

# Bilan du programme de Sélection Assistée par Marqueurs dans les trois principales races bovines laitières françaises et perspectives d'évolution

FRITZ S. (1), DRUET T. (2), GUILLAUME F. (3), MALAFOSSE A. (1), BOSCHER M.Y. (4), EGGEN A. (5), GAUTIER M. (5), COLLEAU J.J. (2), BOICHARD D. (2).

(1) Union Nationale des Coopératives agricoles d'Elevage et d'Insémination Animale, Service Génétique, 149 rue de Bercy, 75595 Paris cedex 12, France

(2) INRA Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas cedex, France

(3) Institut de l'Elevage, 149 rue de Bercy, 75595 Paris cedex 12, France

(4) GIE LABOGENA, 78352 Jouy-en-Josas cedex, France

(5) INRA Laboratoire de Génétique Biochimique et Cytogénétique, 78352 Jouy-en-Josas cedex, France

**RESUME** – Différents types de programmes de sélection assistée par marqueurs (SAM) peuvent être mis en œuvre en fonction du degré de connaissance des gènes impliqués dans le déterminisme génétique des caractères. Le programme français de SAM mis en place depuis 2001 chez les bovins laitiers est un programme de première génération (SAM1) s'appuyant sur les premiers résultats de détection de QTL (région chromosomique ayant un effet significatif sur un caractère quantitatif) obtenus entre 1996 et 1999. Les efforts de génotypages entre 2001 et 2007 ont permis d'une part de confirmer l'existence des QTL initialement choisis et d'autre part d'augmenter la précision des index des jeunes animaux : évolution moyenne des CD de 0,18 pour la fertilité ou 0,33 pour la quantité de lait en sélection classique (sans marqueurs) à respectivement 0,29 et 0,44 en SAM. Ainsi les entreprises de sélection peuvent espérer des suppléments de progrès génétique par une utilisation de la SAM dans les programmes de sélection. A partir de 2008, les résultats de cartographie fine de QTL devraient permettre au programme d'évoluer vers un programme de seconde génération (SAM2) et de gagner encore en efficacité grâce à l'utilisation de marqueurs très proches des QTL sélectionnés. Dans le même temps, l'intérêt et l'efficacité de la sélection génomique seront étudiés.

## Assessment of Marker-Assisted Selection in the three main French breeds of dairy cattle and future developments

S. FRITZ (1), T. DRUET (2), F. GUILLAUME (3), A. MALAFOSSE (1), M.Y. BOSCHER (4), A. EGGEN (5), M. GAUTIER (5), J.J. COLLEAU (2), D. BOICHARD (2).

(1) Union Nationale des Coopératives agricoles d'Elevage et d'Insémination Animale, Service Génétique, 149 rue de Bercy, 75595 Paris Cedex 12, France

**SUMMARY** – Several types of Marker-Assisted Selection (MAS) programs can be implemented according to the knowledge of the genes involved in genetic variation. Since 2001, the French MAS program in dairy cattle is a first generation of MAS (MAS1) based on first QTL (chromosomal regions with a significant effect on a quantitative trait) detection results obtained between 1996 and 1999. On the one hand, additional genotypes between 2001 and 2007 were used to confirm previously selected QTL. On the other hand, they improve MAS-EBV of young animals: in average, the reliabilities for fertility or milk yield increased from 0.18 and 0.33 to 0.29 and 0.44, respectively. Breeding companies can expect higher genetic progress by using MAS in their breeding schemes. In 2008, QTL fine mapping results will be used to change the MAS program to a second generation program (MAS2) which will be more efficient due to the use of closely linked markers. At the same time, the efficiency and the interest of genomic selection will be studied.

