

Estimation de paramètres génétiques et recherche de QTL affectant la composition du lait en acides gras dans trois races bovines laitières françaises.

GOVIGNON-GION A. (1), FRITZ S. (2), LARROQUE H. (3), CHANTRY-DARMON C. (4), LAHALLE F. (5,6), BROCHARD M. (5), BOICHARD D. (1)

(1) INRA UMR 1313 GABI, 78 352 Jouy en Josas,

(2) UNCEIA 149 rue de Bercy, 75595 Paris cedex 12,

(3) INRA UR631 SAGA, B.P. 52627, 31326 Castanet Tolosan cedex,

(4) LABOGENA 78 352 Jouy en Josas,

(5) Institut de l'Élevage, 149 rue de Bercy, 75 595 Paris cedex 12,

(6) CNIEL, 42 rue de Chateaudun, 75009 Paris.

RESUME

Cette étude présente les principaux résultats d'estimation de paramètres génétiques et de détection de QTL obtenus dans le cadre du projet PhenoFinlait sur les caractères de composition en acides gras (AG) du lait dans les trois races bovines Holstein, Montbéliarde et Normande. Les paramètres génétiques sont estimés à partir de 101 858 contrôles élémentaires de 21 947 vaches en première lactation. Chaque contrôle est caractérisé par son profil en AG, estimé par spectrométrie moyen infra rouge. Les paramètres génétiques dépendent beaucoup du mode d'expression des caractères : acides gras exprimés en proportion du lait ou de la matière grasse. Exprimés en teneur dans le lait, les AG saturés (SAT) présentent une héritabilité plus élevée que les insaturés (INSAT) mais l'écart est plus faible quand ils sont exprimés en pourcentage d'AG totaux. 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 SAT, positives entre INSAT. Entre SAT et INSAT, le signe des corrélations diffère selon le mode d'expression. Les SAT sont très fortement corrélés au taux butyreux du lait. 7500 vaches ont été génotypées avec la puce SNP Bovine 50k Beadchip et 500 autres ont été imputées depuis un génotypage sur la puce SNP bovine 7k, permettant une recherche de QTL par LDLA (analyse combinant liaison intra famille et déséquilibre de liaison). Au total, sur 31 caractères, 544 QTL ont été détectés dans l'ensemble des trois races, soit environ 9 par caractère et par race. Les QTL les plus importants ont été trouvés sur les chromosomes 14 (gène DGAT1), 26 (gène SCD), 13, 16, 17, 19 et 27. Les QTL les plus significatifs sont observés sur le chromosome 5 dans les 3 races, une région non décrite jusqu'à présent. On observe une forte colocalisation de QTL entre caractères, en particulier entre acides gras du même type, reflétant leur origine métabolique commune. Une fraction notable, au moins un tiers, de ces QTL semble partagée entre races. Ces résultats favorables ouvrent la voie à la prise en compte du profil en AG en sélection mais aussi à la recherche des mutations responsables de ces variations.

Estimation of genetic parameters and detection of QTL affecting milk fatty acid composition in three dairy cattle breeds.

GOVIGNON-GION A. (1), FRITZ S. (2), LARROQUE H. (3), CHANTRY-DARMON C. (4), LAHALLE F. (5,6), BROCHARD M. (5), BOICHARD D. (1)

(1) INRA UMR 1313 GABI, 78 352 Jouy en Josas

SUMMARY

This study presents the main genetic results obtained in the PhenoFinlait projects with regards to genetic parameters and QTL detection for milk composition in fatty acids (FA) in the French Montbeliarde (MO), Normande (NO) and Holstein (HO) cattle breeds. Genetic parameters were estimated from 101,858 test-day records from 21,947 cows in first lactation. Each test-day was characterized by its profile in FA in milk, as estimated from Mid-Infrared spectrometry. Genetic parameters were found to be very sensitive to the mode of expression of the traits, in % of milk or in % of fat. Expressed in % of milk, test-day saturated FA (SAT) had higher heritability estimates than unsaturated FA (UNSAT) but this difference was smaller when the traits were in % of fat. FA measurements were highly genetically correlated across different stages of lactation except in the beginning of lactation. Genetic correlations between FA in milk were positive across SAT, and also positive across UNSAT. Between SAT and UNSAT, their sign changed according to the mode of expression. SAT were strongly and positively correlated to fat content in milk. Seven-thousand five-hundred cows were genotyped with the Bovine 50k Beadchip and 500 other cow genotypes were imputed from their Bovine LD Beadchip (7k). A QTL detection analysis was carried out by Linkage and Linkage Disequilibrium Analysis (LDLA). For 31 traits, 544 QTL were detected in the three breeds, corresponding to an average of 9 per trait and breed. The most important QTL were found on chromosomes 14 (DGAT1 gene), 26 (SCD gene), 13, 16, 17, 19 and 27. The most significant QTL were found on chromosome 5, a region not yet described until now. QTL often colocalized across traits, especially for fatty acids of the same type, reflecting their common metabolic origin. A relatively high fraction of the QTL, more than one third, was found to be shared across breeds. These results pave the way for including FA in selection and also for the discovery of the underlying causative mutations.

