	2	86%	1,3E-03
MTR	MTR_OAR13_1a	13	26	61 544 570	63 888 090	2,34	8,1%	38	6	84%	2,4E-07

l'hypothèse qu'ils sont associés à des mutations létales (embryonnaires, périnatales ou néonatales). De l'autre côté, on observe des haplotypes présentant un déficit en homozygotie partiel (compris entre 81% et 96%) où quelques individus homozygotes sont observés, on peut faire l'hypothèse que ces haplotypes sont en déséquilibre de liaison partiel avec la mutation létale ou qu'ils peuvent conduire à une létalité plus tardive compromettant la survie du jeune ou associés à des défauts congénitaux (animal viable mais écarté de la sélection).

Quatre haplotypes en races Lacaune (LL_OAR3_1a, LL_OAR3_4a, LL_OAR4_1a et LL_OAR13_1a, couvrant des régions chromosomiques de 1,0 à 3,0 Mb) et quatre haplotypes en race Manech à Tête Rousse (MTR_OAR1_2a, MTR_OAR1_3a, MTR_OAR1_4a et MTR_OAR20_1a, couvrant des régions chromosomiques de 0,7 à 2,3 Mb) semblent être associés à des mutations létales puisqu'aucun individu homozygote est observé pour ces haplotypes alors que le nombre d'animaux homozygotes attendus en fonction des informations parentales et populationnelles était de 10 à 21 en race Lacaune et de 10 à 13 en race Manech à Tête Rousse. Les fréquences de ces haplotypes sont plutôt faibles comprises entre 2,2% et 3,3% en race Lacaune et 4,3% à 5,2% en race Manech à Tête Rousse.

Huit haplotypes en races Lacaune (LL_OAR3_3a, LL_OAR17_1, LL_OAR23_1a, LL_OAR24_1a, LL_OAR2_1a, LL_OAR18_1a, LL_OAR18_2a et LL_OAR_3_2a, couvrant des régions chromosomiques de 0,4 à 1,8 Mb) et 3 haplotypes en race Manech à Tête Rousse (MTR_OAR1_1a, MTR_OAR9_1a et MTR_OAR13_1a, couvrant des régions chromosomiques de 0,5 à 2,3 Mb) présentent quelques individus homozygotes. Le nombre d'individus homozygotes observés varie de 1 à 4 en race Lacaune et de 2 à 6 en race Manech à Tête Rousse alors que le nombre d'animaux homozygotes attendus était de 12 à 72 en race Lacaune et de 14 à 38 en race Manech à Tête Rousse. Leurs fréquences sont comprises entre 1,4% et 5,4% en race Lacaune et de 4,3% à 8,1% en race Manech à Tête Rousse.

2. Étude du lien entre les haplotypes présents sur un même chromosome

La recherche de déficit en homozygotie met en évidence l'existence de plusieurs haplotypes présents sur un même chromosome (**Figure 24**). Pour chacune des races, on regarde si ces haplotypes sont en déséquilibre de liaison afin de savoir s'ils pourraient être associés à la même mutation. Les haplotypes identifiés sur le chromosome 3 en race Lacaune ne présentent aucune association entre eux (**Tableau 12-A**). Les haplotypes LL_OAR18_1a et LL_OAR18_2a ne présentent qu'une association partielle (**Tableau 12-B**). Ainsi, les haplotypes découverts en Lacaune ne semblent pas être associés à la même mutation causale. Ils seront donc traités indépendamment dans la suite de l'étude. Il en est de même en Manech à Tête Rousse pour les haplotypes présents sur le chromosome 1. A l'exception des haplotypes

Tableau 12. Calcul du déséquilibre de liaison entre les différents haplotypes en déficit présents sur un même chromosome en races Lacaune (A. OAR3 ; B. OAR18) et Manech à Tête Rousse (C. OAR1). Le déséquilibre de liaison (r^2) est indiqué en pourcentage. La distance entre chaque haplotype est indiquée en mégabases (Mb). Les calculs sont réalisés à partir des génotypes phasés des 19 102 animaux Lacaune et des 4 900 animaux Manech à Tête Rousse.

A.				B.	
OAR3	LL_OAR3_2a	LL_OAR3_3a	LL_OAR3_4a	OAR18	LL_OAR18_2a
LL_OAR3_1a	0% (95,1 Mb)	1% (101,9 Mb)	0% (98,1 Mb)	LL_OAR18_1a	40% (5,4 Mb)
LL_OAR3_2a		0% (5,5 Mb)	0% (1,7 Mb)		
LL_OAR3_3a			0% (2,8 Mb)		

C.			
OAR1	MTR_OAR1_2a	MTR_OAR1_3a	MTR_OAR1_4a
MTR_OAR1_1a	0% (87,9 Mb)	0% (219,5 Mb)	0% (223,0 Mb)
MTR_OAR1_2a		0% (130,9 Mb)	0% (134,5 Mb)
MTR_OAR1_3a			76% (1,2 Mb)

Tableau 13. Assignment des génotypes aux marqueurs constitutifs des différents haplotypes présents dans les loci identifiés dans l'ensemble de la population génotypée. Pour chaque haplotype, le nombre d'animaux homozygotes mutés (H/H), hétérozygotes (H/+) et non porteurs (+/+) est indiqué. Remarque : Pour le nombre d'animaux homozygotes mutés, le nombre d'animaux typés seulement sur la puce basse densité (LD) est également indiqué (H/H^{LD}), il s'agit de béliers qui n'ont pas été retenus pour être candidats. (n=19 102 en race Lacaune et n=4 900 en race Manech à Tête Rousse).¹ Le nombre d'animaux entre parenthèse représente le nombre d'animaux homozygotes observés lors de la recherche de déficit à partir des animaux utiles de l'étude. ² Correspond au millésime du premier animal porteur hétérozygote de l'haplotype en déficit d'homozygotie. Les haplotypes présentant un déficit de 100% sont représentés en orange. Les haplotypes pour lesquels on observe des individus homozygotes dans l'ensemble de la population alors qu'aucun individu observé n'avait été repéré dans la sous-population sont représentés en jaune.

Région	+/+		H/+		H/H ¹	dont H/H ^{LD}		Premier [He] ² obs.
LL_OAR3_4	18293	95,80%	809	4,20%	0 (0)	0	0,00%	1998
LL_OAR4_1	17824	93,30%	1278	6,70%	0 (0)	0	0,00%	1998
LL_OAR13_1	17897	93,70%	1205	6,30%	0 (0)	0	0,00%	1996
LL_OAR3_1	18252	95,60%	849	4,40%	1 (0)	1	0,01%	1996
LL_OAR18_1	18046	94,50%	1054	5,50%	2 (2)	0	0,01%	1999
LL_OAR23_1	18361	96,10%	739	3,90%	2 (2)	2	0,01%	2005
LL_OAR2_1	17842	93,40%	1257	6,60%	3 (3)	0	0,02%	1996
LL_OAR3_3	16781	87,80%	2318	12,10%	3 (3)	2	0,02%	1996
LL_OAR17_1	18182	95,20%	917	4,80%	3 (1)	1	0,02%	2001
LL_OAR24_1	17845	93,40%	1254	6,60%	3 (2)	1	0,02%	1997
LL_OAR18_2	18008	94,30%	1090	5,70%	4 (3)	1	0,02%	1999
LL_OAR3_2	17642	92,40%	1452	7,60%	8 (4)	1	0,04%	1996
MTR_OAR1_3	4398	89,80%	502	10,20%	0 (0)	0	0,00%	1999
MTR_OAR20_1	4338	88,50%	562	11,50%	0 (0)	0	0,00%	1998
MTR_OAR1_2	4365	89,10%	534	10,90%	1 (0)	1	0,02%	1998
MTR_OAR9_1	4337	88,50%	559	11,40%	4 (2)	1	0,08%	1999
MTR_OAR1_1	4310	88,00%	585	11,90%	5 (2)	4	0,10%	1999
MTR_OAR13_1	4036	82,40%	854	17,40%	10 (6)	2	0,20%	1999

MTR_OAR1_3a et MTR_OAR1_4a qui sont en fort déséquilibre de liaison ($r^2=76\%$) et semblent donc être associés à la même mutation causale (**Tableau 12-C**). Par la suite, seul MTR_OAR1_3a sera étudié (mais MTR_OAR1_4a sera pris en compte pour la recherche de gènes/variants candidats).

Ainsi, la recherche de déficit en homozygotie permet d'identifier 12 haplotypes en race Lacaune et 6 haplotypes en race Manech à Tête Rousse potentiellement associés à des mutations délétères indépendantes.

3. Détermination des génotypes aux différents loci identifiés pour l'ensemble des individus de la population génotypée

La mise en évidence d'haplotypes en déficit en homozygotie a été réalisée avec une sous-partie des individus génotypés (individus « utiles » associés à des trios). Ainsi, grâce à l'information génotypique il est désormais possible d'assigner les génotypes aux marqueurs constitutifs des haplotypes contenus dans les régions de localisation pour chacun des animaux de la population génotypée (**Tableau 13**). Cela permet également de déterminer la fréquence de l'haplotype dans l'ensemble de la population (prise en compte des deux phases paternelle et maternelle pour chacun des individus). On distingue alors trois situations :

- 1) Aucun individu homozygote est observé dans l'ensemble de la population pour trois haplotypes en race Lacaune (LL_OAR3_4a, LL_OAR4_1a et LL_OAR13_1a) et deux haplotypes en race Manech à Tête Rousse (MTR_OAR1_3a et MTR_OAR20_1a). Cela permet de conforter l'hypothèse que ces haplotypes sont associés à des mutations létales (embryonnaires, périnatales ou néonatales). Les fréquences des porteurs hétérozygotes sont faibles en race Lacaune (4,2% à 6,7%) et plus importante en race Manech (10,2% à 11,5%).
- 2) Pour deux haplotypes (LL_OAR3_1a et MTR_OAR1_2a) on observe la présence d'un individu homozygote alors qu'un déficit de 100% avait été mis en évidence.
- 3) Pour le reste des haplotypes (présentant des déficits partiels d'individus homozygotes), on observe qu'ils ségrégent à des fréquences faibles en race Lacaune (~4-8%), à l'exception de l'haplotype LL_OAR3_3a présent à une fréquence importante (12,1%) dans la population. En Manech, les haplotypes identifiés sont présents à des fréquences importantes (11,4% à 17,4%). Pour l'ensemble de ces haplotypes, on observe quelques individus homozygotes supplémentaires.

B. Recherche de mutations causales dans les loci identifiés

1. Exploration des données de séquences tout génome en race Lacaune

Lorsque l'on recherche une mutation létale ou affectant le jeune, on ne peut disposer facilement de matériel biologique. L'approche de génétique inverse réalisée permet de

Tableau 14. Description des résultats de la recherche de variants dans les 8 régions dont l'on dispose d'animaux porteurs séquencés en race Lacaune. ¹ Les positions sont données sur la version du génome ovin 4.0, ² Chromosome, ³ L'ensemble des variants ont été filtrés à partir des données de séquence des 68 ovins disponibles. Les régions associées aux haplotypes présentant un déficit de 100% sont représentées en orange. *pb* : paires de bases.

Région	Déficit (en %)	Borne gauche (pb) ¹	Borne droite (pb) ¹	OAR ²	Taille (en Mb)	Nombre animaux séquencés porteurs de l'haplotype en déficit	Nombre de variants dans la région	Nombre de variants obtenus après l'application des filtres ³
LL_OAR3_1	100%	32 341 000	34 041 000	3	1,7	2	20 616	95
LL_OAR4_1	100%	43 354 000	46 321 000	4	3,0	2	37 597	24
LL_OAR13_1	100%	44 740 000	46 696 000	13	2,0	1	24 273	43
LL_OAR3_3	96%	135 970 000	137 459 000	3	1,5	2	19 168	16
LL_OAR17_1	92%	140 000	1 338 000	17	1,2	1	22 947	16
LL_OAR23_1	87%	29 299 000	30 351 500	23	1,1	1	20 757	27
LL_OAR24_1	87%	16 573 500	17 856 000	24	1,3	1	24 827	39
LL_OAR3_2	81%	129 113 500	130 474 000	3	1,4	3	17 215	1

Tableau 15. Liste des mutations candidates identifiées dans les régions de localisation des haplotypes en déficit en race Lacaune. ¹ Les positions sont données sur la version du génome ovin 4.0, ² allèles Référence/Alternatif, ³ correspond aux abréviations des acides aminés à 3 lettres. Les régions associées aux haplotypes présentant un déficit de 100% sont représentées en orange.

Région	Déficit (en %)	OAR : position ¹	Réf>Alt ²	Variant candidat	Polymorphisme	Conséquence ³
LL_OAR3_1	100%	3: 34 013 956	G>A	Faux-sens exon gène <i>PREB</i>	c.997C>T	p.Leu333Phe
LL_OAR13_1	100%	13: 45 430 645	del CAAG>C	Délétion exon gène <i>LOC101108851</i>		Délétion d'une Lysine
LL_OAR3_3	96%	3: 136 776 372	G>A	5' UTR gène <i>DHH</i>	c.-193G>A	
LL_OAR3_3	96%	3: 136 933 452	C>A	Non-sens exon (codon stop) gène <i>CCDC65</i>	c.331G>T	p.Glu111*
LL_OAR3_3	96%	3: 137 042 797	C>T	Synonyme exon gène <i>ADCY6</i>	c.1515C>T	p.Ala505Ala

rechercher les variants candidats chez les animaux porteurs hétérozygotes. On fait l'hypothèse que la mutation causale, associée à une anomalie, est en déséquilibre de liaison total avec l'haplotype identifié. Pour cela, des approches de bio-informatique ont été réalisées afin de rechercher des variants candidats dans les données de séquences permettant de confirmer la présence de mutations (létale ou compromettant la survie du jeune animal).

En race Lacaune, on dispose de 28 animaux séquencés tout génome dont 24 animaux sont génotypés dans notre population d'étude. La recherche de polymorphismes candidats est possible pour 8 régions en race Lacaune (**Tableau 14**) pour lesquelles on dispose d'individus porteurs séquencés (1 à 3 individus par haplotype en déficit). Cette recherche ne peut être réalisée pour les régions identifiées en race Manech à Tête Rousse sans individu séquencé. Le nombre de variants (SNP et InDels) présent dans les régions de l'étude varie entre 20 000 à 40 000 (**Tableau 14**). Compte-tenu du nombre d'animaux porteurs hétérozygotes, les différents filtres appliqués résultent à la sélection de moins de 100 variants. Une partie de ces variants sont présents dans les exons (5' et 3' UTR compris), les régions promotrices ou régulatrices (en amont ou en aval des gènes annotés) (**Annexe iii**). Devant ce nombre encore important de variants potentiellement fonctionnels, j'ai restreint cette liste aux seuls variants présents dans les exons, cadre de lecture et régions UTR non traduites.

a. Identification de mutations candidates associées aux haplotypes présentant un déficit total d'individus homozygotes

Les différents filtres cités précédemment permettent de mettre en évidence des polymorphismes candidats dans deux des trois régions (LL_OAR3_1 et LL_OAR13_1) potentiellement associés à des mutations létales dont on dispose d'animaux porteurs à ces haplotypes (**Tableau 15**). En effet, une mutation faux-sens dans le gène *PREB* (*Prolactin Regulatory Element Binding*) est identifiée dans la région LL_OAR3_1. Cette mutation faux-sens en position c.997C>T de l'ADNc est localisée dans l'exon VII (**Figure 25-B**) du gène *PREB* entraînant une substitution d'une Leucine en Phénylalanine (position 333 sur 391 acides aminés (**Figure 25-C**)) affectant un domaine protéique de type WD40). La mutation identifiée est prédite délétère par l'algorithme *PolyPhen-2* (**Figure 25-D**). Le gène *PREB* code pour une protéine régulatrice à motif de liaison WD capable de se fixer sur le promoteur de la prolactine et intervient dans la régulation de la transcription (**Fliss et al., 1999**). De plus, dans la région LL_OAR_13_1, une délétion d'une lysine est observée dans le gène *LOC101108851* qui coderait une protéine de type isopentenyl-diphosphate Delta-isomerase 1 (**NCBI**).