## INTRODUCTION

Depuis 2001, le programme français de sélection assistée par marqueurs (SAM) mené par l'INRA, LABOGENA et l'UNCEIA (Boichard *et al.*, 2002, Fritz *et al.*, 2003) dans les trois principales races bovines laitières françaises (Montbéliarde, Normande et Prim'Holstein) a permis le génotypage de près de 60 000 animaux. Les marqueurs microsatellites (au nombre de 45 depuis 2005) utilisés dans le cadre du programme permettent de suivre 14 régions chromosomiques de génération en génération. L'introduction

de ces informations moléculaires dans l'évaluation génétique a pour objectif d'optimiser les choix des jeunes animaux dans les étapes précoces de sélection afin d'obtenir des suppléments de progrès génétique vis-à-vis de la sélection classique. Les entreprises reçoivent ainsi mensuellement les index SAM des jeunes mâles candidats au testage et des jeunes femelles candidates mères à taureaux pour huit caractères élémentaires, l'INEL (Index Economique Laitier) et un index de synthèse en accord avec les objectifs de sélection.

## 1. LES DIFFERENTS TYPES DE SAM

Différents types de programme de sélection assistée par marqueurs (SAM) peuvent être mis en œuvre en fonction du degré de connaissance des gènes impliqués dans le déterminisme des caractères.

### 1.1. LA PREMIERE GENERATION DE SAM (SAM1)

Dans les premières années de recherches visant à identifier les gènes impliqués dans le déterminisme d'un caractère, l'utilisation d'un nombre limité de marqueurs moléculaires permet de détecter des régions chromosomiques de relativement grande taille (contenant plusieurs centaines de gènes) associées statistiquement au caractère étudié : on parle de QTL (*Quantitative Trait Locus*). Un premier type de SAM (SAM1) peut s'appuyer sur ces premiers résultats où les positions des QTL restent imprécises. Dans un tel programme, les marqueurs ne servent qu'à tracer la transmission des QTL entre générations. D'importants efforts de typages sont donc nécessaires chez les ancêtres ou apparentés des jeunes animaux pour estimer les effets des QTL. Un programme de SAM1 reste intéressant s'il peut générer un supplément de progrès génétique qui compense au moins les coûts liés aux génotypages des animaux.

### 1.2. LA SECONDE GENERATION DE SAM (SAM2)

Des travaux de cartographie fine de QTL utilisant une forte densité en marqueurs permettent de resserrer les positions des QTL dans des régions contenant une dizaine de gènes au maximum. L'utilisation de marqueurs très proches de ces positions permet d'envisager une seconde génération de SAM : la SAM2. Pour chaque région, les animaux ayant hérité d'une combinaison allélique donnée de marqueurs (haplotype de marqueurs) ont une très grande probabilité d'avoir hérité du même allèle au QTL. Les estimations des effets des QTL ne sont plus réalisées intra familles mais globalement au niveau de la population. Cela signifie que le nombre d'effets à estimer est nettement moins important et que la précision de l'estimation est la même pour tout candidat ayant hérité d'un même haplotype donné de marqueurs quelle que soit sa famille d'origine. La difficulté n'est plus alors de suivre la transmission des régions chromosomiques entre générations mais bien de reconstituer correctement les deux haplotypes de marqueurs situés sur chacun des deux chromosomes du candidat. Dans ces conditions, la SAM2 gagne en simplicité et en efficacité, plus particulièrement dans les familles peu représentées et intéressantes pour le maintien de la diversité génétique.

### 1.3. LA SELECTION GENOMIQUE

La baisse récente des coûts de génotypages et l'utilisation d'un nombre élevé de marqueurs permettent également d'envisager un dernier type de SAM : la Sélection Génomique. En s'appuyant sur un nombre élevé de marqueurs couvrant l'intégralité du génome, la sélection génomique cherche à estimer le niveau génétique global d'un jeune animal en estimant simplement les effets de chaque marqueur utilisé. Ce type de SAM n'utilise donc plus a priori la notion de QTL et permettrait de simplifier grandement l'évaluation génétique assistée par marqueurs si son efficacité est avérée. Cette efficacité sera toutefois très dépendante du panel de marqueurs retenus défini sur un échantillon d'ancêtres donné, elle devra être vérifiée par des études en conditions réelles comme celles menées sur la SAM1 et sur la SAM2.

