

# Déterminisme génétique de la composition en acides gras et protéines du lait des ruminants, et potentialités de sélection

D. BOICHARD<sup>1,2</sup>, A. GOVIGNON-GION<sup>1,2</sup>, H. LARROQUE<sup>3,4,5,6</sup>, C. MAROTEAU<sup>3,4,5,6</sup>, I. PALHIÈRE<sup>3,4,5,6</sup>,  
G. TOSSER-KLOPP<sup>3,4,5,6</sup>, R. RÜPP<sup>3,4,5,6</sup>, M.-P. SANCHEZ<sup>1,2</sup>, M. BROCHARD<sup>7,1,2</sup>

<sup>1</sup> INRA, UMR1313 GABI, F-78352 Jouy-en-Josas, France

<sup>2</sup> AgroParisTech, UMR1313 GABI, 16 rue Claude Bernard, F-75321 Paris, France

<sup>3</sup> INRA, UMR1388 GenPhySE, F-31326 Castanet-Tolosan, France

<sup>4</sup> Université de Toulouse INPT ENSAT, UMR1388 GenPhySE, F-31326 Castanet-Tolosan, France

<sup>5</sup> Université de Toulouse INPT ENVT, UMR1388 GenPhySE, F-31076 Toulouse, France

<sup>6</sup> Université de Toulouse INPT, Ecole d'Ingénieurs de Purpan, UMR1388 GenPhySE, F-31076 Toulouse, France

<sup>7</sup> Institut de l'Élevage, 149 rue de Bercy, F-75595 Paris, France

Courriel : didier.boichard@jouy.inra.fr

Reposant sur la prédiction des composants du lait à partir de la spectrométrie dans le moyen infra rouge, le projet *PhénoFinlait* a permis une analyse à grande échelle du déterminisme génétique de la composition en acides gras et protéines du lait dans les trois espèces, bovine, ovine et caprine, ouvrant la possibilité d'une sélection prochaine sur certains de ces critères.

Chez les bovins laitiers, la composition du lait entre dans l'objectif de sélection, explicitement ou implicitement, depuis au moins 40 ans. La mesure du Taux Protéique (TP) vrai s'est généralisée en France au cours des années 1970 plus précocement que dans les autres pays, se substituant à celle du taux azoté. Alors qu'initialement seul le Taux Butyreux (TB) était pris en compte dans la sélection, l'évaluation a été étendue en 1978 au taux de matière utile, combinaison des TP et TB. En 1989, les deux taux (et les quantités de matières correspondantes) ont enfin eu des évaluations génétiques séparées, permettant, dans la limite de leur corrélation génétique, de les sélectionner séparément. Le poids du TP dans l'objectif laitier dit INEL, a toujours été supérieur à sa valeur économique et a conduit à son amélioration génétique lente mais régulière intra-race, rééquilibrant une évolution raciale du cheptel français plutôt défavorable à la composition du lait. Concernant le TB, la sélection réalisée a connu des évolutions variables souvent liées à la conjoncture de court terme et à l'impact des quotas. On a donc connu des phases alternées de sélection pour ou contre le TB, avec une évolution globale à la hausse.

Chez les petits ruminants, la prise en compte de la composition du lait dans la sélection a été plus tardive malgré l'uti-

lisation exclusivement fromagère du lait, du fait du coût élevé de la mesure des taux en contrôle laitier par rapport à la valeur de l'animal. Chez la chèvre, l'accent a été mis sur la quantité de matière protéique et le TP, le TB étant longtemps négligé. Le TP a été fortement augmenté, cette évolution étant également favorisée par l'impact majeur de la caséine alpha-s1 sur le TP (Grosclaude *et al* 1987). Chez les ovins laitiers, la sélection, longtemps orientée sur la production de lait a intégré la composition du lait, au début des années 1990 (Barillet *et al* 1992).

Durant toute cette période, des études ont régulièrement montré l'impact de la nature des constituants du lait sur ses propriétés techno-fonctionnelles. La composition en Acides Gras (AG) influence les qualités nutritionnelles du lait, *via* notamment le rapport acides gras saturés/insaturés. Elle affecte également le point de fusion du beurre et donc son aptitude à être tartiné. La composition protéique joue un rôle sur l'aptitude du lait à coaguler, le rendement fromager et la stabilité à la chaleur. De nombreuses données ont été accumulées sur l'influence de l'alimentation sur la composition en AG du lait (Chilliard *et al* 2007, Schmidely et Bourgeois 2011). En revanche, la composition fine du lait n'a pas fait l'objet de sélection jusqu'à présent, pour deux raisons principales, la difficulté à

mesurer ces composantes en routine à un coût réduit et l'absence de définition claire et partagée de l'objectif de sélection.

En 2007, dans le but de développer des outils de caractérisation, de qualification et de maîtrise de la composition fine du lait, le CNIEL a impulsé des études pluridisciplinaires qui ont conduit à la mise en place du programme *PhénoFinlait*. L'idée centrale reposait sur l'utilisation de la spectrométrie dans le Moyen InfraRouge (MIR) pour caractériser la composition fine du lait au-delà des TB et TP. En effet, les travaux précurseurs de l'Université de Gembloux (Soyeurt *et al* 2006) avaient montré la faisabilité d'une mesure des principaux AG du lait à partir des spectres MIR. L'idée était que si ces résultats étaient confirmés, on pouvait alors accéder à la composition fine du lait à très grande échelle (plus de 20 millions d'analyse MIR de lait par an en France) et à un coût marginal, ouvrant ainsi pour la première fois la possibilité de caractériser individuellement les animaux et de les sélectionner sur des composants spécifiques du lait. Si la priorité initiale concernait les AG, compte tenu de leur valeur santé et des résultats convaincants de la bibliographie, *PhénoFinlait* a été étendu dès l'origine aux protéines, autre composant majeur du lait. Un acquis important du projet *PhénoFinlait* a été d'apporter une solution à la mesure d'un

nombre substantiel de composants avec une précision raisonnable, voire très bonne, ouvrant la voie à des analyses génétiques.

Dans cet article, nous présentons deux types d'analyse de la composition fine du lait, des estimations de paramètres génétiques et des résultats de détection de QTL. Les paramètres génétiques mesurent la part d'origine génétique dans la variabilité génétique, ainsi que les liaisons génétiques entre caractères. Ils constituent l'information de base pour optimiser la sélection et prédire le progrès réalisable. La détection de QTL permet d'affiner la connaissance du déterminisme génétique en mettant en évidence les régions des génomes les plus importantes et en ouvrant la voie à l'identification des mutations impliquées dans ce déterminisme. Estimations de paramètres génétiques et détection de QTL sont présentées à la fois pour les AG et les protéines, dans les trois espèces.

## 1 / Le dispositif et les méthodes utilisées

### 1.1 / Protocole *PhénoFinlait*

Le protocole, décrit en détail dans Gelé *et al* (2014, ce numéro), peut être résumé comme suit.

Pour sa composante bovine, le programme *PhénoFinlait* (Brochard *et al* 2009) concerne 1 224 élevages de 16 départements et les trois races Holstein (Ho), Montbéliarde (Mo) et Normande (No). Entre fin 2009 et fin 2010, en moyenne 6 contrôles expérimentaux ont été réalisés dans ces élevages. Ces contrôles s'intégraient dans le cadre des contrôles laitiers officiels classiques mais ont été complétés par la collecte des spectres MIR lors de l'analyse des laits, ainsi que par une enquête en élevage sur l'alimentation des animaux pour un suivi détaillé de leur ration. 425 000 spectres ont été collectés. Parmi les 90 000 femelles suivies, 12 000 ont été plus particulièrement ciblées en fonction de leur numéro de lactation (surtout primipares), leur date de mise bas et leur origine génétique. Pour ces vaches, une prise de sang a été réalisée par des techniciens des centres d'insémination et envoyée au laboratoire d'analyses génétiques Labogena, pour extraction d'ADN. L'ensemble des données issues de ce protocole a été complété par celles du Système d'Information Génétique et stocké dans la base *PhénoFinlait* développée par le Centre de Traitement de l'Information Génétique (CTIG) de l'INRA.

A noter qu'en race Montbéliarde, la base de spectres MIR a été sensiblement

enrichie par un envoi supplémentaire hors *PhénoFinlait*, permettant de doubler l'information phénotypique de base. Ces données supplémentaires ont été incluses dans l'estimation des paramètres génétiques.

Chez les petits ruminants, le protocole est très similaire. En ovins, 34 000 brebis de races Lacaune et Manech Tête Rousse issues de 160 élevages ont été suivies sur 6 contrôles de troupeau et 118 000 spectres MIR ont été collectés. Chez 3 600 brebis, un échantillon de sang a été prélevé pour extraction d'ADN à Labogena. En caprins, de façon analogue, 70 000 chèvres de races Alpine et Saanen issues de 204 élevages ont été suivies et 4 300 ont fait l'objet d'une prise de sang.

### 1.2 / Estimation du profil en acides gras et protéines

Le profil en AG a été estimé à l'aide d'équations portant sur les spectres MIR collectés. Ces équations ont été établies intra espèce à partir d'échantillons disposant à la fois de spectres et d'analyses de référence par chromatographie en phase gazeuse (Ferrand *et al* 2010, Ferrand-Calmels *et al* 2014). 450 échantillons, choisis pour leur variabilité, ont été utilisés chez les bovins, 200 en ovins et 200 en caprins. Globalement, les équations sont plus précises chez les bovins et ovins que chez les caprins (Gelé *et al* 2014). Les décalages entre analyseurs MIR ou dans le temps pour un même analyseur ont été corrigés grâce à l'analyse d'un lait témoin sur chaque analyseur une à deux fois par mois (Leray *et al* 2011).

Concernant les protéines, le principe a été le même. Les mesures de référence ont été obtenues par chromatographie liquide couplée à la spectrométrie de masse (LC-MS), une technique permettant à la fois l'identification précise des protéines et leur quantification par mesure de la surface sous le pic de chromatographie. Les mesures quantitatives sont disponibles pour les quatre caséines (alpha-s1, alpha-s2, bêta et kappa), l'alpha-lactalbumine et la bêta-lactoglobuline, ainsi que le total des caséines et des protéines du sérum. La précision des prédictions MIR est globalement un peu moins bonne que pour les principaux AG (Gelé *et al* 2014) et particulièrement faible pour l'alpha-lactalbumine en ovins.

### 1.3 / Expression des données

Les AG, tout comme les protéines, peuvent être exprimés de deux façons différentes, soit en proportion de la matière produite (grasse ou protéique), soit en teneur dans le lait. Ces deux expressions sont assez différentes dans

leur signification et leurs propriétés. Concernant les AG, les nutritionnistes humains raisonnent en général en profil d'AG exprimés en teneur dans la matière grasse consommée. Mais le lait ne constitue qu'un élément de la ration humaine et la teneur dans la ration résulte à la fois de la composition des aliments et de la proportion relative de chaque aliment. A l'échelle de la filière laitière, c'est la teneur en AG dans la matière grasse du lait collecté qui fait l'objet d'attention. A l'échelle d'un troupeau ou d'une collecte, Gelé (communication personnelle) a montré que le profil en AG du lait d'un tank, exprimé en % de la matière grasse, dépendait surtout des profils individuels exprimés en % de la matière grasse, bien davantage que des profils individuels exprimés en teneur dans le lait. L'expression en teneur dans la matière grasse semble donc plus intéressante d'un point de vue biologique.

Pour la composition protéique, on peut supposer que l'impact du type de protéine, par exemple sur l'aptitude fromagère, est plus marquant lorsque le composant est exprimé relativement aux protéines totales du lait, de sorte que le choix de l'unité est plus évident.

L'expression en teneur dans la matière présente toutefois un inconvénient important, une liaison mathématique automatique puisque la somme des AG ou des protéines représente quasiment 100%, à la précision des estimations près. Ceci induit des corrélations négatives « artéfactuelles » entre certains AG ou certaines protéines, en particulier les plus importants (par exemple pour les AG, plus il y a de saturés moins il y a d'insaturés) qui n'existent pas dans l'autre mode d'expression.

Dans la suite, l'impact des modes d'expression sera illustré sur les paramètres génétiques bovins. Dans les autres espèces et l'analyse QTL, pour simplifier la présentation, seules les teneurs dans la matière sont présentées.

### 1.4 / Estimation des paramètres génétiques

Les paramètres génétiques ont été estimés séparément dans chacune des races, à partir des données de première lactation de femelles ayant au minimum trois analyses spectrales et des parentés connues. Pour éviter les fortes variations aléatoires, en début et fin de lactation, les contrôles analysés ont été réalisés entre 8 et 350 jours de lactation chez les bovins, 20 et 364 chez les caprins, et après 30 jours pour les ovins. Les questions posées sont les suivantes :

- Les mesures au cours de la lactation reflètent-elles l'expression d'un même

caractère au cours du temps ou bien le déterminisme génétique évolue-t-il au cours de la lactation ? Cette question est étudiée en considérant différentes parties de la lactation, chacune correspondant à un caractère, et en estimant les corrélations génétiques entre caractères. Chez les bovins par exemple, quatre périodes sont définies : T1 = [8-50j], T2 = [51-120j], T3 = [121-200j], T4 = [201-350j].

