

Mutations candidates affectant la composition protéique du lait dans trois races bovines

Candidate mutations affecting major milk proteins in three French dairy cattle breeds

SANCHEZ M.P. (1), GOVIGNON-GION A. (1), FERRAND M. (2), GELE M. (2), POURCHET D. (3), ROSSIGNOL M.N. (4), FRITZ S. (1,5), BOUSSAHA M. (1), CAPITAN A. (1,5), ROCHA D. (1), MIRANDA G. (1), MARTIN P. (1), BROCHARD M. (1,2), BOICHARD D. (1)

(1) INRA, UMR1313 GABI, 78350 Jouy en Josas ; (2) Institut de l'Elevage, 75012 Paris ; (3) ECEL Doubs - Territoire de Belfort, 25640 Roulans ; (4) Labogena, 78350 Jouy en Josas ; (5) UNCEIA, 75012 Paris

INTRODUCTION

La composition protéique du lait de vache, prédite à grande échelle à partir des spectres moyen infra-rouge, combinée au génotypage des vaches, a permis de détecter de nombreux QTL, dont certains avec des effets majeurs sur la composition protéique (PhénoFinlait ; Sanchez *et al.*, 2013). Neuf régions QTL, expliquant de 5 à 50% de la variance génétique du caractère, ont fait l'objet de recherches des mutations causales candidates à partir des séquences des génomes des taureaux.

1. MATERIEL ET METHODES

Dans le cadre du projet PhénoFinLait, de nombreux QTL ont été détectés pour les taux des 6 lactoprotéines majeures (Sanchez *et al.*, 2013) en races Montbéliarde (MO), Normande (NO) et Holstein (HO) à partir des données de 7 654 vaches génotypées avec la puce Illumina 50K. Les 9 régions QTL présentant les effets les plus significatifs ont fait l'objet d'analyse de concordance en 3 étapes. i) Des haplotypes de 15 SNP ont été définis en considérant les 7 SNP situés de part et d'autre du SNP présentant les effets les plus significatifs. ii) L'effet de l'haplotype transmis par le père à ses filles a été testé pour les taureaux ayant au moins 50 filles phénotypées. Un statut au QTL a été attribué à chaque taureau en fonction de la valeur de la probabilité (p) associée à cet effet : hétérozygote ($p \leq 0,05$), homozygote ($p > 0,10$) ou indéterminé ($0,05 < p \leq 0,10$). iii) Pour les taureaux ayant un statut au QTL et dont le génome est séquencé, les polymorphismes ayant un impact sur la séquence en acides aminés de la protéine ont été extraits dans une région de 6Mb (± 3 Mb autour de la position la plus probable du QTL). Parmi ces polymorphismes, ceux dont le génotype était concordant avec le statut au QTL ont été considérés comme des mutations candidates pour expliquer les effets du QTL.

2. RESULTATS ET DISCUSSION

Parmi les taureaux du programme PhénoFinlait, 7 MO, 2 NO et 4 HO avaient à la fois une séquence disponible et un nombre de filles génotypées suffisant (de 53 à 226) pour estimer leur statut au QTL. En supposant que pour un même QTL, le déterminisme génétique sous-jacent était identique pour les 3 races, tous les taureaux ont été analysés conjointement. Selon la région, 9 à 12 taureaux avaient un statut au QTL et 1 à 4 polymorphismes candidats ont pu être détectés dans des gènes (Tableau 1). Dans

les 2 régions QTL 11 et 20_1, des mutations déjà identifiées auparavant ont pu être retrouvées : les 2 mutations dans le gène **LGB** définissant les variants protéiques A et B de la β lactoglobuline (Ganai *et al.*, 2009) et une mutation dans le gène **GHR** associée au taux protéique (Blott *et al.*, 2003), respectivement. Dans 3 autres régions, des gènes candidats pressentis n'ont pas pu être confirmés. Dans la région 6_1 qui renferme le gène **ABCG2**, la mutation décrite par Cohen-Zinder *et al.* (2003) comme associée au taux protéique était fixée et a donc été exclue. Dans cette région, 3 autres mutations sont mises en évidence dans 3 gènes distincts. Dans la région 6_3 qui contient les gènes des caséines, nous avons testé 5 mutations connues et aucune n'était concordante avec le statut des taureaux. Les gènes des caséines étant liés dans une région de 250 kb, s'il existe plusieurs mutations dans ces gènes, leurs combinaisons peuvent expliquer les effets observés, ce qui ne peut pas être détecté avec notre approche qui teste chaque polymorphisme indépendamment. Nous détectons en revanche une mutation candidate située dans un autre gène de cette région. La mutation connue dans le gène **DGAT1** n'est pas retrouvée, du fait de la qualité médiocre de la séquence dans cette région pour le seul taureau hétérozygote. Nous détectons cependant une nouvelle mutation dans un gène proche de **DGAT1**. Enfin, nous détectons de nouvelles mutations candidates dans les 4 autres régions QTL : 3, 4, 3 et 1 mutations dans les régions QTL 2, 6_2, 20_2 et 29, respectivement. Chaque mutation se trouve dans un gène distinct, excepté les 4 mutations de la région 6_2 concentrées dans 300 pb dans le même gène.

CONCLUSION

Cette étude montre l'efficacité des analyses de concordance pour identifier les polymorphismes causaux quand une seule mutation est impliquée. Dans des régions plus complexes, il sera plus pertinent de réaliser des analyses d'association sur les séquences des vaches imputées à partir des génotypes 50k.

Blott S., Kim J.J., Moio S. *et al.* 2003. *Genetics*, 163, 253-266.
Cohen-Zinder M., Seroussi E., Larkin D.M. *et al.* 2003. *Genome Res.*, 15, 936-944.
Ganai N.A., Bovenhuis H., Van Arendonk, J.A.M. *et al.* 2009. *Anim. Genet.*, 40, 127-133.
Sanchez M.P., Govignon-Gion A., Ferrand M. *et al.* 2013. *3R*, 20, 149-152.

Tableau 1 : Résultats des analyses de concordance entre statuts aux QTL et polymorphismes des séquences des taureaux

QTL				SEQUENCE			CONCORDANCE QTL / SEQUENCE	
Nom région	Caractère le plus fortement affecté	Position (Mb)	Nombre taureaux hét/hom	Région étudiée (Mb)	Nombre total de polymorphismes (Np)	Np avec un impact sur la protéine (Npi)	Npi concordant	Gènes
2	Caséine α s2	132,6	3/9	129-135	63 654	338	3	3 gènes
6_1	Caséine α s1	37,7	5/6	34-40	49 553	67	3	3 gènes
6_2	Caséine α s1	46,6	6/5	43-49	46 332	44	4	1 gène
6_3	Caséine κ	87,0	5/5	84-90	49 696	140	1	1 gène
11	β lactoglobuline	103,3	6/3	100-106	50 247	467	2	1 gène LGB
14	Caséine α s1	1,7	1/9	0-4	34 011	248	1	1 gène
20_1	Caséine κ	31,6	1/8	28-34	43 233	88	1	1 gène GHR
20_2	α lactalbumine	58,3	3/7	55-61	58 893	77	3	3 gènes
29	Caséine α s1	9,3	2/10	6-12	60 982	73	1	1 gène