

Identification de QTL associés à de la mortalité embryonnaire chez les bovins laitiers

FRITZ S. (1), CAPITAN A. (1), DJARI A. (2, 4), GROHS C. (3), BOUSSAHA M. (3), BAUR A. (1), BARBAT A. (3), LEFEBVRE R. (3), ESQUERRE D. (4), KLOPP C. (2), ROCHA D. (3), BOICHARD D. (3)

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

(2) INRA, SIGENAE, UR875 Biométrie et Intelligence Artificielle, 31326 Castanet, France

(3) INRA, UMR 1313 Génétique Animale et Biologie Intégrative, 78352 Jouy-en-Josas, France

(4) INRA, UMR 444 Laboratoire de Génétique Cellulaire, 31326 Castanet, France

RESUME – Des régions du génome portant des mutations récessives délétères ont été détectées dans les trois principales races laitières françaises en recherchant des haplotypes relativement fréquents (>1%) dans la population mais présentant un fort déficit en homozygotes. La détection repose sur l'utilisation des données de typages de la puce BovineSNP50[®] d'Illumina utilisées en sélection génomique (47 878 animaux Holstein, 16 833 Montbéliards, et 11 466 Normands). Trente trois régions incluant les régions déjà décrites associées à Brachyspina, CVM, HH1 et HH3 en race Holstein ont été identifiées. La longueur de ces régions varie de 1 à 4,8 Mb et les fréquences des haplotypes de 1,8 à 9%. Un effet significatif négatif sur la fertilité a été observé pour 10 de ces régions dans les accouplements à risque entre taureaux porteurs et filles de taureaux porteurs, confortant leur association avec des mutations létales à l'état embryonnaire. Des effets significatifs ont également été observés sur la mortinatalité pour 7 haplotypes. Ces régions ont été ensuite analysées avec des données de séquence complète du génome de taureaux porteurs et contrôle (45 animaux au total). Des mutations candidates très fortes ont été mises en évidence, dont trois déjà connues dans les gènes FANCI (Brachyspina), SLC35A3 (CVM), et APAF1 (HH1) et quatre nouvelles mutations avec des effets prédits très délétères sur la structure protéique dans les gènes GART, SMC2, SLC37A2, et NOA1. En conclusion, cette étude montre un aspect jusqu'à présent inconnu de la consanguinité dans les races bovines sélectionnées. Alors qu'une éradication systématique serait extrêmement coûteuse financièrement pour les programmes de sélection, une contre-sélection doit être entreprise pour diminuer progressivement la fréquence de ces mutations. Les plans d'accouplement doivent dès à présent éviter les accouplements à risque.

Identification of QTL associated with prenatal mortality in dairy cattle

FRITZ S. (1), CAPITAN A. (1), DJARI A. (2, 4), GROHS C. (3), BOUSSAHA M. (3), BAUR A. (1), BARBAT A. (3), LEFEBVRE R. (3), ESQUERRE D. (4), KLOPP C. (2), ROCHA D. (3), BOICHARD D. (3)

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

SUMMARY – Genomic regions harboring recessive deleterious mutations were detected in three dairy cattle breeds by identifying frequent haplotypes (>1%) showing a deficit in homozygotes among Illumina BovineSNP50[®] Beadchip haplotypic data from the French genomic selection database (47,878 Holstein, 16,833 Montbéliarde, and 11,466 Normande animals). Thirty-three candidate haplotypes were identified including previously reported regions associated with Brachyspina, CVM, HH1, and HH3 in the Holstein breed. Haplotype length varied from 1 to 4.8 Mb and frequencies from 1.8 up to 9%. A significant negative effect on calving rate was observed for 10 of these haplotypes in matings between carrier bulls and daughters of carrier sires, confirming their association with embryonic lethal mutations. Significant effects were also detected on stillbirth for 7 haplotypes. These regions were further investigated using whole genome sequencing data from heterozygous bull carriers and control animals (45 animals in total). Seven strong candidate causative mutations including polymorphisms previously reported in FANCI (Brachyspina), SLC35A3 (CVM), APAF1 (HH1) and four novel mutations with very damaging predicted effect on the protein structure were detected in GART, SMC2, SLC37A2, and NOA1 genes. In conclusion, this study reveals a yet hidden consequence of the important inbreeding rate observed in intensively selected and specialized cattle breeds. Whereas a systematic eradication would be extremely costly and damaging for the breeding scheme, a gradual counter selection must be undertaken to decrease their frequency and an appropriate mating plan must be carried out to avoid matings at risk.

