Mammalian Genome 7, 542–544 (1996).

Mammalian Genome

Restriction map of two yeast artificial chromosomes spanning the murine casein locus

A.M. Tomlinson,[1] R.D. Cox,[2] H.R. Lehrach,[3] M.A. Dalrymple[1]

[1]PPL Therapeutics (Scotland) Ltd., Roslin, Midlothian, EH25 9PP, UK
[2]Physical Mapping and Gene Identification, University of Oxford, the Wellcome Trust Centre for Human Genetics, Windmill Road, Headington Oxford OX3 7BN, UK
[3]Genome Analysis, Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2A 3PX, UK

Received: 19 January 1996 / Accepted: 1 March 1996

The caseins are the major mammalian milk proteins (reviewed by Mercier and Vilotte 1993), constituting a dietary supply of amino acids and calcium to the infant. In the presence of calcium, the so-called calcium-sensitive caseins of the α and β types form loose, non-crystalline aggregates termed micelles, which are stabilized by calcium-insensitive κ casein. In the gut the site-specific cleavage of κ casein by rennin causes the milk to clot and remain in the stomach, facilitating digestion. The caseins are encoded by a small gene family, which in cows and sheep consists of four members, α_{s1}, β, α_{s2}, and κ, and in mice and rats five, α, β, γ, ϵ, and κ (Yu-Lee and Rosen 1983; Jones *et al.* 1985; Thompson *et al.* 1985). The casein genes all map to a single chromosome in rodents, sheep, cows, humans, and pigs (reviewed by Mercier and Villotte, 1993) and are very tightly linked genetically. All four bovine casein genes have been mapped to a single 250-kb locus (Ferretti *et al.* 1990; Threadgill and Womack 1990). This close linkage might be expected in the case of the evolutionarily closely related calcium-sensitive caseins, but there is no evidence that κ casein is evolutionarily related to the other caseins. Both in sequence homology and protein function it appears to be related to γ fibrinogen (Jolles *et al.* 1974; Thompson *et al.* 1985; Alexander *et al.* 1988), which performs a cleavage-induced clotting function in blood similar to the clotting function of κ casein in the stomach. Therefore, the proximity of the bovine κ casein gene to the other casein genes is noteworthy.

Two yeast artificial chromosome (YAC) clones bearing the five murine caseins were obtained by screening the Imperial Cancer Research Fund mouse genomic YAC library, primarily with two genomic probes for murine β casein as described (Cox *et al.* 1993) and secondarily by screening β casein-positive clones with cDNA probes for murine α, γ, ϵ, and κ casein, by colony hybridization and pulsed field gel electrophoresis (PFGE). The two clones are ICRFy902G0781, renamed MP12, and ICRFy902C11116, renamed MP14. MP14 is approximately 380 kb in size, and MP12 is approximately 435 kb in size, as determined by PFGE and Southern blotting (data not shown).

A restriction map of the two YACs was generated by partial digestion with *Pme*I, *Sal*I, *Xho*I and *Cla*I, and probing a single filter with probes specific to the left and right arms of pYAC4 (LA and RA; Fig. 1). The restriction fragments detected by LA and RA are listed in Table 1. When the MP12 and MP14 *Pme*I band patterns are represented as maps, the two maps can be aligned so that the 121-kb MP14 band is lined up with the 105-kb MP12 band. The alignment is almost perfect over the region of overlap, varying by less than 5 kb. By this analysis, the 26-kb MP14 band should fall some 10 kb from the left telomere of MP12 and therefore give a band in the MP12 lanes, but, given that there are 6 kb of vector sequence at this end, it is possible, within the accuracy of the blot below 48.5 kb, that MP12 does not contain this *Pme*I site. Because of the gel conditions and the markers used in this gel (λ concatemers), the accuracy of the blot is compromised below 48.5 kb.