- Quel est le niveau d'héritabilité des différents composants du lait ? Cette question peut être étudiée à l'échelle du contrôle laitier ou à l'échelle de la lactation, c'est-à-dire sur la moyenne des contrôles. On attend évidemment une héritabilité plus élevée à l'échelle de la lactation, du fait de la variance résiduelle plus faible de la moyenne.
- Quelle est l'étendue de la variabilité génétique ? La variabilité génétique (mesurée par exemple par le coefficient de variation génétique  $CV_g$ , c'est-à-dire le rapport de l'écart-type génétique sur la moyenne du caractère) permet de mesurer l'amplitude du changement relatif que l'on peut obtenir par sélection.
- Quelles sont les liaisons entre caractères ? Cette question est abordée en estimant les corrélations génétiques entre les différents caractères, au niveau du contrôle laitier ou à l'échelle de la lactation.

Dans ces analyses, les variances et covariances ont été estimées par REML. Les contrôles élémentaires ont été analysés avec un modèle incluant les effets fixés du troupeau-contrôle, du stade de lactation, du mois de lactation, et éventuellement de l'analyseur et de la traite de l'échantillon, ainsi que les effets aléatoires génétique et d'environnement permanent de la femelle. Les moyennes par lactation ont été construites à partir de données pré-corrigées pour le stade de lactation et analysées avec un modèle incluant les effets fixés du troupeau, de l'âge et du mois de mise bas, de l'analyseur, ainsi que la valeur génétique de la femelle.

### 1.5 / Détection de QTL

Les femelles ont été génotypées par Labogena avec les puces 50k d'Illumina qui contiennent plus de 50 000 marqueurs SNP (« *Single Nucleotide Polymorphism* » : polymorphisme dû à la variation d'un seul nucléotide) répartis sur tout le génome et qui sont disponibles dans les trois espèces. En bovin, pour des raisons de coût, 500 femelles ont été génotypées avec la puce LD (« *Low Density* » : 7 000 marqueurs SNP), les génotypages manquants ont ensuite pu être reconstitués par

imputation de la puce LD vers la puce 50k. Au total, 7 800 vaches, 1 800 brebis et 2 300 chèvres ont été génotypées. Après élimination de marqueurs avec plus de 5% de données manquantes, peu polymorphes, en déséquilibre de Hardy Weinberg ou présentant des difficultés techniques de génotypage, 36 912 SNP ont été conservés en race bovine Montbéliarde, 37 363 en race Normande, 39 683 en race Holstein, 49 663 en caprins, 40 380 en race ovine Lacaune et 40 454 en race Manech Tête Rousse. Les variables analysées étant des données, les données par contrôle ont d'abord été corrigées pour les effets d'environnement et ont ensuite été combinées pour obtenir une moyenne par animal, ainsi qu'une pondération reflétant la précision de cette moyenne, fonction du nombre de contrôles.

Pour détecter des associations entre les marqueurs SNP et les phénotypes (*i.e.* des QTL : « *Quantitative Trait Loci* »), les analyses ont été réalisées intra-race. Chez les bovins, le modèle utilisé est dit « *Linkage Disequilibrium and Linkage Analysis* » (LDLA, Meuwissen et Goddard 2001) et combine toute l'information disponible au sein d'un pedigree complexe, c'est-à-dire l'information dite « de liaison » apportée par les structures familiales (essentiellement les familles de père) et l'information dite « de déséquilibre de liaison » entre familles et d'origine populationnelle. Cette approche est à la fois puissante et résolutive. En ovins et caprins des analyses de liaison et d'association ont été conduites séparément.

Dans ces approches, deux modèles sont comparés pour chacun des marqueurs SNP en considérant des blocs de 6 SNP consécutifs (ou haplotypes). Sous l'hypothèse H0, on ne considère qu'un effet polygénique pour chaque animal (pas de QTL) et on estime donc une variance polygénique et une variance résiduelle. Sous l'hypothèse H1, le modèle inclut aussi l'effet de l'haplotype et on estime donc en plus la variance due à cet haplotype. A la position testée, on conclut à l'existence d'un QTL de variance non nulle lorsque H1 est significativement plus vraisemblable que H0, le test statistique étant celui du rapport de vraisemblance (« *Likelihood Ratio Test* » ou LRT). Pour tenir compte de la multiplicité des tests réalisés, le seuil de LRT est raisonné en fonction du nombre de faux positifs attendus. Ainsi, avec un LRT supérieur à 13, le nombre de pics faux positifs attendus est de l'ordre de 3 par caractère, correspondant à un taux de faux positif d'environ 10%. Pour un LRT supérieur à 20, le nombre de faux positifs attendus est de l'ordre de 0,2 par caractère.

## 2 / Paramètres génétiques des acides gras

### 2.1 / Héritabilité et variabilité génétique

Le tableau 1 présente les estimations d'héritabilité avec le modèle par contrôle. Les données sont en g/100g de lait et en g/100g de Matière Grasse (MG). Intra-espèce, les résultats sont très cohérents entre races.

Chez les bovins, exprimés en % du lait, les AG Saturés (AGS) présentent une héritabilité moyenne à forte, tandis que les Insaturés (AGI) présentent une héritabilité faible à modérée. Ces résultats sont en accord avec la bibliographie (Stoop *et al* 2008) et s'expliquent par l'origine des AG : les AGS courts et une forte fraction des moyens sont synthétisés *de novo* dans la mamelle à partir des précurseurs sanguins. Les AG à chaîne longue, par exemple C18:1 ou C18:2c9c12, sont apportés par la ration, hydrogénés dans le rumen et transportés par le sang. Certains AGS comme C16:0 et les Mono-Insaturés (AGMI) peuvent avoir les deux origines. Dans une certaine mesure, ces différences reflètent sans doute aussi les différences de précision dans les prédictions MIR, généralement meilleure pour les saturés que pour les insaturés. Exprimés en % du gras, les AG saturés et insaturés présentent des héritabilités modérées et davantage comparables.

Chez les ovins, on retrouve les mêmes tendances pour les grandes familles d'AG. Exprimés en % du lait les AG saturés ont des estimations d'héritabilité plus fortes que les AG insaturés, tandis qu'elles sont plus proches lorsque les AG sont exprimés en % de la MG. Les résultats sont très similaires dans les deux races.

Chez les chèvres, les niveaux d'héritabilité sont également modérés (0,18 à 0,49 selon l'AG, la race et le mode d'expression), mais la différence selon le mode d'expression n'apparaît plus et les héritabilités sont en général un peu plus élevées quand les AG sont exprimés en % de la MG.

Ces valeurs d'héritabilité sont logiquement sensiblement plus élevées en moyenne par lactation (tableau 2) et se rapprochent de l'héritabilité du TB pour certains composants. Ces valeurs élevées montrent que la sélection classique serait facile à mettre en œuvre, dès lors qu'une évaluation génétique serait disponible.

Les coefficients de variation génétique sont assez faibles (tableau 2), compris entre 1 et 5% en % de MG, et entre 5 et



**Tableau 1.** Estimations de l'héritabilité des teneurs en acides gras mesurées par contrôle laitier, selon l'espèce, la race et le mode d'expression.

Espèce	Bovine						Caprine				Ovine			
	% du lait			% de matière grasse			% du lait		% de matière grasse		% du lait		% de matière grasse	
Mode d'expression	Ho	Mo	No	Ho	Mo	No	A	S	A	S	MTR	L	MTR	L
TB	0,24	0,27	0,35				0,23	0,20			0,23	0,29		
C4:0	0,27	0,34	0,32	0,31	0,43	0,48	0,22	0,22	0,48	0,49	0,30	0,35	0,36	0,52
C6:0	0,27	0,38	0,34	0,22	0,37	0,31	0,24	0,21	0,37	0,36	0,29	0,34	0,22	0,37
C8:0	0,26	0,38	0,37	0,23	0,33	0,32	0,24	0,23	0,23	0,37	0,31	0,36	0,27	0,35
C10:0	0,26	0,41	0,42	0,23	0,35	0,32	0,30	0,26	0,28	0,26	0,32	0,22	0,31	0,36
C12:0	0,26	0,41	0,45	0,25	0,35	0,34	0,38	0,37	0,32	0,33	0,33	0,24	0,32	0,39
C14:0	0,26	0,36	0,42	0,18	0,27	0,26	0,32	0,26	0,25	0,23	0,3	0,33	0,26	0,33
C16:0	0,27	0,31	0,32	0,22	0,28	0,19	0,24	0,20	0,38	0,30	0,31	0,38	0,29	0,34
C18:0	0,18	0,26	0,16	0,20	0,29	0,15	0,23	0,21	0,47	0,44	0,20	0,20	0,20	0,23
C18:1	0,12	0,12	0,17	0,16	0,27	0,20	0,19	0,18	0,32	0,24	0,17	0,20	0,22	0,23
C18:2 c9t11	0,08	0,16	0,12	0,10	0,16	0,10	0,21	0,23	0,17	0,18	0,21	0,18	0,26	0,16
AGS <sup>1</sup>	0,25	0,34	0,36	0,15	0,28	0,19	0,24	0,21	0,20	0,21	0,26	0,28	0,26	0,32
AGI <sup>2</sup>	0,14	0,13	0,13	0,17	0,28	0,21	0,20	0,20	0,25	0,25	0,24	0,27	0,17	0,2
AGMI <sup>3</sup>	0,14	0,12	0,19	0,17	0,27	0,21	0,21	0,19	0,30	0,25	0,23	0,25	0,17	0,2
AGPI <sup>4</sup>	0,15	0,17	0,23	0,18	0,28	0,21	0,22	0,23	0,30	0,24	0,25	0,25	0,2	0,17
$\omega$ 3	0,28	0,21	0,22	0,26	0,31	0,20	0,28	0,25	0,24	0,23	0,16	0,18	0,20	0,17
$\omega$ 6	0,18	0,25	0,24	0,17	0,27	0,14	0,20	0,23	0,20	0,19	0,19	0,25	0,23	0,33

Ho = Holstein ; Mo = Montbéliarde ; No = Normande ; A = Alpine ; S = Saanen ; MTR = Manech à Tête Rousse ; L = Lacaune.

<sup>1</sup>AGS : Acides Gras Saturés ; <sup>2</sup>AGI : Acides Gras Insaturés ; <sup>3</sup>AGMI : Acides Gras Mono-Insaturés ; <sup>4</sup>AGPI : Acides Gras Poly-Insaturés.

**Tableau 2.** Estimation de l'héritabilité, de l'écart type génétique ( $\sigma_g$ ) et du Coefficient de Variation ( $CV_g$  en %) du Taux Butyreux (TB) et des teneurs moyennes en AG par lactation (exprimés en g/100g de gras) en race bovine Normande.

	$h^2$	$\sigma_g$	$CV_g$
TB (% du lait)	0,56	2,5	5,4
AGS	0,34	1,2	1,8
AGMI	0,38	1,2	4,8
AGPI	0,34	0,15	4,4
C16:0	0,30	1,2	3,6
C18:1	0,37	1,1	5,8

AGS : Acides Gras Saturés ; AGMI : Acides Gras MonoInsaturés ; AGPI : Acides Gras PolyInsaturés.

**Tableau 3.** Corrélations génétiques entre les teneurs en C14:0 mesurées au cours de quatre périodes de la lactation (T1 = [8-50j], T2 = [51-120j], T3 = [121-200j] et T4 = [201-350j]) en races bovines Montbéliarde (Mo) et Holstein (Ho).

	T1-T2	T2-T3	T3-T4
<b>C14:0 (g/100g lait)</b>			
Mo	0,657	0,927	0,917
Ho	0,685	0,747	0,964
<b>C14:0 (g/100g gras)</b>			
Mo	0,966	0,967	0,965
Ho	0,684	0,913	0,963

10% en % de lait (à comparer par exemple aux valeurs de la littérature de 7% pour le TB ou 9% pour la quantité de lait), et sont plus élevés pour les insaturés, particulièrement les AGMI, que pour les saturés. Ces valeurs montrent que certains composants peuvent varier de façon non négligeable (plusieurs points de pourcentage pour certains), mais qu'il n'apparaît pas possible de modifier radicalement la composition du lait qui restera un produit riche en AG saturés.

## 2.2 / Déterminisme génétique au cours de la lactation

L'héritabilité des différentes mesures augmente généralement avec le stade de lactation et elle est maximale au-delà de 200 jours. Ceci peut s'expliquer par la hausse du TB qui s'accompagne d'une meilleure précision de la mesure MIR et d'une diminution de la variance résiduelle. Les corrélations entre stades sont généralement plus élevées quand le caractère est exprimé par rapport à la matière grasse (tableau 3). Elles sont toujours positives et généralement très élevées, reflétant un déterminisme génétique commun, même entre saisons. Seule exception, le premier mois de lactation apparaît différent, avec des corrélations plus faibles avec le reste de la lactation. Ce résultat est sans doute à associer à la forte mobilisation corporelle en début

de lactation et plus limitée après : le taux d'AGMI est donc d'autant plus élevé que la femelle mobilise plus. Ces résultats permettent sans risque de raisonner à l'échelle de la lactation et non du contrôle.