INTRODUCTION

La fertilité est un problème majeur dans les élevages de bovins laitiers. Dans les trois principales races françaises (Holstein, Montbéliarde et Normande), une baisse continue du taux de conception a été observée entre 1995 et 2005 aussi bien sur les génisses que sur les vaches (Barbat *et al.*, 2005). La sélection sur la production laitière n'explique que la moitié de cette évolution, il existe d'autres facteurs explicatifs. La plupart des races sélectionnées de bovins laitiers sont des populations récentes issues d'un faible nombre d'ancêtres fondateurs. Depuis 60 ans, l'utilisation massive d'un nombre limité de taureaux d'élite via l'insémination artificielle pour diffuser le progrès génétique a contribué à réduire la

variabilité génétique dans chaque race. Ainsi, dans chacune des trois races, 50% des gènes sont expliqués par moins de 10 ancêtres (Mattalia *et al.*, 2006).

Dans ce contexte, l'augmentation annuelle de la consanguinité contribue à l'émergence régulière d'anomalies génétiques récessives, en ségrégation dans les populations, conduisant à une mortalité embryonnaire à l'état homozygote. Leur mise en évidence n'est pas facile, car elles ne se traduisent que par une légère baisse de fertilité.

Profitant de l'essor de la sélection génomique et du nombre important d'animaux génotypés chaque année, VanRaden *et al.* (2011) ont proposé de rechercher les régions du génome relativement fréquentes à l'échelle de la population et présentant un déficit d'individus homozygotes. Il est alors fait

l'hypothèse que ces régions sont des QTL (Quantitative Trait Locus) associés à de la mortalité embryonnaire à l'état homozygote.

La disponibilité en France de larges populations d'individus génotypés sur la puce BovineSNP50[®] d'Illumina dans les trois principales races de bovins laitiers a permis la détection de ce type de QTL. Par ailleurs, le séquençage du génome d'un nombre croissant d'individus permet la recherche des mutations causales correspondantes, facilitant ensuite leur éradication.

1. MATERIEL ET METHODES

1.1 CARTOGRAPHIE DE QTL

1.1.1 Populations étudiées

La cartographie de QTL associés à de la mortalité embryonnaire s'est appuyée sur 47 878 individus Holstein, 16 833 individus Montbéliards et 11 466 individus Normands typés sur la puce BovineSNP50[®] et issus d'un père et d'un grand-père maternel génotypés.

1.1.2 Génotypes

Les génotypes ont été contrôlés dans le cadre des indexations génomiques officielles. Seuls les marqueurs SNP (Single Nucleotide Polymorphism) positionnés sur les 29 chromosomes autosomes, avec une position définie sur l'assemblage UMD3.1 du génome, sont utilisés. Les marqueurs présentant une fréquence de l'allèle rare (MAF pour Minor Allele Frequency) inférieure à 3% dans les 3 races ou ne respectant pas l'équilibre de Hardy-Weinberg ($p < 10^{-4}$) sont éliminés. Au final, 43 582 SNP sont conservés dans l'analyse pour chacune des 3 races.

Le logiciel DAGPHASE (Druet et Georges, 2010) a été utilisé pour construire les phases, c'est-à-dire assigner les allèles observés au chromosome d'origine paternelle ou maternelle pour tous les individus typés.

1.1.3 Analyses génétiques

Les mutations recherchées étant probablement récentes, leur détection est réalisée intra race. Cette recherche de régions présentant un déficit d'homozygotes est réalisée par fenêtre de 20 marqueurs (1,5 Mégabases en moyenne) que l'on déplace tout le long du génome. Pour chaque analyse, la fréquence de chaque haplotype observé de 20 marqueurs est calculée sur les phases d'origine maternelle. Seuls les haplotypes présentant une fréquence supérieure à 1% ont été conservés par la suite.

Pour chaque haplotype k , le nombre observé $O(k)$ d'individus homozygotes est comparé au nombre attendu sous l'hypothèse de neutralité. Ce nombre attendu $E(k)$ est estimé par :

$$E(k) = \sum_{i=1}^{n_p} p_{ik} \sum_{j=1}^{n_{gpm}} 0,5(q_{jk} + f_k) n_{ij}$$

$E(k)$ = nombre attendu d'individus homozygotes à l'haplotype k

n_p = nombre de pères

n_{gpm} = nombre de grand-pères maternels

p_{ik} = probabilité de transmission de l'haplotype k par le père i à ses descendants

q_{jk} = probabilité de transmission de l'haplotype k par le grand-père maternel j à ses filles

f_k = fréquence de l'haplotype k dans la population

n_{ij} = nombre de descendants issus du père i et du grand-père maternel j .