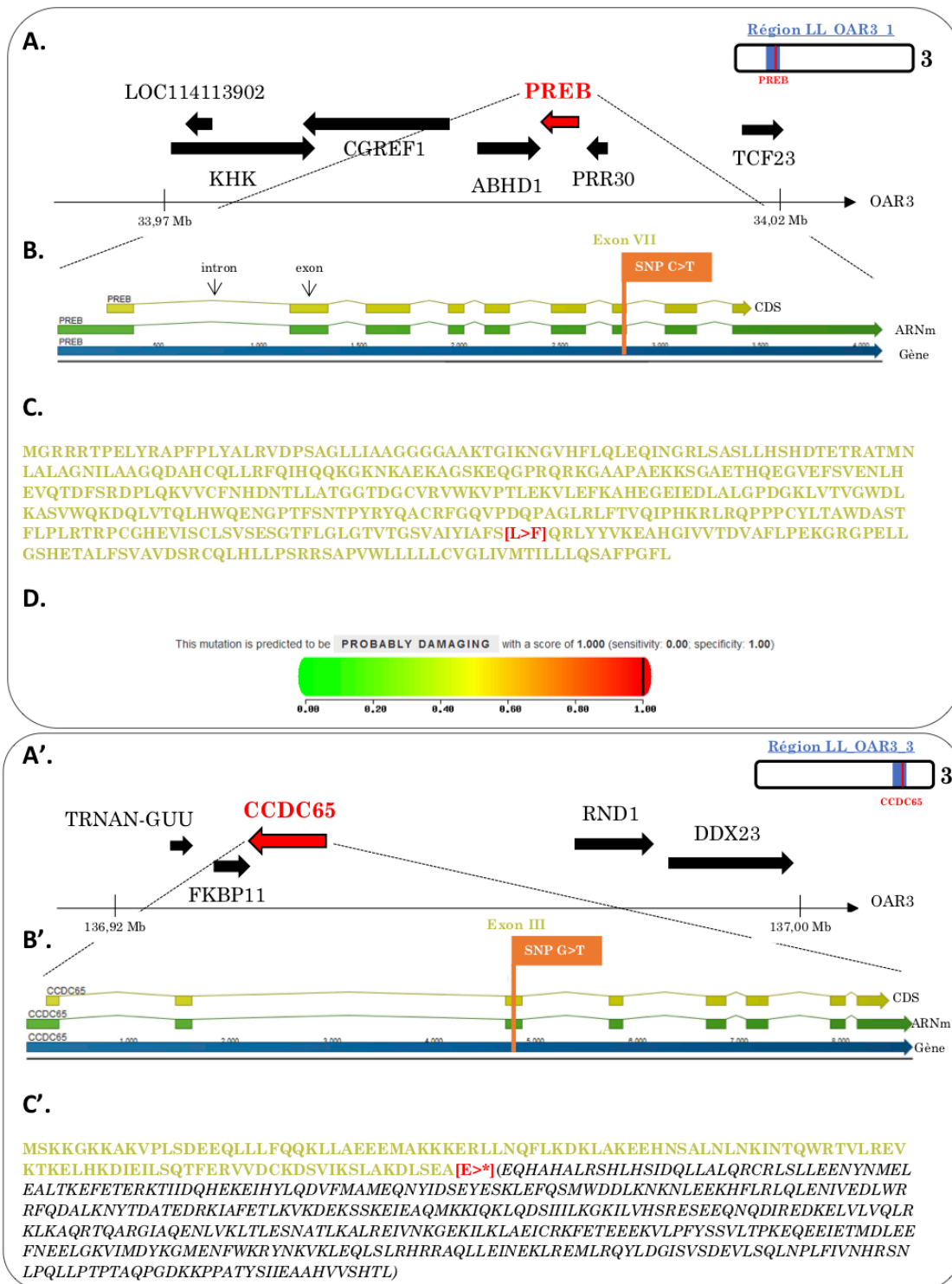


Figure 25. Identification de deux mutations candidates dans les gènes *PREB* (*Prolactin Regulatory Element Binding*) et *CCDC65* (*Coiled-Coil Domain Containing 65*) (*OAR3*) en race Lacaune. **A.** **A'** Extraits des régions entourant les gènes candidats identifiés (version du génome ovine 4.0, d'après NCBI). **B.** **B'** Représentation du gène, de l'ARNm et de la séquence traduite (CDS) **C.** **C'** Séquences protéiques. Les localisations des mutations sont indiquées en rouge (pour le gène *PREB*, substitution d'une Leucine (L) en Phénylalanine (F) en position 333. Pour le gène *CCDC65*, l'acide glutamique (E) est substitué par un codon stop). La séquence traduite est indiquée en vert. **D.** Prédiction de l'effet de la mutation sur la structure de la protéine de *PREB*. Pour cela, un score allant de 0 (mutation neutre) à 1 (mutation délétère) est estimé. (outils PolyPhen2 - Adzhubei et al., 2013)

b. Identification de mutations candidates associées aux haplotypes présentant un déficit partiel d'individus homozygotes

Parmi les 5 régions associées à des déficits en homozygotie partiel, une seule région contient des variants candidats (**Tableau 15**). En effet, la région LL_OAR3_3 contient une mutation non-sens (codon stop) entraînant un raccourcissement de la protéine CCDC65 (*Coiled-Coil Domain Containing 65*). La mutation non-sens (c.331G>T) est localisée dans l'exon III (**Figure 25-B'**) entraînant en théorie la traduction que de 22% de CCDC65 (codon stop en position 111 alors que la protéine traduite fait 498 acides aminés (**Figure 25-C'**)). CCDC65 fait partie du complexe régulateur nexin-dynéine assurant la formation d'organites ciliés (cils et flagelles) (Horani et al., 2013). Les deux autres variants candidats sont localisés dans le 5' UTR du gène *DHH* (*Desert Hedgehog Signaling Molecule*) et dans un exon du gène *ADCY6* (*Adenylate cyclase 6*, variant synonyme). *DHH* code pour des protéines de signalisation jouant un rôle majeur dans la morphogénèse. Le gène *ADCY6* intervient dans la synthèse de l'AMP cyclique (voie de signalisation).

c. Cas particuliers : identification de deux gènes sous sélection dans les régions LL_OAR3_2 et LL_OAR13_1

i. Lien entre LL_OAR3_2a et le gène de résistance aux mammites *SOCS2* en race Lacaune

L'étude des variants présents dans la région LL_OAR3_2 n'a pas mis en évidence de polymorphismes candidats. Cependant, dans cette région sur le chromosome 3 il existe un gène faisant l'objet d'une sélection. Il s'agit du gène de résistance aux mammites *SOCS2* (Rupp et al., 2015). Le gène *SOCS2* présente un variant (C>T entraînant la substitution non conservative R96C) dont l'allèle muté T est associé à des comptages cellulaires élevés (CCS) dans le lait (en lien avec les mammites) en race Lacaune. Actuellement, en sélection, les béliers homozygotes mutés à cet allèle sont écartés. La sélection s'effectue en priorité sur les animaux C/C et une partie des animaux C/T. Ainsi, on peut émettre l'hypothèse que l'haplotype en déficit d'homozygotie identifié (LL_OAR3_2a) est associé à l'allèle muté T de la mutation *SOCS2*. Pour cela, on dispose de 4367 génotypages à ce locus du gène *SOCS2*. En effet, 97% des individus hétérozygotes à l'haplotype détecté sont porteurs hétérozygotes de la mutation *SOCS2* (**Tableau 16**). Cependant, la réciproque n'est pas vraie : être C/T ne veut pas dire que l'animal est porteur de l'haplotype LL_OAR3_2a identifié.

Étant donné que l'haplotype mis en évidence par la recherche de déficit (LL_OAR3_2a) n'explique qu'en partie le lien avec le génotype à la mutation *SOCS2*, on fait l'hypothèse que d'autres haplotypes sont associés à la mutation dans la région LL_OAR3_2. L'ensemble des haplotypes observés dans la région de LL_OAR3_2 (du marqueur 2113 au 2130) ont été

Tableau 16. Lien entre l'haplotype LL_OAR3_2a et l'allèle muté T du gène SOCS2 (OAR3) en race Lacaune. + : correspond aux autres haplotypes de la région LL_OAR3_2. 4367 génotypages au variant SOCS2 sont disponibles (béliers de millésimes 2017-18-19).

Génotypes	C/C	C/T	T/T	Total
LL_OAR3_2a/LL_OAR_2a	0	0	0	0
LL_OAR3_2a/LL_OAR3_2+	2	178	3	183
LL_OAR3_2+/LL_OAR3_2+	3544	632	8	484
Total	3546	810	11	4367

Tableau 17. Liste des haplotypes associés à la mutation SOCS2 (OAR3) présents dans la région de LL_OAR3_2 en race Lacaune. ¹ Chromosome, ² Fréquence de l'haplotype calculée sur les phases d'origine maternelle.

Nom de de l'haplotype	OAR ¹	Nombre marqueurs haplotype	Taille (Mb)	Fréquence ² (%)	Nombre homozygotes		
					Attendus	Observés	Déficit (%)
LL_OAR3_2a	3	18	1,36	3,46%	21	4	81%
LL_OAR3_2b	3	18	1,36	2,77%	13	3	77%
LL_OAR3_2c	3	18	1,36	1,79%	8	0	100%

Tableau 18. Identification de 3 haplotypes liés à l'allèle muté T du gène SOCS2 (OAR3) en race Lacaune. L'ensemble des marqueurs présents dans la zone de LL_OAR3_2 (du marqueur 2113 à 2130) sont indiqués. Le gène SOCS2 est localisé entre le marqueur 2119 et 2120. Le sous-haplotype de LL_OAR3_2 « AGGGGGG » (nommé LL_OAR3_2_T) est retrouvé pour les 3 haplotypes LL_OAR3_2a, LL_OAR3_2b et LL_OAR3_2c.

Haplotypes	Région LL_OAR3_2																Fréquence			
	2113	2114	2115	2116	2117	2118	2119	SOCS2	2120	2121	2122	2123	2124	2125	2126	2127		2128	2129	2130
LL_OAR3_2a	A	G	G	G	G	G	G		G	C	A	G	A	G	C	A	G	A	A	3,46%
LL_OAR3_2b	A	G	G	G	G	G	G		A	A	A	G	G	A	A	G	G	G	G	2,77%
LL_OAR3_2c	A	G	G	G	G	G	G		A	A	A	G	G	A	A	G	G	G	A	1,79%

LL_OAR3_2_T

Tableau 19. Lien entre l'haplotype LL_OAR3_2_T et l'allèle muté T du gène SOCS2 (OAR3) en race Lacaune. + : correspond aux autres haplotypes différents de LL_OAR3_2_T (du marqueur 2113 à 2119). 4367 génotypages au variant SOCS2 sont disponibles (béliers de millésimes 2017-18-19).

Génotypes	C/C	C/T	T/T	Total
LL_OAR3_2_T/LL_OAR3_2_T	0	2	11 (100%)	13
LL_OAR3_2_T/LL_OAR3_2_+	20	776 (96%)	0	796
LL_OAR3_2_+/LL_OAR3_2_+	3526 (99%)	32	0	3558
Total	3546	810	11	4367

identifiés (fréquence >1% sur les phases maternelles) (**Annexe iv**) et le lien entre ces haplotypes et le génotype à *SOCS2* a été réalisé (**Annexe v**). Parmi les 23 haplotypes identifiés, 2 haplotypes supplémentaires (LL_OAR3_2b et LL_OAR3_2c) présentent le même comportement que LL_OAR3_2a (**Tableau 17**). Ils n'ont pas été mis en évidence lors de la première recherche car leur p-value associée est en dessous du seuil de significativité. Cependant, ces deux haplotypes présentent un déficit en homozygotie de plus de 75%. Lorsque l'on regarde les allèles qui composent les trois haplotypes associés à *SOCS2*, on constate qu'ils présentent le même début de séquence « AGGGGGG » (**Tableau 18**). Ainsi, l'exécution du programme de recherche de déficit pour une taille d'haplotype de 7 marqueurs permet de mettre en évidence que le sous-haplotype « AGGGGGG » (nommé LL_OAR3_2_T) est présent à une fréquence de 12% sur les phases maternelles et présente un déficit de 61% (221 Att. vs 77 Obs.). Ce sous-haplotype montre une association parfaite avec le génotype à *SOCS2* (**Tableau 19**). De plus, les 11 animaux T/T présents dans notre étude (**Annexe vi**) sont également homozygotes à l'haplotype LL_OAR3_2_T. Ainsi, on a mis en évidence 3 haplotypes qui sont, ensemble, des bons prédicteurs du génotype à *SOCS2*.

ii. Lien entre LL_OAR13_1a et le gène de résistance/sensibilité à la tremblante *PRNP* en race Lacaune

L'haplotype identifié en race Lacaune LL_OAR13_1a présente dans sa région le gène *PRNP*. Ce gène code pour la protéine prion, à l'origine de la tremblante ovine (**Elsen et al., 1997**). Quatre haplotypes principaux (ARR, AHQ, ARQ et VRQ) au gène *PRNP*, correspondant à 3 polymorphismes présents dans trois codons (136, 154 et 171), sont utilisés en sélection (**Annexe i**). Depuis 2004, tous les béliers sont génotypés au gène *PRNP* à l'âge de 2-3 semaines et seuls les animaux ARR/ARR sont gardés, l'allèle Q étant responsable de la sensibilité à la tremblante. Étant donné que les animaux les plus anciens disponibles dans l'étude sont de millésime 1996, un déficit en homozygotie pourrait être associé à l'allèle Q (exon 171) contre sélectionné dans la population. Cependant, l'haplotype LL_OAR13_1a et l'allèle Q ne montrent pas de lien (**Tableau 20**). De plus, le calcul de la fréquence de l'allèle Q (**Tableau 21**) montre qu'un déficit lié à cet allèle n'aurait pas pu être détecté (fréquence <1%) par cette méthode. Ainsi, LL_OAR13_1a ne semble pas lié à l'allèle Q de l'exon 171 du gène de sensibilité à la tremblante.

2. Identification de gènes candidats dans les régions ne disposant pas d'animaux séquencés porteurs des haplotypes en déficit

Lorsque l'information des séquences tout génome n'était pas disponible, une approche gènes candidats fonctionnels présents dans les régions des haplotypes d'intérêt a été réalisée. On fait

Tableau 20. Lien entre l'haplotype LL_OAR13_1a et l'allèle Q de sensibilité à la tremblante (exon 171) du gène *PRNP* (OAR13) en race Lacaune. + : Correspond aux autres haplotypes présents dans la même région que l'haplotype LL_OAR13_1a. 17 205 génotypages au gène *PRNP* sont disponibles en race Lacaune.

Génotypes	Q/Q	Q/R	R/R	Total
LL_OAR13_1a/LL_OAR13_1a	0	0	0	0
LL_OAR13_1a/LL_OAR13_1+	0	10	1048	1058
LL_OAR13_1+/LL_OAR13_1+	4	133	16010	16147
Total	4	143	17058	17205

Tableau 21. Calcul de la fréquence de l'allèle Q (codon 171 du gène *PRNP*) en utilisant le calcul d'équilibre de Hardy-Weinberg en race Lacaune (**Encadré 1**).

	Q/Q	Q/R	R/R	Total
Nb. Obs.	4	143	17058	17205
Fréq. Obs.	0,02%	0,83%	99,15%	100,0%
Fréq Att.	0,002%	0,874%	99,124%	100,0%
Nb. Att.	0	150	17054	17205

Fréquence allélique Q : 0,4%
Fréquence allélique R : 99,6%

Tableau 22. Identification des principaux gènes candidats fonctionnels localisés dans les régions probablement associées à des mutations létales ne disposant pas d'animaux séquencés (déficit 100%) en races Lacaune et Manech à Tête Rousse.

Région	Gènes candidats fonctionnels et positionnels (d'après GeneCards https://www.genecards.org (Stelzer et al., 2016))
LL_OAR3_4	Cluster gènes <i>HOXC</i> (<i>Homeobox</i> , intervient lors du développement embryonnaire), <i>MAP3K2</i> (<i>Mitogen-Activated Protein Kinase Kinase Kinase 12</i> , différenciation neuronale), <i>AMHR2</i> (<i>Anti-Mullerian Hormone Receptor Type 2</i> , développement organes génitaux), <i>SP1</i> (<i>Sp1 Transcription Factor</i> , réparation de l'ADN, remodelage de la chromatine), <i>SP7</i> (<i>Sp7 Transcription Factor</i> , facteur de transcription spécifique des os), <i>AAS</i> (<i>Aladin WD Repeat Nucleoporin</i> , protéine régulatrice, développement du système nerveux central et périphérique), <i>ESPL1</i> (<i>Extra Spindle Pole Bodies Like 2, Separase</i> , stabilisation des chromatides), <i>RARG</i> (<i>Retinoic Acid Receptor Gamma</i> , régulation de la transcription), <i>ZNF740</i> (<i>Zinc Finger Protein 740</i> , régulation de la transcription), <i>TNS2</i> (<i>Tensin 2</i> , régulation de la migration cellulaire), <i>EIF4B</i> (<i>Eukaryotic Translation Initiation Factor 4B</i> , signalisation cellulaire).
MTR_OAR1_2	Cluster gènes codant pour des histones (<i>H2B type 2-F, H3, H2A type 2-A, H2B type 2-E</i> , constituants protéiques des chromosomes), <i>SF3B4</i> (<i>Splicing Factor 3b Subunit 4</i> , assemblage des complexes spliceosomes).
MTR_OAR1_3/ MTR_OAR1_4	<i>GPR149</i> (<i>G Protein-coupled Receptor 149</i> , transduction cellulaire), <i>DHX36</i> (<i>DEAH-Box Helicase 36</i> , Stabilisation de l'ADN) - <i>MBNL1</i> (<i>Muscleblind Like Splicing Regulator 1</i> , régulation de la transcription, impliqué dans la dystrophie myotonique), <i>IGSF10</i> (<i>Immunoglobulin Superfamily Member 10</i> , migration des neurones), <i>MED12L</i> (<i>Mediator Complex Subunit 12L</i> , régulation de la transcription).
MTR_OAR20_1	<i>CENPQ</i> (<i>Centromere Protein Q</i> , sous-unités des centromères), <i>CRISP2</i> (<i>Cysteine Rich Secretory Protein 2</i> , Capacitation des spermatozoïdes), <i>CRISP3</i> (<i>Cysteine Rich Secretory Protein 3</i> , structure des spermatozoïdes, implantation de l'embryon), <i>CRISP1</i> (<i>Cysteine Rich Secretory Protein 1</i> , fusion spermatozoïde-ovule), <i>TFAP2D</i> (<i>Transcription Factor AP-2 Delta</i> , facteur de transcription), <i>TFAP2B</i> (<i>Transcription Factor AP-2 Beta</i> , facteur de transcription).

l'hypothèse qu'une mutation dans un gène intervenant lors du développement embryonnaire, ou dans la structure de l'ADN peut être à l'origine de la mort de l'animal ou de tares génétiques.

En effet, la recherche de gènes candidats montrent que dans les régions potentiellement associées à des mutations létales (LL_OAR3_4, MTR_OAR1_2a et MTR1_3a/MTR_OAR1_4a et MTR_OAR20_1a), on retrouve des gènes dont leurs fonctions sont essentielles au **développement embryonnaire** (mise en place de structure, modification de la structure de l'ADN..., **Tableau 22**). En race Lacaune, un cluster de gènes *HOXC* est présent dans la région de localisation de l'haplotype LL_OAR3_4. Il s'agit de gènes intervenant pendant le développement embryonnaire lors de l'organogenèse. La région MTR_OAR1_2 contient un cluster de gènes codant pour des histones (constituants protéiques des chromosomes). Au sein de la région MTR_OAR20_1, on observe la présence des gènes *TFAP2D* et *TFAP2B* qui codent pour des facteurs de transcription, jouant un rôle majeur au cours du développement embryonnaire (**Zhao et al., 2003**). Cependant dans ces régions, on trouve également des gènes impliqués dans la **fonction de reproduction**. En effet, on retrouve le gène *AMHR2* intervenant dans le développement des organes génitaux (région LL_OAR3_4). Au sein de la région MTR_OAR20_1, un cluster de gènes *CRISP* est identifié, ces gènes sont importants pour assurer une bonne fertilité mâle.

