SEQUENCE ANALYSIS AND COMPARISON OF BOVINE α_{S1}-CASEIN GENOMIC DNA

C. S. Lin, M. C. Huang[1], K. B. Choo[2] and Y. H. Tseng[3]

Institute of Animal Science, National Chung Hsing University
Taichung 400, Taiwan, ROC

Summary

A phage clone containing the partial α_{s1}-casein gene was isolated from a bovine genomic library by using mixed probes of ovine α_{s1}-, β- and κ-casein cDNAs. Restriction enzyme mapping analysis for 14.6 kb revealed that the map was in conflict with the report of Meade et al. (1990), especially in the 3'-end fragment. Sequence analysis of 12.6 kb revealed a high AT/GC ratio (1.64); we have identified eight exon sequences according to the bovine α_{s1}-casein cDNA sequence. The same exon/intron splice junction sequence was observed between these exons. We suggest that the bovine α_{s1}-casein gene might contain a minimum of 18 exons and the full length is approximately 18-19 kb.
(Key Words: Bovine α_{s1}-Casein Gene, Genomic DNA, Restriction Map, Exon/Intron)

Introduction

The caseins are the major protein of milk which is secreted during lactation as large calcium-containing aggregates termed micelles (Kang et al., 1985). The expression of these genes is regulated by a complex of lactogenic hormones (Eisenstein and Rosen, 1988; Guyette et al., 1979). Thus they are useful model systems for studying multi-hormonal regulation of gene expression and transgenic mammary animals.

The cDNA sequences are available for α_{s1}-, α_{s2}-, β- and κ caseins found in bovine milk (Stewart et al., 1984; Nagao et al., 1984; Stewart et al., 1987; Jimenez-Flores et al., 1987) and for the four homologues found in ovine milk (Mercier et al., 1987). There are more than 90% homologues for casein cDNA of the same type among these species.

In this paper, we describe the isolation and identification of cloned genomic DNA sequences of bovine α_{s1}-casein. The partial nucleotide sequence coding for that portion of 8 exons and exon/intron splice junction sequences is presented and the restriction endonuclease map of the bovine α_{s1}-casein genomic DNA is discussed.

Materials and Methods

Preparation of probes

The partial ovine α_{s1}-, β- and κ-casein cDNAs were provided by Laboratoire de Genetique, INRA, France. The cDNAs were labeled with [α-^{32}P] dCTP (Amersham) by the random priming method with the Random Primed DNA Labeling Kit (Boehringer Mannheim). The specific activity of labeled probes was about 1-3 $\times$ 10^8 cpm/μg.

Genomic library screening

The bovine genomic library was purchased from CLONTECH Lab. Inc., Bovine DNA fragments about 15-17 kb length were constructed into the Sal I site of λEMBL3 at a titer of 1-9 $\times$ 10^9 pfu/ml. The library infected the *Escherchia coli* NM538 to produce plaques which were screened by *in situ* plaque hybridization according to the method described by Maniatis et al. (1982). The nitrocellulose filters fixed with DNA were incubated at 42℃ for 6-8 h with 5X SSC, 5X SSPE pH 7.4, 0.1% SDS and 5X Denhardt's solution for prehybridization. Hybridization was carried out at 42℃ for 8-12 h. Unbound radioactive materials were removed by washing the filters with 2X SSC and 0.1% SDS solution at room temperature and then received

[1]Address reprint requests to Prof. M. C. Huang, Institute of Animal Science, National Chung Hsing University, Taichung 400, Taiwan, ROC.
[2]Recombinant DNA Laboratory, Department of Medical Research, Veterans General Hospital, Taipei, Taiwan, ROC.
[3]Institute of Molecular Biology, National Chung Hsing University, Taichung 400, Taiwan, ROC.
Received November 30, 1992
Accepted July 21, 1993

an additional washing with 1X SSC and 0.1% SDS solution at 55-65℃.

Positive clones were plaque purified (Maniatis et al., 1982), the phage DNA was isolated from plate lysate (Helms et al., 1985) and restriction enzyme digestion was performed (Maniatis et al., 1982).