La probabilité de transmission de l'haplotype k (p_{ik} ou q_{jk}) est égale à 1 pour un taureau homozygote à l'haplotype k , 0,5 pour un taureau hétérozygote et 0 pour un taureau non porteur. L'intervalle final de localisation est obtenu en concaténant toutes les fenêtres présentant les mêmes

estimations. Un QTL est détecté lorsque $O(k)$ est significativement inférieur à $E(k)$ par un test du chi² ($p < 10^{-5}$).

1.2 ESTIMATIONS DES EFFETS DES QTL DETECTES

Pour tous les QTL détectés, les statuts (porteur ou non porteur) de 1 323 taureaux Montbéliards, 1 159 taureaux Normands et 5 696 taureaux Holstein ont été déterminés. Les effets de ces QTL ont été estimés par la différence de fertilité ou de vitalité à la naissance entre les accouplements à risque (définis par les inséminations de vaches issues de taureau porteur avec un taureau également porteur) et les autres accouplements. Dans le cas d'un QTL létal à l'état homozygote, une baisse de $-\square/8$ du taux de conception est attendue dans les accouplements à risque, \square correspondant au taux de conception moyen de la population étudiée.

1.3 RECHERCHE DES POLYMORPHISMES CAUSAUX

1.3.1 Séquences utilisées

La recherche des polymorphismes sous-jacents aux QTL cartographiés a consisté à exploiter les séquences disponibles en 2012, à savoir les séquences du génome complet de 25 taureaux Holstein, de 9 taureaux Normands et de 11 taureaux Montbéliards.

1.3.2 Principe de la recherche

Pour chaque QTL, les polymorphismes de la région d'intérêt sont étudiés de façon à identifier ceux qui respectent les conditions suivantes : jamais présents à l'état homozygote, absents de la séquence de référence, présents à l'état hétérozygote chez tous les taureaux porteurs et enfin absents chez tous les autres taureaux.

Chaque région analysée comprend l'intervalle détecté étendu de 6 Mb de chaque côté, valeur suffisante pour définir un intervalle de confiance fiable.

Dans un second temps, seuls les polymorphismes présents dans les séquences codantes sont retenus. Puis les polymorphismes candidats sont sélectionnés à partir de la sévérité de l'effet prédit sur la protéine, ainsi que de la conservation de la protéine entre espèces.

1.3.3 Validation par génotypages

A titre de première validation, un panel de 3 849 individus Holstein, 2 931 Montbéliards et 909 Normands ont été génotypés sur les polymorphismes candidats identifiés.

A ce stade, ces typages ont pour objectif de vérifier qu'il n'existe aucun individu homozygote, résultat compatible avec l'hypothèse d'un polymorphisme récessif létal. La validation fonctionnelle de ces polymorphismes candidats reste à réaliser.

2. RESULTATS

2.1 LISTE DES QTL DETECTES

Les analyses génétiques ont permis de détecter 33 QTL avec un test du chi² inférieur à 10^{-5} sur l'ensemble des trois races. Les noms des haplotypes des tableaux 1, 2 et 3 respectent les noms publiés et utilisés en sélection (VanRaden, 2011 ; Fritz *et al.*, 2013). Pour plus de la moitié des QTL détectés, des individus homozygotes sont observés, ce qui indique que si le QTL existe, la mutation est apparue récemment dans un haplotype relativement fréquent au sein de la population.

En race Holstein, 18 QTL sont détectés. Les QTL détectés par VanRaden *et al.* sont retrouvés sauf HH2 dont la fréquence semble inférieure à 1% dans la population française. L'haplotype HBY correspond à l'anomalie Brachyspina, quasiment 100% des taureaux porteurs de HBY étant également porteurs de l'anomalie. Dans une analyse complémentaire, il a été montré une forte association entre les haplotypes HH5 et HH6 avec l'anomalie CVM (Complex Vertebral Malformation) située à 43,4 Mb et donc distante de plusieurs Mégabases des QTL détectés. C'est suite à ce constat que des régions élargies ont été utilisées pour la recherche des polymorphismes causaux.