En ce qui concerne les régions où l'on a des déficits partiels en homozygotie, on retrouve également des gènes intervenant dans la structure de l'ADN, ou codant pour des facteurs de transcription, ou encore intervenant dans des voies de signalisation cellulaire (**Annexe vii**). Ainsi, une mutation délétère à l'état homozygote dans l'un de ces gènes pourrait être à l'origine d'une altération de la survie du jeune animal ou d'une réforme précoce des animaux.

C. Effet des haplotypes identifiés sur les performances

1. Étude d'un potentiel avantage des individus hétérozygotes sur les caractères en sélection

Des haplotypes en déficit en homozygotie associés à des mutations létales ségrégent dans les deux populations étudiées à des fréquences non négligeables. La conservation de telles mutations pourrait s'expliquer par un avantage sélectif. Ainsi pour l'ensemble des haplotypes identifiés en races Lacaune et Manech, on évalue l'effet des différents haplotypes sur les caractères en sélection (Production laitière, Quantité Matière Grasse, Quantité Matière Protéique, Taux Butyreux, Taux Protéique, Comptage cellules somatiques dans le lait (CCS) et les caractères de morphologie mammaire (uniquement en race Lacaune)).

Parmi l'ensemble des haplotypes en déficit d'homozygotie, quatre d'entre eux en race Lacaune et un en race Manech présentent un avantage à l'état hétérozygote (**Tableau 23**). En

Tableau 23. Étude d'un potentiel avantage sélectif entre individus porteurs hétérozygotes et non porteurs des haplotypes identifiés en races Lacaune et Manech à Tête Rousse. Seuls les résultats significatifs sont montrés. ¹ Les astérisques indiquent le degré de significativité associés aux seuils $\alpha=5\%$ (*), $\alpha=1\%$ (**) et $\alpha=0,1\%$ (***) corrigés par la correction de Bonferroni, ² He : Hétérozygote, ³ erreur-type. Lait : Production laitière (Litres), QMP : Quantité Matière Protéique (hg), TB : Taux Butyreux (hg), TP : Taux Protéique (g/L), CCS : Comptage Cellules somatiques (exprimé en écart-type génétique). Une différence verte indique un effet favorable sur le caractère pour l'hétérozygote (Respectivement, différence rouge : défavorable). Les régions associées aux haplotypes présentant un déficit de 100% sont représentées en orange.

Haplotype	Caractère	P-value ¹	Effectif non porteurs	Effectif porteurs [He] ²	Moyenne non porteurs (Lsmeans \pm e.t. ³)	Moyenne porteurs [He] (Lsmeans \pm e.t. ³)	Différence porteurs [He]/non porteurs
LL_OAR3_1a	QMP (hg)	***	5159	303	10,50 \pm 0,10	8,67 \pm 0,38	-1,83 hg
LL_OAR3_1a	TP (g/L)	***	5158	302	0,73 \pm 0,02	0,41 \pm 0,07	-0,32 g/L
LL_OAR3_1a	CCS (σ_g)	**	5145	302	-0,012 \pm 0,007	-0,124 \pm 0,028	-0,112 σ_g
LL_OAR3_4a	Lait (L)	***	5260	218	31,22 \pm 0,21	39,06 \pm 0,95	+7,85 L
LL_OAR3_4a	QMP (hg)	***	5242	220	10,31 \pm 0,10	12,32 \pm 0,45	+2,01 hg
LL_OAR3_4a	TB (g/L)	***	5246	221	0,95 \pm 0,03	0,15 \pm 0,15	-0,79 g/L
LL_OAR3_4a	TP (g/L)	***	5239	221	0,73 \pm 0,02	0,24 \pm 0,09	-0,49 g/L
LL_OAR4_1a	TP (g/L)	**	5069	391	0,70 \pm 0,02	0,95 \pm 0,07	+0,26 g/L
LL_OAR17_1a	Lait (L)	***	5348	129	31,40 \pm 0,21	37,18 \pm 1,24	+5,78 L
MTR_OAR9_1a	CCS (σ_g)	*	1376	194	0,031 \pm 0,026	-0,114 \pm 0,044	-0,144 σ_g

particuliers pour les haplotypes en déficit complet, les filles des béliers Lacaune porteurs hétérozygotes de l'haplotype LL_OAR3_1a ont une meilleure santé de la mamelle, associée à moins de cellules somatiques dans le lait que les brebis non porteuses de cet haplotype ($-0,1\sigma_g$). Néanmoins, cela entraîne une dégradation du taux protéique. Les porteurs de LL_OAR3_4a ont des filles qui montrent une meilleure production laitière (+7,8L), mais associée à la dégradation des taux (protéines et lipides). Le lait des brebis porteuses de l'haplotype LL_OAR_4_1a est plus riche en protéines. Seul, l'haplotype LL_OAR17_1, présentant un déficit partiel en homozygotie, conduit à améliorer la production laitière des hétérozygotes (+5,8L). En race Manech, les porteurs de MTR_OAR9_1a ont une meilleure santé de la mamelle associée à un comptage cellulaire plus faible ($-0,14\sigma_g$).

Les individus hétérozygotes porteurs de LL_OAR3_2_T (et donc de l'allèle T du variant *SOCS2*) sont prédisposés aux mammites (concentration de cellules somatiques élevée : $+0,20\sigma_g$) (**Tableau 24**). Cette mutation a un effet pléiotrope sur la production laitière (+4,2 litres) et les quantités (QMG et QMP). Néanmoins, cela conduit à dégrader le TB ($-0,8g/L$) et la morphologie de la mamelle (réduction de la distance plancher-jarret de $-0,12\sigma_g$, associée à des mamelles qui tombent). Lorsque l'on regarde les effets des trois haplotypes qui prédisent le génotype à *SOCS2*, on constate que les hétérozygotes ont un comptage cellulaire élevé, associé à une meilleure production laitière à l'origine de la dégradation des taux et des caractères de morphologie mammaire. On peut également noter que les 3 haplotypes identifiés n'ont pas les mêmes effets (**Tableau 24**).

2. Évaluation de l'impact des haplotypes sur l'efficacité reproductive mâle

a. Étude de la fertilité

La comparaison entre porteurs et non porteurs pour chacun des haplotypes identifiés ne montrent pas de différences significatives pour la fertilité à l'IA et la fertilité globale des béliers.

b. Qualités de semences

L'analyse de la qualité des semences des béliers Lacaune et Manech ne montrent pas de différences significatives entre porteurs et non porteurs des haplotypes. En effet, les valeurs de motilité, volume et concentration observées sont conformes aux objectifs : Motilité ≥ 4 , concentration ≥ 2 milliards spz/mL et volume compris entre 0,5 et 2 mL (**CIIRPO, 2009**).

IV. Discussion et perspectives

La méthode développée par **Fritz et al. (2013)**, permet par une approche « *bottom-up* » de mettre en évidence des anomalies dans les deux races ovines en recherchant des haplotypes en déficit en homozygotie.

Tableau 24. Étude d'un potentiel avantage sélectif entre individus porteurs et non porteurs des haplotypes associés au variant *SOCS2* (OAR3). Seuls les résultats significatifs sont montrés. ¹ Les astérisques indiquent le degré de significativité associés aux seuils $\alpha=5\%$ (*), $\alpha=1\%$ (**) et $\alpha=0,1\%$ (***) corrigés par la correction de Bonferroni, ² He : Hétérozygote, ³ erreur-type. Lait : Production laitière (Litre), QMG : Quantité Matière Grasse (hg), QMP : Quantité Matière Protéique (hg), TB : Taux Butyreux (g/L), TP : Taux Protéique (g/L), CCS : Comptage Cellules somatiques (écart-type génétique). Les caractères de morphologie mammaire (profondeur du sillon, PLJ : distance Plancher-Jarret) sont exprimés en écart-type génétique. Une différence verte indique un effet favorable sur le caractère pour l'hétérozygote (Respectivement, différence rouge : défavorable).

Haplotype	Caractère	P-value ¹	Effectif non porteurs	Effectif porteurs [He] ²	Moyenne non porteurs (Lsmeans \pm e.t. ³)	Moyenne porteurs [He] (Lsmeans \pm e.t. ³)	Différence porteurs [He]/non porteurs
LL_OAR3_2_T	Lait (L)	***	3897	1560	30,29 \pm 0,24	34,44 \pm 0,36	+4,15 L
LL_OAR3_2_T	QMG (hg)	***	3879	1561	12,71 \pm 0,16	14,00 \pm 0,24	+1,28 hg
LL_OAR3_2_T	QMP (hg)	***	3884	1557	9,59 \pm 0,11	12,25 \pm 0,17	+2,65 hg
LL_OAR3_2_T	TB (g/L)	***	3880	1566	1,14 \pm 0,04	0,36 \pm 0,06	-0,78 g/L
LL_OAR3_2_T	CCS (σ_g)	***	3880	1546	-0,079 \pm 0,008	0,122 \pm 0,013	+0,201 σ_g
LL_OAR3_2_T	PLJ (σ_g)	***	3576	1416	0,391 \pm 0,014	0,269 \pm 0,020	-0,123 σ_g
LL_OAR3_2a	Lait (L)	***	4944	533	31,09 \pm 0,22	35,26 \pm 0,58	+4,17 L
LL_OAR3_2a	QMG (hg)	***	4928	532	12,90 \pm 0,14	14,87 \pm 0,40	+1,97 hg
LL_OAR3_2a	QMP (hg)	***	4928	533	10,15 \pm 0,10	12,41 \pm 0,29	+2,26 hg
LL_OAR3_2a	TB (g/L)	***	4933	533	0,97 \pm 0,03	0,46 \pm 0,10	-0,50 g/L
LL_OAR3_2a	CCS (σ_g)	***	4920	526	-0,035 \pm 0,008	0,122 \pm 0,021	+ 0,157 σ_g
LL_OAR3_2a	Sillon (σ_g)	***	4383	471	-0,219 \pm 0,006	-0,307 \pm 0,019	-0,088 σ_g
LL_OAR3_2a	PLJ (σ_g)	***	4528	480	0,356 \pm 0,012	0,174 \pm 0,003	-0,182 σ_g
LL_OAR3_2b	QMP (hg)	***	5134	327	10,27 \pm 0,10	12,25 \pm 0,37	1,98 hg
LL_OAR3_2b	TB (g/L)	***	5134	332	0,97 \pm 0,03	0,03 \pm 0,12	-0,94 g/L
LL_OAR3_2b	CCS (σ_g)	***	5121	325	-0,028 \pm 0,007	0,126 \pm 0,009	+0,154 σ_g
LL_OAR3_2c	Lait (L)	***	5279	199	31,23 \pm 0,21	39,60 \pm 0,99	+8,37 L
LL_OAR3_2c	QMP (hg)	***	5260	202	10,29 \pm 0,09	12,88 \pm 0,47	2,59 hg
LL_OAR3_2c	TB (g/L)	***	5265	202	0,95 \pm 0,03	-0,01 \pm 0,15	-0,96 g/L
LL_OAR3_2c	TP (g/L)	***	5258	202	0,73 \pm 0,02	0,22 \pm 0,09	-0,51 g/L

Tableau 25. Déséquilibre de liaison entre l'haplotype LL_OAR3_4a et les haplotypes associés à la mutation *SOCS2* (OAR3). Les calculs sont réalisés à partir des génotypes phasés des 19 102 animaux Lacaune.

OAR3	LL_OAR3_2_T (Allèle muté T du gène <i>SOCS2</i>)	LL_OAR3_2a	LL_OAR3_2b	LL_OAR3_2c
LL_OAR3_4a	9% (1,7 Mb)	0% (1,7 Mb)	0% (1,7 Mb)	71% (1,7 Mb)

A. Mise en évidence d'haplotypes présentant des déficits en homozygotie

A partir des génotypages phasés, la recherche d'haplotypes en déficit d'homozygotie dans les populations ovines laitières Lacaune et Manech à Tête Rousse a permis de mettre en évidence 12 haplotypes en race Lacaune et 6 haplotypes en race Manech à Tête Rousse. Ces haplotypes sembleraient être associés à des mutations létales (du stade embryonnaire à la phase d'élevage du jeune animal) ou à l'origine de réformes précoces (animal viable, mais écarté de la sélection).

1. Identification de gènes et de polymorphismes candidats pour la létalité embryonnaire

Lors de cette étude, quatre haplotypes en race Lacaune (LL_OAR3_1a, LL_OAR3_4a, LLOAR4_1a et LL_OAR13_1a) et trois haplotypes en race Manech (MTR_OAR1_2a, MTR_OAR1_3a/MTR_OAR1_4a et MTR_OAR20_1a) semblent être associés à des mutations létales liées à un déficit total d'homozygotie dans la région génomique où elles se trouvent.

Le déficit de l'haplotype LL_OAR3_4a (fréquence 4,2%) est probablement associé à une mutation dans un gène *HOXC* affectant le développement embryonnaire. Le maintien de cette mutation s'explique par un avantage sélectif des hétérozygotes lié à une meilleure production laitière. L'étude des caractères enregistrés en sélection montre que l'haplotype LL_OAR3_4a présente les mêmes différences entre porteurs et non porteurs que l'haplotype LL_OAR3_2c associé à *SOCS2*. Les deux régions ne sont distantes que de 1,7 Mb et le calcul de déséquilibre montre une association importante ($r^2=71\%$) entre ces deux haplotypes (**Tableau 25**). On peut émettre l'hypothèse que l'haplotype LL_OAR3_4a est lié à *SOCS2* et que la sélection sur *SOCS2* a contribué par un effet auto-stop à maintenir une mutation récessive délétère à découvrir dans un des gènes *HOXC* dans la région de l'haplotype LL_OAR3_4a.

La recherche de variants dans la région LL_OAR3_1 montre que l'haplotype identifié est associé à une mutation faux-sens dans le gène *PREB* affectant un domaine protéine de type WD40. Ces domaines sont importants pour les interactions entre protéines et participent à la régulation de la transcription ([Marchler-Bauer et al., 2017](#)). En race Holstein, une mutation non-sens a été identifiée dans le gène *APAF1* ([Adams et al., 2016](#)) affectant également un domaine WD40 conduisant à une mortalité embryonnaire à l'état homozygote. [Taylor Clelland, Craciun, et al. \(2000\)](#) montrent que *PREB* est exprimé très tôt lors du développement embryonnaire au niveau du système nerveux périphérique, du foie et de l'ectoderme (différenciation en épiderme et neuroectoderme). Ce gène s'exprime également à un stade plus tardif afin de former l'hypophyse antérieure (lieu de production de la prolactine). En milieu de gestation, *PREB* est associé au développement du squelette fœtal. Sur la base de ces éléments, on peut faire l'hypothèse que la mutation identifiée est une mutation létale embryonnaire en

race ovine Lacaune. Le maintien de la mutation s'expliquerait par un avantage à l'hétérozygote, ces derniers étant moins sujets aux mammites. En effet, les hétérozygotes montrent un comptage des cellules somatiques du lait plus faible que les non porteurs ($-0,1\sigma_g$), ce qui équivaut à un an de progrès génétique (Astruc, communication personnelle). Seul cet haplotype LL_OAR_3_1a montre un lien entre l'évolution de la fréquence des individus hétérozygotes et l'évolution du progrès génétique des caractères (Annexes viii - ix). En effet, la fréquence des hétérozygotes de LL_OAR_3_1a chute de 8% entre les millésimes de 2000 à 2003 (dégradation des taux pour les hétérozygotes, composition de l'ISOL : 70% Lait + 30% Taux) puis est stable à partir de 2003 lorsque l'on a intégré la santé de la mamelle dans l'ISOL (pondération de ¼). Ceci pourrait expliquer le maintien dans la population de cet haplotype.

En race Manech, l'étude de la région MTR_OAR20_1 montrent la présence de deux gènes candidats *TFAP2D* et *TFAP2B* qui codent pour des facteurs de transcription. Ils jouent un rôle majeur au cours du développement embryonnaire (Zhao et al., 2003) intervenant dans le développement de l'ectoderme. Cependant, on retrouve également dans la région un cluster de gènes *CRISP*, importants pour assurer une bonne fertilité mâle. En effet, Lim et al. (2019) montrent le rôle de *CRISP2* dans l'étape de capacitation** du spermatozoïde. Il a également été montré que les individus n'exprimant pas *CRISP2* étaient infertiles. A ce stade, deux hypothèses peuvent expliquer ce déficit : *i*) une mutation dans un des gènes du cluster *CRISP* associée à une infertilité des homozygotes et *ii*) mutation embryonnaire létale dans *TFAP2D* ou *TFAP2B* impactant la mise en place de structures embryonnaires.

La recherche de variants candidats dans la région LL_OAR4_1 n'a pas mis en évidence de mutations affectant le codant. Cependant, parmi les variants, un SNP est identifié dans une séquence régulatrice (*downstream*) du gène *RELN* (*Reelin*) (Annexe vii). Il code pour une protéine de la matrice extracellulaire intervenant dans des interactions entre cellules à l'origine du développement du cerveau. Une délétion dans un exon du gène *RELN* à l'origine de la lissencephalie** a été mise en évidence en race ovine Churra (Suárez-Vega et al., 2013) mais n'a pas été retrouvée dans les variants issus des 68 animaux séquencés disponibles.