INTRODUCTION

Depuis plusieurs années déjà, la composition en acides gras (AG) du lait fait l'objet de recherches importantes pour disposer de méthodes et d'outils permettant à la fois de la

mesurer plus facilement et de la piloter, avec l'objectif de l'orienter pour satisfaire les exigences des consommateurs et les besoins des filières. Le pilotage au niveau de la production est essentiel car la composition en AG des produits laitiers est le reflet direct de la composition du lait

d'origine (Lucas *et al.*, 2006). Les leviers sont potentiellement nombreux mais jusqu'à présent, l'essentiel des efforts s'est porté sur les aspects nutritionnels, en particulier les types de ration ou de complémentation favorisant les acides gras insaturés (INSAT) (Chilliard *et al.*, 2007 ; Schmidely et Bourgeois, 2011)

C'est dans ce contexte que le projet PhénoFinlait a été lancé en 2008 par les acteurs de la profession laitière (CNIEL), de l'élevage (UNCEIA, FCEL, Capgenes, CNBL) et de la recherche-développement (INRA, Institut de l'Élevage, Actilait, Labogena) pour une durée de 5 ans (Brochard *et al.*, 2009).

Le projet PhénoFinlait concerne les trois espèces bovine, ovine et caprine et ses objectifs sont les suivants :

- Disposer de méthodes de mesure des composants élémentaires (AG et protéines) permettant un diagnostic et un pilotage à l'échelle de l'élevage et de l'individu.
- Constituer un référentiel qui permet, par sa taille et sa diversité, de répondre aux questions actuelles relatives à la composition fine du lait mais aussi aux questions futures sur d'autres composants.
- Développer des méthodologies de pilotage de la composition fine du lait, tant par la génétique que la conduite du troupeau.

Concernant les AG, les travaux peuvent être résumés en 4 phases. Les valeurs indiquées concernent le total des trois espèces. a) développements d'équations de prédiction des AG à partir de spectres du moyen infra-rouge (MIR) et définition des conditions de leur validité (Ferrand *et al.*, 2010, Leray *et al.*, 2011) ; b) phase sur le terrain de récupération, en complément des données des bases nationales, de l'ensemble des données (spectres, enquêtes sur les pratiques) et des échantillons biologiques (sang, lait) dans près de 1500 élevages choisis entre autres, pour leur diversité ; c) génotypage d'un total de plus de 12 000 femelles à l'aide de puces à SNP couvrant tout le génome ; d) analyse du déterminisme génétique et des facteurs de milieu de la composition en AG du lait. Pour les femelles ciblées pour le génotypage, un échantillon supplémentaire de lait a été collecté et est conservé à -80°C en « LactoBanque » au Centre de Ressources Biologiques GADIE Jouy-en-Josas, afin de permettre des analyses plus fines ou sur d'autres phénotypes.

Cette synthèse présente les résultats génétiques acquis chez les bovins.

1. MATERIEL ET METHODES

1.1 PROTOCOLE PHENOFINLAIT

Tableau 1. Contenu de la base bovine PhénoFinlait

	MO	NO	HO	Total
Départements	14, 21, 25, 35, 39, 44, 49, 50, 53, 59, 61, 62, 67, 68, 72, 90			
Elevages	1 224 dont 70 en label AB			
Enquêtes	6 432			
Spectres	85 509	113 487	197 122	396 118
Femelles avec au moins 1 spectre	16 434	23 458	34 949	74 841
Sang prélevé	2 616	3 464	3 821	9 901
Femelles génotypées	2 039	2 700	3 080	7 819
Fem. avec éch. en LactoBanque	2 485	3 229	3 466	9 180

Pour sa composante bovine, le programme PhénoFinlait concerne 1224 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'intègrent dans le cadre des contrôles laitiers

officiels classiques mais sont complétés par l'enregistrement 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 collective et individuelle en fourrages, pâturage et concentrés. Parmi les 90 000 femelles suivies, 8000 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 est réalisée par des techniciens des centres d'insémination et envoyée à Labogena, pour une extraction d'ADN. L'ensemble des données issues de ce protocole est complété par celles du Système d'Information Génétique et stocké dans la base PhénoFinlait développée par le le Centre de Traitement de l'Information Génétique de l'INRA.

1.2 ESTIMATION DU PROFIL EN AG

Le profil en AG est prédit à partir d'équations appliquées aux 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). 450 échantillons ont été utilisés chez les bovins. Les décalages entre analyseurs MIR ou dans le temps intra analyseur sont corrigés grâce à l'analyse d'un lait témoin sur chaque analyseur une à deux fois par mois. Chez les bovins, environ 25 AG ou combinaisons d'AG sont prédits avec une prédiction bonne ou très bonne (tableau 2). Ils font partie des AG qui sont analysés dans la suite de cette étude.

Tableau 2. Principaux AG ou groupes d'AG du lait de vache et précision des équations de prédiction. Les AG sont exprimés en % de la matière grasse, les index de désaturation en %.

AG	Nom	Teneur moyenne *	Erreur relative. (%)	R ²
SAT	Total AG saturés	70	1,3	1,00
C4:0	Butyrique	3,8	4,4	0,93
C6:0	Caproïque	2,3	3,6	0,96
C8:0	Caprylique	1,7	4,6	0,96
C10:0	Caprique	3,4	6,5	0,95
C12:0	Laurique	3,8	6,6	0,95
C14:0	Myristique	11,2	5,4	0,94
C16:0	Palmitique	32,2	5,4	0,94
C18:0	Stéarique	11,1	11,9	0,84
INSAT	Total AG insaturés	30	3,8	0,98
MONO	Total AG mono-insaturés	25	4,2	0,97
C14:1c9	Myristoléique	0,5	19,5	0,58
C16:1c9	Palmitoléique	1,5	18,1	0,67
C18:1		23	5,2	0,96
C18:1c9	Oléique	21,7	5,8	0,96
Index 14	C14:1 / [C14:0+C14:1]	4,3	18,8	0,23
Index 16	C16:1 / [C16:0+C16:1]	4,5	14,7	0,75
Index 18	C18:1 / [C18:0+C18:1]	67,5	3,6	0,65
POLY	Total AG poly-insaturés	3 à 5	9,2	0,76
C18:2c9t11	Ruménique (acide linoléique congugué, CLA)	0,4	16,8	0,82
C18:2c9c12	Linoléique (LA)	2,4	11,0	0,77
ω3	Total AG oméga 3	0,6	14,2	0,86
ω6	Total AG oméga 6	2,5	10,8	0,78
ω7	Total AG oméga 7	0,4	11,6	0,84

* Teneur moyenne en proportion de la matière grasse du lait
Qualité des prédictions : validation sur un jeu indépendant de 100 échantillons (com. personnelle, Ferrand M.).