Tableau 1 : Liste des QTL identifiés en race Holstein

Nom	Chr.	Intervalle (Mb)	E(k)	O(k)	Fréquence
HBY	21	20.2-22.3	49	0	3,6%
HH1	5	61.4-66.2	18	0	2,6%
HH3	8	94.0-96.5	21	0	2,5%
HH4	1	1.9-3.3	49	0	3,6%
HH5	3	49.4-52.6	91	38	4,6%
HH6	3	45.8-47.6	68	24	3,9%
HH7	6	51.6-52.6	202	100	6,9%
HH9	10	74.8-77.0	40	13	2,9%
HH10	11	31.5-33.2	24	1	2,2%
HH11	12	2.0-3.6	75	35	3,2%
HH12	15	77.3-79.4	127	54	5,5%
HH13	18	56.4-58.4	32	5	3,7%
HH14	19	42.7-43.9	40	8	2,4%
HH15	20	58.5-59.6	68	25	2,7%
HH16	26	10.4-12.8	26	2	2,0%
HH17	26	24.9-26.0	33	5	1,8%

Tableau 2 : Liste des QTL identifiés en race Montbéliarde

Nom	Chr.	Intervalle (Mb)	E(k)	O(k)	Fréquence
MH1	19	27.6-29.4	131	0	9,0%
MH2	29	27.9-29.1	80	1	7,0%
MH3	2	31.5-32.8	39	9	5,1%
MH4	4	52.0-53.2	21	1	3,5%
MH5	6	73.3-74.4	122	14	7,1%
MH6	7	80.1-81.7	21	3	2,6%
MH7	9	84.6-86.4	126	21	7,1%
MH8	13	76.4-77.6	26	3	3,5%
MH9	20	24.2-25.7	33	6	2,6%
MH10	24	22.9-24.6	26	0	2,5%
MH11	24	33.4-34.6	159	29	7,2%

Tableau 3 : Liste des QTL identifiés en race Normande

Nom	Chr.	Intervalle (Mb)	E(k)	O(k)	Fréquence
NH1	24	38.1-39.2	12	0	1,8%
NH2	1	145.7-146.8	49	14	3,8%
NH3	4	92.3-93.8	41	10	5,9%
NH4	6	37.7-38.9	38	12	5,2%
NH5	7	3.6-4.6	58	20	1,9%
NH6	15	59.8-61.1	45	17	1,9%

En race Montbéliarde, 11 QTL sont détectés. Les deux principaux QTL (MH1 et MH2) présentent un déficit d'homozygotes important avec une fréquence élevée de l'haplotype, respectivement 9% et 7%. Seulement 6 QTL sont détectés en race Normande. Ceci s'explique sans doute par des effectifs plus faibles d'individus génotypés, donc des nombres attendus d'individus homozygotes à un haplotype

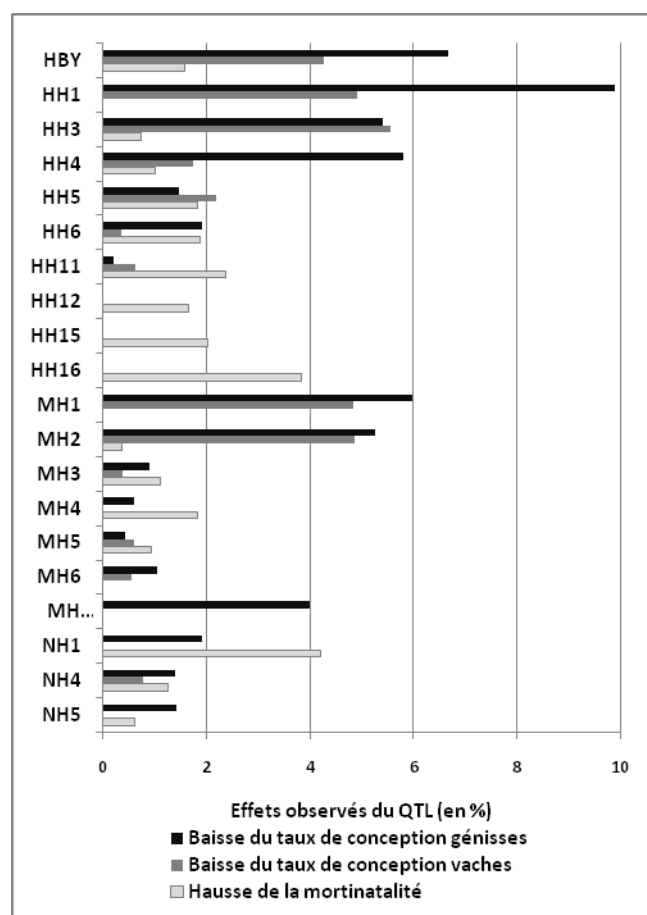
donné plus faibles et donc une puissance de détection plus faible.