The *Sal*I, *Xho*I and *Cla*I patterns were treated the same way, to

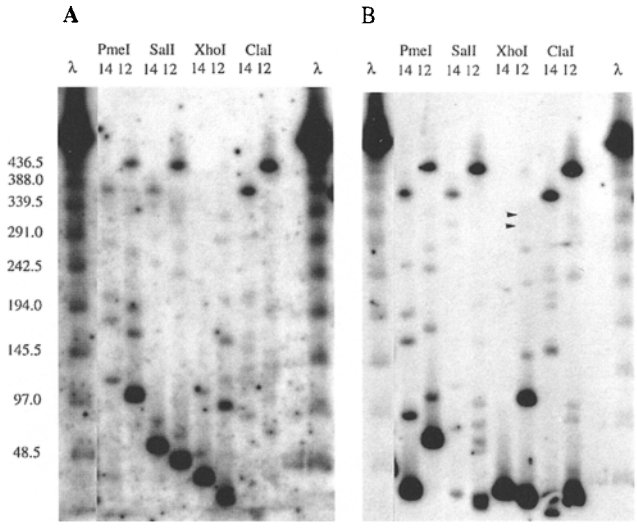

Fig. 1. Southern blot analysis of MP14 and MP12. Yeast genomic DNA was prepared in 0.8% InCert agarose and digested with restriction enzymes essentially according to Hamvas *et al.* (1994), with the exception that 20 units of each enzyme were used in a volume of 200 μl, reactions were carried out for 30 min for partial digests and overnight for complete digests, and reactions were stopped by placing on ice. PFGE was carried out in a CHEF DRII apparatus (Bio-Rad) with 1% Rapid Agarose (Life Technologies) cast and run in 0.5× TBE buffer. Running conditions were 200 V, 14°C, with switching ramped linearly from 5.0 s to 30.0 s. Gels were stained, photographed, blotted, probed, and stripped according to standard procedures. 1200 mJ/cm^2 UV (254 nm) light were used to nick the DNA before blotting. The probes for the left (LA; panel A) and right (RA; panel B) arms of the YACs were the two fragments generated by digesting pBR322 with *Pvu*II and *Eco*RI. β casein probes were generated by polymerase chain reaction (PCR) from genomic murine DNA, with the following primers (upper, then lower): 5′ (MβC5′): GACTTGACAGCCATGAAGGTCT and CGACGTTTATGGAGTCTCCTTC; 3′ (MβC3′), GGAGACACTCCTTAAGAACACT and CAGATTCTCCTAAAGGTCCAAT. Other casein probes were generated from murine lactating mammary gland RNA by reverse transcriptase PCR (RTPCR) with the following primers: α casein, GCAGTTCGCAGTCAAACTCAGC and CACTGACCTGGGAGGTAAGAGG; γ casein, AGCAATACATCTCCAGTGAGGA and TGACAGAAGTGAAGACGAGGGT; κ casein, CAGATTCAAACTGCCGTGGTGA and GTCTAGAAAGAGCAGAAGGGAA.

Correspondence to: M. Dalrymple

Table 1. Restriction fragments detected by LA and RA probes.

Probe	YAC	*Pme*I	*Sal*I	*Xho*I[b]	*Cla*I
LA	MP14	26[a], 121, 179, 206, 279, 350[c], 381	59, 85, 251, 293, 350, 381	28[a], 114	23[ac], 92, 124, 137, 154, 167, 210, 382
	MP12	105, 164, 193, 265, 336, 381, 437	41[a], 70, 236, 276, 343, 437	3[a], 102, 124, 159, 206, 286, 327	78, 109, 122, 139, 152, 195, 437
RA	MP14	22[a], 98, 164, 196, 251, 365[c], 381	20[a], 88, 126, 293, 322, 381	32[a]	3[a], 23, 161, 172, 206, 221, 240, 253, 289, 355[c], 382
	MP12	77, 116, 179, 249, 281, 343, 437	8[a], 65, 74, 90, 109, 168, 206, 377[c], 402[c], 437	23[a], 119, 156, 233, 283, 323[c], 351[c], 437	23[a], 100, 114, 242, 255, 295, 307, 320, 340, 370[c], 437

[a] Fragments denoted as less than 50 kb are below the limit of resolution of the gel.
[b] *Xho*I digest of MP14 was excessive, leading to incomplete partial band pattern.
[c] Bands that are faint in Fig. 1 but clearly discernible on the original autoradiograph.