Confirmation de la létalité

La confirmation définitive de la létalité des haplotypes passe par l'étude de la fertilité des accouplements à risque (béliers porteurs x brebis porteuses) *versus* les accouplements non à risque (autres accouplements possibles). L'organisation de tels accouplements n'a pas été possible lors de ce stage, mais sont envisagés pour la suite du travail. Nous avons tenté d'approcher ce problème par l'analyse de la fertilité globale et à l'IA des populations Lacaune et Manech à Tête Rousse, mais sans montrer de différence significative entre les mâles porteurs vs non porteurs des haplotypes en déficit identifiés dans notre étude. Cependant, la fréquence

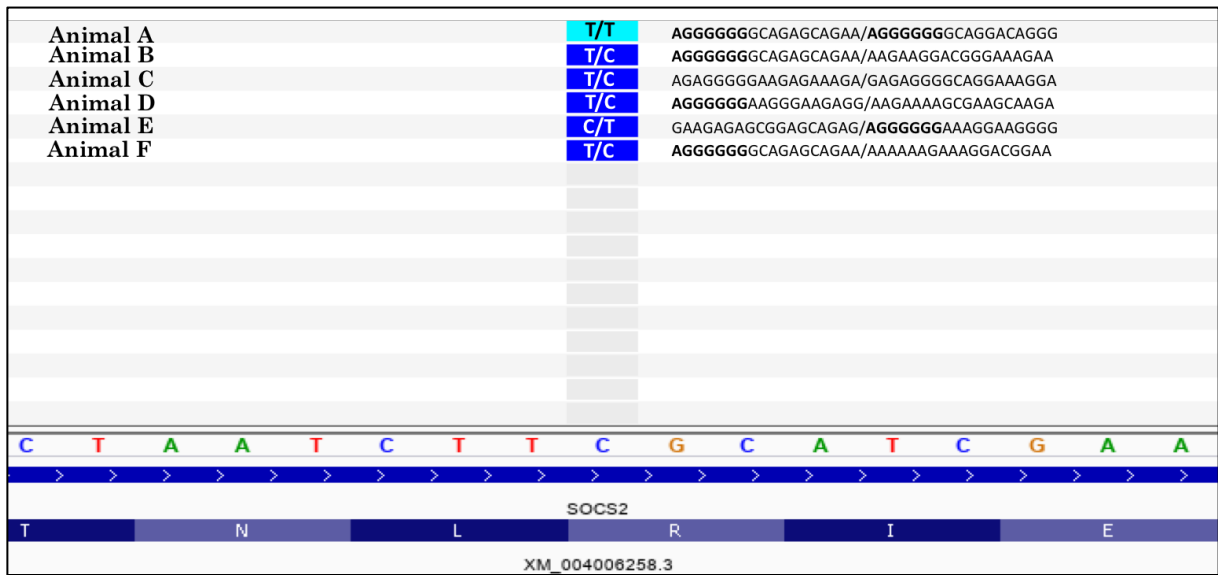


Figure 26. Génotypes au variant (C>T entraînant la substitution non conservative R96C) du gène *SOCS2* (OAR3) des 6 animaux séquencés tout génome, porteurs de l'allèle T (issus des séquences tout génome des 68 animaux disponibles au laboratoire). La visualisation des données de séquences est effectuée sous IGV (*Integrative Genome Viewer*, v.2.5.3, Thorvaldsdottir et al., 2013). Les haplotypes associés à ces animaux sont indiqués (présents dans la région LL_OAR3_2).

des animaux porteurs de ces haplotypes est faible et nous n'avions pas le génotypage des brebis pour se focaliser sur les accouplements à risque.

2. Identification de mutations affectant le jeune animal

La recherche de déficit en homozygotie a permis de mettre en évidence 8 haplotypes en race Lacaune et 3 haplotypes en race Manech présentant des déficits partiels d'animaux homozygotes (81 à 96%). La recherche de gènes/variants candidats montre la localisation dans ces régions de gènes assurant des fonctions liées à la structure de l'ADN ou intervenant dans la mise en place de tissus au cours du développement de l'animal. On peut penser que les jeunes béliers homozygotes présentent des défauts morphologiques expliquant leur réforme (pouvant également conduire à une mortalité tardive). Par exemple, l'haplotype le plus significatif LL_OAR3_3a ségrège à une forte fréquence dans la population (12,1%) et semble être associé à une mutation non-sens conduisant très certainement à une perte de fonction totale de la protéine du gène *CCDC65*. Horani et al. (2013) montrent que le gène *CCDC65* est associé à la dyskinésie ciliaire primitive. Il s'agit d'une maladie génétique pulmonaire affectant la régulation du complexe nexine-dynéine. Les patients homozygotes pour la mutation identifiée présentent des détresses respiratoires, une mauvaise implantation des organes et peuvent être infertiles (motilité réduite des spermatozoïdes). La fonction ciliaire du gène a été confirmée en utilisant un modèle animal (*zebrafish*) qui démontre le rôle du gène *CCDC65* dans la structure/assemblage du complexe nexine-dyéneine, associé à des battements de cils ralentis suite au *knockdown* du gène (Austin-Tse et al., 2013). La protéine *CCDC65* est également très fortement exprimée dans le testicule et plus particulièrement dans les queues des spermatozoïdes (Zheng et al., 2006) et intervient lors de la capacitation des spermatozoïdes. Ainsi, la mutation associée à l'haplotype LL_OAR3_3a pourrait être liée à une mutation affectant le jeune animal. En effet, on fait l'hypothèse que les agneaux homozygotes à la mutation sont viables mais présenteraient des structures ciliaires (poumons, voies nasales) malformées et ne seraient pas fertiles (motilité réduite des spermatozoïdes).

3. Identification d'une mutation en contre-sélection : cas du gène de résistance aux mammites *SOCS2*

La recherche de déficit en homozygotie a permis de mettre en évidence une mutation qui est contre-sélectionnée volontairement à l'état homozygote dans les schémas de sélection. A partir de 3 haplotypes présents dans la région LL_OAR3_2, on a pu mettre en évidence un haplotype plus court associé à la mutation R96C dans *SOCS2*. L'analyse des données de séquence montre que les animaux porteurs du variant *SOCS2* sont porteurs également de l'haplotype LL_OAR3_2_T (Figure 26). En effet, on retrouve les résultats obtenus par Rupp et al. (2015).

Le génotype à *SOCS2* prédit par l'haplotype permet d'expliquer 11% de la variance du caractère cellules somatiques (12% dans la publication). Les individus hétérozygotes ont un comptage cellulaire dans le lait important. Cependant, ils ont une meilleure production laitière expliquant le maintien de l'allèle mutée dans la population. Un comptage cellulaire élevé entraîne également une dégradation des trois postes de morphologie de la mamelle (Astruc, 2017).

B. Réflexion sur la méthode de recherche de déficit employée

1. Recherche de déficits en homozygotie

L'efficacité de la méthode de détection de segments en déficit en homozygotie repose sur *i*) un nombre important de génotypages, *ii*) une fréquence de l'haplotype relativement élevée dans la population pour pouvoir être détecté (>1%), *iii*) fort déséquilibre de liaison entre l'haplotype et la mutation délétère (Fritz et al., 2016).

Populations d'étude

Cette méthode initialement développée en bovins, est désormais possible dans d'autres races en sélection génomique générant des milliers de génotypages de type SNP. Les animaux de l'étude sont composés d'ovins laitiers génotypés sur les puces SNP 50K et/ou 15K. La qualité d'imputation des marqueurs de la puce basse densité 15K vers la 50K est bonne, permettant ainsi de rechercher des déficits en homozygotie dans l'ensemble de la population génotypée. Néanmoins, l'imputation n'est pas fiable à 100%. En effet, dans les cas des haplotypes LL_OAR3_1a et MTR_OAR1_2a on observe un individu homozygote typé sur la puce LD, on fait l'hypothèse qu'il résulte d'une erreur d'imputation des génotypes des marqueurs de la puce 50K à partir de la puce LD.

En petits ruminants, la sélection génomique est également pratiquée dans les populations laitières caprines (races Saanen et Alpine). Il est donc envisageable de rechercher des mutations létales dans ces populations en utilisant cette méthode.

Paramètres d'exécution : Taille de la fenêtre glissante et pourcentage de déficit

Le choix de la taille de la fenêtre dans la recherche est un élément crucial, il ne doit être *i*) ni trop court car l'haplotype peut exister à une forte fréquence dans la population sans être en déséquilibre de liaison avec la mutation, il devient alors peu informatif, *ii*) ni trop long car la mutation peut se retrouver par le biais de recombinaisons associée à plusieurs haplotypes (Fritz et al., 2016) rendant la détection moins efficace. Une taille de fenêtre comprise entre 10 et 20 marqueurs (fenêtre de 0,5 à 2Mb) est le plus pertinent pour mettre en œuvre cette recherche lorsque l'on prend en compte la densité de marqueurs le long du génome (un marqueur SNP informatif tout les 62kb sur la puce 50K) et l'information populationnelle. Seuls les haplotypes obtenus de 10 et 20 marqueurs présentant un déficit supérieur à 75% ont été gardés. En effet,

cela a permis de mettre en évidence des mutations létales (déficit total) mais aussi des anomalies compromettant la survie du jeune ou conduisant à des réformes précoces (déficit partiel).

Limites de la méthode employée

En bovins, les génotypages sont pratiqués sur les taureaux et les vaches. Cela permet de gagner en puissance dans l'analyse puisque l'on augmente le nombre d'animaux homozygotes potentiellement observables (trio : descendant, père et mère). Or, les individus génotypés pour la sélection génomique en ovins sont exclusivement des mâles. Ainsi, le calcul du nombre d'animaux homozygotes attendus repose principalement sur le trio descendant, père et grand-père maternel. Le statut de la mère est prédit par son père et la fréquence de l'haplotype dans la population totale. Le génotypage de femelles intervenant dans les accouplements raisonnés permettrait d'améliorer la puissance de l'analyse.

La méthode de recherche employée montre que l'on peut passer à côté de mutations associées à des haplotypes plus court. En effet, seul LL_OAR3_2a avait été mise en évidence, partiellement associé à la mutation *SOCS2* (pénétrance incomplète de l'haplotype). Cependant, l'étude approfondie de la région met en évidence une association parfaite entre l'allèle muté T du variant *SOCS2*, et un haplotype (LL_OAR3_T) d'uniquement 7 marqueurs.

De plus, par construction la méthode ne permet pas de mettre en évidence des haplotypes ayant une fréquence <1%. En effet, la sélection pratiquée sur le gène de résistance à la tremblante aurait dû mettre en évidence un déficit en homozygotie mais la fréquence de l'allèle associé à la sensibilité à la tremblante est trop faible pour pouvoir signer le déficit observé dans la région LL_OAR13_1.

2. Recherche de mutations causales

Animaux séquencés disponibles dans l'étude

La recherche de variants s'est faite à partir des 68 animaux séquencés disponibles au laboratoire. En Lacaune, 24 animaux sont séquencés et génotypés, ce qui a permis d'étudier 8 haplotypes. Malgré le faible nombre d'animaux porteurs séquencés, la disposition d'un nombre important d'autres ovins séquencés (non porteurs) a permis de restreindre le nombre de variants candidats respectant les différents filtres appliqués (<100 polymorphismes par région). Pour les autres haplotypes pour lesquels nous ne disposons pas d'individus porteurs séquencés, l'identification des mutations causales passera par le séquençage complet du génome d'animaux porteurs (1 ou 2 animaux). Cela permettrait de confirmer ou non, par exemple, la présence d'une mutation létale dans un gène *HOXC* (région LL_OAR3_4). En race Manech à Tête Rousse, aucune séquence étant disponible, il s'agira de la même façon de séquencer des animaux porteurs et non porteurs des haplotypes d'intérêt sur la base des gènes candidats présents dans les régions (notamment dans les régions MTR_OAR20_1 et MTR_OAR1_2).

Filtres appliqués pour la recherche de variants

Les filtres utilisés sont restrictifs expliquant que seules 3 régions sur 8 ont mis en évidence des variants candidats. En effet, comme aucun des haplotypes identifiés dans les deux races convergent vers les mêmes régions, on fait l'hypothèse que les mutations délétères sont propres à chacune des races. Néanmoins, une même mutation peut se retrouver dans des races différentes. C'est le cas de l'haplotype *PRNP* de la tremblante ovine, mais aussi des mutations de prolificité (Bodin et al., 2011). Par la suite, il sera intéressant de filtrer de nouveau les fichiers de variants qu'à partir des animaux Lacaune et de regarder si l'on trouve des variants candidats communs à d'autres races.

De plus, on fait l'hypothèse que l'haplotype identifié est en parfait déséquilibre de liaison avec la mutation causale. Or ceci n'est pas toujours vrai en cas de pénétrance incomplète. Par exemple Fritz et al. (2018), en recherchant des déficits en homozygotie, ont mis en évidence une mutation dans le gène *SDE2* (intervenant dans la maintenance des télomères). Parmi l'ensemble des variants présents dans l'intervalle de 1,1Mb en déficit d'homozygotie, seul un polymorphisme présente une corrélation supérieure à 0,5 entre le statut des animaux à l'haplotype et le génotype des variants.

De même, dans notre étude, les bornes de recherches des variants candidats sont celles de l'haplotype. Or dans certains cas, la mutation peut être en dehors de l'haplotype. Par exemple, l'haplotype associé à l'anomalie *CVM* est distant de 2 Mb de la mutation causale (Fritz et al., 2013). Par la suite, il conviendra donc d'élargir les zones de recherches.

C. Perspectives

Au cours du stage, des régions en déficit en homozygotie et des variants candidats (associés à ces haplotypes) ont pu être repérés. La poursuite de ce sujet en thèse, me permettra d'étudier plus précisément ces mutations d'un point de vue moléculaire et d'étudier les autres régions pour lesquelles on manque actuellement d'informations de séquences tout génome.

Génétique moléculaire

Une nouvelle recherche de variants va être réalisée, prenant en compte les différents paramètres cités ci-dessus, afin de rechercher des polymorphismes candidats non mis en évidence par les différents filtres appliqués au cours de mon stage. Dans un premier temps, nous nous sommes intéressés seulement aux mutations présentes dans les régions transcrites. Dans, un second temps, il s'agira d'étudier les mutations pouvant impacter les séquences promotrices ou régulatrices et donc l'expression de gènes candidats.

Par la suite, l'objectif est de tester fonctionnellement toutes les mutations d'intérêt que l'on a mises en évidence pour prouver leur causalité par des tests *ad hoc in vitro* (création de mini-

gène, mutagène dirigée, transfection, protéines recombinantes, etc...) et d'analyser l'expression spatio-temporelle, dans les embryons ou le jeune animal, des gènes impliqués et évaluer l'influence des mutations (RT-PCR en temps réel, immunohistochimie).

Zootecnie : élaboration d'accouplements à risque

Pour pouvoir étudier les effets de ces mutations, il est nécessaire de repérer des individus porteurs afin de générer les croisements adéquats en élevage ou en installation expérimentale et génotyper et phénotyper leurs produits. Pour les mutations déjà identifiées (dans *PREB* et *CCDC65*), nous pouvons développer dès à présent des tests de génotypage à façon (type *RFLP***) et sur la base des haplotypes connus, génotyper les béliers encore présents et leurs descendants pour trouver des mâles et des femelles porteurs. Pour les mutations non encore identifiées, cette stratégie pourra être menée sur la base des marqueurs constituant les haplotypes, mais elle sera plus lourde à mettre en place d'un point de vue génotypage.

Par la suite, l'objectif est également d'aider à établir et proposer des modes de gestion des mutations dans les populations de petits ruminants concernées (transfert vers la filière).

Gestion des mutations découvertes : intégration dans les schémas de sélection

L'impact économique dans la population d'une anomalie génétique dépend de l'âge d'apparition de l'anomalie et de la fréquence de cette anomalie. Le coût économique engendré par l'anomalie est d'autant plus important lorsqu'elle ségrége dans la population à une fréquence élevée (Hoze, Fritz, 2018). Ainsi, trois modes de gestions principales peuvent être mises en place : *i*) Contre-sélection de l'anomalie : on ne rentre plus de béliers porteurs de l'anomalie en CIA, *ii*) on évite les accouplements à risque en se basant sur les génotypes des mâles (seuls les béliers collectifs sont génotypés) en interdisant les accouplements entre mâles porteurs et filles de mâles porteurs, *iii*) on génotype les femelles (non réalisé en ovins pour le moment) pour permettre d'éviter les accouplements à risque. De tels outils existent en bovins. L'outil pANO, qui a été développé en race Montbéliarde (Brochard et al., 2018), a permis de diminuer le nombre d'accouplements à risque programmés de 25% à 1,04% (prenant en compte 4 anomalies).

Conclusion

Le développement de la sélection génomique en race Lacaune et à Manech donne accès à de nombreux génotypages utilisables pour identifier des mutations létales par un raisonnement de génétique inverse.

La recherche de déficit en homozygotie a permis de mettre en évidence 18 régions du génome présentant des déficits en animaux homozygotes. Ces régions varient de 0,5 à 3 Mb et les haplotypes associés ségrégent à des fréquences de 3,9 à 17,4%.

Parmi ces régions, 4 haplotypes en race Lacaune et 3 haplotypes en race Manech présentent un déficit total d'animaux homozygotes, potentiellement associés à des mutations létales affectant l'embryon (stade embryonnaire) ou le jeune (période néonatale). Les régions pour lesquelles on dispose d'individus porteurs ont été analysées pour recherche des variants candidats. En race Lacaune, une mutation faux-sens dans le gène *PREB* (intervenant lors du développement embryonnaire) laisse supposer que l'on a identifié un très bon polymorphisme candidat. De plus, la recherche de gènes candidats dans les autres régions en déficit total montrent la présence de gènes d'intérêts (gènes homéotiques, gènes intervenant lors du développement embryonnaire, gènes codant pour des histones ...).

Cependant, cette méthode n'est pas spécifique des mutations létales. En effet, 8 régions en race Lacaune et 3 régions en race Manech sont associées à des déficits partiels d'animaux homozygotes. Ces déficits peuvent être en déséquilibre de liaison partiel avec une mutation létale ou être associés à des mutations post-néonatales compromettant la survie du jeune ou présentant des défauts congénitaux à l'origine de réformes précoces. En effet, la fouille de données de séquence tout génome en race Lacaune a permis de mettre en évidence une mutation candidate non-sens dans le gène *CCDC65*, intervenant dans la mise en place de structures ciliaires (poumons, spermatozoïdes ...). Parmi ces régions, un haplotype associé à la mutation *SOCS2* a été découvert (déficit partiel lié à la contre-sélection d'un allèle en sélection).