2.2 EFFETS OBSERVÉS DES QTL

Des estimations des effets de ces QTL sur le taux de conception (pourcentage de veaux nés par IA) et sur le taux de mortalité ont été réalisées pour conforter la détection de QTL. De par la méthode de détection, l'hypothèse est faite que ces QTL sont récessifs et ne s'expriment qu'à l'état homozygote. Donc une baisse du taux de conception ou une hausse de la mortalité n'est attendue que dans les accouplements à risques.

Théoriquement, dans le cas d'un déséquilibre complet de l'haplotype et du polymorphisme, l'effet attendu dans les accouplements à risque (insémination d'un taureau porteur sur une vache issu d'un taureau porteur) est de 0,125 où est le niveau moyen du caractère étudié dans la population étudiée. Pour le taux de conception, la baisse attendue dans les accouplements à risque est de l'ordre de 5 à 6% et pour la mortalité la hausse attendue peut atteindre 11 à 12%. Cependant ces estimations sont des maximums car ils supposent une association complète entre QTL et haplotype et une mortalité totale, à l'état embryonnaire ou à la naissance.

Figure 1 : Effets observés des QTL sur la baisse du taux de conception des génisses et des vaches (en %) et sur la hausse de la mortalité (en %) dans les accouplements à risque.



2.3 POLYMORPHISMES CANDIDATS

Les polymorphismes déjà connus ont été retrouvés par cette approche (cf. Tableau 4). Cinq autres polymorphismes ont été identifiés, 2 en race Holstein associés à HH3 et HH4 et 3 en race Montbéliarde associés à MH1, MH2 et MH5. Ces mutations touchent des fonctions essentielles à la vie et on imagine aisément qu'elles ne soient pas tolérées à l'état homozygote. Ainsi, par exemple, GART code pour une enzyme impliquée dans la synthèse des purines et SMC2

code pour une protéine essentielle pour la condensation des chromosomes et donc pour la mitose.

Tableau 4 : Liste des polymorphismes candidats identifiés dans les régions QTL étudiées.

QTL	Gène	Changement Acide Aminé
HBV	FANCI	p.V877Lfs27X (*)
HH1	APAF1	p.Q579X (*)
HH3	SMC2	p.F1135S
HH4	GART	p.N290T
HH5	SLC35A3	p.V180F (*)
HH6		
MH1	SHBG	p.Q52X
MH2	SLC37A2	p.R12X
MH5	NOA1	p.D363Rfs9X

(*) Polymorphismes déjà connus avant cette étude

La validation par génotypage d'un large panel d'individus a permis de vérifier que ces mutations ne sont présentes que dans une seule race. Aucun individu homozygote n'a été observé pour 5 des 6 mutations candidates étudiées sur le panel d'individus typés. Compte tenu des effectifs, ce résultat est hautement significatif pour les 5 mutations citées dans le tableau 5, renforçant la présomption de leur causalité.

Pour MH1, les génotypages sur la mutation candidate dans SHBG ont permis d'identifier 13 individus homozygotes, ce qui l'invalide. De nouveaux travaux sont en cours pour identifier d'autres polymorphismes candidats.

Tableau 5 : Pourcentage d'individus porteurs des mutations candidates testées obtenu à partir d'un échantillon de 3849 individus Holstein, 909 Normands et 2931 Montbéliards génotypés.

QTL	Gène	Polymorphisme	HOL	MON	NOR
HH1	APAF1	p.Q579X (*)	3,5%	0,0%	0,0%
HH3	SMC2	p.F1135S	6,7%	0,0%	0,0%
HH4	GART	p.N290T	8,0%	0,0%	0,0%
MH2	SLC37A2	p.R12X	0,0%	12,1%	0,0%
MH5	NOA1	p.D363Rfs9X	0,0%	17,3%	0,0%

Les fréquences de ces polymorphismes observées sur cet échantillon sont toutes supérieures aux fréquences des haplotypes associés. Les individus génotypés étant des individus très récents, cela suggère que les fréquences de ces polymorphismes sont actuellement plutôt en augmentation.