### 2.3 / Corrélations génétiques

Chez les bovins, les corrélations entre la quantité de lait et les différents AG exprimés en % du gras sont faibles. Avec le TB, elles sont positives pour les saturés et négatives pour les insaturés (tableau 4). Ces estimations sont en accord avec Soyeurt *et al* (2007). Quand les AG sont exprimés en % du lait, ces corrélations

sont plus marquées. Elles sont toutes positives avec le TB, très élevées (> 0,8) pour les saturés et modérées (0,35 à 0,6) pour les insaturés, tandis qu'elles sont toutes modérées et négatives avec la quantité de lait. On mesure ici l'effet du mode d'expression : les vaches plus productives produisent un lait globalement moins riche en matière grasse, donc moins riche globalement en AGS et AGI, mais avec un rapport AGI/AGS plus élevé.

Le tableau 4 présente les gammes de corrélations génétiques estimées entre les différents AG selon les deux types d'expression. Elles sont globalement en

accord avec la bibliographie (Soyeurt *et al* 2007, Stoop *et al* 2008). Lorsque les AG sont exprimés en teneur dans le lait, toutes les corrélations génétiques entre AG sont positives, montrant que plus la teneur en gras augmente, plus la teneur de chaque composant augmente. Les saturés sont fortement corrélés entre eux (0,58 à 0,96) et plus modérément avec les insaturés (0,24 à 0,73).

Lorsque les AG sont exprimés % du gras, on observe une opposition quasi-totale entre saturés et mono-insaturés (en grande partie par construction mathématique, leur somme étant proche de 100%),

**Tableau 4.** Corrélations génétiques entre acides gras et avec la production laitière et le taux butyreux (moyenne des corrélations observées en Holstein, Normande et Montbéliarde ; pour les acides gras les variables sont exprimées en pourcentage du lait au-dessus de la diagonale et en pourcentage de la matière grasse en dessous).

	C4:0	C6:0	C8:0	C10:0	C14:0	C16:0	C18:0	C14:1 c9	C16:1 c9	C18:1	CLA C18:2 c9f11	C18:2	ω3	ω6	AGS	AGMI	AGPI	AGI	Lait	TB
C4:0		0,91	0,78	0,58	0,64	0,70	0,47	0,41	0,31	0,38	0,34	0,35	0,59	0,37	0,77	0,43	0,50	0,46	-0,33	0,83
C6:0	0,74		0,96	0,85	0,87	0,78	0,44	0,63	0,35	0,32	0,33	0,48	0,63	0,61	0,90	0,41	0,62	0,45	-0,31	0,90
C8:0	0,41	0,87		0,95	0,93	0,72	0,32	0,70	0,30	0,25	0,31	0,50	0,57	0,64	0,86	0,35	0,63	0,40	-0,27	0,83
C10:0	-0,01	0,59	0,87		0,96	0,68	0,31	0,68	0,32	0,23	0,24	0,53	0,55	0,71	0,84	0,33	0,64	0,37	-0,29	0,79
C14:0	-0,04	0,53	0,79	0,91		0,80	0,33	0,73	0,49	0,33	0,31	0,55	0,59	0,70	0,91	0,43	0,67	0,47	-0,32	0,89
C16:0	0,03	0,05	-0,03	-0,07	0,05		0,43	0,57	0,69	0,49	0,43	0,52	0,46	0,54	0,94	0,57	0,58	0,60	-0,35	0,94
C18:0	0,09	-0,15	-0,34	-0,32	-0,47	-0,38		-0,09	0,26	0,48	0,02	0,25	0,38	0,50	0,56	0,45	0,36	0,45	-0,42	0,50
C14:1 c9	-0,12	0,26	0,45	0,41	0,50	0,00	-0,71		0,48	0,30	0,47	0,64	0,24	0,66	0,62	0,44	0,58	0,47	-0,05	0,52
C16:1 c9	-0,45	-0,68	-0,65	-0,54	-0,39	0,06	-0,21	0,05		0,74	0,55	0,67	0,34	0,48	0,62	0,78	0,64	0,78	-0,34	0,66
C18:1	-0,15	-0,60	-0,60	-0,60	-0,60	-0,60	0,23	-0,22	0,44		0,57	0,73	0,45	0,58	0,49	0,99	0,71	0,98	-0,53	0,68
CLA C18:2 c9f11	0,01	-0,22	-0,18	-0,29	-0,26	-0,30	-0,24	0,15	0,30	0,57		0,73	0,34	0,35	0,38	0,60	0,65	0,62	-0,17	0,35
C18:2	-0,18	-0,33	-0,23	-0,22	-0,26	-0,68	-0,02	0,14	0,31	0,75	0,74		0,48	0,81	0,58	0,78	0,92	0,81	-0,37	0,62
ω3	0,25	0,15	0,09	0,07	0,13	-0,42	0,10	-0,25	-0,10	0,30	0,31	0,36		0,44	0,56	0,46	0,75	0,49	-0,38	0,66
ω6	-0,34	-0,21	-0,06	0,07	-0,03	-0,77	0,21	0,21	0,00	0,49	0,25	0,75	0,20		0,69	0,64	0,79	0,68	-0,32	0,63
AGS	0,23	0,60	0,57	0,55	0,53	0,58	-0,13	0,08	-0,45	-0,97	-0,60	-0,76	-0,30	-0,54		0,57	0,68	0,60	-0,39	0,98
AGMI	-0,19	-0,61	-0,59	-0,58	-0,57	-0,57	0,13	-0,11	0,49	0,99	0,58	0,76	0,25	0,49	-0,98		0,75	1,00	-0,52	0,76
AGPI	-0,08	-0,22	-0,15	-0,13	-0,14	-0,71	0,02	0,00	0,22	0,71	0,69	0,93	0,65	0,67	-0,72	0,69		0,79	-0,45	0,76
AGI	-0,18	-0,58	-0,55	-0,55	-0,53	-0,60	0,12	-0,09	0,47	0,99	0,61	0,79	0,30	0,53	-0,98	1,00	0,74		-0,53	0,78
Lait	0,06	0,04	0,10	0,07	0,16	-0,06	-0,10	0,00	-0,12	0,03	0,08	0,16	0,03	0,12	-0,04	0,03	0,17	0,05		-0,45
TB	-0,08	0,23	0,23	0,24	0,19	0,53	-0,16	0,16	-0,13	-0,65	-0,49	-0,70	-0,41	-0,49	0,59	-0,62	-0,74	-0,64		

**Tableau 5.** Corrélations génétiques entre acides gras et avec la production laitière et le taux butyreux en race caprine Alpine (pour les acides gras les variables sont exprimées en pourcentage du lait au-dessus de la diagonale et en pourcentage de la matière grasse en dessous).

	C4:0	C6:0	C8:0	C10:0	c12:0	C14:0	C16:0	C18:0	C18:1	CLA C18:2 c9f11	ω3	ω6	AGS	AGMI	AGPI	AGI	Lait	TB
C4:0		0,66	0,63	0,37	0,20	0,20	0,50	0,62	0,31	0,36	0,73	0,52	0,13	0,06	0,32	0,13	-0,21	0,83
C6:0	0,39		0,89	0,69	0,67	0,41	0,75	0,84	0,74	0,56	0,81	0,73	0,64	0,84	0,52	0,64	-0,31	0,90
C8:0	0,13	0,15		0,68	0,80	0,84	0,89	0,85	0,31	0,69	0,77	0,43	0,56	0,67	0,29	0,56	-0,37	0,83
C10:0	-0,42	0,34	0,60		0,86	0,48	0,85	0,26	0,31	0,62	0,70	0,77	0,85	0,79	0,47	0,85	-0,31	0,79
C12:0	-0,26	0,12	0,33	0,79		0,84	0,74	0,42	0,54	0,72	0,70	0,79	0,80	0,70	0,65	0,80	-0,27	0,79
C14:0	-0,64	-0,26	-0,09	0,57	0,69		0,69	0,18	0,70	0,65	0,65	0,73	0,44	0,71	0,61	0,44	-0,17	0,89
C16:0	-0,15	-0,12	-0,31	-0,17	-0,38	0,19		0,53	0,84	0,53	0,74	0,73	0,59	0,83	0,59	0,59	-0,23	0,94
C18:0	0,57	0,64	0,24	-0,26	-0,51	-0,43	0,01		0,45	0,31	0,74	0,59	0,70	0,70	0,55	0,70	-0,24	0,50
C18:1	0,50	-0,37	-0,31	-0,78	-0,63	-0,42	-0,48	0,81		0,70	0,54	0,86	0,71	0,79	0,30	0,71	-0,38	0,68
CLA C18:2 c9f11	-0,08	-0,60	-0,06	0,02	0,36	0,24	-0,40	-0,64	0,25		0,33	0,67	0,48	0,61	0,60	0,48	-0,43	0,35
ω3	0,59	0,14	0,03	-0,10	0,14	0,10	-0,22	0,32	0,04	0,08		0,60	0,43	0,70	0,73	0,43	-0,31	0,66
ω6	-0,09	0,44	-0,25	-0,21	0,38	-0,03	-0,58	-0,25	0,40	0,65	0,45		0,80	0,66	0,84	0,80	-0,45	0,63
AGS	0,26	0,46	0,33	0,24	0,16	0,16	0,20	0,02	-0,48	-0,32	-0,08	-0,40		0,93	0,62	0,81	-0,31	0,63
AGMI	0,09	0,16	-0,03	-0,34	-0,47	-0,56	-0,08	0,46	0,61	-0,13	0,03	0,10	-0,34		0,64	0,93	-0,33	0,88
AGPI	0,04	-0,27	-0,30	-0,23	0,25	0,10	-0,26	0,05	0,26	0,29	0,59	0,75	-0,39	0,16		0,62	-0,26	0,67
AGI	0,00	-0,36	-0,36	-0,20	-0,19	-0,11	-0,13	0,01	0,68	0,34	0,07	0,44	-0,51	0,54	0,43		-0,34	0,79
Lait	0,21	0,07	-0,05	-0,03	-0,04	0,17	0,12	0,16	-0,01	-0,07	0,24	-0,01	0,10	-0,05	0,13	-0,04		-0,38
TB	-0,54	0,30	0,27	0,17	0,12	-0,17	0,15	0,01	0,00	-0,40	-0,37	-0,56	0,16	0,21	-0,49	-0,05		

**Tableau 6.** Estimations de l'héritabilité des teneurs en protéines mesurées par contrôle laitier, selon l'espèce, la race et le mode d'expression.

Espèce	Bovine						Ovine			
	g/100 g de lait			g/100 g de matière protéique			g/100 g de lait		g/100 g de matière protéique	
Mode d'expression	Ho	Mo	No	Ho	Mo	No	MTR	L	MTR	L
Race	Ho	Mo	No	Ho	Mo	No	MTR	L	MTR	L
TP	0,20	0,44	0,35	-	-	-	0,38	0,39	-	-
Caséine alpha-s1	0,22	0,46	0,36	0,42	0,44	0,39	0,40	0,38	0,33	0,37
Caséine alpha-s2	0,21	0,41	0,32	0,21	0,33	0,23	0,29	0,43	0,33	0,44
Caséine bêta	0,20	0,44	0,35	0,29	0,37	0,32	0,31	0,34	0,20	0,31
Caséine kappa	0,24	0,44	0,36	0,39	0,48	0,44	0,25	0,29	0,20	0,32
Total caséines	0,21	0,44	0,35	0,36	0,47	0,30	0,40	0,39	0,17	0,22
Alpha lactalbumine	0,21	0,43	0,33	0,32	0,47	0,40	0,03	0,06	0,30	0,37
Bêta lactoglobuline	0,44	0,56	0,50	0,55	0,62	0,58	0,29	0,36	0,39	0,47
Protéines sériques	0,30	0,43	0,36	0,37	0,41	0,40	0,28	0,38	0,31	0,46

Ho = Holstein ; Mo = Montbéliarde ; No = Normande ; MTR = Manech à Tête Rousse ; L = Lacaune.

mais aussi entre saturés et polyinsaturés (- 0,7). Ces corrélations négatives se retrouvent à un niveau moins élevé mais toujours marqué entre les composants individuels. Les  $\omega 3$  constituent une exception avec une corrélation négative forte avec le C16:0, mais une corrélation proche de 0 ou légèrement positive avec les AG saturés de chaîne plus courte. Les corrélations entre saturés présentent un pattern particulier, résultant sans doute des flux de molécules. Ainsi, les corrélations sont fortes entre AG substrats et produits, mais décroissent rapidement quand la différence de nombre d'atomes de carbone augmente. Par ailleurs, les corrélations entre le C16:0 et les autres saturés (y compris C14:0) est très faible. Ceci s'explique sans doute par le fait que le C16:0 est le produit d'élongation terminal et résulte d'une accumulation, tandis que les autres sont des substrats dont la concentration résulte d'un flux.

En fonction de ces résultats, on déduit qu'une sélection contre le TB diminuerait le taux de saturés, très fortement lié au TB, et augmenterait donc le taux relatif d'AGI dans le gras, mais sans changement important de leur teneur absolue dans le lait.

Chez les ovins, les corrélations entre grandes familles d'AG (AGS, AGMI, AGI) sont analogues à celles des bovins.