1.3 EXPRESSION DES DONNEES

Les AG peuvent être exprimés de deux façons différentes, soit en proportion de la matière grasse produite, soit en teneur dans le lait. Ces deux expressions sont assez différentes dans leur signification et leurs propriétés. 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 du régime humain et la teneur dans le régime résulte à la fois de la composition des aliments et de la proportion relative de chacun d'eux. 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, M Gelé a montré que le profil en AG 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 (communication personnelle).

Par contre, les teneurs exprimées en % de la matière grasse présentent un inconvénient important, une liaison mathématique automatique puisque la somme des AG représente quasiment 100%. Ceci induit des corrélations négatives artificielles entre certains AG, en particulier les plus importants (par exemple, 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 de cette synthèse, les deux modes d'expression sont analysés pour les paramètres génétiques. Cependant, dans l'analyse QTL, pour simplifier la présentation, seules les teneurs dans la matière grasse sont présentées.

1.4 ESTIMATION DES PARAMETRES GENETIQUES

Les paramètres génétiques ont été estimés séparément dans chacune des trois races, à partir des données de première lactation de vaches ayant au minimum deux analyses spectrales (5 en moyenne) et des parentés connues. Les contrôles analysés sont réalisés entre 8 et 305 jours de lactation.

Plusieurs modèles ont été utilisés selon les paramètres génétiques à estimer, et les calculs ont été réalisés avec le logiciel WOMBAT (Meyer, 2006). Les données incluent 21 947 vaches et 101 858 analyses avec spectre.

Le modèle 1 s'applique aux données de chaque contrôle supposé être, pour un AG, l'expression d'un même caractère répété. Il permet d'estimer l'héritabilité moyenne par contrôle ainsi que la répétabilité (ou corrélation moyenne) entre contrôles. Ce modèle inclut les effets fixés du troupeau-contrôle, du stade de lactation, du mois de vêlage intra année, et les effets aléatoires de la valeur génétique de la vache et de son effet d'environnement permanent. Ce modèle est appliqué à chaque AG.

Le modèle 2 s'applique également aux données de chaque contrôle. La lactation est découpée en 4 périodes (T1=[8-50j], T2=[51-120j], T3=[121-200j], T4=[200-305j]), chacune correspondant à un caractère différent. Le modèle 2 est multicaractère et permet d'estimer les corrélations génétiques entre périodes. Les effets du modèles sont les mêmes que dans le modèle 1.

Le modèle 3 s'applique à des données moyennes par lactation. Il analyse plusieurs AG simultanément et donc estime des héritabilités par lactation et des corrélations génétiques entre AG. Les données à chaque contrôle sont d'abord ajustées pour le stade de lactation (estimé avec proc GLM de SAS) avant d'en faire la moyenne sur la lactation, en pondérant chaque mesure par le taux butyreux du contrôle. Le modèle 3 inclut les effets du troupeau, du mois de mise bas intra année et la valeur génétique de l'animal.

1.5. DETECTION DE QTL

L'analyse de détection de QTL est réalisée intra race en recherchant une relation entre les performances et les marqueurs génétiques couvrant tout le génome. Les performances sont des données par lactation. Dans un premier temps, un modèle d'évaluation génétique similaire au modèle 1, mais étendu à tous les animaux du dispositif, est mis en œuvre, permettant de corriger les mesures pour les effets d'environnement. Pour chaque AG, ces mesures par contrôle sont ensuite combinées pour obtenir une moyenne par animal, ainsi qu'une pondération reflétant la précision de cette moyenne, fonction du nombre de contrôles.

Les performances des animaux génotypés sont ensuite analysées à l'aide d'un modèle dit « Linkage Disequilibrium and Linkage Analysis (LDLA, Meuwissen et Goddard, 2001) combinant 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 plus puissante et résolutive qu'une analyse de liaison intra famille seule, qui n'utilise principalement que l'information paternelle. Elle est aussi plus puissante et plus robuste, mais généralement moins résolutive qu'une analyse d'association pure qui ne prend pas en compte l'information familiale. Pour approcher au plus près les gènes sous-jacents, il conviendra dans un second temps de compléter cette analyse avec les méthodes les plus résolutes.

Dans l'approche LDLA, deux modèles sont comparés. Sous H0, on ne considère qu'un effet polygénique pour chaque animal. On estime donc une variance polygénique et une variance résiduelle. Sous H1, le modèle inclut aussi l'effet d'un QTL supposé porté par les deux haplotypes paternel et maternel de chaque individu à une position donnée de son génome et on estime la part de variance due à ce QTL. 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 (ou LRT). Un seuil de LRT de 13 (correspondant à une probabilité critique théorique de $3 \cdot 10^{-4}$ à chaque test) a été considéré pour conclure à l'existence d'un QTL, compte tenu de la multiplicité des tests et de leur corrélation.