Lorsque les différents polymorphismes d'un gène impliqué dans la variabilité d'un caractère sont identifiés, le gène lui-

même peut servir de marqueur au programme. Cependant, il convient de ne pas oublier que cette situation favorable devra se coupler en général à un programme de SAM2 ou de sélection génomique pour la prise en compte de tous les autres gènes intervenant sur la variabilité du caractère sélectionné.

## 2. BILAN DU PROGRAMME FRANÇAIS DE SAM CHEZ LES BOVINS LAITIERS (2001-2007)

Depuis 2001, le programme français de SAM chez les bovins laitiers est un programme de SAM1 s'appuyant sur les résultats du programme français de détection de QTL : 1554 taureaux issus de 14 pères (9 Prim'Holstein, 3 Normands et 2 Montbéliards) génotypés sur 169 marqueurs (Boichard *et al.*, 2003).

### 2.1. PRINCIPALES EVOLUTIONS

Entre 2001 et 2007, près de 60 000 animaux ont été génotypés dont près de 18 000 jeunes candidats au testage. L'efficacité du programme pour ces jeunes candidats est fortement liée aux évolutions des connaissances sur les QTL et aux efforts de génotypages réalisés à l'intérieur de chaque race.

#### 2.1.1. Evolution des connaissances sur les QTL

Les efforts de génotypages réalisés dans le cadre du programme permettent de vérifier la pertinence des choix initiaux sur de larges échantillons de taureaux pour chacune des 3 races : 1 030 taureaux issus de 19 pères en race Montbéliarde, 1 225 taureaux issus de 24 pères en race Normande et 4 623 taureaux issus de 62 pères en race Prim'Holstein.

**Tableau 1** : Significations des détections de QTL de juin 2007 à l'échelle du chromosome (BTA) en race Montbéliarde (\*\* : 1 % ; \* : 5 % et vide si non significative), confirmation des choix initiaux de 2001 indiquée par un fond gris.

BTA	2	3	6	7	14	19	20	21	26
INEL	**	*	**	*	**	**	*	**	
CEL	*		*						
FER						*			*
DE	**				*				

**Tableau 2** : Significations des détections de QTL de juin 2007 à l'échelle du chromosome (BTA) en race Normande (\*\* : 1 % ; \* : 5 % et vide si non significative), confirmation des choix initiaux de 2001 indiquée par un fond gris.

BTA	1	3	6	7	14	15	20	21	23	26
INEL		**	**		**		*	**	*	**
CEL						*	*	*		
FER	**	**		*						
DPJ			*							**

**Tableau 3** : Significations des détections de QTL de juin 2007 à l'échelle du chromosome (BTA) en race Prim'Holstein (\*\* : 1 % ; \* : 5 % et vide si non significative), confirmation des choix initiaux de 2001 indiquée par un fond gris.

BTA	1	2	3	6	7	14	15	20	21	26
INEL			**	**		**		**		**
CEL					*		**			*
FER	**	*	**							
DPJ		*		**				**	**	

Les résultats de détection de QTL obtenus sur ces échantillons avec le logiciel QTLMAP de l'INRA (tableaux 1, 2 et 3) confirment en grande partie les choix initiaux. On note toutefois plus de différences sur les caractères fonctionnels peu héréditaires (CEL, FER). De nombreux gènes candidats ont été décrits pour expliquer le polymorphisme causal de ces QTL comme ABCG2 sur le chromosome 6 (Cohen-Zinder *et al.*, 2005), DGAT1 sur le chromosome 14 (Grisart *et al.*, 2002) ou encore GHR sur le chromosome 20 (Blott *et al.*, 2003). Des travaux sont en cours à l'INRA pour vérifier leur pertinence dans les populations françaises. Par exemple, la mutation K232A de DGAT semble effectivement avoir un impact important sur les caractères de production laitière en race Prim'Holstein et Normande. En revanche, elle n'est pas la cause du QTL observé à la même position en race Montbéliarde (Gautier *et al.*, 2007).