Subcloning and restriction mapping

Bovine genomic DNA fragments cut from phage clones were subcloned into the pUC plasmid. Recombinant plasmid DNA was prepared by the alkaline lysis method (Maniatis et al., 1982). All restriction enzymes and modification enzymes were purchased from Bethesda Research Laboratories, New England Biolabs or Boehringer Mannheim and used according to the specifications of the suppliers.

A series of deletion subclones was obtained by the method according to which DNA was treated with exonuclease Ⅲ for constant time interval, followed by nuclease S1, klenow fragment and ligase (Xie and Potts, 1989).

DNA sequencing

The DNA sequence analysis was performed by the dideoxy method for dsDNA using a Sequenase® Kit (United States Biochemical Corp.). Plasmid DNA for sequencing was prepared by the boiling method with TELT solution (2.5 M LiCl, 50 mM Tris. Cl pH 8.0, 62.5 mM EDTA and 0.4% Triton X-100) (Chen and Seeburg, 1985). Sequence data were analyzed using PC/GENE software from IntelliGenetics, Inc.

Results and Discussion

Identification of probes and screening

More than 90% homologues exist for casein cDNA sequences of the same type among the different species (Bonsonng and Mackinlay, 1987). However few homologues were observed for casein cDNAs of different types except the leading peptide sequences.

The partial sequences of ovine casein cDNAs were used as hybridization probes for the screening of the bovine casein gene from the bovine genomic library. The ovine α_{s1}-, β- and κ-casein cDNA did not cross hybridized each other (figure 1). Therefore the mixed probes containing 32p-labeled ovine α_{s1}-, β- and κ-casein cDNA fragments were used for the initial screening. The positive plaque from screening was purified and identified by using individual cDNA probes.

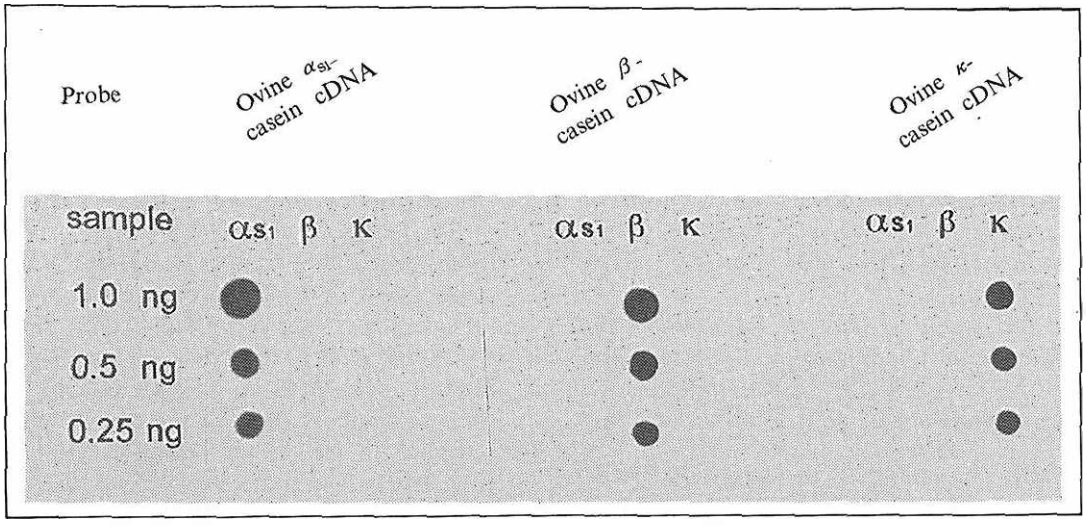

Figure 1. Dot-hybridization among ovine α_{s1}-, β- and κ-casein cDNAs.

On the first screening of 500,000 recombinant phages, three independently derived clones bearing bovine α_{s1}-casein sequences were identified and plaque-purified by dot-hybridization using individual ovine cDNA probes. We designate these clones λBo 13-2, λBo 12-4 and λBo 6-4. These positive phage clones appear to contain only the bovine α_{s1}-casein DNA sequence (figure 2).