Les QTL étant inconnus, on les suit entre chromosomes des animaux du dispositif par l'intermédiaire des marqueurs génétiques qui permettent d'estimer la probabilité d'identité entre les QTL. Pour maximiser l'information disponible, chaque marqueur n'est pas considéré isolément mais par haplotype de 6 marqueurs consécutifs. Le génome est analysé complètement en déplaçant l'haplotype considéré le long des chromosomes.

2. ESTIMATIONS DE PARAMETRES GENETIQUES EN RACES BOVINES

2.1 HERITABILITE ET VARIABILITE GENETIQUE

Les tableaux 3 et 4 résument les estimations d'héritabilité avec le modèle 1 pour les trois races. Les données sont en g/100g de lait et en g/100g de gras. Les écarts types d'erreur sont toujours inférieurs à 0,05. Dans les trois races et pour la majorité des SAT, les taux dans le lait sont plus héritables que les taux dans le gras. Des observations similaires sont décrites dans la bibliographie.

Tableau 3 : Estimation de l'héritabilité des AG par contrôle (exprimés en g/100g de lait) dans chacune des trois races.

	MO	NO	HO
SAT	0,32	0,35	0,28
MONO	0,13	0,18	0,15
POLY	0,21	0,22	0,22
INSAT	0,11	0,19	0,14
C16:0	0,32	0,32	0,28
C14:1c9	0,40	0,41	0,32
C18:1c9	0,13	0,15	0,11
C18:2c9t11	0,17	0,14	0,11

Les résultats sont très cohérents entre races. Exprimés en teneur dans le lait, les SAT présentent une héritabilité

nettement supérieure à celles des INSAT. Ceci s'explique par l'origine des AG : les SAT courts et une forte fraction des moyens sont synthétisés de novo dans la mamelle à partir des précurseurs sanguins. Les AG à chaîne longue sont apportés par la ration, hydrogénés dans le rumen et transportés par le sang. Certains SAT comme C16:0 et les monoinsaturés (MONO) peuvent avoir les deux origines. Les différences de ration ont plus d'impact sur les AG longs que sur les courts.

Lorsque les AG sont exprimés en % du gras, les différences sont tassées, du fait de la liaison mathématique entre variables. Ceci est particulièrement vrai pour les catégories les plus importantes en proportion. Ainsi entre SAT et INSAT, par exemple, leur somme étant proche de 100%, il existe donc une relation quasi linéaire entre eux et leur héritabilité est donc similaire.

Tableau 4 : Estimation de l'héritabilité des AG par contrôle (exprimés en g/100g de gras) dans chacune des trois races.

	MO	NO	HO
SAT	0,24	0,18	0,18
MONO	0,22	0,21	0,19
POLY	0,27	0,21	0,25
INSAT	0,23	0,21	0,20
C16:0	0,27	0,18	0,32
C14:1c9	0,37	0,29	0,31
C18:1c9	0,23	0,21	0,20
C18:2c9t11	0,15	0,10	0,12

Enfin, il convient de noter que ces résultats sont par contrôle et que la moyenne par lactation, non présentée ici, est plus héritable que les mesures individuelles de chaque contrôle.

Les coefficients de variation génétique (σ_g / μ) sont assez faibles (tableau 5 pour la race Normande), compris entre 1 et 5% (en % de MG, entre 5 et 10 % en % de lait), à comparer à 5% pour le taux butyreux. Ces valeurs sont en général plus élevées pour les INSAT, particulièrement les MONO, que pour les SAT.

Ces résultats permettent de déduire que la composition du lait en acides gras peut être modifiée par la génétique dans une certaine mesure mais que la marge de manœuvre est tout de même réduite.

Tableau 5 : Estimation de l'écart type génétique (σ_g) et du coefficient de variation (CV_g en %) des AG (exprimés en g/100g de gras) en race Normande

	σ_g	CV_g
TB (% du lait)	2,5	5,4
SAT	1,2	1,8
MONO	1,2	4,8
POLY	0,15	4,4
C16:0	1,2	3,6
C18:1	1,1	5,8

2.2 DETERMINISME GENETIQUE AU COURS DE LA LACTATION

L'héritabilité des différentes mesures généralement augmente 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 6). Elles sont toujours positives et généralement très élevées, reflétant un déterminisme génétique commun, même entre saisons. Cependant, 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 qui est plus limitée après le pic de lactation. Ces résultats permettent sans risque de raisonner à l'échelle de la lactation et non du contrôle.

Tableau 6 : Corrélations génétiques entre périodes au cours de la lactation

	C14:0	T1-T2	T2-T3	T3-T4
g/100g lait				
MO		0,657	0,927	0,917
HO		0,685	0,747	0,964
g/100g gras				
MO		0,966	0,967	0,965
HO		0,684	0,913	0,963

Tableau 7 : Corrélations génétiques avec la quantité de lait et le taux butyreux

	SAT		INSAT	
	%gras	%lait	%gras	%lait
Lait	-0,12	-0,15 à -0,40	+0,33	-0,65
TB	+0,68	+0,77 à +0,95	-0,64	+0,51

2.3 CORRELATIONS GENETIQUES ENTRE AG

Les corrélations sont modérées entre quantité de lait et les différents AG exprimés en % du gras, négatives avec SAT et positives avec INSAT (tableau 7). Elles sont dans le sens inverse avec le taux butyreux. Ces estimations sont en accord avec Soyeurt et al. (2007). Les estimations sont dans le même sens mais un peu plus marquées pour les SAT exprimés en % du lait. Par contre, elles changent de sens pour les INSAT qui sont négativement corrélées avec le lait et positivement avec le TB. 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 SAT et INSAT, mais avec un rapport INSAT/SAT plus élevé.