Certaines de ces mutations sont conservées dans la population car elles procurent un avantage sélectif des hétérozygotes (confirmation d'un avantage pour six haplotypes).

Ces résultats sont très encourageants pour identifier et caractériser de nouvelles mutations délétères en ovins. On peut voir cependant que cette approche n'est pas spécifique des mutations délétères et que l'identification de la mutation et de sa fonction/ses fonctions est primordiale avant l'utilisation de ces nouvelles informations en sélection.

L'identification de mutations létales par un raisonnement de génétique inverse s'inscrit dans un projet de plus grande envergure sur les anomalies génétiques en petits ruminants. En effet, un projet de création d'un observatoire des anomalies génétiques des petits ruminants en France est en réflexion au sein de l'UMT GGPR (Gestion Génétique et génomique des Petits Ruminants), similaire à l'observatoire existant en bovins « l'ONAB ». De plus, le développement des projets de séquençage complet du génome en petits ruminants (Projet VarGoats : www.goatgenome.org, SEQOccin en ovins et caprins : <https://get.genotoul.fr/seqoccin/>) permettra d'avoir accès à plusieurs centaines de séquences tout génome. L'analyse bio-informatique de ces données pourrait permettre d'identifier des variants fortement délétères par une autre méthode de génétique inverse, basée sur l'étude des données de séquence tout génome.

Références bibliographiques

1. ADAMS, H.A., SONSTEGARD, T.S, VANRADEN, P.M., NULL, D.J., VAN TASSELL, C.P., LARKIN, D.M. et LEWIN, H.A., 2016. Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in reproductive efficiency in Holstein dairy cattle. In : *Journal of Dairy Science*. 2016. Vol. 99, n° 8, p. 6693-6701.
2. ADZHUBEI, I., JORDAN, D.M. et SUNYAEV, S.R., 2013. Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. In : *Current protocols in human genetics*. janvier 2013. Vol. 07, p. Unit7.20. DOI 10.1002/0471142905.hg0720s76.
3. AGRESTE, 2019. 2018-035 : *Infos Rapides Ovins : Entre 2011 et 2018, repli des effectifs ovins en France* [en ligne]. S.I. Ovins – Cheptel novembre 2018 – Résultats français et européens. Disponible à l'adresse : http://www.agreste.agriculture.gouv.fr/IMG/pdf/2019_035_InforapOvinsv2.pdf.
4. ASTRUC, J-M., 2017. *Bilan de 10 ans de sélection sur la résistance aux mammites en ovins laitiers* [en ligne]. Journée de restitution du projet MAMOVICAP. ENV Toulouse. [Consulté le 15 juillet 2019]. Disponible à l'adresse : http://idele.fr/fileadmin/medias/Documents/Docs_UMT_SPR/MAMOVICAP_fevrier2017_Bilan_de_10ans_de_selection_sur_la_resistance_aux_mammites_en_ovins_laitiers_JM_Astruc.pdf.
5. ASTRUC, J-M., BALOCHE, G., BUISSON, D., LABATUT, J., LAGRIFFOUL, G., LARROQUE, H., ROBERT-GRANIÉ, C., LEGARRA, A. et BARILLET, F., 2016. La sélection génomique des ovins laitiers en France. In : *INRA Productions Animales*. 2016. Vol. 29, n° 1, p. 41-56.
6. ASTRUC, J-M. et BUISSON, D., 2018a. La stratégie de génotypage en ovins laitiers. In : *Dossier : Les ovins laitiers résolument tournés vers la sélection génomique*. décembre 2018. p. 5. Disponible à l'adresse : idele.fr/no_cache/recherche/publication/idelestor/recommends/evaluation-genomique.html
7. ASTRUC, J-M. et BUISSON, D., 2018b. Les éléments clés des schémas de sélection ovins laitiers en France. In : *Dossier : Les ovins laitiers résolument tournés vers la sélection génomique*. décembre 2018. p. 3. Disponible à l'adresse : idele.fr/no_cache/recherche/publication/idelestor/recommends/evaluation-genomique.html
8. ASTRUC, J-M. et BUISSON, D., 2018c. Les nouveaux caractères. In : *Dossier : Les ovins laitiers résolument tournés vers la sélection génomique*. décembre 2018. p. 5. Disponible à l'adresse : idele.fr/no_cache/recherche/publication/idelestor/recommends/evaluation-genomique.html
9. ASTRUC, J-M. et BUISSON, D., 2018d. Les populations de référence en ovins laitiers. In : *Dossier : Les ovins laitiers résolument tournés vers la sélection génomique*. décembre 2018. p. 3. Disponible à l'adresse : idele.fr/no_cache/recherche/publication/idelestor/recommends/evaluation-genomique.html
10. AUSTIN-TSE, C., HALBRITTER, J., ZARIWALA, M.A., GILBERTI, R.M., GEE, H.Y., HELLMAN, N., PATHAK, N., LIU, Y., PANIZZI, J.R., PATEL-KING, R.S., TRITSCHLER, D., BOWER, R., O'TOOLE, E., PORATH, J.D., HURD, T.W, CHAKI, M., DIAZ, K.A., KOHL, S., LOVRIC, S., HWANG, D., BRAUN, D.A., SCHUELER, M., AIRIK, R., OTTO, E.A., LEIGH, M.W., NOONE, P.G., CARSON, J.L., DAVIS, S.D., PITTMAN, J.E., FERKOL, T.W., ATKINSON, J.J., OLIVIER, K.N., SAGEL, S.D., DELL, S.D., ROSENFELD, M., MILLA, C.E., LOGES, N., OMRAN, H., PORTER, M.E., KING, S.M., KNOWLES, M.R., DRUMMOND, F et HILDEBRANDT, F., 2013. Zebrafish Ciliopathy Screen Plus Human Mutational Analysis Identifies C21orf59 and CCDC65 Defects as Causing Primary Ciliary Dyskinesia. In : *The American Journal of Human Genetics*. octobre 2013. Vol. 93, n° 4, p. 672-686. DOI 10.1016/j.ajhg.2013.08.015.
11. BARBEZANT, Maurice, BOICHARD, Didier et GASTINEL, Pierre-Louis, 2010. *La révolution génomique animale*. S.I. : France agricole.

12. BARILLET, F., LAGRIFFOUL, G., RUPP, R., PORTES, D. et BOCQUIER, F., 2016. Objectifs de sélection et stratégie raisonnée de mise en œuvre à l'échelle des populations de brebis laitières françaises. In : *INRA Productions Animales*. 2016. Vol. 29, n° 1, p. 19-40.
13. BODIN, L., DEMARS, J., DROUILHET, L., MULSANT, P., SARRY, J., TABEL, C., TOSSER-KLOPP, G., FABRE, S., BOSCHER, M.Y., TIPHINE, L., BERTRAND, C., BOUQUET, P.M., MATON, C., TEYSSIER, J., JOUANNAUX, C., HALLAUER, J., CATHALAN, D., GUEUX, J. et POCACHARD, M., 2011. Etat des lieux et gestion pratique des gènes d'ovulation détectés dans les races ovines françaises. In : *journées 3R* [en ligne]. 2011. [Consulté le 13 août 2019]. Disponible à l'adresse : http://www.journees3r.fr/IMG/pdf/Texte1_genetique_Bodin.pdf.
14. BOICHARD, D., CROISEAU, P., FRITZ, S. et DUCROCQ, V., 2015. Quel futur pour l'amélioration génétique chez les espèces animales domestiques ? In : *Académie d'agriculture de France*. 16 avril 2015. p. 15.
15. BOICHARD, D., GROHS, C., DANCHIN-BURGE, C. et CAPITAN, A., 2016. Les anomalies génétiques : définition, origine, transmission et évolution, mode d'action. In : *INRA Productions Animales*. 2016. Vol. 29, n° 5, p. 297-306.
16. BROCHARD, M., BOICHARD, D., CAPITAN, A., FAYOLLE, G., FRITZ, S., NICOD, L. et VACELET, J-M., 2018. pANO, le risque d'anomalie létale pour les produits d'accouplements : principe et utilisation en race Montbéliarde sur la zone GEN'IAtest. In : *24èmes Renc. Rech. Ruminants - Génétique* [en ligne]. 2018. [Consulté le 19 mars 2019]. Disponible à l'adresse : http://journees3r.fr/IMG/pdf/texte_7_genetique_m-brochard.pdf.
17. BUISSON, D., ASTRUC, J-M. et BARILLET, F., 2018. Bilan et perspectives de la gestion de la variabilité génétique des ovins laitiers en France. In : *INRA Productions Animales*. 11 juin 2018. Vol. 31, n° 1, p. 1-12. DOI 10.20870/productions-animales.2018.31.1.2202.
18. CHARLESWORTH, D. et WILLIS, J.H., 2009. The genetics of inbreeding depression. In : *Nature Reviews. Genetics*. novembre 2009. Vol. 10, n° 11, p. 783-796. DOI 10.1038/nrg2664.
19. CHARLIER, C., AGERHOLM, J.S., COPPIETERS, W., KARLSKOV-MORTENSEN, P., LI, W., DE JONG, G., FASQUELLE, C., KARIM, L., CIRERA, S., CAMBISANO, N., AHARIZ, N., MULLAART, E., GEORGES, M. et FREDHOLM, M., 2012. A Deletion in the Bovine FANCI Gene Compromises Fertility by Causing Fetal Death and Brachyspina. In : VEITIA, R.A. (éd.), *PLoS ONE*. 29 août 2012. Vol. 7, n° 8, p. e43085. DOI 10.1371/journal.pone.0043085.
20. CHARLIER, C., COPPIETERS, W., ROLLIN, F., DESMECHT, D., AGERHOLM, J.S., CAMBISANO, N., CARTA, E., DARDANO, S., DIVE, M., FASQUELLE, C., FRENNET, J-C., HANSET, R., HUBIN, X., JORGENSEN, C., KARIM, L., KENT, M., HARVEY, K., PEARCE, B.R., SIMON, P., TAMA, N., NIE, H., VANDEPUTTE, S., LIEN, S., LONGERI, M., FREDHOLM, M., HARVEY, R.J. et GEORGES, M., 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. In : *Nature Genetics*. avril 2008. Vol. 40, n° 4, p. 449-454. DOI 10.1038/ng.96.
21. CHARLIER, C., LI, W., HARLAND, C., LITTLEJOHN, M., COPPIETERS, W., CREAGH, F., DAVIS, S., DRUET, T., FAUX, P., GUILLAUME, F., KARIM, L., KEEHAN, M., KADRI, N., TAMMA, N., SPELMAN, R. et GEORGES, M., 2016. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. In : *Genome Research*. octobre 2016. Vol. 26, n° 10, p. 1333-1341. DOI 10.1101/gr.207076.116.
22. CIIRPO, 2009. Insémination animale : du bélier à la paillette. In : www.inn-ovin.fr [en ligne]. 29 octobre 2009. [Consulté le 3 juillet 2019]. Disponible à l'adresse : http://www.inn-ovin.fr/wp-content/uploads/2017/03/paillettes_4.pdf.
23. CINGOLANI, P., PLATTS, A., WANG, L.L., COON, M., NGUYEN, T., WANG, L., LAND, S.J., LU, X. et RUDEN, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. In : *Fly*. avril 2012. Vol. 6, n° 2, p. 80-92. DOI 10.4161/fly.19695.

24. CORAM, 2019. Les Races de Massif - CORAM - Collectif des Races locales des Massifs (ovins, bovins). In : [en ligne]. 2019. [Consulté le 1 avril 2019]. Disponible à l'adresse : <https://www.races-montagnes.com/index.php>.
25. DAETWYLER, H.D., CAPITAN, A., PAUSCH, H., STOTHARD, P., VAN BINSBERGEN, R., BRØNDUM, R.F., LIAO, X., DJARI, A., RODRIGUEZ, S.C., GROHS, C., ESQUERRÉ, D., BOUCHEZ, O., ROSSIGNOL, M-N., KLOPP, C., ROCHA, D., FRITZ, S., EGGEN, A., BOWMAN, P.J., COOTE, D., CHAMBERLAIN, A.J., ANDERSON, C., VANTASSELL, C.P., HULSEGGE, I., GODDARD, M.E., GULDBRANDTSEN, B., LUND, M.S., VEERKAMP, R.F., BOICHARD, D., FRIES, R. et HAYES, B.J., 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. In : *Nature Genetics*. août 2014. Vol. 46, n° 8, p. 858-865. DOI 10.1038/ng.3034.
26. DERKS, M.F.L., GJUUSLAND, A.B., BOSSE, M., LOPES, M.S., VAN SON, M., HARLIZIUS, B., TAN, B.F., HAMLAND, H., GRINDFLEK, E., GROENEN, M.A.M. et MEGENS, H-J., 2019. Loss of function mutations in essential genes cause embryonic lethality in pigs. In : LEEB, T. (éd.), *PLOS Genetics*. 15 mars 2019. Vol. 15, n° 3, p. e1008055. DOI 10.1371/journal.pgen.1008055.
27. DERKS, M.F.L., LOPES, M.S., BOSSE, M., MADSEN, O., DIBBITS, B., HARLIZIUS, B., GROENEN, M.A.M. et MEGENS, H-J., 2018. Balancing selection on a recessive lethal deletion with pleiotropic effects on two neighboring genes in the porcine genome. In : LEEB, T. (éd.), *PLOS Genetics*. 19 septembre 2018. Vol. 14, n° 9, p. e1007661. DOI 10.1371/journal.pgen.1007661.
28. DERKS, M.F.L., MEGENS, H-J., BOSSE, M., VISSCHER, J., PEETERS, K., BINK, M.C.A.M., VEREIJKEN, A., GROSS, C., DE RIDDER, D., REINDERS, M.J.T. et GROENEN, M.A.M., 2018. A survey of functional genomic variation in domesticated chickens. In : *Genetics Selection Evolution*. décembre 2018. Vol. 50, n° 1, p. 17. DOI 10.1186/s12711-018-0390-1.
29. DUCHESNE, A., GROHS, C., MICHOT, P., BERTAUD, M., BOICHARD, D., FLORIOT, S. et CAPITAN, A., 2016. Du phénotype à la mutation causale : le cas des anomalies récessives bovines. In : *INRA Productions Animales*. 2016. Vol. 29, n° 5, p. 319-328.
30. DUCOS, A., EGGEN, A., DARRE, R. et BOICHARD, D., 2002. Les anomalies génétiques dans l'espèce bovine. In : *9èmes Rencontres Recherches Ruminants, Paris*. 4 décembre 2002. p. 85-91.
31. DUPONT, J-Y., 2001. Les mutations. In : *ENS Lyon* [en ligne]. 14 août 2001. [Consulté le 4 juin 2019]. Disponible à l'adresse : <http://acces.ens-lyon.fr/biotic/genetic/mutation/html/mutation.htm>.
32. ELSEN, J.M., VU TIEN KHANG, J., SCHELCHER, F., AMIGUES, Y., LAPLANCHE, J.L., POIVEY, J.P. et EYCHENNE, F., 1997. Génétique de la sensibilité à la tremblante ovine : recherches en cours et perspective. In : *INRA Productions Animales*. 1997. Vol. 10, n° 2, p. 133-140.
33. FGE, 2019. France génétique Elevage - Races ovines. In : [en ligne]. 2019. [Consulté le 1 avril 2019]. Disponible à l'adresse : <http://fr.france-genetique-elevage.org/-Races-ovines-.html>.
34. FLISS, M. S., HINKLE, P. M. et BANCROFT, C., 1999. Expression cloning and characterization of PREB (prolactin regulatory element binding), a novel WD motif DNA-binding protein with a capacity to regulate prolactin promoter activity. In : *Molecular Endocrinology (Baltimore, Md.)*. avril 1999. Vol. 13, n° 4, p. 644-657. DOI 10.1210/mend.13.4.0260.
35. FRITZ, S., CAPITAN, A., DJARI, A., RODRIGUEZ, S., BARBAT, A., BAUR, A., GROHS, C., WEISS, B., BOUSSAHA, M., ESQUERRÉ, D., KLOPP, C., ROCHA, D. et BOICHARD, D., 2013. Detection of Haplotypes Associated with Prenatal Death in Dairy Cattle and Identification of Deleterious Mutations in GART, SHBG and SLC37A2. In : *PLoS ONE*. 7 juin 2013. Vol. 8, n° 6, p. e65550. DOI 10.1371/journal.pone.0065550.
36. FRITZ, S., HOZE, C., REBOURS, E., BARBAT, A., BIZARD, M., CHAMBERLAIN, A., ESCOUFLAIRE, C., VANDER JAGT, C., BOUSSAHA, M., GROHS, C., ALLAIS-BONNET, A., PHILIPPE, M., VALLÉE, A., AMIGUES, Y., HAYES, B., BOICHARD, D. et CAPITAN, A.,

