Determination of fatty acid composition in milk of individual animals by Fourier-Transform Mid-Infrared Spectrometry

Marion FERRAND
Institut de l'Elevage (IE) – Service Biométrie

B. HUQUET, M. BROCHARD, F. FAUCON (IE),
O. LERAY (Actilait)
M.R. AUREL (UE La Fage), S. BARBEY (UE Pin-au-Haras),
F. BOUVIER (UE Bourges), J.M. TROMMENSCHLAGER (UE Mirecourt)
F.BARILLET, H. CAILLAT, H. LARROQUE, R. LEFEBVRE,
I. PALHIERE (INRA-SAGA et GABI)

www.phenofinlait.fr

phenofinlait@inst-elevage.asso.fr

PhénoFinlait

Outline

Context and motivations

Material and methods

Results

Conclusions and perspective

Context

- Consumers are aware of the food impact on their health, especially FA
- In France, more and more farmers are paid on the FA composition of their milk

But…

⇒ No reference method to routinely analyze milk FA composition

⇒ No tool (animal genetic and feeding strategy) to adapt fine milk composition to consumers demand

Phénofinlait: How ?

Partners from every rings of dairy industry

CNIEL (Dairy industries and farmers)

France Génétique Elevage
(France Livestock Genetics)

UNCEIA, ANIO and CapGène
(about 10 breeding companies)

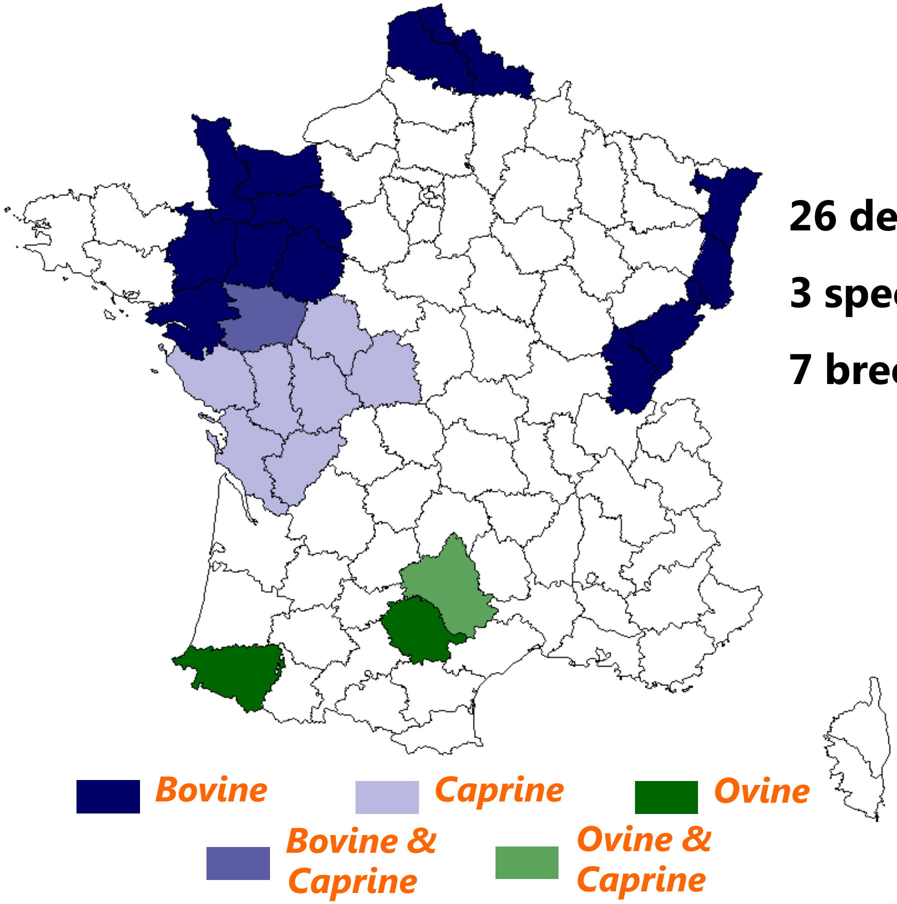
FCL and CNBL
(Milk recording organizations)

Actilait and regional laboratories

INRA
(French National Institute for Agricultural Research)
(4 labo, 4 exp. units, 2 dép.)

Institut de l'Elevage
(French Livestock Institute)
(4 teams, 3 dép.)

And about 1500 farms !