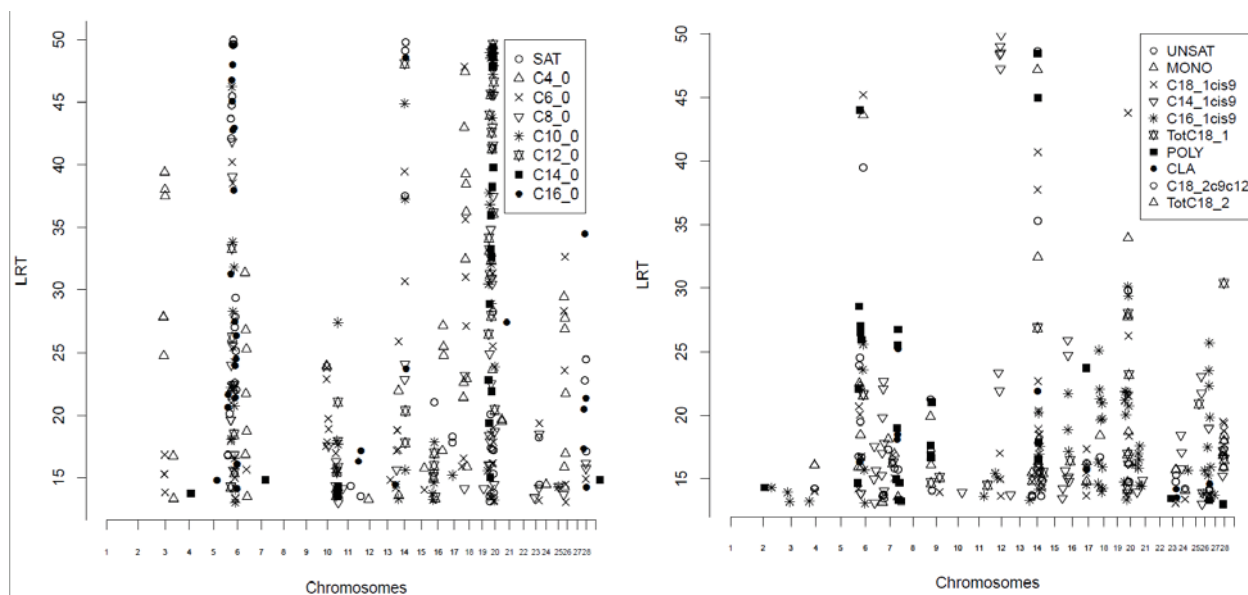
Le tableau 8 présente les gammes de corrélations génétiques estimées entre les différents AG selon les deux types d'expression. Les corrélations génétiques sont fortement positives entre saturés et sont en accord avec la bibliographie (Soyeurt, 2007). Des corrélations positives sont également observées entre insaturés, pour les deux modes d'expression. Du fait de leur origine hétérogène, ces corrélations ne s'interprètent pas simplement, même si de nombreux insaturés sont le produit de la même enzyme, la Stéaroyl-CoA désaturase (SCD) et sont logiquement corrélés. Ces corrélations sont plus élevées quand les AG sont exprimés en % du gras, probablement du fait de la corrélation négative entre INSAT et TB.

Tableau 8 : Corrélations génétiques entre AG saturés et insaturés

(g/100g lait)	Saturés	Insaturés
Saturés	0,19 à 0,95	
Insaturés	0,16 à 0,78	0,69
(g/100g gras)		
Saturés	0,33 à 0,90	
Insaturés	-0,12 à -0,75	0,10 à 0,66

Les corrélations entre saturés et insaturés sont fortement négatives quand les AG sont exprimés en % du gras, en accord avec Soyeurt et al (2007). Ces corrélations sont affectées par le mode d'expression puisqu'elles sont modérément positives quand les AG sont exprimés en teneur dans le lait.

Figure 1 : Localisation des QTL à l'échelle du génome pour les différents AG. A gauche, AG saturés en race Holstein, A droite, AG insaturés en race Normande.



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'INSAT dans la matière grasse laitière, mais sans changement important dans leur teneur absolue dans le lait.

3. DETECTION DE QTL

3.1. RESULTATS GENERAUX

Les résultats suivants portent sur 31 caractères (acides gras individuels, différents totaux et les trois index de désaturation) exprimés en teneur dans la matière grasse. Les analyses sont réalisées intra race. Deux QTL sont considérés distincts lorsque le seuil LRT de 13 est atteint et lorsque la distance entre deux pics LRT est au minimum de 10 Mb.

Les analyses réalisées permettent la détection d'un grand nombre de QTL, pour tous les caractères. Ainsi sur l'ensemble des trois races, 29 QTL sont détectés pour SAT, 33 pour MONO, 20 pour la somme des polyinsaturés (POLY). Pour les saturés, le nombre tend à diminuer avec la longueur de la chaîne, de 36 QTL pour C4:0 à 13 QTL pour C14:0, avec une remontée à 20 pour le C16:0, qui est la fin de la chaîne. 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 respectivement pour les C14:1cis9, C16:1cis9 et C18:1cis9, 17 pour les CLA, 24 pour les ω 3, 19 pour les ω 6, 12 pour le C18:2cis9cis12.

Les résultats sont généralement très significatifs, le seuil de LRT choisi étant élevé. Même si la méthodologie utilisée (LDLA) n'est pas la plus précise en terme 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 montre une forte colocalisation des QTL (chromosomes 5, 6, 13, 14, 15, 19...) entre les différents AG saturés, reflétant une forte origine métabolique commune. La même remarque est valable pour les AG insaturés (chromosomes 5, 7, 14, 20, 27). Par contre, les régions sont plus souvent différentes entre saturés et insaturés.

