

Structure des ADN complémentaires des lactoprotéines : application à la recherche des gènes et à leur localisation chromosomique

P. GAYE (*), J.-C. MERCIER, G. PÉTRISSANT (*), J.-L. VILOTTE, P. POPESCU (**)

Laboratoire de Génétique biochimique,
(*) Laboratoire de Physiologie de la Lactation,
(**) Laboratoire de Cytogénétique,
I.N.R.A., 78350 Jouy-en-Josas, France.

L'expression des gènes des lactoprotéines est un processus coordonné et limité dans le temps, qui survient dans la glande mammaire au moment de la parturition sous l'action combinée de différentes hormones. Chez les ruminants, les protéines synthétisées par les cellules épithéliales mammaires (caséines α_{s1} , α_{s2} , β , κ , β -lactoglobuline et α -lactalbumine) représentent plus de 90 % des protéines sécrétées dans le lait. L'analyse détaillée de la structure et de l'organisation des gènes des lactoprotéines est une étape essentielle pour la compréhension du déterminisme génétique et hormonal de la cellule mammaire. C'est dans cette perspective que les travaux suivants ont été entrepris.

1) *Construction de banques ovine et bovine d'ADN complémentaires (ADNc). Sélection et identification des clones recombinants contenant les ADN complémentaires des ARNm des 6 principales lactoprotéines.* Les deux banques ont été construites à partir d'ARN poly (A)⁺ totaux isolés de glande mammaire en lactation, par le procédé classique à la nucléase S1 (banque ovine) ou par la méthode de Gubler et Hoffman (1983) qui combine l'action de la Rnase H et de la DNA polymérase I (banque bovine). Dans les deux cas, les ADNc ont été insérés dans le site Pst I du pBR 322, puis clonés dans *E. coli* (Souche C600). Environ 3×10^5 transformants par μg d'ADN ont été obtenus par la méthode classique et 6×10^4 avec le procédé de Gubler et Hoffman.

Les clones contenant les ADN complémentaires d'ARN messagers (ARNm) prédominants (forte hybridation avec une sonde monocaténaire d'ADN obtenue par transcription inverse d'ARN poly (A)⁺ mammaire) ont été classés en 5 séries en utilisant successivement comme sonde l'ADNc d'un clone non homologue. L'ADN recombinant d'un clone de chaque série a été identifié par hybridation-traduction de l'ARNm complémentaire puis immunoprécipitation et électrophorèse de la protéine synthétisée. Les clones « α -lactalbumine » ont été identifiés directement en utilisant comme sonde un fragment d'ADNc correspondant à l' α -lactalbumine de cobaye (R. K. Craig, Institut Courtauld, Londres). 180, 165, 86, 80, 50 et 12 clones représentatifs des caséines α_{s1} , β , κ , α_{s2} , de la β -lactoglobuline et de l' α -lactalbumine ont été sélectionnés parmi le millier de clones examinés.

2) *Caractérisation et analyse structurale des ARNm codant pour les principales lactoprotéines ovines.* La taille des ARNm des caséines α_{s1} , α_{s2} , β , κ , de la β -lactoglobuline et de l' α -lactalbumine ovines, déterminée par la technique dite de « Northern » est respectivement de 1,2 ; 1,1 ; 1,25, 0,9 ; 0,85 et 0,8 Kb. Six ARNm de longueurs similaires ont été identifiés dans l'espèce porcine, ce qui suggère une très grande homologie de structure entre les ARNm homologues de ces deux espèces. Les séquences des ARNm ont été déterminées complètement (caséines α_{s1} et α_{s2} : Mercier *et al.*, 1985 ; Boissard et Pétrissant, 1985 ; β -

lactoglobuline : Gaye *et al.*, 1985) ou sont presque achevées (caséines χ : Furet *et al.* ; α -lactalbumine : Gaye *et al.*).