Chez les caprins (Maroteau *et al* 2014), les corrélations présentent plusieurs différences notables avec les autres espèces (tableau 5). Exprimés en % du lait,

comme pour les bovins, les 18 AG ou combinaisons d'AG analysés sont fortement et positivement corrélés entre eux. Ils sont tous positivement liés au TB et négativement corrélés avec la quantité de lait. En revanche, exprimés en % du gras, ils présentent tous une corrélation proche de 0 avec le lait. Avec le TB, la corrélation est positive avec les saturés de C6:0 à C12:0 et surtout avec le C16:0 mais négative pour le C14:0 et surtout le C4:0. Elle est légèrement positive pour les AGMI et clairement négative avec les polyinsaturés. Le pattern de corrélation entre saturés est différent de celui des bovins, sans doute lié à un arrêt d'élongation plus précoce. On observe toujours des corrélations plus fortes entre AG substrats et produits, mais les AG de C6:0 à C12:0 sont positivement corrélés entre eux et généralement négativement corrélés aux autres. La corrélation entre saturés et insaturés est négative (- 0,5) et du même niveau pour les mono- et polyinsaturés.

### 3 / Paramètres génétiques des protéines chez les bovins et les ovins

#### 3.1 / Héritabilité et variabilité génétique

Le tableau 6 présente les estimations d'héritabilité des mesures de concentration en protéines à chaque contrôle pour les races bovines et ovines. Pour l'ensemble des caractères, les estimations

d'héritabilité ( $h^2$ ) sont modérées à fortes (0,17 à 0,62) et du même ordre que dans Schopen *et al* (2009). Seule exception notable chez les ovins, l'héritabilité de l'alpha-lactalbumine apparaît très faible, reflétant la mauvaise précision de l'équation de prédiction. En teneur dans le lait, les coefficients sont variables entre races bovines, tandis que les estimations sont plus homogènes quand les caractères sont exprimés en % des protéines. Suivant la race, les estimations d'héritabilité sont du même ordre ou plus élevées pour les caractères exprimés en % des protéines qu'en teneur dans le lait.

Le tableau 7 présente les estimations d'héritabilité en moyenne par lactation dans les 3 races bovines. Les estimations sont donc par construction plus élevées que par contrôle. Elles varient de 0,27 à 0,86. Elles sont sensiblement plus élevées en race Montbéliarde qu'en Holstein et en Normande, ce qui reflète le nombre de contrôles disponibles par lactation plus élevé dans cette race. Ces résultats montrent ainsi que l'augmentation du nombre de contrôles, l'existence d'un déterminisme génétique unique tout au long de la lactation et une corrélation non génétique assez faible entre contrôles permettent de réduire fortement la variance résiduelle. Les valeurs extrêmement élevées (0,61 à 0,86) indiquent que l'essentiel de la variabilité de la composition relative en protéines du lait est d'origine génétique. Quelle que soit la race ou l'unité d'expression du caractère, le caractère le plus héritable est la teneur en bêta-lactoglobuline.



**Tableau 7.** Estimations de l'héritabilité des concentrations moyennes par lactation en différentes protéines dans les trois races bovines exprimées en g/100g de lait (% lait) et en g/100g de protéines (% prot).

	Race	TP	$\alpha$ Lac.	$\beta$ Lg	Caséine				Protéolyse
					$\alpha$ s1	$\alpha$ s2	$\beta$	$\kappa$	
%lait	Mo	0,57	0,57	0,86	0,54	0,54	0,66	0,48	-
	No	0,41	0,42	0,61	0,42	0,38	0,43	0,42	-
	Ho	0,27	0,31	0,61	0,30	0,29	0,27	0,32	-
%prot.	Mo	-	0,72	0,79	0,67	0,58	0,42	0,61	0,80
	No	-	0,53	0,72	0,57	0,25	0,39	0,55	0,39
	Ho	-	0,44	0,71	0,53	0,31	0,39	0,54	0,49

Mo = Montbéliarde ; No = Normande ; Ho = Holstein.

$\alpha$  Lac. = alpha lactalbumine ;  $\beta$  Lg = bêta lactoglobuline.

Les coefficients de variation génétique ( $CV_g$ ) sont équivalents pour les 3 races bovines et varient de 2,9 à 8,6% pour les teneurs dans le lait (à comparer à 4% pour le TP) et de 0,5 à 7,2% pour les teneurs en pourcentage des protéines. Chez les ovins, ces valeurs sont assez comparables, les coefficients de variation variant de 1 à 7% en % de la matière protéique et de 4 à 7% en % du lait. Les taux d'alpha-lactalbumine ont les  $CV_g$  les plus élevés (chez les bovins) alors que les taux des caséines alpha-s1 et bêta, qui sont les protéines les plus abondantes, ont les  $CV_g$  les plus faibles.

### 3.2 / Corrélations génétiques entre protéines

Les estimations des corrélations génétiques ( $r_g$ ) sont cohérentes entre races bovines (tableau 8). Le TP est corrélé plus fortement avec les caséines qu'avec les protéines sériques. Les corrélations génétiques sont globalement beaucoup plus fortes entre les teneurs dans le lait qu'entre teneurs exprimées en % de protéines. Les taux de protéines exprimés en % du lait sont toujours corrélés positivement, les valeurs les plus fortes sont

obtenues entre les différentes caséines (0,83 à 0,98) et les valeurs les plus modérées entre l'alpha-lactalbumine et la bêta-lactoglobuline (- 0,01 à 0,27). Les corrélations entre protéines sériques et caséines sont plus fortes avec l'alpha-lactalbumine (0,68 à 0,80) qu'avec la bêta-lactoglobuline (0,07 à 0,59).

Au contraire, les taux exprimés en pourcentage de protéines présentent des corrélations génétiques nulles ou négatives dans la plupart des cas. La bêta-lactoglobuline est négativement corrélée avec toutes les caséines, et très fortement avec la caséine kappa (- 0,55 à - 0,75). Entre les caséines, il apparaît un antagonisme entre les caséines bêta d'une part, et les caséines alpha-s1 et alpha-s2 d'autre part (- 0,26 à - 0,51). Cet antagonisme est quelque peu « artéfactuel » compte tenu du mode d'expression et de la part de ces protéines dans le total.

Chez les ovins, les résultats sont homogènes et très comparables à ceux des bovins. On note cependant un comportement spécifique de la caséine alpha-s2, en corrélation faible avec les autres caséines (0,16 à 0,45) exprimée en teneur dans le lait, et en corrélation négative exprimée en proportion des protéines totales.

**Tableau 8.** Estimation des corrélations génétiques entre les teneurs en protéines et avec le Taux Protéique (TP) du lait pour les trois races bovines Montbéliarde, Normande et Holstein et la race ovine Lacaune (les teneurs en protéines sont exprimées en % du lait au-dessus de la diagonale ou en % des protéines au-dessous de la diagonale).

		$\alpha$ Lac.	$\beta$ Lg	CAS $\alpha$ s1	CAS $\alpha$ s2	CAS $\beta$	CAS $\kappa$
TP	Montbéliarde	0,72	0,52	0,98	0,97	0,99	0,85
	Normande	0,70	0,61	0,99	0,99	0,98	0,90
	Holstein	0,73	0,45	0,98	0,98	0,98	0,87
	Lacaune	- 0,03	0,93	0,95	0,41	0,77	0,93
$\alpha$ Lac.	Montbéliarde		0,18	0,69	0,72	0,70	0,80
	Normande		0,27	0,71	0,69	0,68	0,74
	Holstein		- 0,01	0,72	0,71	0,75	0,80
	Lacaune		- 0,02	- 0,08	0,05	- 0,16	- 0,06
$\beta$ Lg	Montbéliarde	- 0,27		0,51	0,46	0,47	0,10
	Normande	- 0,33		0,59	0,57	0,58	0,31
	Holstein	- 0,52		0,39	0,39	0,39	0,07
	Lacaune	- 0,01		0,88	0,33	0,72	0,85
CAS $\alpha$ s1	Montbéliarde	0,01	- 0,13		0,94	0,96	0,84
	Normande	0,11	- 0,10		0,98	0,95	0,89
	Holstein	- 0,01	- 0,19		0,96	0,94	0,87
	Lacaune	- 0,28	0,10		0,33	0,82	0,94
CAS $\alpha$ s2	Montbéliarde	0,05	- 0,08	- 0,33		0,95	0,88
	Normande	- 0,04	- 0,14	- 0,07		0,96	0,92
	Holstein	- 0,09	- 0,22	0,04		0,95	0,88
	Lacaune	0,37	0,00	- 0,31		0,16	0,31
CAS $\beta$	Montbéliarde	0,06	- 0,33	- 0,26	- 0,37		0,83
	Normande	0,01	- 0,25	- 0,44	- 0,51		0,85
	Holstein	0,25	- 0,20	- 0,48	- 0,37		0,84
	Lacaune	- 0,01	- 0,09	0,43	- 0,35		0,79
CAS $\kappa$	Montbéliarde	0,45	- 0,75	0,02	0,38	- 0,11	
	Normande	0,39	- 0,55	0,02	0,49	- 0,36	
	Holstein	0,46	- 0,66	0,14	0,36	- 0,31	
	Lacaune	0,07	- 0,28	0,47	- 0,12	0,25	

$\alpha$  Lac. = alpha lactalbumine ;  $\beta$  Lg = bêta lactoglobuline.

## 4 / Détection de QTL d'acides gras

### 4.1 / Résultats généraux chez les bovins

Les analyses réalisées permettent la détection d'un grand nombre de QTL, pour tous les caractères. Ainsi sur l'ensemble des trois races bovines, 29 QTL sont détectés pour le taux d'AGS, 33 pour le taux d'AGMI, 20 pour le taux d'AGPI. Pour les saturés, le nombre tend à diminuer avec la longueur de la chaîne, de 36 QTL pour le C4:0 à 13 QTL pour le C14:0, avec une remontée à 20 pour le C16:0, qui est la fin de la chaîne d'élongation. Bien que moins bien estimés au travers des spectres MIR (car à une teneur plus faible), les AG insaturés présentent un nombre de QTL analogue, voire supérieur : 24 pour les C18:1, 27, 28 et 28 pour les C14:1c9, C16:1c9 et C18:1c9, 17 pour les CLA, 24 pour les  $\omega$ 3, 19 pour les  $\omega$ 6, 12 pour le C18:2c9c12.

Les résultats sont souvent très significatifs, les seuils de LRT étant élevés. Même si la méthodologie utilisée (analyse LDLA mono-QTL) n'est pas la plus précise en termes de localisation, les pics sont souvent assez fins et pointent vers des régions réduites, permettant de bonnes inférences sur les gènes candidats possibles.

La figure 1 présente les « *Manhattan Plot* » de 8 caractères d'AG en race Holstein. Elle montre ainsi les effets majeurs des chromosomes 5, 14 et 19 sur de nombreux caractères, ainsi que des chromosomes 1, 6, 11, 13, 17, 20, 25, 26 et 27 pour certains d'entre eux. Les co-localisations de QTL sont fréquentes entre les différents saturés d'une part, et les différents insaturés d'autre part, reflétant les voies métaboliques communes dans chaque grande catégorie. En revanche, les régions sont fréquemment différentes entre saturés et insaturés. Concernant les saturés, les régions impliquées évoluent graduellement avec la longueur de la chaîne.

### 4.2 / Confirmation de régions connues, découverte de nouvelles régions majeures

Parmi les nombreuses régions mises en évidence, on retrouve logiquement celles déjà connues comme des régions très significatives. Bouwman *et al* (2011) décrivent trois régions majeures dans la population Holstein hollandaise sur les chromosomes 5, 14 et 19. Ces trois régions sont retrouvées dans notre étude.

L'extrémité centromérique du chromosome 14 porte le gène DGAT1 qui code

pour l'enzyme branchant le 3<sup>ème</sup> AG sur les triglycérides et qui a un effet majeur sur le TB (Grisart *et al* 2002), retrouvé dans les trois races françaises (Guillaume, 2009). On montre ici que l'allèle favorable au TB est aussi favorable aux AG saturés et tout particulièrement au C16:0. Son effet sur les AG saturés courts est moins net et moins constant entre races. L'analyse étant réalisée sur les variables exprimées en % de gras, ce gène affecte un grand nombre de caractères. En race Holstein, un SNP est en déséquilibre presque total avec la mutation K232A de DGAT1. Après correction pour ce SNP, l'effet de la région n'est plus significatif. Mais en l'absence de typage sur la mutation causale, il est difficile d'exclure la possibilité d'un autre polymorphisme, en particulier dans les races autres que la Holstein.

De même, plusieurs régions sont mises en évidence sur le chromosome 19. Plusieurs gènes candidats comme GH, ACLY, STAT5 ou surtout FASN sont souvent cités, mais aucun n'a été validé jusqu'à présent. Le chromosome 19 a un rôle majeur sur les quantités produites (lait et matière), mais son effet sur le TB est plus faible, et uniquement en Holstein (Guillaume 2009). Dans cette étude, ce chromosome joue un rôle majeur sur les AG saturés C4:0 à C18:0, sur l'index de désaturation du C16:0 et sur le taux d'AGMI.

L'étude *PhénoFinlait* met en évidence deux régions du chromosome 5 qui semblent très importantes pour le déterminisme de nombreux AG dans le lait et que nous présentons plus en détail ensuite.