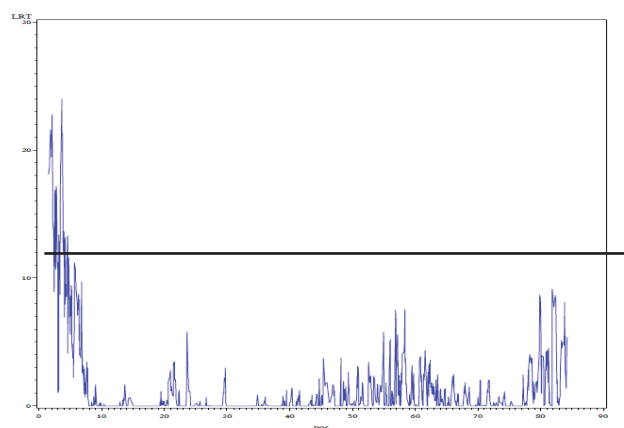
3.2. CONFIRMATION DE REGIONS CONNUES, DECOUVERTE DE NOUVELLES REGIONS MAJEURES

Parmi les nombreuses régions mises en évidence, celles déjà connues parmi les régions très significatives sont fort logiquement retrouvées dans cette étude. Bouwman et al (2011) décrivent trois régions majeures dans la population

Holstein hollandaise sur les chromosomes 14, 19 et 26. Ces trois régions sont retrouvées dans notre étude.

C'est le cas par exemple de l'extrémité centromérique du chromosome 14. Cette région porte le gène DGAT1 qui a un effet majeur sur le taux butyreux, retrouvé dans les trois races françaises (Guillaume, 2009). On montre ici que l'allèle favorable au TB est aussi favorable à l'ensemble des AG saturés. A titre d'exemple, la figure 2 montre le profil observé pour un acide gras intermédiaire, C8:0, en race Holstein. Le profil relativement « haché » est spécifique des analyses LDLA, reflétant la composante déséquilibre de liaison peu conservée à courte distance et à l'origine de la bonne précision de localisation.

Figure 2 : Profil de vraisemblance de C8:0 sur le chromosome 14 en race Holstein.



Le chromosome 26 montre un pic très net en race Normande pour différents AG insaturés, à la position du gène SCD1 qui code pour la delta9-désaturase, une enzyme de désaturation des AG exprimée dans la mamelle. Bouwman et al (2011) avaient déjà montré l'effet majeur de ce gène sur le taux d'insaturés. Sur le taux butyreux global, l'effet de ce gène n'est pas confirmé, même si le chromosome 26 est globalement très important dans le déterminisme de la quantité de matière grasse et du TB (Guillaume, 2009) particulièrement en race Holstein.

De même, plusieurs régions sont mises en évidence sur le chromosome 19 portant plusieurs gènes cités en tant que candidats comme GH (codant pour l'hormone de croissance), ACLY (codant pour l'ATP-citrate synthase) ou FASN (codant pour la fatty acid synthase) mais dont l'implication n'est pas confirmée jusqu'à présent. Le chromosome 19 a un rôle majeur sur les quantités produites (lait et matière) mais son effet sur le taux butyreux est plus faible, et uniquement en Holstein (Guillaume, 2009).

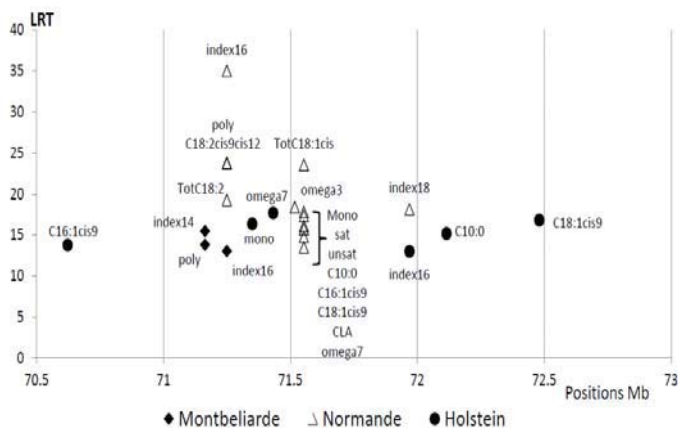
Bouwman et al (2011) mentionnent aussi des régions sur les chromosomes 13, 17 et 27 que nous retrouvons dans notre étude. Le chromosome 27 était déjà connu pour affecter le TB en race Holstein, ce qui n'est pas le cas pour les deux autres chromosomes.

Mais l'étude PhénoFinlait met en évidence d'autres régions qui semblent très importantes pour le déterminisme des AG dans le lait, par exemple sur les chromosomes 5 et 16 que nous présentons plus en détail (figures 3 et 4).

3.3. COLOCALISATION ENTRE CARACTERES

Comme indiqué précédemment, les colocalisations sont fréquentes entre caractères. Ainsi, par exemple, sur le chromosome 16, des QTL sont détectés pour 15 caractères entre les positions 68.98 et 71.96 Mb et la plupart sur un intervalle de 500 kbases, et parfois dans les trois races (figure 3). Il est probable que certains de ces QTL correspondent au même gène. Parmi les gènes présents dans cette région du génome, aucun n'apparaît comme impliqué de façon évidente dans le déterminisme des acides gras et une cartographie plus précise sera nécessaire.

Figure 3 : Colocalisation de QTL sur le chromosome 16 dans la région [70,5 – 73 Mbases].