26 departments

3 species,

7 breeds

With the collaboration of Foss and Bentley Instrument / Milk analyzer

PhénoFinLait: aims

- **Develop and control methods to analyze fine milk composition**
- High scale phenotyping and genotyping → 20 000 animals in 1500 farms with fine milk composition, detailed feeding and genotyping
- Understand how genetic and feeding strategies impact fine milk composition
- Create tools (genetics + feeding strategies) to face evolving consumer demands including health requirements

Outline

Context and motivations

Material and methods

Results

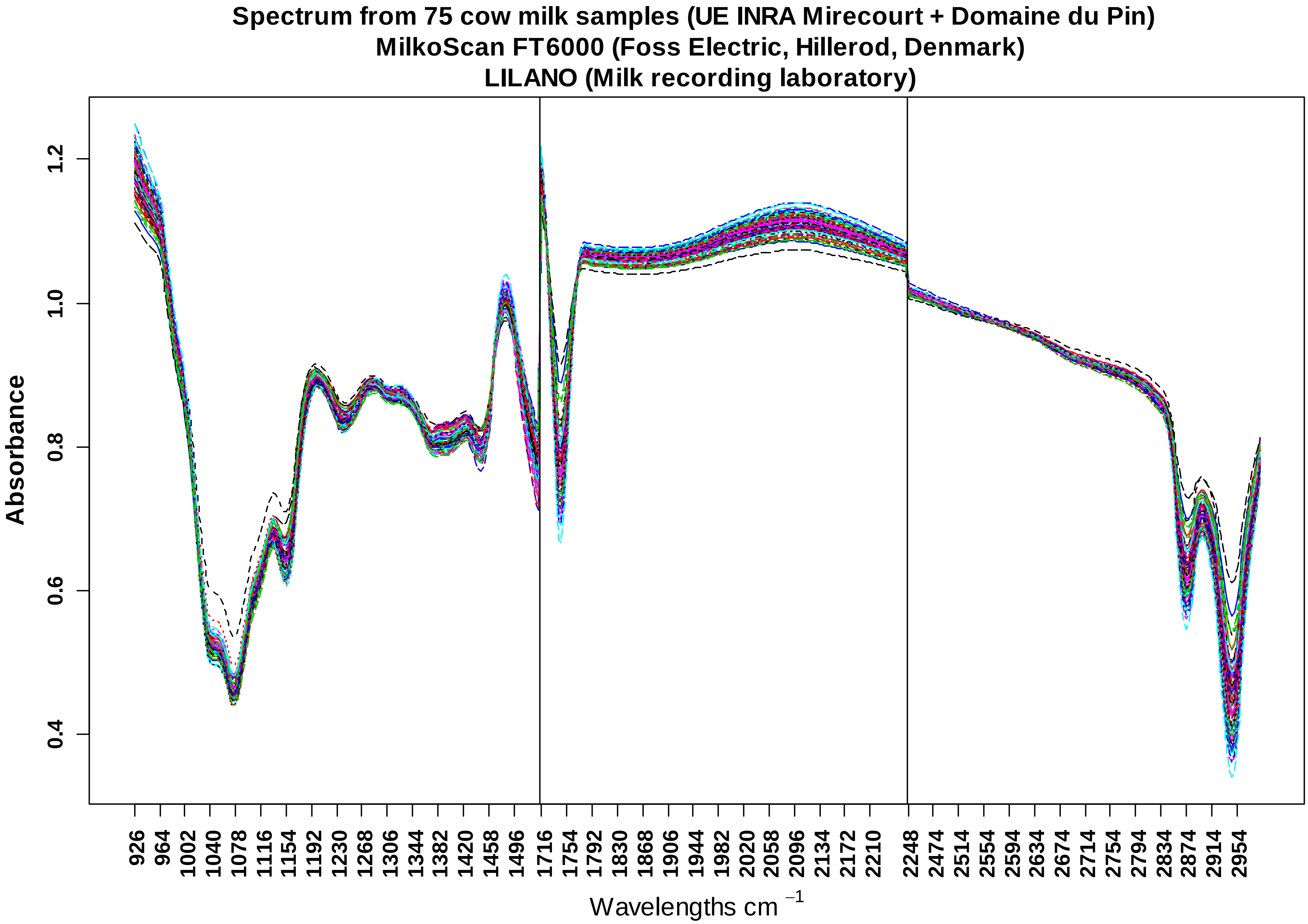
Conclusions and perspective

How to measure the fine composition in routine?