generate a four-enzyme map of the locus from the LA autoradiograph. The same was done for the RA band patterns, and then the two maps were aligned about a central site; the *Pme*I site 206 kb from the left arm of MP14 was chosen arbitrarily and aligned with the *Pme*I site 249 kb from the right arm of MP12. The agreement between the two maps is almost perfect with an error of about 5–10 kb, but with the following ambiguities: the *Xho*I site at 220 kb was positioned at 230 kb by the RA map, but at 215 kb by the LA map (this was the only site that diverged to this extent between the two maps); the pattern of MP14 bands cut with *Xho*I was incomplete in this gel owing to over-digestion, but the MP12 *Xho*I digest allowed the map to be generated, with good agreement between the LA and RA band patterns; RA detected a *Cla*I doublet in MP14 and MP12, at 200 kb and 210 kb in Fig. 2, but only a single diffuse band could be discerned in the LA pattern. The RA autoradiograph is clearer than the LA autoradiograph, so this doublet is considered to be genuine. The collinearity of MP12 and MP14 precludes the possibility that either YAC is rearranged or chimeric.

The same filter was then probed with the probes for the other caseins, in order to map their locations within the YACs. Another blot was generated (not shown) from complete single and double digests to confirm the partial digest results.

α and β casein probes map to the same 65-kb *Cla*I fragment in partial digests. In complete digests, α casein maps to a 25-kb *Sal*I fragment, while MβC5′ maps to a 165-kb *Sal*I fragment. MβC3′ detects both of these fragments. Taken together, these data indicate that α casein lies between the *Sal*I sites at 60 and 85 kb in Fig. 2 and that β casein spans the *Sal*I site at 85 kb and is transcribed from right to left in the YAC.

γ casein was positioned primarily on the basis of partial *Pme*I digests, but these data alone were inconclusive. Confirmation was provided by the detection by the γ probe of identical *Pme*I and *Pme*I/*Sal*I complete digestion fragments of 65 kb; the minimal fragment detected by γ is a 15-kb *Cla*I–*Pme*I fragment which maps between the sites at 165 and 180 kb in Fig. 2.

The partial mapping data were inconclusive with respect to ε casein, but did allow assignment in the middle of the map. Complete *Pme*I, *Cla*I, *Pme*I/*Cla*I, PmeI/*Sal*I and *Pme*I/*Xho*I digests position ε on a 10-kb *Cla*I fragment between the sites at 200 and 210 kb in Fig. 2, spanning the *Pme*I site at 205 kb.

On the basis of partial *Pme*I data and complete *Xho*I, *Pme*I/*Xho*I, and *Sal*I restriction fragments, κ casein maps to a 34-kb fragment between the *Cla*I sites at 300 and 335 in Fig. 2.

The bovine genomic casein region has been restriction mapped (Ferretti *et al.* 1990; Threadgill and Womack 1990; Fig. 2). The spatial arrangement of the genes is largely conserved between the two species (bovine α_{s1}, β, α_{s2}, and κ caseins are homologous to murine α, β, ε, and κ caseins respectively; cattle have no known γ casein). The murine calcium-sensitive caseins are present as two pairs, separated by approximately 75 kb of intergenic DNA. The κ casein gene, although some 90–155 kb from the other genes, is nevertheless remarkably close considering its evolutionary unrelatedness. The ratio of κ to the calcium-sensitive caseins in milk is believed to play a significant role in determining micelle size, milk volume, and milk quality (Lin *et al.* 1989). As has been suggested (Alexander *et al.* 1988; Mercier and Vilotte 1993), the presence of the κ casein gene close to the other casein genes may indicate that coordination of casein gene expression is effected at the transcription level, and thus may be indicative of the existence of casein locus control.