2018. An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. In : *Journal of Dairy Science*. juillet 2018. Vol. 101, n° 7, p. 6220-6231. DOI 10.3168/jds.2017-14119.
37. FRITZ, S., MICHOT, P., HOZE, C., GROHS, C., BARBAT-LETERRIER, A., BOUSSAHA, M., BOICHARD, D. et CAPITAN, A., 2016. Anticiper l'émergence d'anomalies génétiques grâce aux données génomiques. In : *INRA Productions Animales*. 2016. Vol. 29, n° 5, p. 339-350.
 38. GEORGES, M., CHARLIER, C. et HAYES, B., 2019. Harnessing genomic information for livestock improvement. In : *Nature Reviews Genetics*. mars 2019. Vol. 20, n° 3, p. 135-156. DOI 10.1038/s41576-018-0082-2.
 39. GROHS, C., DUCHESNE, A., FLORIOT, S., DUCOS, A. et DANCHIN-BURGE, C., 2016. L'Observatoire National des Anomalies Bovines, son action et ses résultats pour une aide efficace à la gestion des anomalies génétiques. In : *INRA Productions Animales*. 2016. Vol. 29, n° 5, p. 307-318.
 40. HILL, W. G., 1981. Estimation of effective population size from data on linkage disequilibrium. In : *Genetical Research*. décembre 1981. Vol. 38, n° 3, p. 209-216. DOI 10.1017/S0016672300020553.
 41. HILL, W. G. et ROBERTSON, A., 1968. Linkage disequilibrium in finite populations. In : *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*. juin 1968. Vol. 38, n° 6, p. 226-231. DOI 10.1007/BF01245622.
 42. HORANI, A., BRODY, S.L., FERKOL, T.W., SHOSEYOV, D., WASSERMAN, M.G., TA-SHMA, A., WILSON, K.S., BAYLY, P.V., AMIRAV, I., COHEN-CYMBERKNOH, M., DUTCHER, S.K., ELPELEG, O. et KEREM, E., 2013. CCDC65 Mutation Causes Primary Ciliary Dyskinesia with Normal Ultrastructure and Hyperkinetic Cilia. In : *PLoS ONE*. 26 août 2013. Vol. 8, n° 8, p. e72299. DOI 10.1371/journal.pone.0072299.
 43. HOZE, C. et FRITZ, S., 2018. Prise en compte des gènes d'intérêt dans les objectifs de sélection en bovins laitiers. In : *journées 3R* [en ligne]. 2018. [Consulté le 21 mai 2019]. Disponible à l'adresse : http://journées3r.fr/IMG/pdf/texte_6_genetique_s-fritz.pdf.
 44. IDELE, 2015. Guide de compréhension des indicateurs de variabilité génétique. In : *idele.fr*. mars 2015.
 45. IDELE, 2017. *Conformation de la mamelle et concentrations cellulaires chez la brebis laitière : Typologie des morphologies des mamelles* [en ligne]. MAMOVICAP : Outils d'intervention et d'aide à la décision pour la maîtrise des mammites en élevages de petits ruminants laitiers. S.I. [Consulté le 20 juin 2019]. Disponible à l'adresse : http://idele.fr/fileadmin/medias/Documents/Docs_UMT_SPR/Typologie_des_morphologies_des_mamelles.pdf.
 46. IDELE, 2018. Indicateurs de variabilité génétique - races bovines - Edition 2018. In : *idele.fr* [en ligne]. 2018. [Consulté le 15 août 2019]. Disponible à l'adresse : http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/indicateurs-de-variabilite-genetique-races-bovines-edition-2018-1.html.
 47. IDELE et CNBL, 2018. *Résultats contrôle laitier ovin - France 2018* [en ligne]. S.I. Résultats. Disponible à l'adresse : <http://idele.fr/domaines-techniques/ameliorer-le-troupeau/performances-et-phenotypes/publication/idelesolr/recommends/resultats-de-controle-laitier-france-2018.html>.
 48. IDELE, CNBL et INRA, 2016. 0015202046 : *Évaluation génétique des ovins laitiers en France - Caractères laitiers - Cellules somatiques - Morphologie de la mamelle* [en ligne]. S.I. [Consulté le 29 août 2019]. Disponible à l'adresse : http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/comment-evaluer-genetiquement-les-ovins-laitiers-en-france.html.
 49. ILLUMINA, 2010. Illumina Sequencing Technology. In : *770-2007-002*. 2010. p. 5.

50. ILLUMINA, 2015. *OvineSNP50 Genotyping BeadChip* [en ligne]. S.l. [Consulté le 20 mars 2019]. Disponible à l'adresse : https://www.illumina.com/documents/products/datasheets/datasheet_ovinesnp50.pdf.
51. KADRI, N., SAHANA, G., CHARLIER, C., ISO-TOURU, T., GULDBRANDTSEN, B., KARIM, L., NIELSEN, U., PANITZ, F., AAMAND, G., SCHULMAN, N., GEORGES, M., VILKKI, J., LUND, M. et DRUET, T., 2014. A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the Common Occurrence of Balancing Selection in Livestock. In : LEEB, T. (éd.), *PLoS Genetics*. 2 janvier 2014. Vol. 10, n° 1, p. e1004049. DOI 10.1371/journal.pgen.1004049.
52. LIM, S., KIERZEK, M., O'CONNOR, A.E., BRENKER, C., MERRINER, D.J., OKUDA, H., VOLPERT, M., GAIKWAD, A., BIANCO, D., POTTER, D., PRABHAKAR, R., STRÜNKER, T. et O'BRYAN, M.K., 2019. CRISP2 Is a Regulator of Multiple Aspects of Sperm Function and Male Fertility. In : *Endocrinology*. 1 avril 2019. Vol. 160, n° 4, p. 915-924. DOI 10.1210/en.2018-01076.
53. MARCHINI, Jonathan et HOWIE, Bryan, 2010. Genotype imputation for genome-wide association studies. In : *Nature Reviews. Genetics*. juillet 2010. Vol. 11, n° 7, p. 499-511. DOI 10.1038/nrg2796.
54. MARCHLER-BAUER, A., BO, Y., HAN, L., HE, J., LANCZYCKI, C.J., LU, S., CHITSAZ, F., DERBYSHIRE, M.K., GEER, R.C., GONZALES, N.R., GWADZ, M., HURWITZ, D.I., LU, F., MARCHLER, G.H., SONG, J.S., THANKI, N., WANG, Z., YAMASHITA, R.A., ZHANG, D., ZHENG, C., GEER, L.Y. et BRYANT, S.H., 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. In : *Nucleic Acids Research*. 4 janvier 2017. Vol. 45, n° D1, p. D200-D203. DOI 10.1093/nar/gkw1129.
55. MCCLURE, M., KIM, E., BICKHART, D., NULL, D., COOPER, T., COLE, J., WIGGANS, G., AJMONE-MARSAN, P., COLLI, L., SANTUS, E., LIU, G., SCHROEDER, S., MATUKUMALLI, L., VAN TASSELL, C. et SONSTEGARD, T., 2013. Fine Mapping for Weaver Syndrome in Brown Swiss Cattle and the Identification of 41 Concordant Mutations across NRCAM, PNPLA8 and CTTNBP2. In : MOORE, S. (éd.), *PLoS ONE*. 20 mars 2013. Vol. 8, n° 3, p. e59251. DOI 10.1371/journal.pone.0059251.
56. MCCLURE, M.C., BICKHART, D., NULL, D., VANRADEN, P., XU, L., WIGGANS, G., LIU, G., SCHROEDER, S., GLASSCOCK, J., ARMSTRONG, J., COLE, J.B., VAN TASSELL, C.P. et SONSTEGARD, T.S., 2014. Bovine Exome Sequence Analysis and Targeted SNP Genotyping of Recessive Fertility Defects BH1, HH2, and HH3 Reveal a Putative Causative Mutation in SMC2 for HH3. In : *PLoS ONE*. 25 mars 2014. Vol. 9, n° 3, p. e92769. DOI 10.1371/journal.pone.0092769.
57. MESBAH-UDDIN, M., HOZE, C., MICHOT, P., BARBAT, A., LEFEBVRE, R., BOUSSAHA, M., SAHANA, G., FRITZ, S., BOICHARD, D. et CAPITAN, A., 2019. A missense mutation (p.Tyr452Cys) in the CAD gene compromises reproductive success in French Normande cattle. In : *Journal of Dairy Science*. 2 mai 2019. DOI 10.3168/jds.2018-16100.
58. MICHOT, P., 2017. *Utilisation des séquences de génome complet pour l'identification de mutations délétères responsables d'anomalies génétiques récessives chez le bovin*. Thèse. S.l. : Université Paris-Saclay.
59. MICHOT, P., FANTINI, O., BRAQUE, R., ALLAIS-BONNET, A., SAINTILAN, R., GROHS, C., BARBIERI, J., GENESTOUT, L., DANCHIN-BURGE, C., GOURREAU, J-M., BOICHARD, D., PIN, D. et CAPITAN, A., 2015. Whole-genome sequencing identifies a homozygous deletion encompassing exons 17 to 23 of the integrin beta 4 gene in a Charolais calf with junctional epidermolysis bullosa. In : *Genetics Selection Evolution* [en ligne]. décembre 2015. Vol. 47, n° 1. [Consulté le 19 mars 2019]. DOI 10.1186/s12711-015-0110-z. Disponible à l'adresse : <http://www.gsejournal.org/content/47/1/37>.
60. MICHOT, P., FRITZ, S., BARBAT, A., BOUSSAHA, M., DELOCHE, M-C., GROHS, C., HOZE, C., LE BERRE, L., LE BOURHIS, D., DESNOES, O., SALVETTI, P., SCHIBLER, L.,

- BOICHARD, D. et CAPITAN, A., 2017. A missense mutation in PFAS (phosphoribosylformylglycinamide synthase) is likely causal for embryonic lethality associated with the MH1 haplotype in Montbéliarde dairy cattle. In : *Journal of Dairy Science*. octobre 2017. Vol. 100, n° 10, p. 8176-8187. DOI 10.3168/jds.2017-12579.
61. MINVIELLE, F., 2010. *Principes d'amélioration génétique des animaux domestiques*. S.l. : Quae. Mieux Comprendre.
 62. PAUSCH, H., SCHWARZENBACHER, H., BURGSTALLER, J., FLISIKOWSKI, K., WURMSER, C., JANSEN, S., JUNG, S., SCHNIEKE, A., WITTEK, T. et FRIES, R., 2015. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. In : *BMC Genomics*. décembre 2015. Vol. 16, n° 1. DOI 10.1186/s12864-015-1483-7.
 63. PETIT, M., 2014. *Détection de mutations létales chez les ovins. Mise en évidence de segments chromosomiques homozygotes à partir d'information moléculaire haute densité (54 000 SNPs)*. INRA Toulouse : Mémoire, Ecole Nationale Supérieure des Sciences Agronomiques de Bordeaux-Aquitaine.
 64. PRACHE, S., CAILLAT, H. et LAGRIFFOUL, G., 2018. Diversité dans la filière petits ruminants : une source de résilience ? In : *Innovations Agronomiques*. 2018. n° 68, p. 171-191.
 65. RACES DE FRANCE, 2019. Moutons - Races de France. In : [en ligne]. 2019. [Consulté le 12 avril 2019]. Disponible à l'adresse : <https://www.racesdefrance.fr/moutons>.
 66. RUPP, R., MUCHA, S., LARROQUE, H., MCEWAN, J. et CONINGTON, J., 2016. Genomic application in sheep and goat breeding. In : *Animal Frontiers*. 1 janvier 2016. Vol. 6, n° 1, p. 39-44. DOI 10.2527/af.2016-0006.
 67. RUPP, R., SENIN, P., SARRY, J., ALLAIN, C., TASCA, C., LIGAT, L., PORTES, D., WOLOSZYN, F., BOUCHEZ, O., TABOURET, G., LEBASTARD, M., CAUBET, C., FOUCRAS, G. et TOSSER-KLOPP, G., 2015. A Point Mutation in Suppressor of Cytokine Signaling 2 (Socs2) Increases the Susceptibility to Inflammation of the Mammary Gland while Associated with Higher Body Weight and Size and Higher Milk Production in a Sheep Model. In : KIJAS, James (éd.), *PLOS Genetics*. 11 décembre 2015. Vol. 11, n° 12, p. e1005629. DOI 10.1371/journal.pgen.1005629.
 68. SARGOLZAEI, M., CHESNAIS, J.P. et SCHENKEL, F.S., 2014. A new approach for efficient genotype imputation using information from relatives. In : *BMC Genomics*. 2014. Vol. 15, n° 1, p. 478. DOI 10.1186/1471-2164-15-478.
 69. SCHÜTZ, E., WEHRHAHN, C., WANJEK, M., BORTFELD, R., WEMHEUER, W.E., BECK, J. et BREINIG, B., 2016. The Holstein Friesian Lethal Haplotype 5 (HH5) Results from a Complete Deletion of TFB1M and Cholesterol Deficiency (CDH) from an ERV-(LTR) Insertion into the Coding Region of APOB. In : BARENDSE, W. (éd.), *PLOS ONE*. 29 avril 2016. Vol. 11, n° 4, p. e0154602. DOI 10.1371/journal.pone.0154602.
 70. SCHWARZENBACHER, H., BURGSTALLER, J., SEEFRIED, F.R., WURMSER, C., HILBE, M., JUNG, S., FUERST, C., DINHOPL, N., WEISSENBOECK, H., FUERST-WALTL, B., DOLEZAL, M., WINKLER, R., GRUETER, O., BLEUL, U., WITTEK, T., FRIES, R. et PAUSCH, H., 2016. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. In : *BMC Genomics*. décembre 2016. Vol. 17, n° 1, p. 400. DOI 10.1186/s12864-016-2742-y.
 71. SERVIN, B., ROGEL, C., MILAN, D., BIDANEL, P. et RIQUET, J., 2014. *Le séquençage du génome du porc : apport des nouvelles technologies de séquençage à la génomique et à la génétique porcines*. Paris. Journées Recherche Porcine en France.
 72. SHUSTER, D. E., KEHRLI, M. E., ACKERMANN, M. R. et GILBERT, R. O., 1992. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. In : *Proceedings of the National Academy of Sciences*. 1 octobre 1992. Vol. 89, n° 19, p. 9225-9229. DOI 10.1073/pnas.89.19.9225.

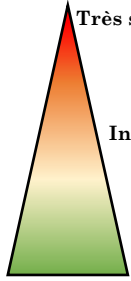
73. SONSTEGARD, T., COLE, J., VANRADEN, P., VAN TASSELL, C., NULL, D., SCHROEDER, S., BICKHART, D. et MCCLURE, M., 2013. Identification of a Nonsense Mutation in CWC15 Associated with Decreased Reproductive Efficiency in Jersey Cattle. In : PAS, Marinus FW. te. (éd.), *PLoS ONE*. 22 janvier 2013. Vol. 8, n° 1, p. e54872. DOI 10.1371/journal.pone.0054872.
74. STELZER, G., ROSEN, N., PLASCHKES, I., ZIMMERMAN, S., TWIK, M., FISHILEVICH, S., STEIN, T.I., NUDEL, R., LIEDER, I., MAZOR, Y., KAPLAN, S., DAHARY, D., WARSHAWSKY, D., GUAN-GOLAN, Y., KOHN, A., RAPPAPORT, N., SAFRAN, M. et LANCET, D., 2016. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. In : *Current Protocols in Bioinformatics*. 20 2016. Vol. 54, p. 1.30.1-1.30.33. DOI 10.1002/cpbi.5.
75. SUÁREZ-VEGA, A., GUTIÉRREZ-GIL, B., BENAVIDES, J., PEREZ, V., TOSSER-KLOPP, G., KLOPP, C., KEENNEL, S. et ARRANZ, J., 2015. (Combining GWAS and RNA-Seq Approaches for Detection of the Causal Mutation for Hereditary Junctional Epidermolysis Bullosa in Sheep). In : WICHE, G. (éd.), *PLOS ONE*. 8 mai 2015. Vol. 10, n° 5, p. e0126416. DOI 10.1371/journal.pone.0126416.
76. SUÁREZ-VEGA, A., GUTIÉRREZ-GIL, B., CUCHILLO-IBÁÑEZ, I., SÁEZ-VALERO, J., PÉREZ, V., GARCÍA-GÁMEZ, E., BENAVIDES, J. et ARRANZ, J.J., 2013. Identification of a 31-bp Deletion in the RELN Gene Causing Lissencephaly with Cerebellar Hypoplasia in Sheep. In : BROWNING, G.F. (éd.), *PLoS ONE*. 19 novembre 2013. Vol. 8, n° 11, p. e81072. DOI 10.1371/journal.pone.0081072.
77. TAGU, D., JAUBERT-POSSAMAI, S. et MÉREAU, A., 2018a. Fiche 42 : Séquencer un génome : généralités. In : *Principes des techniques de biologie moléculaire et génomique*. Quae. S.l. : s.n. Savoir faire. p. 184-185.
78. TAGU, D., JAUBERT-POSSAMAI, S. et MÉREAU, A., 2018b. Fiche 43 : Séquençage d'ADN par la technique Illumina. In : *Principes des techniques de biologie moléculaire et génomique*. Quae. S.l. : s.n. Savoir faire. p. 186-189.
79. TAGU, D., JAUBERT-POSSAMAI, S. et MÉREAU, A., 2018c. Fiche 63 : Génotypage SNP (Single Nucleotide Polymorphism). In : *Principes des techniques de biologie moléculaire et génomique*. Quae. S.l. : s.n. Savoir faire. p. 258-265.
80. TAYLOR CLELLAND, C. L., CRACIUN, L., BANCROFT, C. et LUFKIN, T., 2000. Mapping and developmental expression analysis of the WD-repeat gene Preb. In : *Genomics*. 1 février 2000. Vol. 63, n° 3, p. 391-399. DOI 10.1006/geno.1999.6089.
81. TESTONI, S., DIANA, A., OLZI, E. et GENTILE, A., 2008. Brachyspina syndrome in two Holstein calves. In : *Veterinary Journal*. juillet 2008. Vol. 177, n° 1, p. 144-146. DOI 10.1016/j.tvjl.2007.03.011.
82. THOMSEN, B., HORN, P., PANITZ, F., BENDIXEN, E., PERTERSEN, A.H., HOLM, L.E., NIELSEN, V.H., AGERHOLM, J.S., ARNBJERG, J. et BENDIXEN, C., 2006. A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. In : *Genome Research*. 2006. Vol. 16, n° 1, p. 97-105. DOI 10.1101/gr.3690506.
83. THORVALDSDOTTIR, H., ROBINSON, J. T. et MESIROV, J. P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. In : *Briefings in Bioinformatics*. 1 mars 2013. Vol. 14, n° 2, p. 178-192. DOI 10.1093/bib/bbs017.
84. TORTEREAU, F., 2016. Conséquences de la sélection pour la résistance à la tremblante sur la diversité génétique du génome ovin. In : *Séminaire Ressources Génétiques Animales* [en ligne]. Paris. 12 mai 2016. [Consulté le 20 avril 2019]. Disponible à l'adresse : https://www6.inra.fr/anr/content/download/3185/32405/version/1/file/2016_05_13_tremblante_seminaire_ressources_genetiques.pdf.