La région du gène codant pour la bêta-lactoglobuline (environ 103,3 Mb sur le chromosome 11) semble affecter plusieurs caractères, les taux d' $\omega$ 3 et  $\omega$ 6 ainsi que le taux de certains AG saturés courts. La nature de cet effet n'est pas connue, mais suggère que la bêta-lactoglobuline a un rôle fonctionnel dans la cellule mammaire.

Le chromosome 27, déjà décrit par Bouwman *et al* (2011) est également retrouvé ici, affectant tout particulièrement le taux de C16:0 et la proportion de saturés, mono et polyinsaturés. Cette région est connue pour affecter le TB en race Holstein, avec le gène AGPAT6 comme candidat fort. Dans cette étude, le QTL est partagé par les trois races. AGPAT6 est une enzyme essentielle pour la synthèse de triacylglycerol dans la mamelle, un précurseur des lipides.

Bouwman *et al* (2011) décrivent aussi des régions des chromosomes 13 et 17 que nous retrouvons dans notre étude avec des effets très forts sur la proportion d'AG courts C4:0 à C10:0.

Une région du chromosome 26, portant le gène SCD qui code pour la delta9-désaturase, l'enzyme de désaturation des AG exprimée dans la mamelle, était fortement attendue pour les caractères de saturation et les taux d'insaturés. Bien que Bouwman *et al* (2011) aient déjà montré l'effet majeur de ce gène sur le taux d'insaturés en race Holstein, ce QTL est observé, avec un pic très net pour différents AG insaturés et index de désaturation, mais surtout en race Normande.

### 4.3 / Co-localisation entre caractères

Comme indiqué précédemment, les co-localisations sont fréquentes entre caractères. Ainsi, par exemple, le chromosome 5 porte des QTL nombreux et très significatifs (figure 2). Ce chromosome porte deux régions proches mais distinctes, impliquées dans le déterminisme des AG, l'une autour de 92-95 Mégabases (Mb) partagée entre les trois races, l'autre autour de 112-115 Mb partagée entre la Normande et la Montbéliarde mais apparemment pas par la Holstein. Dans la première région, le gène PIK3C2G, une kinase mal connue mais essentielle dans le maintien cellulaire, est un gène candidat très attractif qui va faire l'objet d'investigations complémentaires. La seconde région est particulièrement riche en gènes potentiellement impliqués, mais il ne ressort pas de gène candidat évident.

Ces co-localisations fortes entre caractères s'expliquent très probablement par les voies métaboliques communes et illustrent les corrélations génétiques élevées entre certains caractères. Elles montrent aussi que le nombre élevé de caractères n'est sans doute pas un handicap très important en sélection, du fait de leur forte redondance biologique. En effet, il est sans doute possible de définir quelques indicateurs peu nombreux sur lesquels la sélection pourrait s'appliquer, dès lors que l'objectif recherché est clairement défini.

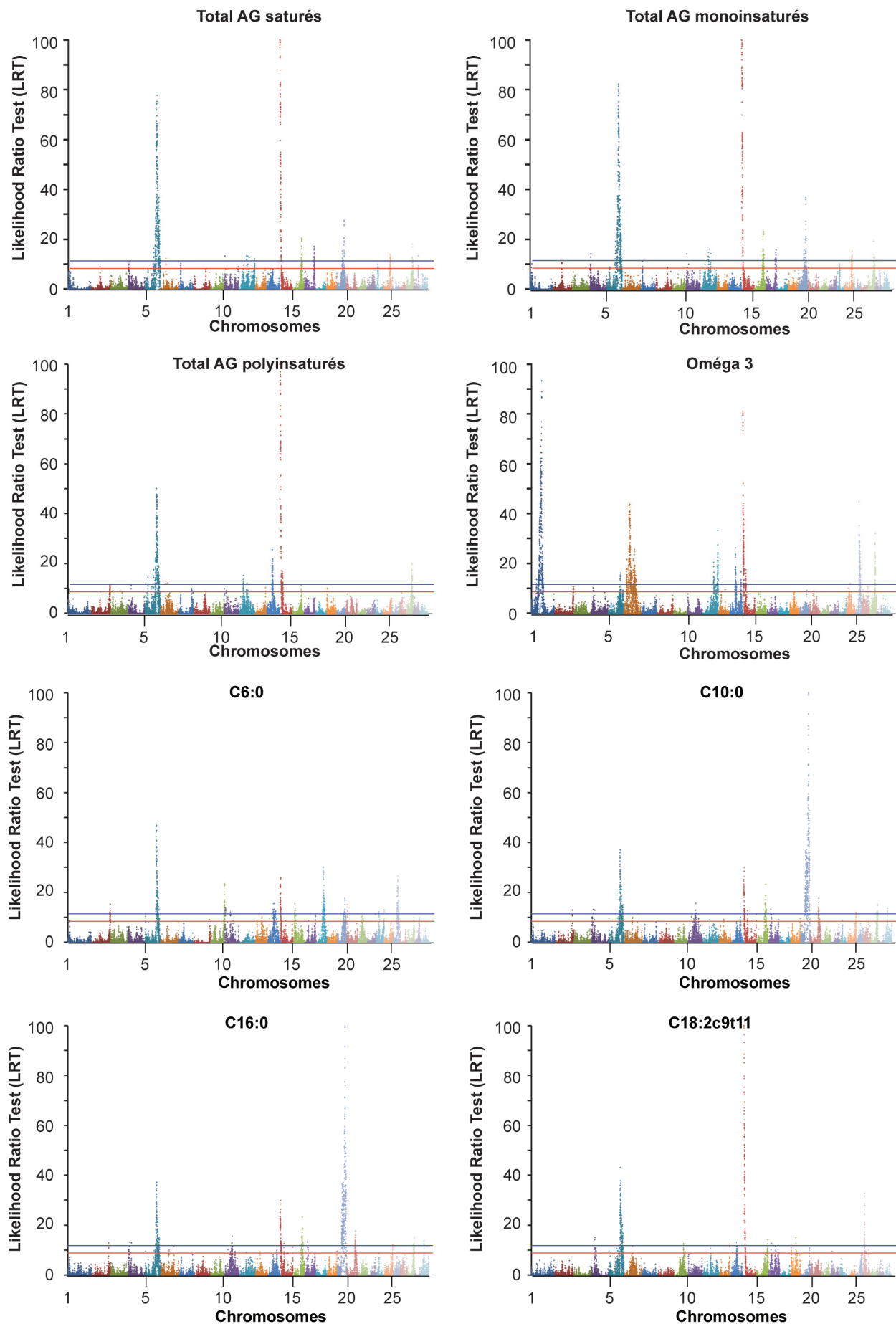
### 4.4 / Co-localisation entre races

On observe fréquemment, dans des dispositifs de détection de QTL dans des populations différentes, une proportion assez faible de QTL communs. Ceci est dû généralement à plusieurs raisons : une puissance de détection insuffisante de sorte que des QTL communs ne sont pas détectés dans toutes les races ; des échantillons ne permettant pas de mettre en évidence des variants rares ; mais aussi des QTL fixés dans certaines populations.

Dans le cas présent, on observe au contraire un taux de co-localisation relativement élevé, comme le montre la



**Figure 1.** Localisation des QTL à l'échelle du génome pour les différents AG en race Holstein.



**Figure 2.** Co-localisation de QTL sur le chromosome 5 dans les régions [91-97 Mb] et [107-112 Mb].

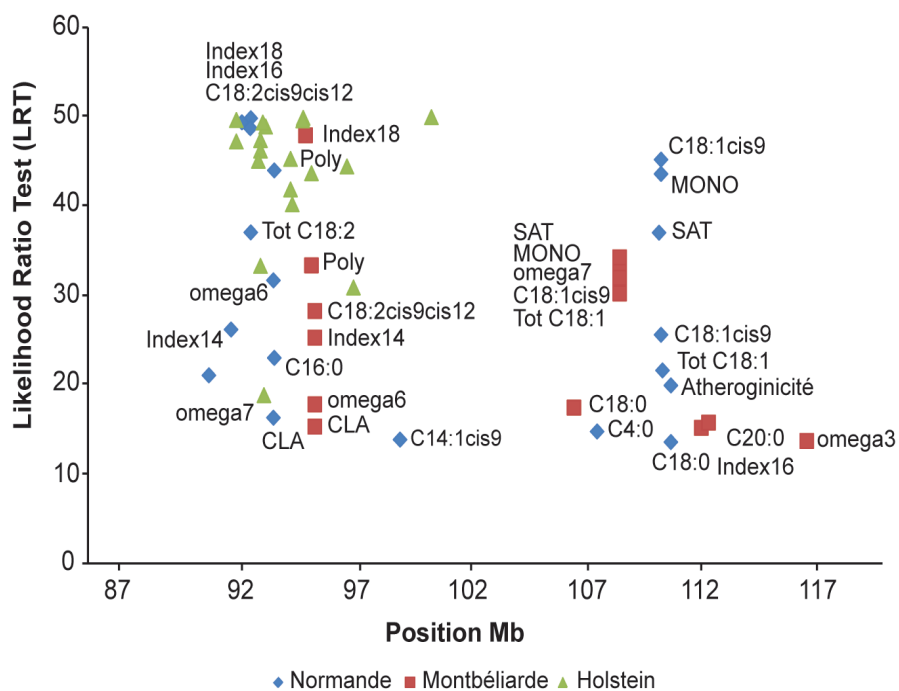
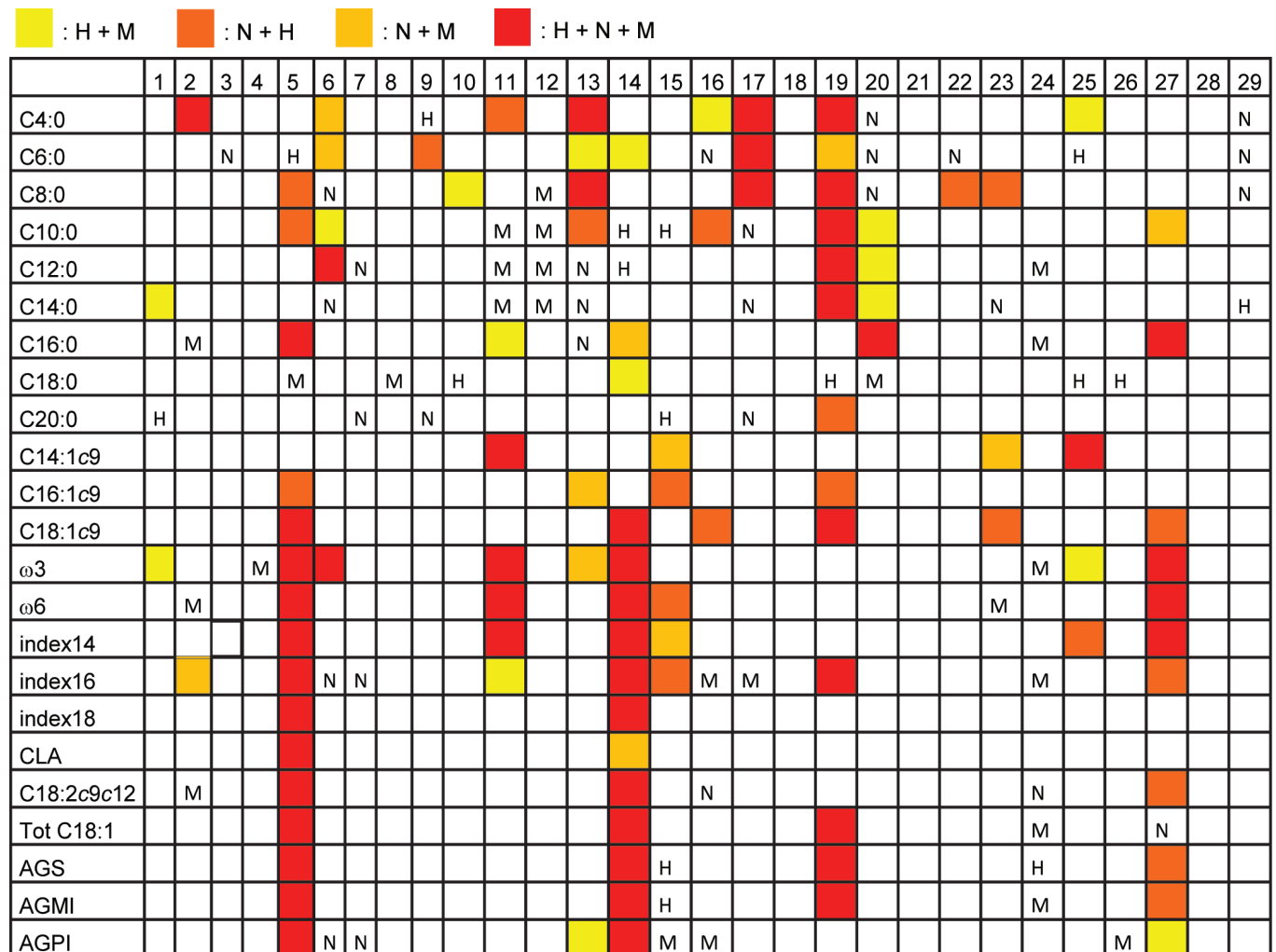


figure 3. Dans cette étude, des QTL sont considérés comme communs si leurs pics sont distants de moins d'une Mb. Ainsi, des QTL semblent communs à deux races, voire aux trois races, pour les chromosomes 5, 11, 13, 14, 16, 17, 19 et 27, et à chaque fois pour plusieurs caractères. Plusieurs autres régions peuvent être ajoutées si on ne considère qu'un ou deux caractères seulement.