### 2.1.2. Efforts de génotypage

Dans un programme de SAM1, le suivi des QTL de génération en génération et la précision des estimations des effets des QTL sont des éléments clef de l'efficacité du programme. C'est pourquoi les efforts de génotypage et leur accumulation au cours du temps permettent au programme de gagner progressivement en efficacité. Le pourcentage de candidats ayant au moins 3 grands-parents génotypés était de 45 % pour les veaux nés en 2002 alors qu'il est de 71 % pour les veaux nés en 2006. Dans le même temps, l'estimation des effets des QTL qui s'appuyait respectivement sur 1 600, 1 631 et 4 901 animaux génotypés et avec performances en 2002 en race Montbéliarde, Normande et Prim'Holstein s'appuie en 2007 sur 7 636, 5 310 et 20 021 animaux. La situation est donc incomparablement plus favorable aujourd'hui qu'elle ne l'était en 2001 pour estimer les valeurs génétiques des candidats.

### 2.1.3. Augmentation de la précision des index

En intégrant les QTL dans le modèle de l'évaluation génétique, la SAM permet d'augmenter la précision (CD) des index des jeunes animaux sans performances. Les gains de CD calculés dans le cadre du programme SAM renseignent sur les précisions gagnées grâce au suivi des QTL par rapport à un modèle dit polygénique (sans QTL). Ils sont proches de zéro lorsque les QTL ne sont pas suivis ou que les effets des QTL ne peuvent pas être estimés ; ils sont maximaux lorsque tous les QTL sont suivis et que les

estimations des effets sont précises. De plus, pour un caractère donné et pour une race donnée, les gains de CD sont d'autant plus élevés que la part de variance génétique expliquée par les QTL est importante.

Le tableau 4 présente les gains de CD observés lors de l'évaluation SAM de juillet 2007 pour les jeunes candidats nés en 2006. Pour tous les caractères, ces gains de CD sont quasiment deux fois plus élevés que les gains de CD observés dans les toutes premières années du programme confirmant l'augmentation de son efficacité entre 2001 et 2007.

## 2.2. UTILISATION DE LA SAM DANS LES PROGRAMMES DE SELECTION

### 2.2.1. Utilisation de la SAM en ferme

Pour assurer un gain de progrès génétique vis-à-vis de la sélection classique, les candidats sélectionnés par la SAM doivent posséder plus d'allèles favorables aux QTL que la moyenne des candidats. Les entreprises de sélection se doivent donc d'élargir leurs bases de sélection (listes initiales de candidats) pour repérer ces candidats tout en maintenant d'autres critères de sélection (aptitude bouchère, morphologie, anomalie génétique, etc.). Il est apparu très rapidement que cette condition ne pourrait être remplie que si la SAM était réalisée très précocement, c'est-à-dire en ferme, dans les deux ou trois premiers mois de vie des animaux. De gros efforts d'organisation ont donc été nécessaires entre les différents acteurs pour faire baisser de façon significative les délais entre la naissance d'un candidat et l'évaluation génétique de celui-ci. Plus de 95 % des 1 300 candidats nés en 2006 et pour lesquels un prélèvement de sang est parvenu à LABOGENA dans leur premier mois de vie étaient indexés à trois mois d'âge dans le cadre de la SAM.

### 2.2.2. Gestion de la variabilité génétique

Dans le souci d'optimiser la gestion de la variabilité génétique dans les étapes précoces de sélection (de la naissance à la mise en testage des jeunes mâles) en lien avec le programme SAM, un index de synthèse tenant compte des parentés des animaux (parenté entre candidats et parentés avec les taureaux testés) est distribué mensuellement aux entreprises de sélection (Colleau *et al.*, 2006). Ces index permettent de repérer les meilleurs jeunes de la SAM qui minimisent les parentés tout en respectant les objectifs de progrès génétique, les fils des taureaux déjà très utilisés sont donc désavantagés.