85. VANRADEN, P.M., OLSON, K.M., NULL, D.J. et HUTCHISON, J.L., 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. In : *Journal of Dairy Science*. décembre 2011. Vol. 94, n° 12, p. 6153-6161. DOI 10.3168/jds.2011-4624.
86. VENHORANTA, H., PAUSCH, H., FLISIKOWSKI, K., WURMSER, C., TAPONEN, J., RAUTALA, H., KIND, A., SCHNIEKE, A., FRIES, R., LOHI, H. et ANDERSSON, M., 2014. In frame exon skipping in UBE3B is associated with developmental disorders and increased mortality in cattle. In : *BMC Genomics*. 2014. Vol. 15, n° 1, p. 890. DOI 10.1186/1471-2164-15-890.
87. VIGNAL, A., 2011. Etat actuel du séquençage et de la connaissance du génome des espèces animales. In : *INRA Productions Animales*. 2011. Vol. 24, n° 4, p. 18.
88. ZHAO, F., LUFKIN, T. et GELB, B., 2003. Expression of Tfp2d, the gene encoding the transcription factor Ap-2 delta, during mouse embryogenesis. In : *Gene expression patterns: GEP*. mai 2003. Vol. 3, n° 2, p. 213-217.
89. ZHENG, Y., ZHANG, J., WANG, L., ZHOU, Z., XU, M., LI, J. et SHA, J-H., 2006. Cloning and characterization of a novel sperm tail protein, NYD-SP28. In : *International Journal of Molecular Medicine*. décembre 2006. Vol. 18, n° 6, p. 1119-1125.

Annexes

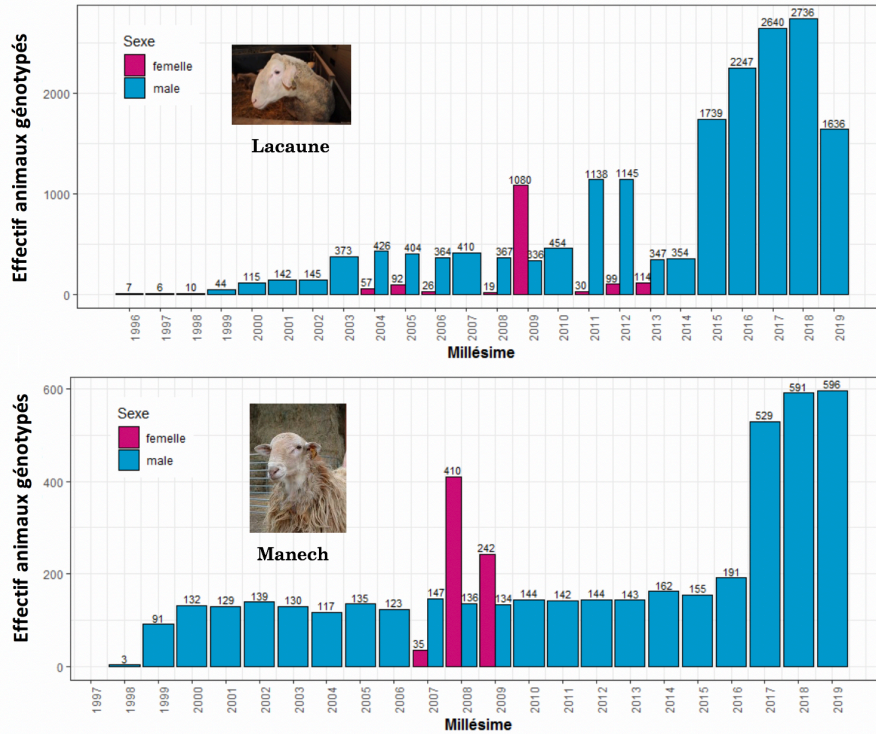
Liste des annexes

ANNEXE I. LIEN ENTRE LES DIFFERENTS HAPLOTYPES UTILISES EN SELECTION AU GENE <i>PRNP</i> (OAR13) ET DEGRE DE SENSIBILITE A LA TREMBLANTE	46
ANNEXE II. DESCRIPTION DES DEUX POPULATIONS D'ETUDE. 19 102 ANIMAUX LACAUNE ET 4 900 ANIMAUX MANECH A TETE ROUSSE GENOTYPES SONT DISPONIBLES.	47
ANNEXE III. NOMBRE DE VARIANTS CANDIDATS PRESENTS DANS LES REGIONS CODANTES, PROMOTRICES OU REGULATRICES OBSERVES DANS LES REGIONS DE LOCALISATION PROBABLE DE MUTATIONS CAUSALES POUR LESQUELLES ON DISPOSE D'ANIMAUX PORTEURS SEQUENCES EN RACE LACAUNE.	48
ANNEXE IV. ENSEMBLE DES HAPLOTYPES OBSERVES (FREQUENCE >1%) DANS LA REGION EN DEFICIT D'HOMOZYGOTIE LL_OAR3_2 (MARQUEUR 2113 A 2130) EN RACE LACAUNE.	49
ANNEXE V. LIEN ENTRE LE GENOTYPE DES HAPLOTYPES LIES AU GENOTYPE SOCS2 (LL_OAR3_2A, LL_OAR3_2B, LL_OAR3_2C) ET LA MUTATION SOCS2 (OAR3) EN RACE LACAUNE.....	50
ANNEXE VI. ENSEMBLE DES HAPLOTYPES OBSERVES (FREQUENCE >1%) DES ANIMAUX T/T A LA MUTATION SOCS2 (OAR3) DANS LA REGION EN DEFICIT D'HOMOZYGOTIE LL_OAR3_2 (MARQUEUR 2113 A 2130) EN RACE LACAUNE.....	51
ANNEXE VII. RECHERCHE DE GENES CANDIDATS DANS LES REGIONS PRESENTANT UN DEFICIT EN HOMOZYGOTIE DONT L'ON NE DISPOSE PAS D'ANIMAUX PORTEURS SEQUENCES EN RACES LACAUNE ET MANECH A TETE ROUSSE.....	52
ANNEXE VIII. ÉVOLUTION DE LA VALEUR DES DYD (« DAUGHTER YIELD DEVIATION ») POUR LES BELIERS EN SELECTION GENOTYPES DE L'ETUDE EN RACES LACAUNE (A) ET MANECH A TETE ROUSSE (B) EN FONCTION DES MILLESIMES.....	53
ANNEXE IX. ÉVOLUTION DE LA FREQUENCE DES BELIERS HETEROZYGOTES EN SELECTION EN RACES LACAUNE (A) ET MANECH A TETE ROUSSE (B) EN FONCTION DES MILLESIMES. C. ÉVOLUTION DE LA FREQUENCE DES BELIERS HETEROZYGOTES EN SELECTION PORTEURS DES HAPLOTYPES ASSOCIES A LA MUTATION DU GENE SOCS2 EN RACE LACAUNE.....	54

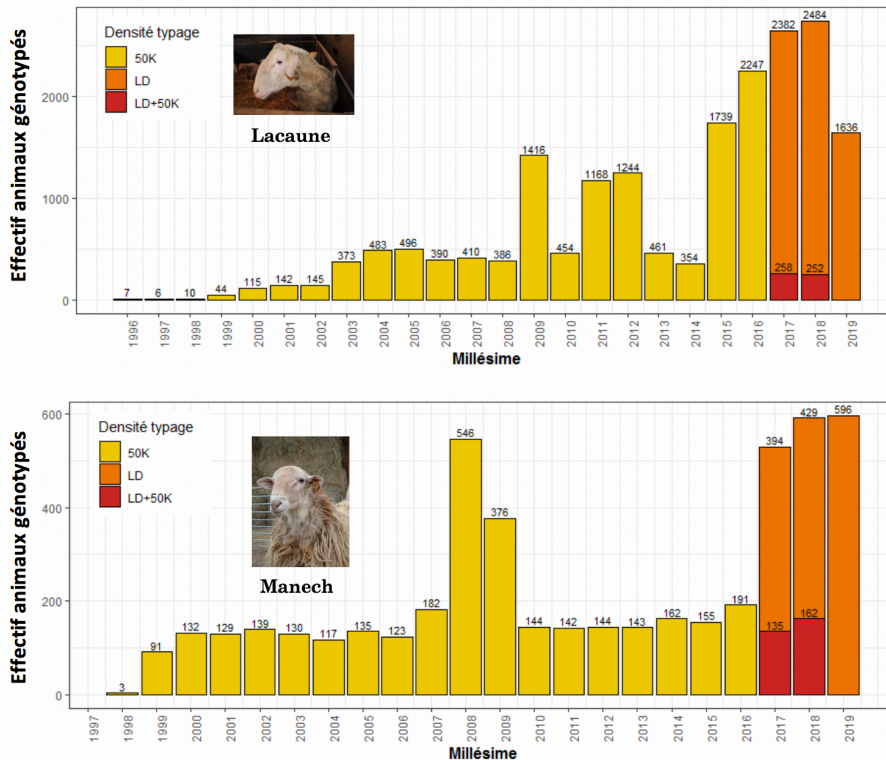
Génotypes	Codon 136	Codon 154	Codon 171	Sensibilité
Forme mutée	V	R	Q	
Forme ancestrale	A	R	Q	
Forme mutée	A	H	Q	
Forme mutée	A	R	R	

Annexe i. Lien entre les différents haplotypes utilisés en sélection au gène *PRNP* (OAR13) et degré de sensibilité à la tremblante d'après [Elsen et al. \(1997\)](#) et [Tortereau \(2016\)](#). La tremblante ovine est une maladie à prion (Encéphalopathie Spongiforme Subaiguës Transmissibles ESST), affectant le système nerveux central conduisant à la mort de l'animal. Le gène *PRNP* codant pour la protéine à prion (*PrP*) a été identifié et une dizaine polymorphismes ont été mis en évidence dans ce gène. Les trois plus importants utilisés en sélection sont les polymorphismes présents dans les codons : codon 136 : Alanine (A) > Valine (V), codon 154 : Arginine (A) > Histidine (H) et codon 171 : Glutamine (Q) > Arginine (R). Au vu des dégâts causés par cette maladie dans les années 1990, un Programme National d'Amélioration Génétique pour la Résistance à la Tremblante (PNAGRT) a été mis en place en 2001 pour éliminer les animaux ARQ (forme ancestral) et sélectionner les animaux homozygotes ARR/ARR. La sensibilité à la tremblante est essentiellement liée à l'allèle Q. Aujourd'hui seuls les béliers ARR/ARR rentrent en centre d'élevage.

A.



B.



Annexe ii. Description des deux populations d'étude. 19 102 animaux Lacaune et 4 900 animaux Manech à Tête Rousse génotypés sont disponibles (Crédit photo : Diane Buisson, CDEO Ordiap). A. Répartition des animaux génotypés suivant leur sexe et millésime. **B.** Répartition des animaux génotypés suivant leur densité de typage et millésime. Remarque : au moment de l'acquisition des données pour cette analyse (mars 2019 en Lacaune et mai 2019 en Manech), tous les génotypages et choix des béliers reproducteurs n'ont pas été réalisés pour l'année 2019. LD : typage basse densité, LD+50K : typages basse densité puis sur la 50K, 50K : typage 50K.

Région	Déficit de l'haplotype (en %)	SNP présents dans les séquences codantes ¹ , régulatrices ou promotrices ²						InDels présents dans les séquences codantes
		Faux-sens ¹	Non-sens ¹	Synonyme ¹	3'UTR ¹	5'UTR ¹	Upstream/Downstream ²	
LL_OAR3_1	100%	1	-	-	-	-	5	-
LL_OAR4_1	100%	-	-	-	-	-	1	-
LL_OAR13_1	100%	-	-	-	-	-	3	1
LL_OAR3_3	96%	-	1	1	-	1	1	-
LL_OAR17_1	92%	-	-	-	-	-	1	-
LL_OAR23_1	87%	-	-	-	-	-	1	-
LL_OAR24_1	87%	-	-	-	-	-	4	-
LL_OAR3_2	81%	-	-	-	-	-	-	-

Annexe iii. Nombre de variants candidats présents dans les régions codantes, promotrices ou régulatrices observés dans les régions de localisation probable de mutations causales pour lesquelles on dispose d'animaux porteurs séquencés en race Lacaune. Les régions potentiellement associées à des mutations létales sont indiquées en orange (déficit de 100%).

Région LL_OAR3_2																					Fréquence
Haplotype	2113	2114	2115	2116	2117	2118	2119	SOCS2	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130		
LL_OAR3_2a	A	G	G	G	G	G	G		G	C	A	G	A	G	C	A	G	A	A	3,46%	
LL_OAR3_2b	A	G	G	G	G	G	G		A	A	A	G	G	A	A	G	G	G	G	2,77%	
LL_OAR3_2c	A	G	G	G	G	G	G		A	A	A	G	G	A	A	G	G	G	A	1,79%	
LL_OAR3_2d	A	A	A	A	A	A	G		G	A	G	G	A	G	A	A	G	G	G	10,65%	
LL_OAR3_2e	G	G	A	G	A	A	A		A	A	G	A	G	G	C	A	G	G	G	9,05%	
LL_OAR3_2f	A	A	G	A	A	A	A		G	C	G	A	G	G	A	G	G	A	A	7,02%	
LL_OAR3_2g	G	A	A	A	G	G	G		G	A	G	G	A	G	A	G	G	G	G	3,49%	
LL_OAR3_2h	G	G	A	G	A	A	A		A	A	A	G	G	A	C	A	G	A	G	3,35%	
LL_OAR3_2i	A	A	G	A	A	G	G		A	C	G	G	G	A	A	A	G	G	G	3,13%	
LL_OAR3_2j	A	A	A	G	A	G	A		G	C	G	A	G	G	C	A	G	A	A	2,97%	
LL_OAR3_2k	G	A	A	G	A	G	A		G	C	G	G	A	G	A	A	G	G	A	2,90%	
LL_OAR3_2l	A	A	A	A	A	G	G		G	C	G	A	G	A	A	G	A	G	G	2,49%	
LL_OAR3_2m	G	G	A	G	A	A	A		A	A	G	G	G	A	A	G	A	G	G	2,36%	
LL_OAR3_2n	A	A	A	A	A	A	G		A	A	A	G	G	A	C	G	G	A	A	1,98%	
LL_OAR3_2o	A	A	A	A	A	A	G		A	A	A	G	G	A	C	G	G	G	A	1,97%	
LL_OAR3_2p	A	A	A	A	A	A	G		A	A	A	G	G	A	A	G	A	G	G	1,81%	
LL_OAR3_2q	G	A	A	G	A	G	A		G	C	G	G	A	G	C	A	G	A	G	1,80%	
LL_OAR3_2r	G	A	A	A	G	G	G		G	A	A	G	G	A	A	A	G	G	A	1,79%	
LL_OAR3_2s	A	A	A	G	A	G	A		G	C	G	G	A	G	A	A	G	G	G	1,71%	
LL_OAR3_2t	A	G	G	G	A	A	A		G	C	A	G	A	G	A	G	G	G	A	1,70%	
LL_OAR3_2u	A	A	G	A	A	A	G		A	C	G	A	G	G	C	G	G	G	G	1,54%	
LL_OAR3_2v	A	A	A	A	A	G	G		A	C	A	G	G	G	C	A	G	G	G	1,52%	
LL_OAR3_2w	A	A	A	G	A	A	G		G	A	A	G	A	G	C	G	G	G	A	1,37%	

Annexe iv. Ensemble des haplotypes observés (fréquence >1%) dans la région en déficit d'homozygotie LL_OAR3_2 (marqueur 2113 à 2130) en race Lacaune. Le gène *SOCS2* est localisé entre le marqueur 2119 et 2120. Le sous-haplotype de LL_OAR3_2 « AGGGGGG » (nommé LL_OAR3_2_T) est retrouvé pour les 3 haplotypes LL_OAR3_2a, LL_OAR3_2b et LL_OAR3_2c.

Génotypes	C/C	C/T	T/T	Total
LL_OAR3_2a/LL_OAR3_2a	0	0	0	0
LL_OAR3_2a/LL_OAR3_2+	2	178	3	183
LL_OAR3_2+/LL_OAR3_2+	3544	632	8	4184
Total	3546	810	11	4367

Génotypes	C/C	C/T	T/T	Total
LL_OAR3_2b/LL_OAR3_2b	0	0	0	0
LL_OAR3_2b/LL_OAR3_2+	6	252	5	263
LL_OAR3_2+/LL_OAR3_2+	3540	558	6	4104
Total	3546	810	11	4367

Génotypes	C/C	C/T	T/T	Total
LL_OAR3_2c/LL_OAR3_2c	0	0	0	0
LL_OAR3_2c/LL_OAR3_2+	0	150	6	156
LL_OAR3_2+/LL_OAR3_2+	3546	660	5	4211
Total	3546	810	11	4367

Annexe v. Lien entre le génotype des haplotypes liés au génotype *SOCS2* (LL_OAR3_2a, LL_OAR3_2b, LL_OAR3_2c) et la mutation *SOCS2* (OAR3) en. Race Lacaune. + : Correspond à des haplotypes différents des haplotypes en déficit. 4367 génotypages au variant *SOCS2* sont disponibles (bélisiers de millésimes 2017-18-19).