- MIR spectra routinely obtained by milk recording laboratories for fat and protein percentage measurements
- Possibility to estimate several fatty acids, lactoferrine and some minerals in cow milk
  (Université de Gembloux - Soyeurt et al.)

How to estimate fine milk composition from spectra ?

- Analytical technique based on radiation absorption by molecules
- Linked to the chemical composition : bonds between atoms absorbs infrared energy at characteristic frequencies.
  - ➢**MIR-Spectrum reflect the biochemical composition**



Spectrum from 75 cow milk samples (UE INRA Mirecourt + Domaine du Pin)
MilkoScan FT6000 (Foss Electric, Hillerod, Denmark)
LILANO (Milk recording laboratory)

Retained wavelengths :
964-1543 cm^{-1}, 1716-2273 cm^{-1}, 2435-2971 cm^{-1} (Manuel technique Foss)

How to estimate fine milk composition from spectra ?

- Problem : peaks don't match to fatty acids but to bonds
- Use of PLS regression : allow to take into account the collinearity in the spectrum and to extract the information

➢**Coefficients are applied to wavelentghs to estimate the fatty acids**

How to improve equations accuracy ?

- Several authors have suggested **to apply a selection of variables before PLS regression** to improve results (Leardi 1998, Hoskuldsson 2001)
- Genetic algorithms already successfully used on IR data (Leardi R. 1998, Gomez-Carracedo 2007)
- Previous study in cow milk with good results (Ferrand, 2009)

Samples analyzed by CPG and MIR

- **Cow:** Pin-Au-Haras, Mirecourt and Méjusseaume experimental farms
  Crossed Holstein X Normande (F2) (150+100), Prim'Holstein and Montbéliard (100), Holstein (96)
- **Goat:** Domaine de Galles farm for Alpine goat (150) + 1 private flock for Saanen goat(50 à 80)
- **Ewe:** Domaine de La Fage for Lacaune ewe (150) + 3 private flock for Manech red faced (30) and Basco-Bearnaise (20)

→ **Almost 75 FA and ratio for each species**

Outline

- Context and motivations
- Material and methods
- Results
- Conclusions and perspective

Results in cow milk

Validation dataset Mirecourt+le Pin	N	Mean	Sd	Relative error (%)	R^2
Fat content	70	3,816	0,637	0,32	1,00
C4:0	72	0,149	0,025	5,71	0,88
C6:0	70	0,087	0,015	3,97	0,95
C8:0	70	0,050	0,010	5,00	0,94
C10:0	71	0,111	0,029	6,92	0,93
C12:0	71	0,126	0,037	11,12	0,86
C14:0	72	0,435	0,088	6,10	0,91
C16:0	71	1,271	0,282	6,41	0,92
C18:0	71	0,342	0,099	12,58	0,81
Total 18:1	69	0,780	0,203	6,70	0,93
Saturated	72	2,766	0,510	2,09	0,99
Monounsat.	69	0,889	0,220	5,80	0,95
Polyunsat.	69	0,107	0,019	8,06	0,80
Omega 3	70	0,029	0,010	16,24	0,77
Omega 6	70	0,075	0,016	11,23	0,72

INSTITUT DE L'ELEVAGE

Estimations of ratios with techno functional or nutritional interests

Validation dataset Mirecourt+le Pin	N	Mean	Sd	Relative error (%)	R^2	Remarks on residuals
index 14	72	7,05	1,461	19,66	0,11	bad fitting
index 16	70	4,31	1,013	12,43	0,72	
index 18	72	67,0	4,100	4,72	0,41	bad fitting
LA/ALA	73	2,79	1,804	18,71	0,92	2 blocks of data; error more important when the ratio is high
OMEGA 6 / OMEGA 3	73	2,92	1,522	19,28	0,86	2 blocks of data; error more important when the ratio is high
C18:1/C16:0	71	0,68	0,252	17,28	0,79	
Atherogenicity index	73	3,17	0,986	10,35	0,89	