Il n'en demeure pas moins qu'un nombre important de résultats restent spécifiques d'une race. Globalement, la moitié des QTL détectés ne sont pas partagés entre races.

Cette relative homogénéité de résultats est très favorable à la fois pour les perspectives de sélection, mais aussi pour la compréhension des mécanismes et pour la recherche des mutations causales. En effet, cette convergence entre races, sous réserve qu'elle soit due aux mêmes mutations dans les mêmes gènes, permet d'augmenter très sensiblement la précision de cartographie et donc de faciliter leur découverte. Elle montre aussi que les mutations sous-jacentes préexistaient largement à la constitution des races.

**Figure 3.** Bilan des co-localisations des QTL d'AG (LRT>20) entre races bovines sur chacun des 29 chromosomes. H = Holstein ; N = Normande ; M = Montbéliarde.



#### 4.5 / Détection de QTL chez les caprins

Un total de 6 (19), 10 (186), 5 (231) QTL a été détecté en analyse de liaison (en analyse d'association), respectivement, pour les caractères laitiers, les AG en g/100g de lait et les AG en g/100g de MG (Maroteau *et al* 2013). Le QTL le plus significatif est localisé sur le chromosome 14 et affecte le TB. Cette région comporte le gène candidat positionnel et fonctionnel, DGAT1. Ce QTL affecte également tous les AG saturés. A noter que pour le TP, il ressort également un QTL majeur correspondant au cluster des caséines sur le chromosome 6.

Tout comme chez les bovins, il existe une forte co-localisation des QTL entre les différents caractères, et en particulier entre AG. Cette co-localisation coïncide avec les corrélations génétiques élevées entre caractères et s'explique très probablement par l'existence de voies métaboliques communes.

En revanche, les co-localisations de QTL entre races (Alpine *versus* Saanen) sont peu nombreuses, en dehors des chromosomes 6, 8, 14, et 19. Ce résultat peut illustrer un manque de puissance du dispositif, mais aussi des différences de fréquences alléliques des QTL entre races.

#### 4.6 / Détection de QTL chez les ovins

Selon les méthodes (analyse de liaison ou d'association), 9 à 83 QTL sont détectés en race Lacaune pour les caractères en % de la matière grasse dont 11 sont très significatifs, et 8 à 73 en race Manech Tête Rousse dont 1 seul est très significatif. Les QTL les plus significatifs se trouvent sur les chromosomes 2, 4, 8, 13, 22, et 26 et affectent les AGS comme les AGI. Deux QTL sur les chromosomes 5 et 22 (proche du gène SCD) n'affectent que les AGPI. L'importance du gène SCD avait déjà été mise en évidence par García-Fernández *et al* 2010). Les régions proches des gènes candidats les plus connus (ABCG2, DGAT1, FASN, GHR) présentent des QTL suggestifs. Le chromosome 11 qui porte FASN (situé aussi sur le chromosome 19 bovin) apparaissait comme le plus significatif dans l'étude de Carta *et al* (2008). La figure 4 présente un QTL du chromosome 4 en race Lacaune, affectant le taux d'AGPI, confirmé par les deux méthodes d'analyse.

### 5 / Détection de QTL de protéines laitères

#### 5.1 / Régions QTL détectées chez les bovins

Les QTL détectés présentent un niveau de significativité très élevé, généralement

Figure 4. Illustration d'un QTL du chromosome ovin 4 en race Lacaune affectant le taux d'AG polyinsaturés, détecté par analyse de liaison (LA) et d'association (LD).

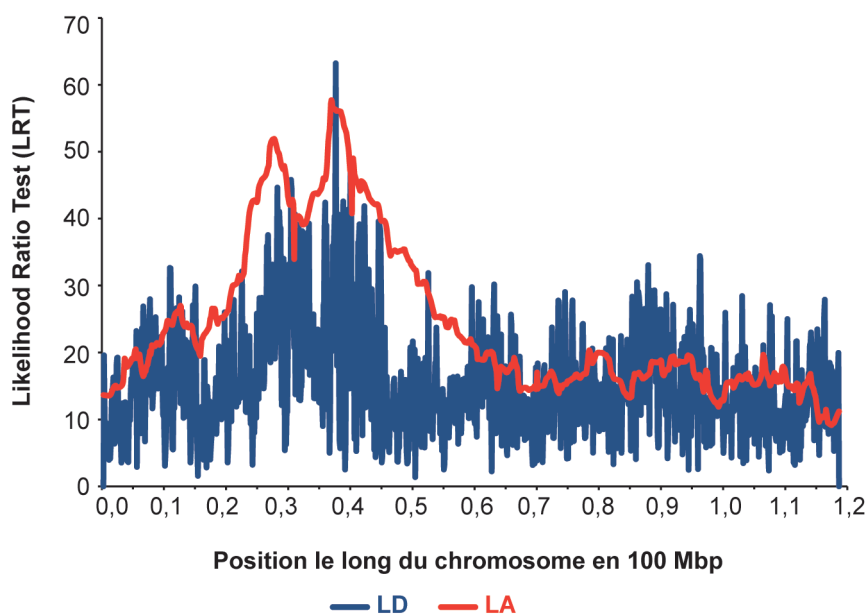
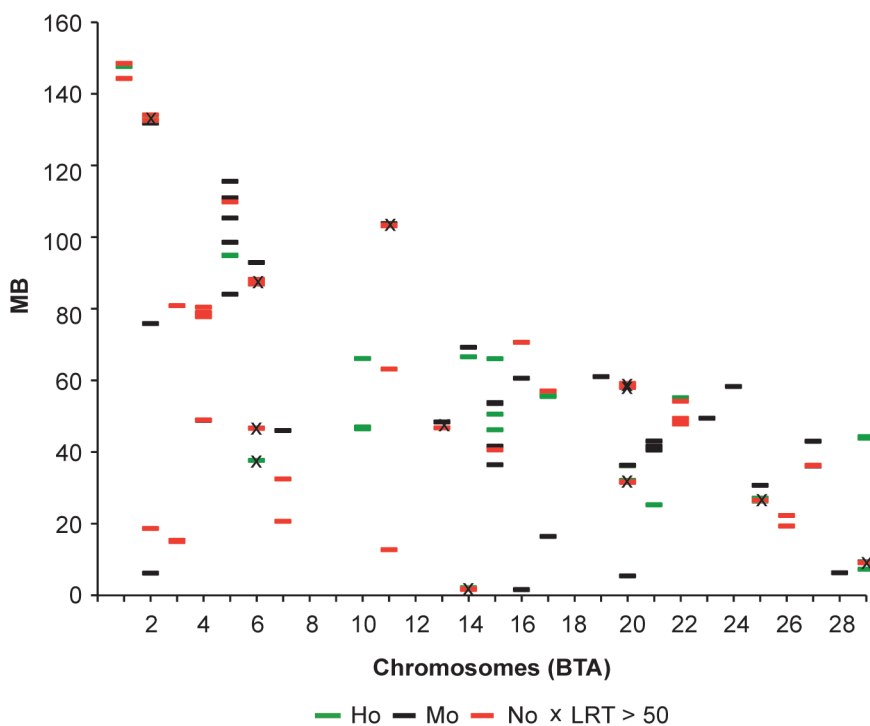


Figure 5. Positions des SNP présentant une valeur de LRT maximale  $\geq 20$  pour les QTL de protéines laitères dans les trois races Holstein (Ho), Montbéliarde (Mo) et Normande (No).



bien plus élevé que pour les AG. Ces résultats montrent que les prédictions MIR sont tout à fait suffisantes pour des applications de génétique, même si la précision des équations de prédiction reste modérée.

La figure 5 montre les résultats significatifs à très significatifs ( $LRT \geq 20$ ) sur l'ensemble du génome. Leur nombre varie de 67 à 81 selon la race, soit plus de 10 par race et caractère. A ce niveau de seuil, le nombre de résultats faux

positifs est inférieur à 1 par caractère. Dix de ces régions QTL sont très hautement significatives ( $LRT > 50$ ) et quatre d'entre elles qui présentent les effets les plus significatifs sont retrouvées dans les trois races. Certains de ces résultats sur les chromosomes 5, 6, 11 et 14 sont également décrits par Schopen *et al* (2011).

A ce stade, on peut conclure que les teneurs en protéines du lait sont soumises à des QTL à effet fort, en accord avec l'héritabilité élevée des caractères. Les



mutations sous-jacentes, sous réserve qu'elles soient les mêmes, semblent conservées entre races et sont donc sans doute anciennes.

## 5.2 / Gènes candidats

La figure 6 montre quatre exemples de QTL à effets très forts, conservés entre races.

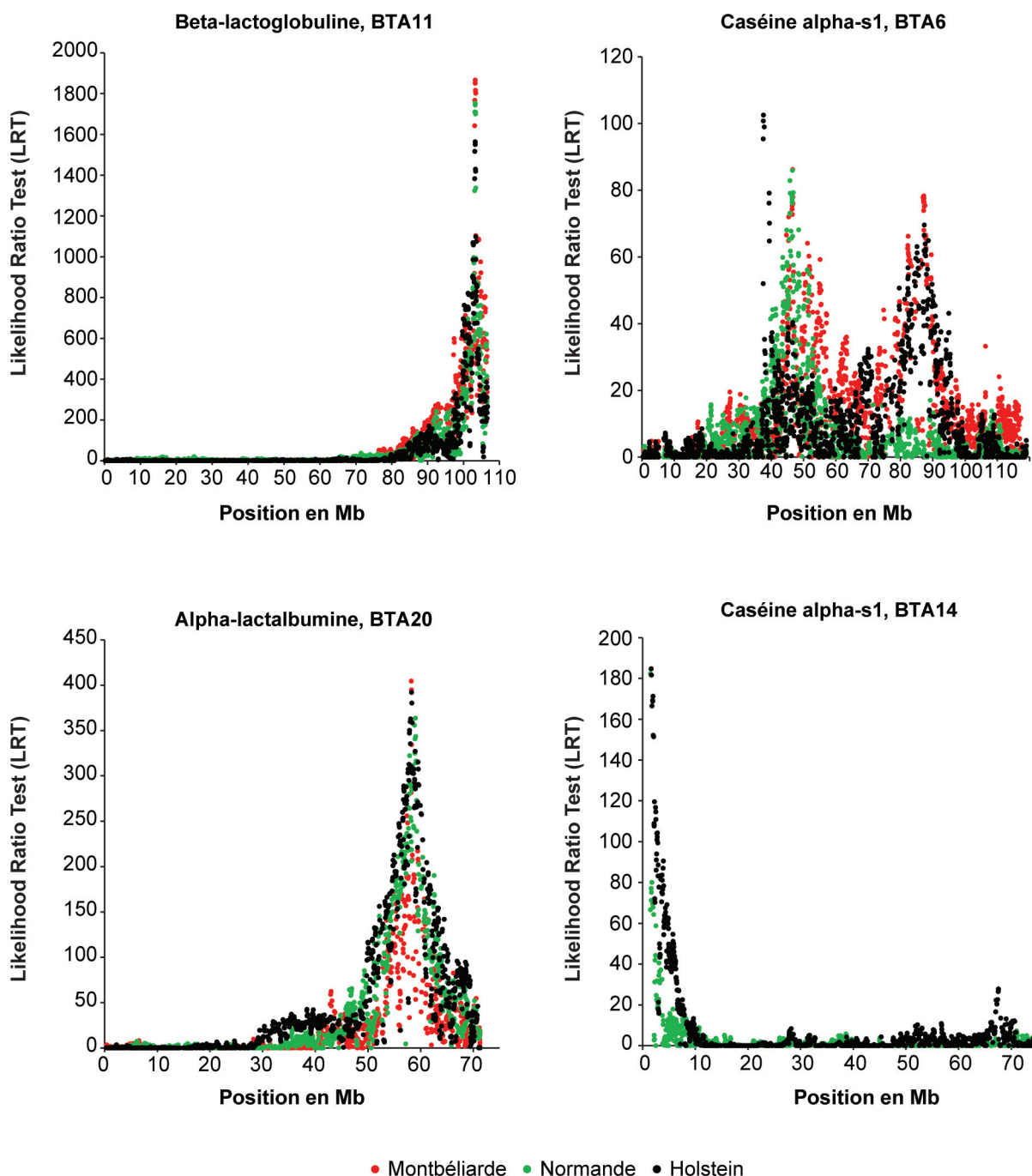
La région BTA6a (entre 37 et 46 Mb) affecte la caséine alpha-s1. Le SNP le plus significatif de cette région présente des caractéristiques un peu différentes

en race Holstein par rapport aux deux autres races : il se situe à environ 9 Mb en amont et son effet est plus important, il explique environ 30% de la variabilité génétique contre environ 7% en race Montbéliarde et 10% en race Normande. Il est probable que les mutations sous-jacentes soient différentes. Vers 37 Mb, deux gènes candidats sont attractifs : ABCG2 code pour un transporteur d'ATP membranaire et il est connu pour affecter la production laitière (Cohen-Zinder *et al* 2005) ; OPN code pour l'ostéopontine, impliquée dans la régulation du calcium. Vers 46 Mb, aucun gène candidat évident n'est identifié.