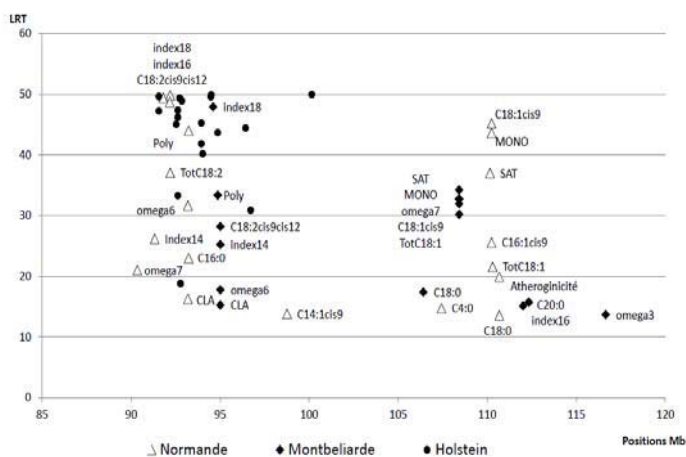


De même, le chromosome 5 porte les QTL les plus nombreux et significatifs (figure 4). Il s'agit d'une région nouvelle, jamais mentionnée jusqu'à présent. Ce chromosome porte deux régions proches mais distinctes impliquées dans le déterminisme des AG, l'une autour de 92-95 Mbases partagée entre les trois races, l'autre autour de 112-115 Mbases partagée entre Normande et Montbéliarde mais apparemment pas par la Holstein. Dans la première région, le gène PIK3C2G codant pour la phosphatidylinositol-4-phosphate 3-kinase est un gène candidat attractif qui va faire l'objet d'investigations complémentaires. La seconde région est particulièrement riche en gènes, potentiellement impliqués dans le métabolisme des lipides, mais il ne ressort pas de gène candidat évident.

Ces colocalisations 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, du fait de leur forte redondance biologique. 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.

Figure 4 : Colocalisation de QTL sur le chromosome 5 dans les régions [91-97 Mb] et [107-112 Mb].



3.3. COLOCALISATION 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 du généralement à plusieurs raisons : une puissance de détection insuffisante de sorte que des QTL communs ne sont pas détectés dans toutes les races ; mais aussi des QTL fixés ou rares dans certaines populations.

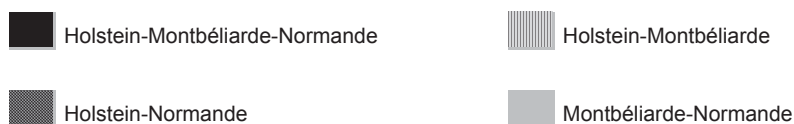
Dans le cas présent, on observe au contraire un taux de colocalisation relativement élevé, comme le montre le tableau 9. Dans ce tableau qui représente une partie des caractères analysés, les QTL sont considérés communs entre races si leurs pics sont distants de moins de deux mégabases. Ainsi, des QTL 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 rajoutées si on ne considère seulement qu'un ou deux caractères.

Il n'en demeure pas moins qu'un nombre important de résultats restent spécifiques de race. Le cas le plus visible est la delta9-desaturase (chromosome 26) dont l'effet n'est clairement visible qu'en race Normande. Globalement, plus de la moitié des QTL détectés ne sont pas trouvés 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.

Tableau 9 : Bilan des colocalisations de QTL entre races pour une partie des AG analysés.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
C4:0		■									■		■				■			■									
C6:0									■				■	■			■			■									
C8:0					■					■			■				■			■		■							
C10:0													■						■										
C12:0													■						■										
C14:0																			■										
C16:0														■														■	
C20:0																			■	■									
C14:1cis9											■															■			
C16:1cis9					■								■		■	■				■								■	
C18:1cis9					■								■	■		■				■								■	
w3	■				■								■	■													■		
w6					■								■	■													■		
index14					■								■	■													■		
index16					■								■	■													■		
index18					■								■	■													■		
CLA					■								■	■															
C18:2c9c12					■								■	■														■	
TotC18:1					■								■	■														■	



DISCUSSION ET 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 du lait de vache. Le projet a d'abord permis de développer des équations de prédiction de la composition en AG à partir des spectres MIR, ce qui ouvre la voie à leur obtention à très grande échelle dans le cadre du contrôle laitier. Le recalage des spectromètres reste un obstacle important mais il est maîtrisable, tout au moins pour les aspects génétiques où on s'intéresse aux différences intra troupeau. Les paramètres génétiques obtenus valident ces équations.

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 artefactuelles 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 paramètres génétiques estimés sont comparables entre races. Cette conclusion n'est pas étonnante, elle reflète l'homogénéité du déterminisme génétique d'un caractère donné dans les différentes populations et les conclusions opposées sont plutôt les exceptions. Ils sont également comparables à ceux de Soyeurt et al (2007), la bibliographie étant encore rare sur ce sujet. Ces paramètres génétiques confortent la bonne précision des équations de prédiction. Les héritabilités à l'échelle de chaque contrôle sont modérées et donc relativement élevées au niveau de la lactation complète combinant plusieurs contrôles, d'autant que le déterminisme génétique reste très homogène en cours de lactation. Les AG sont nombreux mais sont souvent fortement corrélés entre eux de sorte que l'objectif de sélection devrait être relativement simple, dès lors que la direction générale est bien définie. D'une façon générale, si l'objectif est d'augmenter la fraction insaturée de la matière grasse totale, cela devrait induire une

baisse du taux butyreux. Si l'objectif est d'augmenter la teneur d'AG insaturés dans le lait (indépendamment du taux d'AG saturés), alors les conclusions sont sensiblement différentes. Le choix de l'objectif de sélection est relativement urgent et il est souhaitable qu'il soit en accord avec les politiques prochaines de paiement du lait.