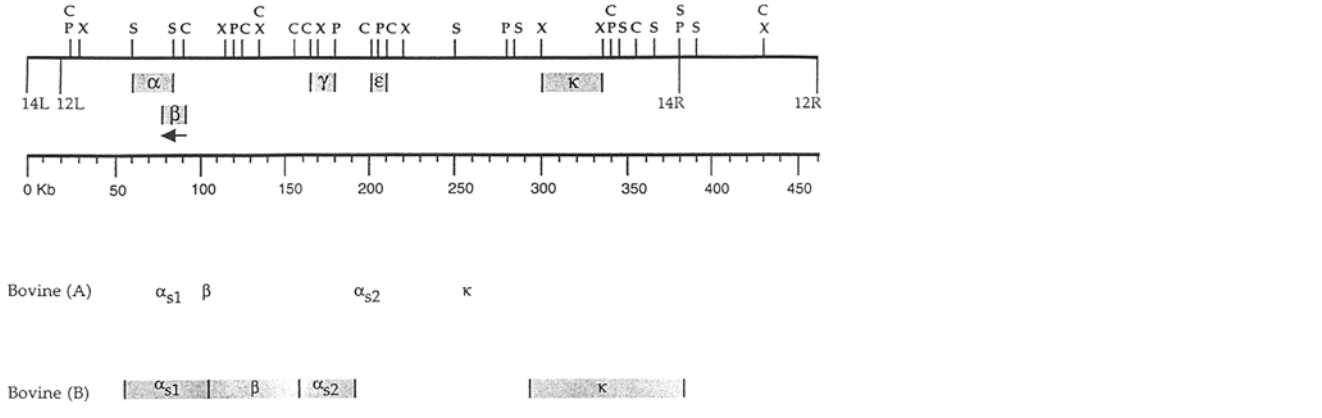

Fig. 2. Restriction map of the murine casein locus. Scale in kb is shown below the map. C, *Cla*I; P, *Pme*I; S, *Sal*I; X, *Xho*I; 14L and 14R, left and right ends of MP14; 12L and 12R, left and right ends of MP12. Casein genes are represented by Greek letters; shaded boxes delimit the minimal restriction fragment that each gene has been mapped to. Below the murine map are represented the two published maps of the bovine casein locus, to scale and aligned at α/α_{s1}. **A** from Threadgill and Womack 1990; **B** from Ferretti *et al.* 1990.

Acknowledgements. The authors would like to thank Ian Garner and Alan Colman for their support. This work was partly funded by the E.U.

References

Alexander, L.J., Stewart, A.F., Mackinlay, A.G., Kapelinskaya, T.V., Tkack, T.M., Gorodetsky, S.I. (1988). Isolation and characterization of the bovine κ-casein gene. Eur. J. Biochem. 178, 395–401.

Cox, R.D., Meier-Ewert, S., Ross, M., Larin, Z., Monaco, A.P., Lehrach, H. (1993). Genomic mapping and cloning of mutations using yeast artificial chromosomes. Methods Enzymol. 225, 637–653.

Ferretti, L., Paolo, L., Sgamarella, V. (1990). Long range restriction analysis of the bovine casein genes. Nucleic Acids Res. 18, 6829–6833.

Hamvas, R.J., Francis, F., Cox, R.D., Nizetic, D., Goldsworthy, M.E., Brown, S.M., Lehrach, H.R. (1994). Rapid restriction analysis of YAC clones. Nucleic Acids Res. 22, 1318–1319.

Jolles, J., Fiat, A.M., Schoentgen, F., Alais, C., Jolles, P. (1974). The amino acid sequence of sheep κ-A-casein. II. Sequence studies concerning the κA-casein glycopeptide and establishment of the complete primary sequence of the protein. Biochim. Biophys. Acta 365, 335.

Jones, W.K., Yu-Lee, L.Y., Clift, S.M., Brown, T.L., Rosen, J.M. (1985). The rat casein multigene family. Fine structure and evolution of the β-casein gene. J. Biol. Chem. 260, 7042–7048.

Lin, C.Y., McAllister, A.J., Ng-Kwai-Hang, K.F., Hayes, J.F., Batra, T.R., Lee, A.J., Roy, G.L., Vesely, J.A., Wauthy, J.M., Winter, K.A. (1989). Relationships of milk protein types to lifetime performance. J. Dairy Sci. 72, 3085–3090.

Mercier, J.-C., Vilotte, J.-L. (1993). Structure and function of milk protein genes. J. Diary Sci. 76, 3079–3098.

Thompson, M.D., Dave, J.R., Nakhasi, H.L. (1985). Molecular cloning of mouse mammary gland κ-casein: comparison with rat κ-casein and rat and human γ-fibrinogen. DNA 4, 263–271.

Threadgill, D.W., Womack, J.E. (1990). Genomic analysis of the major bovine milk protein genes. Nucleic Acids Res. 18, 6935–6942.

Yu-Lee, L.Y., Rosen, J. (1983). The rat casein multigene family 1. Fine structure of the γ casein gene. J. Biol. Chem. 258, 10794–10804.