La région BTA6b (vers 87 Mb) a un effet sur de nombreux caractères surtout en races Montbéliarde et Normande, le caractère le plus affecté étant la caséine kappa dans les trois races (16 à 20% de la variance génétique en % dans le lait en Montbéliarde et en % des protéines en Normande et Holstein). Cette localisation coïncide parfaitement avec le cluster des quatre gènes des caséines et là où les mutations causales doivent être recherchées dans cette région d'environ 250 kb.

BTA11 (~103Mb) est la région QTL la plus significative avec un LRT qui

**Figure 6.** Illustration de quelques QTL majeurs sur les chromosomes 6, 11 et 20, retrouvés dans les différences races bovines.



atteint la valeur astronomique de 1866 en race Montbéliarde. Plusieurs caractères exprimés surtout en % des protéines sont affectés par cette région, les caséines alpha-s1, bêta et kappa. Mais le principal effet est observé sur la bêta-lactoglobuline (41 à 51% de la variance génétique selon la race). Les SNP présentant les résultats les plus significatifs dans cette région se situent à proximité immédiate du gène codant pour la bêta-lactoglobuline ( $\beta$  Lg), situé entre 103,301 et 103,306Mb. Le même SNP a un niveau de LRT très élevé dans les 3 races et se situe à 103,289Mb, soit un peu en amont du gène  $\beta$  Lg.

BTA20b (~58-59Mb) est la deuxième région QTL la plus significative, elle affecte l'alpha-lactalbumine exprimée en % du lait et surtout en % des protéines (expliquant 25 à 32% de la variance génétique selon la race). Aucun gène candidat n'est identifié à ce jour.

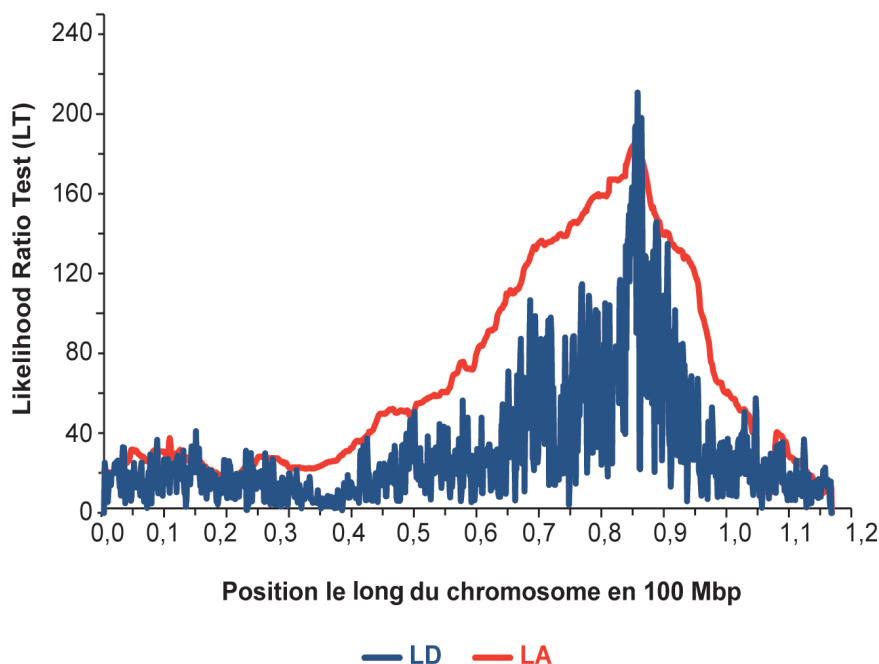
Une autre région se distingue également sur BTA14 (~1,7-1,8Mb). Elle est très significative dans les races Normande et Holstein avec des effets sur de nombreux caractères exprimés essentiellement en % du lait (TP, alpha-lactalbumine et toutes les caséines) et significative en race Montbéliarde sur la caséine alpha-s2 en % des protéines. Les effets de cette région sont toutefois un peu plus modérés que les effets des 4 régions décrites plus haut (3 à 11% de la variance génétique selon le caractère et la race). Le gène candidat le plus attractif est DGAT1 dont l'effet sur le TP est important, même si son effet le plus marquant est sur le TB.

Enfin, cinq autres régions détectées dans une seule race ont des effets très significatifs sur un caractère : BTA2 (caséine alpha-s2 en % des protéines) et BTA13 (protéolyse) en race Normande, BTA25 (protéolyse) et BTA29 (caséine alpha-s1 en % des protéines) en race Montbéliarde. La région BTA20a affecte le taux de caséine kappa en race Holstein ainsi que le taux d'alpha-lactalbumine. Elle se situe à proximité immédiate du gène GHR (31,890-32,200 Mb) codant pour le récepteur à l'hormone de croissance et dont une mutation est connue pour fortement affecter le TP (Blott *et al* 2003).

### 5.3 / Détection de QTL de protéines laitières chez les ovins

En ovins comme en bovins les QTL détectés ont des niveaux de significativité très élevés. Pour les caractères exprimés en % de la matière protéique, en analyse de liaison une vingtaine de QTL ont été détectés en race Lacaune et seulement quatre en race MTR du fait d'une puissance moindre. Ce nombre passe à plus d'une centaine dans chacune des

**Figure 7.** Illustration d'un QTL du chromosome ovin 6 en race Lacaune affectant le taux de caséine alpha-s2, détecté par analyse de liaison (LA) et d'association (LD). Le pic correspond au cluster des caséines.



deux races en analyse d'association. Parmi les régions les plus significatives, la région du cluster des gènes des caséines sur le chromosome 6 affecte toutes les caséines mais tout particulièrement la caséine alpha-S2 (figure 7). Une région du chromosome 13 affecte la bêta-lactoglobuline dans les deux races, ainsi que les caséines bêta et kappa en race Lacaune. En race Lacaune, une région du chromosome 11 est significative pour les caséines et la bêta-lactoglobuline. Les chromosomes 8, 9 et 15 portent également des zones significatives dans les 2 races et pour plusieurs protéines.

### Conclusion

Cette analyse est la première en France, et l'une des plus importantes au monde sur le déterminisme génétique de la composition en acides gras et en protéines du lait. Les résultats obtenus montrent que les prédictions des composants du lait à l'aide de la spectrométrie MIR sont d'une précision suffisante pour des analyses génétiques et des utilisations en sélection. Même quand la précision n'est que modérée, elle permet une caractérisation individuelle suffisante et la mise en évidence des régions du génome responsables des variations de composition. Un atout important est la répétition des mesures au cours de la lactation qui permet une réduction des erreurs aléatoires et donc une bonne caractérisation individuelle. La précision des équations des acides gras est plus élevée que pour les protéines, mais le déterminisme génétique apparaît plus fort pour les protéines, que

l'on se base sur le niveau des estimations d'héritabilité ou sur la clarté de la détection des QTL.

L'étude a mis en évidence l'effet majeur du mode d'expression des caractères. Il est généralement recommandé de travailler en pourcentage du total de la matière grasse. Mais ce mode d'expression induit des corrélations négatives « artificielles » entre AG saturés et insaturés, tandis qu'une expression en teneur dans le lait rend les mesures d'AG plus indépendantes les unes des autres et plus proches de leur métabolisme propre.

Les résultats de paramètres génétiques sont globalement très cohérents entre espèces et entre races. Une différence notable est observée chez les caprins avec une opposition nette entre AG saturés courts et longs, expliquée sans doute par l'arrêt de la chaîne d'élongation plus précoce que chez les bovins. Il est possible également que certaines différences s'expliquent par la précision plus faible des équations caprines.

De nombreux QTL affectant la composition du lait en AG et en protéines ont pu être identifiés grâce au dispositif *PhénoFinlait*. Certains de ces QTL présentent des seuils de significativité extrêmement élevés, sans doute parce que les caractères d'intérêt sont des phénotypes fins, résultats directs de l'action d'un nombre limité de gènes. De façon similaire, les valeurs d'héritabilité confirment la composante génétique très forte des taux de protéines et des AG saturés. Plusieurs QTL, parmi les plus importants,

sont retrouvés dans deux, voire trois espèces. Notons par exemple l'effet de la région des caséines sur les différentes protéines, l'effet de DGAT1 principalement sur les acides gras, l'effet du gène de la SCD sur les acides gras insaturés. Le QTL bovin du chromosome 20 vers 58 Mb est également retrouvé chez les caprins. Pour les QTL à effet moins important, le taux de concordance est moins élevé, mais il est difficile de conclure à des QTL spécifiques de race ou plus simplement à une puissance de détection insuffisante. Cette analyse reste à approfondir.

Autre aspect favorable, les QTL localisent fréquemment entre composants, reflétant les voies métaboliques, mais aussi entre races. Ils ont fréquemment des effets très importants. Ces deux points sont très favorables à l'identification des mutations sous-jacentes. Plusieurs d'entre elles sont déjà connues : DGAT1 ou AGPAT6 pour les AG, le locus des caséines, de la bêta-lactoglobuline, DGAT1 ou GHR pour les protéines. La plupart de ces gènes ont d'ailleurs des effets sur le TB ou le TP et ont été mis en évidence dans l'analyse de ces deux caractères. D'autres régions sont plus originales, soit parce qu'elles n'étaient pas décrites auparavant (chromosome 1 pour les  $\omega$ 3, chromosome 17 pour les AG courts, lien entre bêta-lactoglobuline et AG), soit parce que les mutations causales ne sont pas encore identifiées (chromosomes 5 ou 19, par exemple). Les prochaines étapes consisteront à mobiliser les données de séquençage des taureaux pour rechercher, parmi les polymorphismes mis en évidence, ceux qui expliquent les QTL détectés.

Les caractères sont nombreux mais ils sont souvent fortement corrélés, de sorte que le nombre de caractères indépendants reste assez faible. Par ailleurs, ils n'ont pas tous un sens biologique ou technologique susceptible de justifier un traitement individuel spécifique. Il apparaît, donc, que si l'on souhaite sélectionner ces caractères, on pourra se limiter à un nombre raisonnable. Il n'en demeure pas moins que le choix de l'objectif reste à faire. Par ailleurs, tous les choix ne sont pas biologiquement possibles. Par exemple, dissocier les concentrations entre acides gras saturés ou bien les concentrations en caséines est un problème difficile, nécessitant une forte pression de sélection pour un résultat limité et sans doute aléatoire. Concernant les acides gras, la question posée est principalement l'équilibre entre saturés et insaturés. Augmenter les insaturés en proportion de la matière grasse est possible dans une certaine mesure, mais sans doute au prix d'une diminution du TB. En effet, plus le TB est fort, plus sa fraction saturée est importante en proportion. Modifier un composant par sélection est possible, mais au prix de variations corrélées, conduisant à un nouvel équilibre global qu'il convient de prédire et prendre en compte. Pour les protéines, le contexte est sans doute plus simple : il est possible d'augmenter l'ensemble des caséines en proportion des protéines du lait. Par ailleurs, augmenter le TP du lait conduit à augmenter la fraction caséinique, traduisant une relation favorable. Enfin, on peut sans doute fortement diminuer la bêta-lactoglobuline compte tenu de sa forte héritabilité et de sa dépendance à un gène majeur, ce qui conduit aussi à augmenter la fraction caséinique.

Ces résultats montrent qu'il est tout à fait possible de modifier la composition fine du lait par sélection dans les populations laitières en utilisant les données des spectres MIR collectées en routine. Compte tenu des paramètres génétiques, la sélection peut être rapide et efficace pour une partie des composants au moins. Elle peut bénéficier de données massives, de mesures héritables, voire de gènes à gros effets. Les conditions techniques sont favorables, tant avec les méthodes classiques polygéniques bénéficiant de phénotypes nombreux et héritables, qu'en sélection génomique avec une population de référence potentiellement large et des QTL à effets très significatifs, autorisant donc des précisions élevées des prédictions de valeurs génétiques. Dès aujourd'hui, avec une population femelle génotypée encore réduite, les prédictions génomiques ont une précision honorable, du fait des gènes à effet important mis en évidence.

## Remerciements

Le programme *PhénoFinlait* est cofinancé par l'ANR, Apis-gène, FGE, FranceAgriMer, le Ministère en charge de l'Agriculture, dont le compte d'affectation spéciale CASDAR, et l'Union Européenne. Les auteurs remercient les partenaires du consortium *PhénoFinlait*, les acteurs de terrain et les éleveurs. Ces travaux ont été conduits dans le cadre des UMT 3G et GGPR. Les auteurs remercient J.M. Astruc, C. Carillier, S. Esvan, M. Ferrand, G. Miranda et S. Fritz pour leurs contributions. Les génotypes ont été réalisés par Labogena.