Results in ewe milk

Validation dataset	N	Mean	Sd	Relative error (%)	R2
Fat content	54	6,802	1,398	0,40	1,00
C4:0	52	0,233	0,035	5,88	0,85
C6:0	54	0,177	0,033	4,21	0,95
C8:0	54	0,175	0,037	4,83	0,95
C10:0	54	0,574	0,147	5,90	0,95
C12:0	54	0,339	0,103	8,57	0,92
C14:0	54	0,821	0,214	6,98	0,93
C16:0	54	1,650	0,345	6,70	0,90
C18:0	55	0,511	0,143	12,62	0,80
Total 18:1	54	1,276	0,414	4,40	0,98
Saturated	54	4,825	0,994	2,31	0,99
Monounsat.	54	1,389	0,443	3,83	0,99
Polyunsat.	55	0,238	0,075	7,03	0,95
Omega 3	52	0,069	0,016	13,65	0,66
Omega 6	55	0,137	0,036	12,13	0,79

INSTITUT DE
L'ELEVAGE

Results in goat milk

Cross-validation	N	Mean	Sd	Relative error (%)	R^2
Fat content	150	3,310	0,666	0,48	1,00
C4:0	150	0,092	0,025	9,23	0,87
C6:0	150	0,078	0,020	8,97	0,87
C8:0	150	0,080	0,022	12,36	0,78
C10:0	150	0,264	0,071	12,48	0,77
C12:0	150	0,134	0,041	13,36	0,79
C14:0	150	0,307	0,077	9,17	0,85
C16:0	150	0,996	0,197	5,14	0,93
C18:0	150	0,282	0,099	18,14	0,73
Total 18:1	150	0,756	0,176	8,84	0,85
Saturated	150	2,351	0,485	3,55	0,97
Monounsat.	150	0,798	0,184	8,92	0,84
Polyunsat .	150	0,128	0,028	12,47	0,65
Omega 3	150	0,100	0,031	19,15	0,58
Omega 6	150	0,018	0,005	19,58	0,44

Outline

Context and motivations

Material and methods

Results

Conclusions and perspective

Conclusions

- Ambitious multispecies program with a lot of stakes
- Importance to produce robust and accurate equations to estimate milk FA content
- Accurate estimations for saturated fatty acids and some monounsaturated
- It needs to improve accuracy for polyunsaturated fatty acids

Conclusions

Contributions of genetic algorithms

- Accuracy gain of 15% on average
- Notable improvement for FA of a crucial interest regarding human nutrition
- Stabilization of the equations over the time

Prospects

- Validation of the equations by an external organism
- A priori, possibility to work with multispecies equations

Projects

- Estimation of C18:1 trans 10/ C18:1 trans 11 ratio
- Estimations of protein content → 20 000 milks analyzed both MIR and HPLC-MS method
- Improvement of equations by wavelet use

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
MINISTÈRE
DE L'ALIMENTATION
DE L'AGRICULTURE
ET DE LA PÊCHE

APIS-GENE

INSTITUT DE
L'ELEVAGE

INRA

Actilait

The authors thank Apis-Gène and the French Ministry of Agriculture for the financial support of this part of PhénoFinlait programme and every partners of this project

Thank you for your attention !