La comparaison des séquences des ARNm des caséines α_{s1} et α_{s2} ovines avec leurs homologues de rat (Hobbs et Rosen, 1982) et de cobaye (Hall *et al.*, 1984a, b) révèle une plus grande homologie dans les régions non codantes 3' et surtout 5'. Dans la partie codante, les seules régions réellement conservées sont celles correspondant à la séquence signal et aux sites de phosphorylation (près de 80 % d'homologie). Ceci suggère que les régions correspondant à la partie 5' non codante et au peptide signal ont acquis au cours de l'évolution une conformation fonctionnelle optimale, permettant une synthèse efficace et une sécrétion abondante de ces protéines. La conservation des sites de phosphorylation traduit bien le rôle important que ceux-ci jouent dans la stabilité des micelles de caséines.

La caséine α_{s2} ovine existe sous deux formes apparemment non alléliques qui se différencient par une délétion interne. L'analyse des ARNm correspondants a permis d'identifier un type mineur d'ARNm caractérisé par une délétion de 27 nucléotides, localisée dans la partie 5' de la séquence codante. Cette délétion pourrait avoir des répercussions sur la stabilité de la caséine délétée qui ne possède plus les deux résidus cystéinyle de la forme majeure.

3) *Isolement des gènes des lactoprotéines : caractérisation préliminaire.* Plusieurs clones réagissant positivement avec des sondes ADNc « α -lactalbumine, caséines α_{s2} et χ » ont été isolés de deux banques λ d'ADN génomique bovin. Trois clones « α -lactalbumine » ont été identifiés, ce qui a permis de dresser la carte de restriction du gène correspondant, et d'en déterminer la séquence partielle (plus des 2/3 sont maintenant connus). La partie codante de ce gène (426 paires de bases (pb)) est séparée en 4 régions de 133, 159, 76 et 58 pb par 3 introns de tailles respectives 321, 473 et au moins 520 pb. L'un des fragments génomiques (2 kb), qui commence 750 pb en amont du site de transcription, contient la totalité des signaux de transcription et probablement les séquences régulatrices.

4) *Etude de la localisation chromosomique des gènes des lactoprotéines ovines par hybridation « in situ ».* Dans ce type d'analyse, des sondes ADNc marquées par « nick-translation » avec des nucléotides triphosphates thioesters ^{35}S sont hybridées directement avec des chromosomes en métaphase préalablement dénaturés. Après autoradiographie et photographie, les caryotypes de chaque métaphase sont reconstitués et les grains d'argent localisés sur chaque chromosome sont dénombrés. L'analyse de plus de 80 caryotypes semble indiquer que les gènes des 4 caséines sont localisés sur le grand bras du chromosome métacentrique n° 2.

Conclusion. Ce type d'analyse, appliqué aux gènes des lactoprotéines des ruminants, devrait permettre non seulement d'identifier les éléments impliqués dans leur régulation, mais également, à plus ou moins long terme, d'intervenir sélectivement sur la structure et le fonctionnement de ces gènes : une application intéressante, tant du point de vue nutritionnel que technologique, serait la réduction du taux de lactose du lait, par l'introduction de copies de gènes « α -lactalbumine » anti-sens.

Boisnard M., Pétrissant G., 1985. *Biochimie*, **67**, 1043-1051.

Gaye P., Hue-Delahaie D., Mercier J. C., Soulier S., Vilotte J. L., Furet J. P., 1985 soumis à *Biochimie*.

Gubler U., Hoffman B. J., 1983. *Gene*, **25**, 263-269.

Hall L., Laird J. E., Craig R. K., 1984a. *Biochem. J.*, **222**, 561-570.

Hall L., Laird J. E., Pascall J. C., Craig R. K., 1984b. *Eur. J. Biochem.*, **138**, 585-589.

Hobbs A. A., Rosen J. M., 1982. *Nucleic Acids Res.*, **10**, 8079-8098.

Mercier J. C., Gaye P., Soulier S., Hue-Delahaie D., Vilotte J. L., 1985. *Biochimie*, **67**, 959-971.