## Références

- Barillet F., Boichard D., Barbat A., Astruc J.M., Bonaiti B., 1992. Use of animal model for genetic evaluation of the Lacaune dairy sheep. *Livest. Prod. Sci.*, 31, 287-299.
- Blott S., Kim J.J., Moio S., Schmidt-Küntzel A., Cornet A., Berzi P., Cambisano N., Ford C., Grisart B., Johnson D., Karim L., Simon P., Snell R., Spelman R., Wong J., Vilkkii J., Georges M., Farnir F., Coppeters W., 2003. Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics*, 163, 253-266.
- Bouwman A.C., Bovenhuis H., Visker M.H.P.W., van Arendonk J.A.M., 2011. Genome-wide association of milk fatty acids in Dutch dairy cattle. *BMC Genetics*, 12, 43.
- Brochard M., Faucon F., Barillet F., Bolard M., Brunschwig P., Duhem K., Eggen A., Esvan S., Ferrand M., Fritz S., Gastinel P.L., Guerin J.L., Journaux L., Krychowski T., Lagriffoul G., Larroque H., Lecomte C., Leray O., Leroux C., Leverrier S., Martin P., Mattalia S., Miranda G., Palhière I., Peyraud J.L., Boichard D., 2009. *PhénoFinlait* : un programme national français de détection de QTL et/ou de gènes majeurs affectant la composition fine du lait des ruminants laitiers. *Renc. Rech. Rum.*, 16, 423.
- Carta A., Casu S., Usai M.G., Addis M., Fiori M., Fraghi A., Miari S., Mura L., Piredda G., Schibler L., Sechi T., Elsen J.M., Barillet F., 2008. Investigating the genetic component of fatty acids composition in sheep milk. *Small Rum. Res.*, 79, 22-28.
- Chilliard Y., Glasser F., Ferlay A., Bernard L., Rouel J., Doreau M., 2007. Diet, rumen biohydrogenation and nutritional quality of cow and goat milk fat. *Eur. J. Lipid Sci. Technol.*, 109, 828-855.
- Cohen-Zinder M., Sroussi E., Larkin D.M., Looor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I., Ron M., 2005. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.*, 15, 936-944.
- Ferrand M., Huquet B., Barbey S., Barillet F., Faucon F., Larroque H., Leray O., Trommenschlager J.M., Brochard M., 2010. Determination of fatty acid profile in cow's milk using mid-infrared spectrometry: Interest of applying a variable selection by genetic algorithms before a PLS regression. *Chemometr. Intell. Lab. Syst.*, 106, 183-189.
- Ferrand-Calmels M., Palhière I., Brochard M., Leray O., Astruc J.M., Aurel M.R., Barbey S., Bouvier F., Caillat H., Douguet M., Faucon-Lahalle F., Thomas G., Trommenschlager J.M., Larroque H., 2014. Estimation of fatty acid profile in cow, ewe and goat milk by mid-infrared spectrometry and standardization of the results. *J. Dairy Sci.*, 97, 17-35.
- García-Fernández M., Gutiérrez-Gil B., García-Gómez E., Sánchez J.P., Arranz J.J., 2010. Detection of quantitative trait loci affecting the milk fatty acid profile on sheep chromosome 22: role of the stearoyl-CoA desaturase gene in Spanish Churra sheep. *J. Dairy Sci.*, 93, 348-357.



- Gelé M., Minery S., Astruc J.M., Brunschwig P., Ferrand M., Lagriffoul G., Larroque H., Legarto J., Martin P., Miranda G., Palhière I., Trossat P., Brochard M., 2014. Phénotypage et génotypage à grande échelle de la composition fine des laits dans les filières bovine, ovine et caprine. In *PhénoFinlait* : Phénotypage et génotypage pour la compréhension et la maîtrise de la composition fine du lait. Brochard M., Boichard D., Brunschwig P., Peyraud J.L. (Eds). Dossier, INRA Prod. Anim., 27, 255-268.
- Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M., Snell R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with a major effect on milk yield and composition. *Genome Res.*, 12, 222-231.
- Grosclaude F., Mahé M.F., Brignon G., Di Stasio L., Jeunet R., 1987. A Mendelian polymorphism underlying quantitative variations of goat  $\alpha$ S1-casein. *Genet. Sel. Evol.*, 19, 399-412.
- Guillaume F., 2009. Intégration de l'information moléculaire dans l'évaluation génétique. Thèse AgroParisTech, 145p.
- Leray O., Ferrand M., Larroque H., Astruc J.M., Douguet M., Brochard M., Duhem K., 2011. Harmonisation of milk analysers for fatty acid determination by FTMIR: An essential step prior to collective data use. ICAR workshop 2011, Bourges-en-Bresse, France, 22-24.
- Maroteau C., Palhière I., Larroque H., Clément V., Tosser-Klopp G., Rupp R., 2013. QTL detection for traits of interest for the dairy goat industry. *Fed. Eur. Zootech.*, Nantes, France.
- Maroteau C., Palhière I., Larroque H., Clément V., Ferrand M., Tosser-Klopp G., Rupp R., 2014. Genetic parameter estimation for major milk fatty acids in Alpine and Saanen primiparous goats. *J. Dairy Sci.*, 97, 3142-3155.
- Meuwissen T.H.E., Goddard M.E., 2001. Prediction of identity by descent probabilities from marker-haplotypes. *Genet. Sel. Evol.*, 33, 605-634.
- Schmidely P., Bourgeois E., 2011. Impact des pratiques en alimentation animale sur la composition en acides gras des produits animaux destinés à l'Homme. *ANSES Éd. Sci.*, 274p.
- Schopen G.C.B., Heck J.M.L., Bovenhuis H., Visker M.H.P.W., van Valenberg H.J.F., van Arendonk J.A.M., 2009. Genetic parameters for major milk proteins in Dutch Holstein-Friesians. *J. Dairy Sci.*, 92, 1182-1191.
- Schopen G.C.B., Visker M.H.P.W., Koks P.D., Mullaart E., van Arendonk J.A.M., Bovenhuis H., 2011. Whole-genome association study for milk protein composition in dairy cattle. *J. Dairy Sci.*, 94, 3148-3158.
- Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko G., Marlier M., Bertozzi C., Mayeres P., Gengler N., 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.*, 89, 3690-3695.
- Soyeurt H., Gillon A., Vanderick S., Mayeres P., Bertozzi C., Gengler N., 2007. Estimation of heritability and genetic correlations for the major fatty acids in bovine milk. *J. Dairy Sci.*, 90, 4435-4442.
- Stoop W.M., van Arendonk J.A.M., Heck J.M.L., van Valenberg H.J.F., Bovenhuis H., 2008. Genetic parameters for major milk fatty acids and milk production traits of dutch Holstein-Friesians. *J. Dairy Sci.*, 91, 385-394.

## Résumé

Cette étude présente les principaux résultats d'estimation de paramètres génétiques et de détection de QTL obtenus dans le cadre du programme *PhénoFinlait* sur les caractères de composition en Acides Gras (AG) et protéines du lait dans trois races bovines (Holstein, Montbéliarde et Normande), deux races ovines (Lacaune et Manech Tête Rousse) et deux races caprines (Alpine et Saanen). La composition du lait est estimée à partir de la spectrométrie dans le moyen infrarouge. Les paramètres génétiques sont estimés à partir des données de 102 000 contrôles laitiers de 22 000 vaches en première lactation, 67 000 contrôles de 20 000 brebis, et 45 000 contrôles de 13 700 chèvres. Ils sont très homogènes entre espèces et entre races. En revanche, ils dépendent beaucoup du mode d'expression des caractères, exprimés en proportion du lait ou de la matière. Exprimés en teneur dans le lait, les AG saturés présentent une héritabilité plus élevée que les insaturés chez les bovins et les ovins, mais l'écart est plus faible quand ils sont exprimés en teneur dans le gras. Chez les caprins, les estimations d'héritabilité sont plus élevées pour les caractères exprimés en teneur dans la matière grasse. Les mesures d'AG sont fortement corrélées entre stades de lactation, à l'exception du premier mois qui apparaît comme un caractère assez différent. Les corrélations génétiques sont positives entre AG saturés et entre AG insaturés. Entre AG saturés et insaturés, les corrélations sont positives pour les AG exprimés en teneur dans le lait mais négatives quand les AG sont exprimés en pourcentage de la matière grasse. Les AG saturés sont très fortement corrélés au taux butyreux du lait. Concernant les protéines, les estimations d'héritabilité sont très élevées pour la bêta-lactoglobuline, assez élevées pour les caséines, plus modérées pour l'alpha-lactalbumine. Concernant les corrélations, il existe une forte analogie entre AG et protéines. Ainsi, les caséines sont fortement corrélées entre elles et fortement liées au taux protéique. Leur corrélation avec les protéines sériques est positive quand les protéines sont exprimées en teneur dans le lait, mais très négatives quand elles sont exprimées en teneur dans les protéines.

Les analyses de détection de QTL reposent sur les données de 7 800 vaches, 1 800 brebis et 2 300 chèvres génotypées avec des puces SNP pangénomiques. En moyenne, 9 QTL d'AG ont été détectés par caractère et par race bovine. Les QTL les plus importants ont été trouvés sur les chromosomes 14 (gène DGAT1), 5, 19, 27, 17, 11 et 13. On observe une forte co-localisation de QTL entre AG du même type, reflétant leur origine métabolique commune. Une fraction notable de ces QTL semble partagée entre races. 22 à 29 QTL sont détectés en moyenne pour chaque taux de protéine. Les plus significatifs se situent sur les chromosomes 6 (2 régions QTL, régions des gènes ABCG2 et des caséines), 11 (gène de la bêta-lactoglobuline) et 20 (gène GHR vers 32 Mb, mais aussi vers 58Mb). Le gène DGAT1 affecte également de nombreuses protéines exprimées en teneur dans le lait.

Ces résultats indiquent que la composition fine du lait pourrait être modifiée par sélection, même si les grands équilibres entre composants peuvent difficilement être bouleversés. Il est ainsi possible d'augmenter la fraction de caséines dans les protéines. Il est aussi possible d'augmenter la fraction d'AG insaturés dans le lait, mais sans doute au prix d'une diminution du taux butyreux.

## Abstract

### *Genetic determinism of milk composition in fatty acids and proteins in ruminants, and selection potential*

This study presents the main genetic results obtained from the *PhénoFinlait* project with regards to genetic parameters and QTL detection for milk composition in fatty acids (FA) and proteins in three dairy cattle breeds (Montbeliarde, Normande, and Holstein), two goat breeds (Alpine and Saanen) and two sheep breeds (Lacaune and Manech Tete Rousse). Milk composition was estimated from mid-infrared spectrometry. Genetic parameters were estimated from about 102,000 test-day records from 22,000 cows in first lactation, 67,000 records from 20,000 ewes and 45,000 records from 13,700 goats. Genetic parameter results were very homogeneous across species and breeds. They were found to be sensitive to the mode of expression of the traits, in % of milk or in % of fat or protein.

Expressed in % of milk, test-day saturated FA (SAT) had higher heritability estimates than unsaturated FA (UNSAT) but this difference was smaller when traits were in % of fat. In goats, the results were markedly different with higher heritability estimates found for traits expressed in % of fat. FA measurements were highly genetically correlated across different stages of lactation except in the beginning of the lactation. Genetic correlation estimates were found to be positive across saturated FA, and also across unsaturated FA. Between saturated and unsaturated FA, correlation estimates were positive when FA were expressed in % milk but negative when FA were expressed in % fat. Saturated FA were strongly correlated with fat content. With regards to proteins, heritability estimates were very high for beta-lactoglobulin, moderate to high for caseins, moderate for alpha-lactalbumin. Correlation pattern showed a strong similarity between FA and proteins. Indeed, caseins were strongly correlated with each other and with protein content. Their correlation with whey proteins was positive or very negative, whether proteins were expressed in % milk or in % proteins.

QTL detection analyses were based on pangenomic genotyping data of 7800 cows, 1800 ewes, and 2300 goats. On average, 9 QTL were detected per FA trait and cattle breed. The most important QTL were found on chromosome 14 (DGAT1 gene), 5, 19, 27, 17, 11, and 13. A strong co-location of QTL was observed for FA sharing a common metabolic origin. A large proportion of the QTL seems to be shared across breeds. Twenty-two to 29 QTL were detected for each protein. The most significant QTL were found on chromosome 6 (2 regions close to ABCG2 gene and to casein cluster), 11 (beta-lactoglobulin gene), and 20 (2 regions, around GHR gene and around 58 Mb). The DGAT1 gene (BTA14, around 1.8 Mb) was also found to affect many proteins when expressed in % of milk.

These results show that milk composition in FA or proteins can be significantly modified by selection, even if the major characteristics cannot be changed. For instance, it is possible to increase the casein percentage in total proteins. It is also possible to increase the unsaturated FA fraction in fat, but at the expense of a decrease in fat content in milk.

BOICHARD D., GOVIGNON-GION A., LARROQUE H., MAROTEAU C., PALHIÈRE I., TOSSER-KLOPP G., RUPP R., SANCHEZ M.-P., BROCHARD M., 2014. Déterminisme génétique de la composition en acides gras et protéines du lait des ruminants, et potentialités de sélection. In *PhénoFinlait* : Phénotypage et génotypage pour la compréhension et la maîtrise de la composition fine du lait. Brochard M., Boichard D., Brunschwig P., Peyraud J.-L. (Eds). Dossier, INRA Prod. Anim., 27, 283-298.