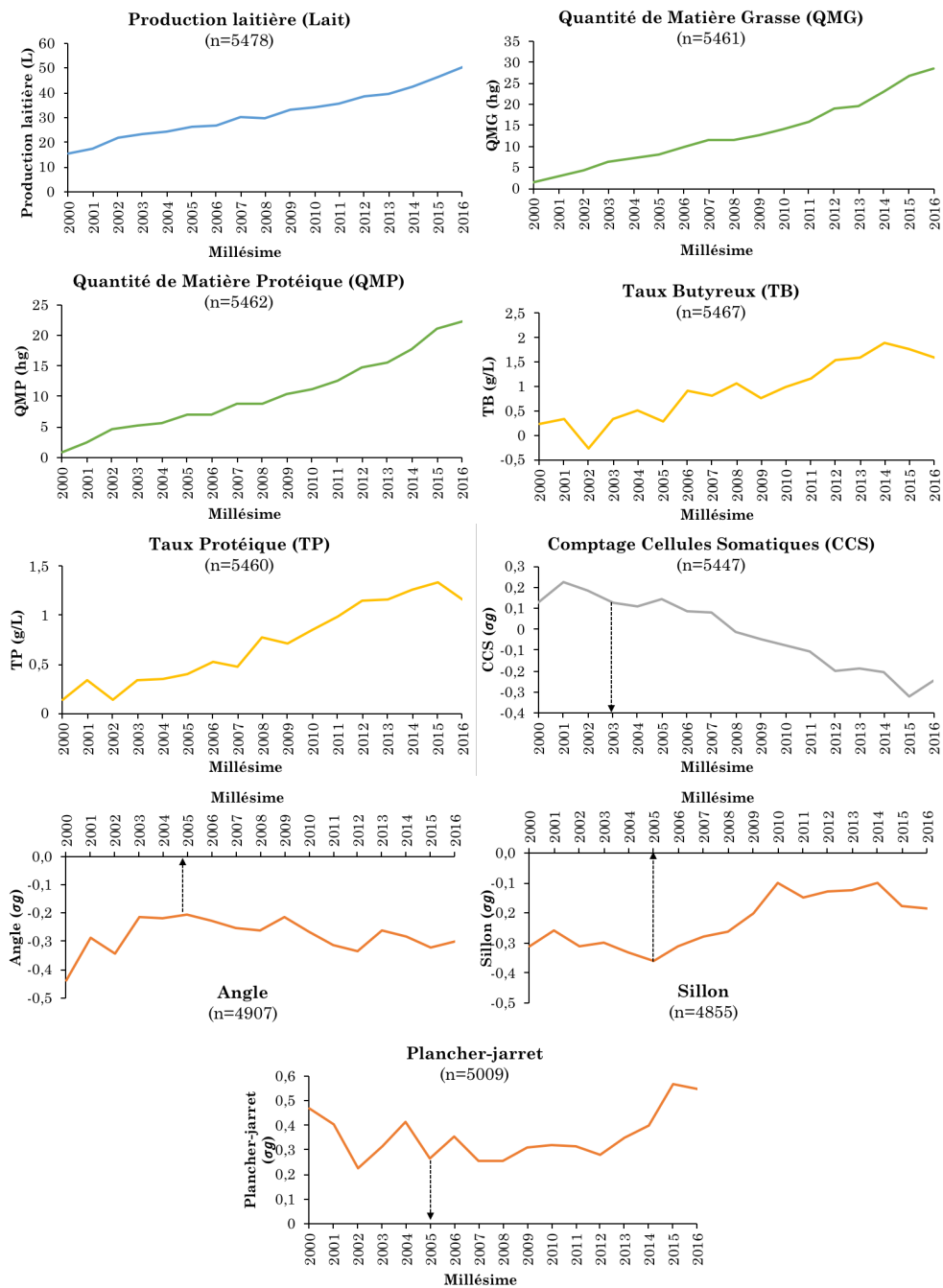
		Région LL_OAR3_2																		
N° animal	Phase	2113	2114	2115	2116	2117	2118	2119	SOCS2	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130
		LL_OAR3_2 T																		
Animal 1	1	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	G	G	A	A
Animal 1	2	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	A	G	A	A
Animal 2	1	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	G	G	A	A
Animal 2	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	G
Animal 3	1	A	G	G	G	G	G	G	T	G	C	G	A	G	G	A	G	A	G	A
Animal 3	2	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	A	G	A	A
Animal 4	1	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	G	G	A	A
Animal 4	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	A
Animal 5	1	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	A	G	G
Animal 5	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	A
Animal 6	1	A	G	G	G	G	G	G	T	G	C	G	A	G	G	A	G	A	G	A
Animal 6	2	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	A	G	A	A
Animal 7	1	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	A
Animal 7	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	G
Animal 8	1	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	A
Animal 8	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	G
Animal 9	1	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	G
Animal 9	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	A
Animal 10	1	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	A
Animal 10	2	A	G	G	G	G	G	G	T	G	C	A	G	G	A	C	A	G	G	G
Animal 11	1	A	G	G	G	G	G	G	T	G	C	A	G	A	G	C	G	G	A	A
Animal 11	2	A	G	G	G	G	G	G	T	A	A	A	G	G	A	A	G	G	G	G

Annexe vi. Ensemble des haplotypes observés (fréquence >1%) des animaux T/T à la mutation SOCS2 (OAR3) dans la région en déficit d'homozygote LL_OAR3_2 (marqueur 2113 à 2130) en race Lacaune. Le gène SOCS2 est localisé entre le marqueur 2119 et 2120. Le sous-haplotype de LL_OAR3_2 « AGGGGG » (nommé LL_OAR3_2_T) est retrouvé sur les deux phases (paternelle (1) et maternelle (2)) des 11 animaux porteurs homozygotes du variant SOCS2.

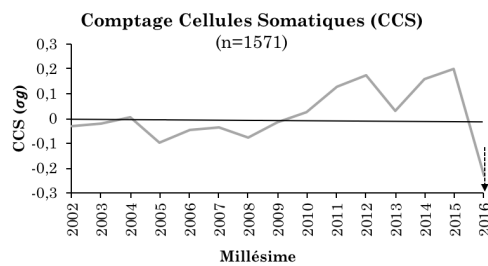
Région	Gènes candidats fonctionnels et positionnels (d'après GeneCards https://www.genecards.org (Stelzer et al., 2016))
LL_OAR3_4	Cluster gènes HOXC (<i>Homeobox</i> , intervient lors du développement embryonnaire), MAP3K2 (<i>Mitogen-Activated Protein Kinase Kinase Kinase 12</i> , différenciation neuronale), AMHR2 (<i>Anti-Mullerian Hormone Receptor Type 2</i> , développement organes génitaux), SP1 (<i>Sp1 Transcription Factor</i> , réparation de l'ADN, remodelage de la chromatine), SP7 (<i>Sp7 Transcription Factor</i> , facteur de transcription spécifique des os), AAS (<i>Aladin WD Repeat Nucleoporin</i> , protéine régulatrice, développement du système nerveux central et périphérique), ESPL1 (<i>Extra Spindle Pole Bodies Like 2, Separase</i> , stabilisation des chromatides), RARG (<i>Retinoic Acid Receptor Gamma</i> , régulation de la transcription), ZNF740 (<i>Zinc Finger Protein 740</i> , régulation de la transcription), TNS2 (<i>Tensin 2</i> , régulation de la migration cellulaire), EIF4B (<i>Eukaryotic Translation Initiation Factor 4B</i> , signalisation cellulaire).
LL_OAR18_1	CEMIP (<i>Cell Migration Inducing Hyaluronidase 1</i> , migration cellulaire (surdité)), MESDC1 (<i>Tail Rod Domain 1/Mesoderm Development Candidate 1</i> , différenciation du mésoderme), CFAP161 (<i>Cilia And Flagella Associated Protein 161</i> , fonction des cils mobiles), TMC3 (<i>Transmembrane Like 3</i> , surdité, épidermodysplasie), APBA2 (<i>Amyloid Beta Precursor Protein Binding Family A Member 2</i> , transduction de signaux), NSMCE3 (<i>NSE3 Homolog, SMC5-SMC6 Complex Component</i> , Complexe de réorganisation de la chromatine).
LL_OAR18_2	LOXL1 (<i>Lysyl Oxidase Like 1</i> , formation tissu conjonctif), STXBP6 (<i>Syntaxin Binding Protein 6</i> , régulation de la formation du complexe SNARE).
MTR_OAR1_1	Cluster gènes ZMYM (<i>Zinc Finger MYM</i> , associé au syndrome de myasthénie entraînant une fragilité musculaire), KIAA0319L (<i>KIAA0319 Like</i> , guidance axonal), TFAP2E (<i>Transcription Factor AP-2 Epsilon</i> , régulation de la transcription), CLSPN (<i>Claspin</i> , régulation cycle cellulaire).
MTR_OAR1_2	Cluster gènes codant pour des histones (<i>H2B type 2-F, H3, H2A type 2-A, H2B type 2-E</i> , constituants protéiques des chromosomes), SF3B4 (<i>Splicing Factor 3b Subunit 4</i> , assemblage des complexes spliceosomes).
MTR_OAR1_3/ MTR_OAR_4	GPR149 (<i>G Protein-coupled Receptor 149</i> , transduction cellulaire), DHX36 (<i>DEAH-Box Helicase 36</i> , Stabilisation de l'ADN) - MBNL1 (<i>Muscleblind Like Splicing Regulator 1</i> , régulation de la transcription, impliqué dans la dystrophie myotonique), IGSF10 (<i>Immunoglobulin Superfamily Member 10</i> , migration des neurones), MED12L (<i>Mediator Complex Subunit 12L</i> , régulation de la transcription)
MTR_OAR9_1	ZPM2 (<i>Zinc Finger Protein, FOG Family Member 2</i> , facteurs de transcription), DCSTAMP (<i>Dendrocyte Expressed Seven Transmembrane Protein</i> , ostéoclastogenèse, différenciation de la myéloïde).
MTR_OAR13_1	CDK5RAP1 (<i>CDK5 Regulatory Subunit Associated Protein 1</i> , modification de l'ADN), SNTA1 (<i>Syntrophin Alpha 1</i> , jonctions cellulaires des muscles, associé à de l'arythmie cardiaque), RALY (<i>RALY Heterogeneous Nuclear Ribonucleoprotein</i> , épissage de l'ARN pré-messager, développement embryonnaire), MAP1LC3A (<i>Microtubule Associated Protein 1 Light Chain 3 Alpha</i> , interaction entre microtubules et composés du cytosquelette), PIGU (<i>Phosphatidylinositol Glycan Anchor Biosynthesis Class U</i> , contrôle division cellulaire).
MTR_OAR20_1	CENPQ (<i>Centromere Protein Q</i> , sous-unités des centromères), CRISP2 (<i>Cysteine Rich Secretory Protein 2</i> , Capacitation des spermatozoïdes), CRISP3 (<i>Cysteine Rich Secretory Protein 3</i> , structure des spermatozoïdes, implantation de l'embryon), CRISP1 (<i>Cysteine Rich Secretory Protein 1</i> , fusion spermatozoïde-ovule), TFAP2D (<i>Transcription Factor AP-2 Delta</i> , facteur de transcription), TFAP2B (<i>Transcription Factor AP-2 Beta</i> , facteur de transcription).

Annexe vii. Recherche de gènes candidats dans les régions présentant un déficit en homozygotie dont l'on ne dispose pas d'animaux porteurs séquencés en races Lacaune et Manech à Tête Rousse.

A.

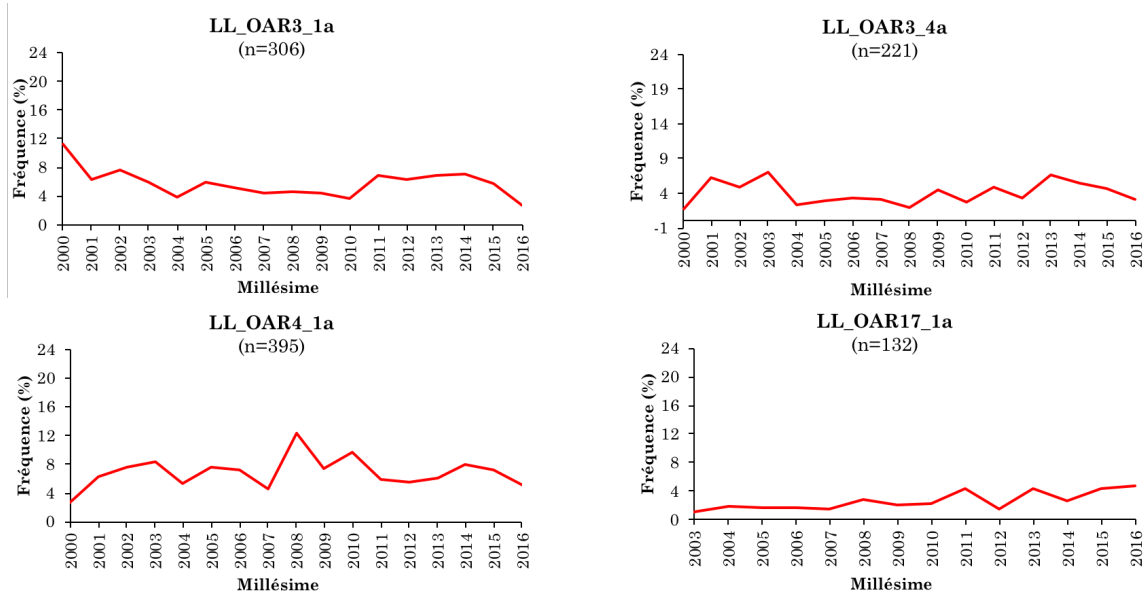


B.

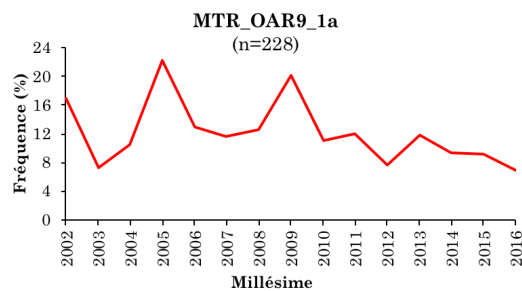


Annexe viii. Évolution de la valeur des DYD (« *Daughter Yield Deviation* ») pour les béliers en sélection génotypés de l'étude en races Lacaune (A) et Manech à Tête Rousse (B) en fonction des millésimes. La flèche représente l'année où le caractère intègre l'ISOL (si elle n'est pas représentée, cela signifie que le caractère a été intégré à l'ISOL avant 2000).

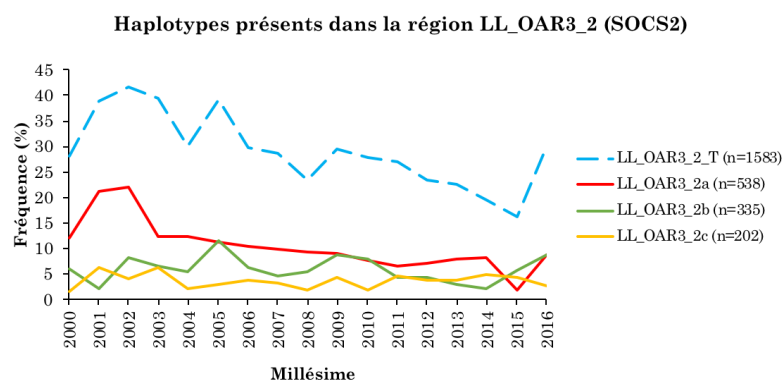
A.



B.



C.



Annexe ix. Évolution de la fréquence des béliers hétérozygotes en sélection en races Lacaune (A) et Manech à Tête Rousse (B) en fonction des millésimes. C. Évolution de la fréquence des béliers hétérozygotes en sélection porteurs des haplotypes associés à la mutation du gène *SOCS2* en race Lacaune.

Résumé

L'amélioration de la reproduction et la fertilité est essentielle chez les ovins. Des mutations récessives homozygotes délétères peuvent émerger sous l'effet de la consanguinité et engendrer des conséquences néfastes sur le développement embryonnaire ou la survie du jeune (mutations létales). Pour repérer de telles mutations, une approche de génétique inverse a été réalisée pour identifier des déficits en homozygotes dans certaines régions du génome à partir de données de génotypage haut-débit venant de la sélection génomique pour les races ovines Lacaune et Manech à Tête Rousse. Dix-huit régions haplotypiques de 0,5 à 3 Mb présentent entre 81 et 100% de déficit en homozygotie. Les fréquences de ces haplotypes en déficit sont comprises entre 3,9 et 17,4%. Parmi ces régions, quatre haplotypes en race Lacaune et trois haplotypes en race Manech présentent un déficit total d'animaux homozygotes, associés à des mutations létales. L'exploration des données de séquence tout génome en race Lacaune a permis de mettre en évidence deux mutations candidates fortement délétères dans les gènes *PREB* (intervenant lors du développement embryonnaire) et *CCDC65* (mise en place de structures ciliaires). Cependant, la sélection génétique peut aussi engendrer des déficits d'homozygotie liés à un gène en sélection, le gène *SOCS2* de résistance aux mammites. L'intégration de ces anomalies dans les schémas de sélection pourra être prise en compte afin de limiter les accouplements à risques lorsque la causalité de ces mutations sera confirmée.

Mots-clés : mutation létale récessive, anomalie génétique, sélection génomique, génétique inverse, haplotype en déficit d'homozygotie, ovins (*Ovis aries*)

Abstract

Reproduction and fertility are two main concerns in dairy sheep. Deleterious homozygous recessive mutations may emerge as a result of consanguinity and may have consequences on embryo development or survival of the young animal (so-called lethal mutations). To identify such mutations, a reverse genetic design was conducted to identify homozygous deficit in some chromosome segments. High-density genotyping data from genomic selection were used in two dairy sheep breeds, Lacaune and Manech à Tête Rousse. Eighteen regions with haplotypes covering 0.5 to 3 Mb exhibit deficit in homozygosity (from 81 to 100%). The frequencies of these haplotypes are between 3.9 and 17.4%. Among these regions, four Lacaune haplotypes and three Manech haplotypes show a total deficit of homozygous animals probably associated with lethal mutations. In Lacaune dairy sheep, the screening of whole genome sequence data revealed two highly deleterious variants in the *PREB* (involved during embryonic development) and *CCDC65* genes (establishment of ciliary structures). However, genetic selection can also lead to homozygosity deficits linked to a gene under selection (such as *SOCS2* gene associated to resistance to mastitis). If the causality of these variants is confirmed as lethal mutations, their integration in selection schemes could be considered in order to avoid mating between animals carrying the same mutation.

Keywords : recessive lethal mutations, genetic disorder, genomic selection, reverse genetic screens, homozygous haplotype deficit, sheep (*Ovis aries*)