**Tableau 4** : CD polygéniques moyens et gains de CD moyens et maximaux observés lors de l'indexation SAM de juillet 2007 sur 445 candidats Montbéliards (MO), 473 candidats Normands (ND) et 2 074 candidats Prim'Holstein (PH) nés en 2006.

Caractère	MO			ND			PH		
	CD pol	Gain CD SAM		CD pol	Gain CD SAM		CD pol	Gain CD SAM	
		moy	max		moy	max		moy	max
Quantité de lait (Lait)	0,35	0,11	0,19	0,34	0,10	0,21	0,33	0,11	0,23
Quantité de matières grasses (QMG)	0,35	0,13	0,25	0,34	0,09	0,17	0,33	0,13	0,25
Quantité de matières protéiques (QMP)	0,35	0,13	0,25	0,34	0,10	0,18	0,33	0,11	0,23
Taux butyreux (TB)	0,38	0,11	0,21	0,37	0,10	0,19	0,35	0,20	0,39
Taux protéique (TP)	0,38	0,11	0,20	0,37	0,05	0,12	0,35	0,14	0,29
Cellules (CEL)	0,31	0,08	0,17	0,30	0,07	0,15	0,20	0,08	0,19
Fertilité vache (FER)	0,22	0,07	0,15	0,22	0,05	0,12	0,18	0,11	0,22
Distance plancher jarret (DPJ ou DE)	0,36	0,04	0,08	0,33	0,07	0,16	0,34	0,09	0,19

### 2.2.3. Estimation de l'efficacité actuelle de la SAM

Des études scientifiques menées par l'INRA sur les bases de données de la SAM ont mis en évidence que les connaissances sur les QTL et les génotypages accumulés étaient suffisantes pour réaliser un progrès génétique supplémentaire vis-à-vis de la sélection classique sur la série de candidats de 2004 pour les caractères de production laitière (Druet *et al.*, 2006 ; Guillaume *et al.*, 2007). Ces études ont d'ailleurs montré que cette supériorité s'est accentuée entre 2004 et 2006. On peut exprimer ce supplément de progrès génétique par le nombre de taureaux nécessaires à tester dans le cadre de la SAM pour obtenir le même progrès génétique qu'en sélection classique avec 150 taureaux Montbéliards, 130 taureaux Normands et 500 taureaux Prim'Holstein (tableau 5).

**Tableau 5** : Nombre de taureaux testés par race (MO : Montbéliarde, ND : Normande, PH : Prim'Holstein) pour obtenir le même progrès génétique qu'en sélection classique avec l'efficacité de la SAM de 2006 (et réduction en %).

	MO		ND		PH	
	Nb	%	Nb	%	Nb	%
Lait	138	-8,0	109	-16,2	415	-17,0
QMG	130	-13,3	108	-16,9	404	-19,2
QMP	128	-14,7	111	-14,6	438	-12,4
TB	120	-2 0,0	102	-21,5	421	-33,0
TP	131	-12,7	114	-12,3	436	-18,0

Même s'il reste à vérifier l'efficacité de la SAM sur les caractères fonctionnels, les résultats encourageants accumulés depuis 2001 mettent tous en évidence l'intérêt du programme. De plus, le programme mené entre 2001 et 2007 a permis aux différents acteurs de s'organiser, d'accumuler de l'expérience, des connaissances et des données, ce qui va s'avérer essentiel dans les années à venir étant donné l'intérêt grandissant apporté à la SAM et à ses perspectives d'évolutions au niveau international.