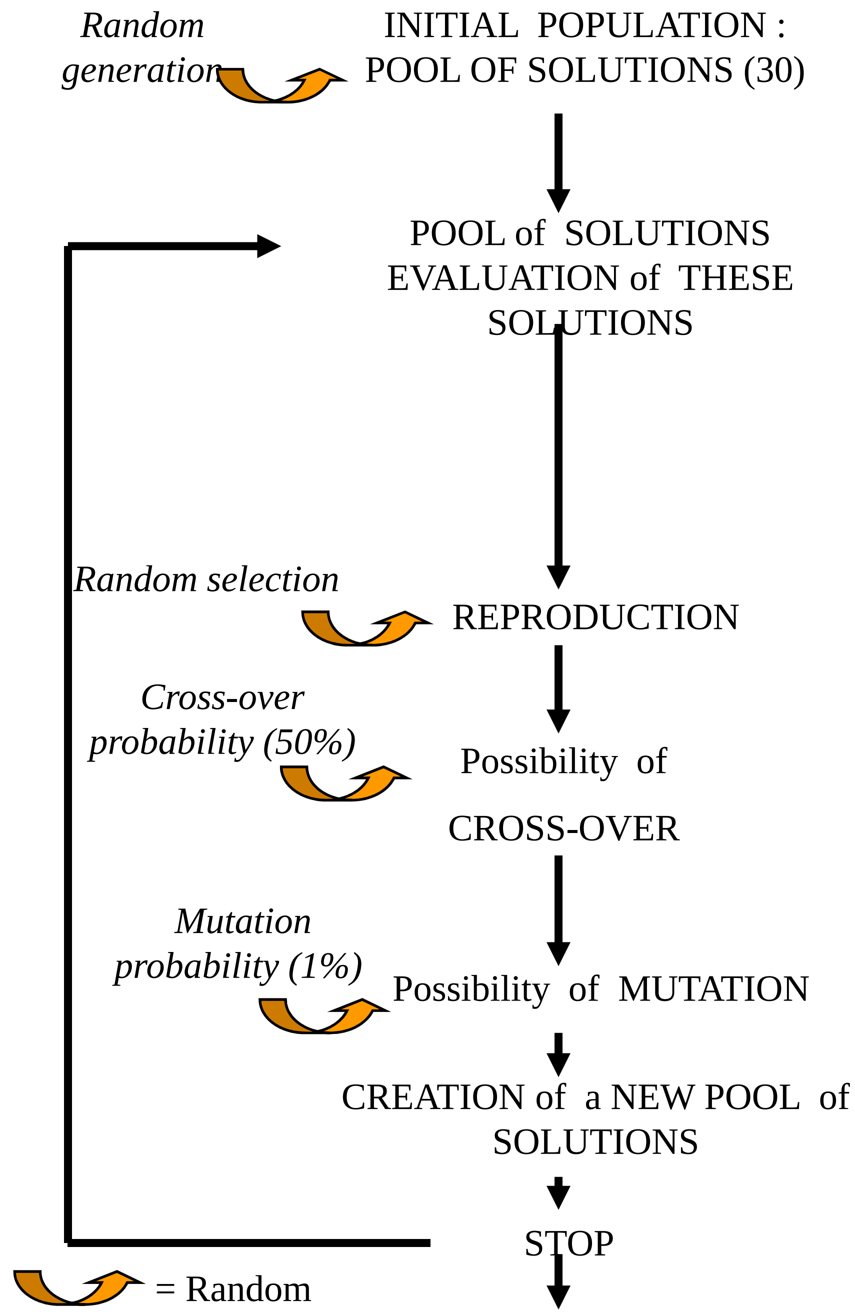
www.phenofinlait.fr

phenofinlait@inst-elevage.asso.fr

PhénoFinlait

Genetic algorithms method

- Based on evolutionary biology
- **Principle**: evolution of a population of solutions using genetic operators like reproduction. mutation and selection
- **Objective**: obtain a population with the best solutions



Random generation

INITIAL POPULATION : POOL OF SOLUTIONS (30)

N solutions generated at random

POOL of SOLUTIONS EVALUATION of THESE SOLUTIONS

Evaluation

	Var1	Var2…		Var446	$R2_{CV}$
Solution 1	1	1	…	1	
Solution 2	1	0	…	1	
…					
Solution N	0	1	…	0	

Variable i takes value of 1 if selected , else 0. $R2_{CV}$ is obtained by PLS regression on selected variables.

Random selection

REPRODUCTION

Selection of 2 solutions
The better a solution is, the highest the probability of being chosen is

Cross-over probability (50%)

Possibility of CROSS-OVER

Combination of 2 solutions
Objective : to obtain 2 better solutions
Limit : variability of solutions decreases

Mutation probability (1%)

Possibility of MUTATION

Each variable has a mutation probability of x% (1 no selected variable become selected and conversely)
Objective : avoid having a pool of uniform solutions

CREATION of a NEW POOL of SOLUTIONS

Substitution of the 2 worst solutions by new solutions

STOP

When quality of solutions is constant, algorithm is stopped.

FINAL RESULT

Getting N solutions among the bests

= Random

adapted from Haupt (2004) and Leardi (1998)

References

Haug A. Bovine milk in human nutrition – a review. *Lipids in Health and Disease* 2007. **6**:25. 2007.

Hoskuldsson A. Variable and subset selection in PLS regression. *Chemometrics and Intelligent Laboratory Systems.* 55 :23-38. 2001.

Leardi R. and Lupiañez G. Genetic algorithms applied to feature selection in pls regression : how and when to use them. *Chemometrics and Intelligent Laboratory Systems.* 41:195-208.1998.

Legrand P. Interêt nutritionnel des principaux acides gras des lipides du lait. *Cholé-doc.*105:1-4. 2008.

Schennink and al.. Genome-wide scan for bovine milk-fat composition**.** *J. Dairy Sci.* 92 :4676–4682. 2009.