Les QTL mis en évidence sont très clairs. Le niveau de signification choisi garantit l'absence de faux positif, même s'il limite la puissance et donc le nombre de QTL détectés. La colocalisation entre QTL est également une indication de la bonne fiabilité de la détection.

La part réelle de variance expliquée par chaque QTL reste à confirmer. Comme c'est fréquemment le cas lors d'une phase de détection mono-QTL, les parts de variance sont élevées (10 à 15 % pour les plus grands effets) mais sont fortement surestimées. Une estimation conjointe de l'effet de l'ensemble des QTL sera prochainement réalisée, nous permettant d'estimer la part totale de variance génétique estimée et celle attribuable à chaque QTL.

Les régions mises en évidence sont fréquemment pléiotropes, agissant sur plusieurs caractères. C'est sans doute vrai pour les AG issus de la même voie métabolique. C'est plus surprenant pour d'autres, mais cela s'explique sans doute par la définition des caractères, induisant des corrélations entre AG dès lors qu'ils ont une teneur non négligeable. Une analyse complémentaire sera utile avec des AG exprimés en teneur dans le lait.

A titre d'exemple, DGAT1 est connu pour fortement moduler le taux butyreux. On confirme ici les résultats de Bowman et al (2011) indiquant que cette augmentation de taux butyreux est essentiellement due aux AG saturés dans leur ensemble, quelle que soit la longueur de la chaîne. De même, la région contenant le gène SCD est favorable à l'ensemble des AG insaturés. Mais nous montrons des phénomènes analogues sur les chromosomes 5 et 16, sans connaître les gènes en cause.

Dans la chaîne des saturés, une question reste ouverte. Existe-t-il une variabilité de l'élongation de C14:0 en C16:0. Ce point est à clarifier car il nous semble qu'il s'agisse d'un facteur non négligeable de la teneur en saturés. En effet, nous observons d'une façon générale une relation forte entre taux butyreux, AG saturés, C16:0 et rapport C16:0 / C14:0.

Les régions les plus importantes sont mises en évidence avec une précision tout à fait satisfaisante, de l'ordre du cM ou du million de bases. Cela ouvre des perspectives très intéressantes de recherche des gènes en cause, même si ce travail n'a que tout juste débuté. Par ailleurs, l'effort de reséquençage en cours concernant de nombreux taureaux des trois races, sera une ressource essentielle pour l'identification des mutations responsables.

Enfin, la population PhenoFinLait constitue une première base de référence pour la mise en place d'une sélection génomique. Avec environ 2500 femelles par race, c'est une population encore insuffisante (le CD attendu n'étant que de l'ordre de 0,3), mais on peut supposer qu'avec la généralisation du stockage et de l'utilisation des spectres MIR et l'extension du génotypage des femelles, on disposera rapidement d'une population beaucoup plus grande, permettant des prédictions précises des valeurs génétiques dans ces trois races.

Le projet PhenoFinLait est cofinancé par l'ANR, Apis-gène, FGE, FranceAgriMer, le Ministère de l'Agriculture et l'Union Européenne. Les auteurs remercient les partenaires du consortium PhenoFinLait, les acteurs de terrain et les éleveurs. Concernant la partie bovine, ce travail a été conduit dans le cadre de l'UMT 3G.

Bouwman AC, Bovenhuis H, Visker MHPW, van Arendonk, JAM. 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 PL, Guerin JL, Journaux L, Krychowski T, Lagriffoul G, Larroque H, Lecomte C, Leray O, Leroux C, Leverrier C, Martin P, Mattalia S, Miranda G, Palhiere I, Peyraud JL et Boichard D. 2009. PhenoFinLait : un programme national français de détection de QTL et/ou de gènes majeurs affectant la composition fine du lait des ruminants laitiers. *16^{èmes} Renc. Rech. Rum.*, Paris, France, Decembre 2-3, 423.

Chilliard, Y., Glasser, F., Enjalbert, F., Ferlay, A., Bocquier, F. and Schmidely, P. 2007. Données récentes sur les effets de l'alimentation sur la composition en acides gras du lait de vache, de chèvre et de brebis. *14^{èmes} Renc. Rech. Rum.*, Paris, France, 321-328.

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. *European Journal of Lipid Science and Technology*, 109, 828-855.

Ferrand M, Huquet B, Barbey S, Barillet F, Faucon F, Larroque H, Leray O, Trommenschlager J-M and 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

Guillaume F., 2009. Intégration de l'information moléculaire dans l'évaluation génétique. Thèse AgroParisTech, 145p.

Leray O, Ferrand M, Larroque H, Astruc JM, 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, Bourgen-Bresse, France, June 22-24.

Lucas, A., Rock, E., Chamba, J. F., Verdier-Metz, I., Brachet, P., Coulon, J. B. 2006. Respective effects of milk composition and the cheese-making process on cheese compositional variability in components of nutritional interest. *Lait* 86, 21-41.

Meuwissen, THE, Goddard, ME. 2001. Prediction of identity by descent probabilities from marker-haplotypes. *Genet Sel Evol*, 33, 605-634

Meuwissen THE, Goddard ME. 2001. Prediction of identity by descent probabilities from marker-haplotypes. *Genet Sel Evol*, 33, 605-634

Meyer K. 2006. Wombat – Digging deep for quantitative genetic analyses by restricted maximum likelihood. *Proc. 8th World Congr. Genet. Appl. Livest. Prod.*, 27-14.

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.

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. *J. Dairy Sci*, 91, 385-394