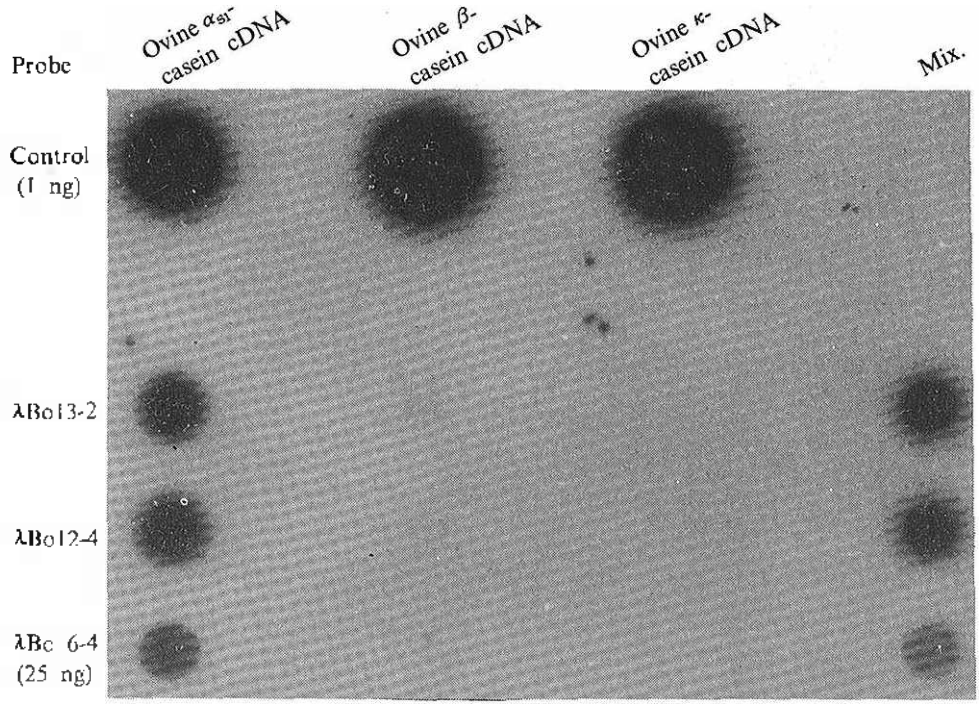

Figure 2. Autoradiograms of recombinant plaque DNAs containing bovine genomic DNA sequences. Mix.: mixture with three kind of probes.

Restriction mapping

The bovine α_{s1}-casein genomic DNA fragment of λBo 13-2 was reconstructed into pUC18; we designate this clone pBo 13-2.

To establish the DNA orientation of the pBo 13-2, 5'-versus 3'-specific cDNA probes were generated by digesting the ovine α_{s1}-casein cDNA at two Pvu II sites. The sizes of 5'- and 3'-specific cDNA probes were 274 bp and 562 bp, respectively. It was observed that the 5'-specific cDNA probe hybridized with a 8.0 kb BamH I fragment of the pBo13-2 and the 3'specific cDNA probe hybridized with 5.1 kb BamH I -Sal I fragment (figure 3), thereby establishing the orientation of the pBo 13-2. In further studies, we identified that at 5'specific cDNA probe could hybridize with a 1.1 kb EcoR I -Sal I DNA fragment (figure 4). As a result, these two sequences are 54 bp (exon 11) and 156 bp (exon 17) located on different regions of the pBo 13-2.

A detailed restriction map of the pBo 13-2 which involved 14.6 kb is shown in figure 4. It is suggested that the pBo 13-2 might lack about 2 kb of bovine α_{s1}-casein upstream and downstream sequences, repectively, according to the bovine α_{s1}-casein cDNA sequence. Therefore the full length of this gene may be approximately 18-19 kb in contrast to 16-17 kb reported by Meade et al. (1990). There is also a different restriction cutting site between them, especially in the 3'-end fragment. There are two possible ways to explain the result. (1) Both of them are different genetic variants of the α_{s1}-casein gene (Eigel et al., 1984). (2) Variants of genomic DNA sequences exist between different individuals. In particular the α_{s1}-casein gene has a more rapid rate of divergation, due to loose functional constraints, than the other caseins (Bonsing and Mackinlay, 1987).