3. DISCUSSION ET PERSPECTIVES

3.1 DISCUSSION

Cette étude montre la puissance du dispositif de sélection génomique et du séquençage de génome complet pour mettre en évidence et caractériser complètement plusieurs QTL de fertilité jusqu'à la mutation causale. Ces résultats laissent penser qu'un nombre croissant de mutations sous-jacentes aux QTL va être identifié, une étape attendue depuis longtemps en sélection et qui commence à se réaliser. Identifier les mutations causales permet une sélection plus facile et une meilleure compréhension des mécanismes impliqués.

Les QTL mis en évidence utilisent une information massive. Cette méthode de détection demande une quantité

importante d'individus génotypés, elle n'est donc applicable qu'aux grandes populations. Elle ne met en évidence que des mutations déjà diffusées dans la population depuis plusieurs générations. Ces travaux sont donc très complémentaires de l'observatoire national des anomalies bovines (ONAB) mais ne visent pas à le remplacer.

La recherche de déficit en homozygotes conduit à la mise en évidence d'un nombre élevé de régions, dont une partie seulement se confirme avec un effet sur la fertilité ou la mortinatalité. Il sera particulièrement intéressant de comprendre les raisons de ces déficits, peut-être dus à un effet sur d'autres phénotypes induisant une élimination précoce avant génotypage. D'autres raisons, plus artéfactuelles, ne peuvent pas être éliminées à ce stade.

3.2 UTILISATION EN ELEVAGE

Les QTL détectés ont un effet faible (peu visible) sur le taux de conception à l'échelle de la population, parce qu'ils sont récessifs et que leur fréquence est généralement basse. Eradiquer tous les animaux porteurs de ces QTL ne permettra pas d'augmenter de façon marquée le taux de conception moyen de la race. En revanche, les QTL détectés étant, de par la méthodologie employée, des QTL portés par des taureaux très diffusés, éradiquer brutalement ces QTL reviendrait à éliminer une grande partie de la descendance des taureaux présentant les niveaux génétiques les plus élevés sur les objectifs de sélection. Une éradication brutale entraînerait donc une baisse du progrès génétique sur les objectifs de sélection pour un gain faible de fertilité.

Le plus efficace est donc de baisser progressivement la fréquence de ces QTL dans les programmes de sélection et d'optimiser les accouplements en évitant tout accouplement à risque. Cette stratégie est déjà employée en race Montbéliarde pour les QTL MH1 et MH2. Cela nécessite la disponibilité de ces informations (liste des individus porteurs) et leur utilisation dans les plannings d'accouplements. Le développement de puces basse densité incluant toutes les mutations candidates identifiées comme la puce EuroG10K® d'Illumina doit permettre la mise en place de cette stratégie dans les principales races de bovins laitiers.

CONCLUSION

Le développement de la sélection génomique dans les élevages bovins devrait franchir un nouveau palier avec les premières découvertes de mutations causales de QTL d'intérêt. En ce qui concerne les QTL de mortalité embryonnaire, une éradication systématique serait extrêmement coûteuse pour les programmes de sélection et aurait un impact négatif sur le progrès génétique des objectifs de sélection. Une contre-sélection progressive est recommandée accompagnée d'une optimisation des plans d'accouplements pour éviter les accouplements à risque.

Le projet CARTOSEQ (ANR-10-GENM-0018) est un programme financé conjointement par l'ANR et par APIS-GENE. Les typages utilisés proviennent de l'activité de sélection génomique et de divers projets de recherche (CARTOFINE, AMASGEN, LACTOSCAN) financés par l'ANR et APIS-GENE.

Barbat A., Druet T., Bonaïti B., Guillaume F., Colleau J.J., Boichard D., 2005. Renc. Rech. Rum., 12, 137-140.
Druet T. et Georges M., 2010. Genetics, 184, 789-798.
Fritz S., Capitan A., Djari A., Rodriguez S., Barbat A., Baur A., Grohs C., Weiss B., Boussaha M., Esquerré D., Klopp C., Rocha D., Boichard D., 2013. Plos One, 8, e65550.
Mattalia S., Barbat A., Danchin C., Brochard M., Le Mezec P., Minery S., Jansen G., Van Doormal B., Verrier E., 2006. Renc. Rech. Rum., 13, 239-246.
VanRaden PM., Olson K.M., Null D.J., Hutchison J.L., 2011. J. Dairy Sci., 94, 6153-6161.