### 3. PERSPECTIVES D'EVOLUTION

L'expérience et les connaissances accumulées au cours des sept premières années du programme combinées à l'abondance et au faible coût des marqueurs de type SNP (pour *Single Nucleotide Polymorphism*) permettent une évolution vers une SAM de seconde génération (SAM2) dès 2008.

#### 3.1. LE PROJET CARTOFINE

Le projet CARTOFINE soutenu par l'ANR dans le cadre du GIS AGENAE envisage des travaux de cartographie fine de QTL à partir du génotypage de 3 200 taureaux issus du programme SAM sur une puce de 60 000 marqueurs SNP. A court terme, les analyses prévues début 2008 doivent permettre d'identifier des haplotypes d'une dizaine de marqueurs contenant les QTL sélectionnés dans le cadre du programme pour évoluer vers une SAM2. A terme, ce programme permettra également d'intégrer de nouvelles régions et de nouveaux caractères au programme et d'étudier la faisabilité et l'efficacité de la sélection génomique.

### 3.2. EVOLUTIONS POUR 2008

A partir de 2008, on envisage de sélectionner les QTL en SAM2 à l'aide d'une dizaine de marqueurs SNP situés dans des régions chromosomiques d'environ 1cM (contenant une dizaine de gènes). Dans ces conditions, cette SAM va globalement gagner en efficacité et plus particulièrement dans les familles françaises originales ou dans les familles étrangères peu connues en France.

#### 3.3. ESTIMATION DE L'EFFICACITE DES FUTURES EVOLUTIONS

Parmi les 3 200 taureaux génotypés dans le cadre du projet CARTOFINE, 1 000 jeunes taureaux (600 Prim'Holstein, 200 Montbéliards et 200 Normands) nés entre 2000 et 2002 et testés sur descendance serviront à estimer l'efficacité des évolutions envisagées. Comme ces jeunes taureaux ont déjà été évalués dans le cadre de la SAM1 et que toutes les données de l'époque ont été conservées, il sera possible de se placer dans les conditions réelles de l'époque et comparer l'efficacité de la sélection classique, de la SAM1, de la SAM2 et de la sélection génomique. Ces études rendues possibles grâce au programme actuel permettront aux entreprises de sélection d'utiliser ces évolutions dès 2008 sans avoir besoin d'attendre que les candidats nés en 2008 terminent leur période de testage sur descendance.

### CONCLUSION

L'utilisation des informations moléculaires pour optimiser les choix des jeunes animaux dans les étapes précoces de sélection engendre inévitablement des années de recherche sur le sujet et une accumulation des données pour garantir une efficacité optimale aux entreprises de sélection. En France, on estime que l'utilisation de la SAM dans les schémas permet effectivement aux entreprises de sélection d'en diminuer globalement les coûts tout en assurant la continuité du progrès génétique. En 2008, la SAM devrait une nouvelle fois évoluer avec l'utilisation de SNP très proches des QTL et ainsi gagner considérablement en efficacité. Enfin, dans les années à venir, les résultats du projet CARTOFINE devraient permettre d'ajouter des caractères supplémentaires voire d'évoluer vers la sélection génomique si son efficacité est avérée.

**Blott S., 2003.** *Genetics*, 163, 253-266

**Boichard D. et al., 2002.** 7<sup>th</sup> WCGALP, N°22-03

**Boichard D. et al., 2003.** *Genet. Sel. Evol.*, 35, 77-101

**Cohen-Zinder M. et al., 2005.** *Genome Research*, 15, 936-944

**Colleau J.J. et al., 2006.** *Renc. Rech. Rum.*, 13, 235-238

**Druet T. et al., 2006.** *BTIA*, 120, 11-13

**Fritz S. et al., 2003.** *Renc. Rech. Rum.*, 10, 53-56

**Guillaume F. et al., 2007.** *Genet. Sel. Evol.*, accepté

**Gautier M. et al., 2007.** *J. Dairy Sci.*, 90, 2980-2988

**Grisart B. et al., 2002.** *Genome Research*, 12, 222-231