DNA sequencing and analysis

65 subclones in total of pBo 13-2 were used for sequencing and 12.5 kb have been sequenced for the pBo 13-2. Comparison of the sequences of pBo 13-2 with bovine α_{s1}-casein cDNA published by McKnight et al.(1989) showed that eight exons were included in the derived genomic sequences (figure 5). We suggest that these exons might be the 5th, 6th, 7th, 9th, 11th, 12th, 13th and 17th exons. There is a base which is located

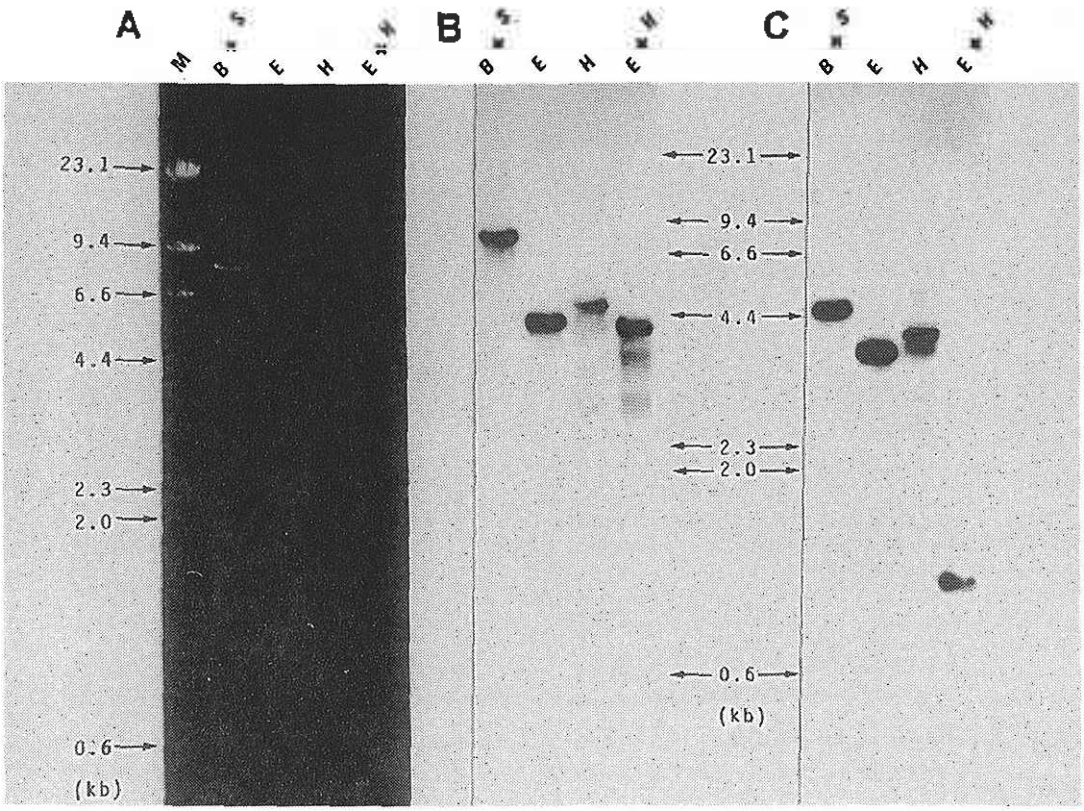

Figure 3. Analysis of pBo 13-2 DNA fragments (A) hybridized with ovine α_{S1}-casein cDNA 5′ specific probe (B) and 3′-specific probe (C). M, λ marker (Hind Ⅲ); B, BamH Ⅰ; S, Sal Ⅰ; E, EcoR Ⅰ; Hi, Hinc Ⅱ; H, Hind Ⅲ.

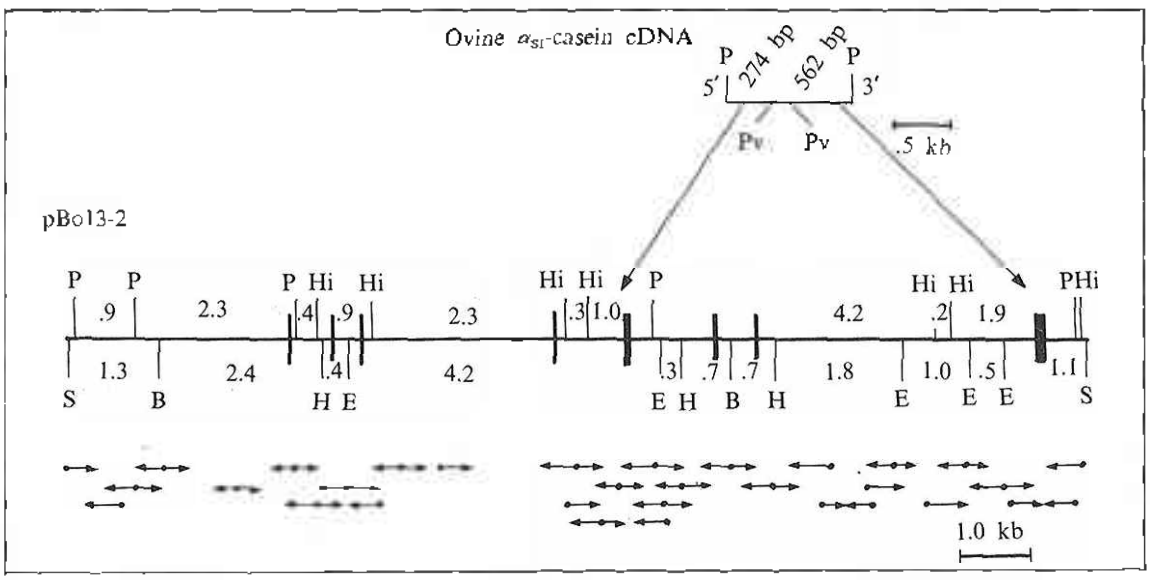

Figure 4. Restriction map of pBo 13-2 and relation of hybridization to ovine α_{S1}-casein cDNA. The exon regions are represented by blackened boxes; arrows indicate the sequencing strategy; numbers indicate the size of DNA fragments. B, BamH Ⅰ; E, EcoR Ⅰ; Hi, Hinc Ⅱ; H, Hind Ⅲ; P, Pst Ⅰ; Pv, Pvu Ⅱ; S, Sal Ⅰ.

```
exon 5
      172                                  195
      CCT TTT CCA GAA GTG TTT GGA AAG ---- bovine αs1-casein cDNA
  tag CCT TTT CCA GAA GTG TTT GGA AAG gta -- pBo 13-2

exon 6
      196                                  219
      GAG AAG GTC AAT GAA CTG AGC AAG
  aag GAG AAG GTC AAT GAA CTG AGC AAG gta

exon 7
      220                                  243
      GAT ATT GGG AGT GAA TCA ACT GAG
  cag GAT ATT GGG AGT GAA TCA ACT GAG gta

exon 9
      267                                                  298
      CAA ATG GAA GCT GAA AGC ATT TCG TCA AGT GAG
  tag CAA ATG GAA GCT GAA AGC ATT TCG TCA AGT GAG gta

exon 11
      325
      CAG AAG CAC ATT CAA AAG GAA AGT GTG CCC TCT GAG CGT TAC CTG
  tag CAG AAG CAC ATT CAA AAG GAA AGT GTG CCC TCT GAG CGT TAC CTG
                 378
      GGT TAT CTG
      GGT TAT CTG gta

exon 12
      379                                                                  420
      GAA CAG CTT CTC AGA CTG AAA AAA TAC AAA GTC CCC CAG CTG
  tag GAA CAG CTT CTC AGA CTG AAA AAA TAC AAA GTA CCC CAG CTG gta

exon 13
      421                                  444
      GAA ATT GTT CCC AAT AGT GCT GAG
  aag GAA ATT GTT CCC AAT AGT GCT GAG gtg

exon 17
      537
      CTT TTC AGA CAA TTC TAC CAG CTG GAT GCC TAT CCA TCT GGT GCC
  tag CTT TTC AGA CAA TTC TAC CAG CTG GAT GCC TAT CCA TCT GGT GCC

      TGG TAT TAC GTT CCA CTA GGC ACA CAA TAC ACT GAT GCC CCA TCA
      TGG TAT TAC GTT CCA CTA GGC ACA CAA TAC ACT GAT GCC CCA TCA

      TTC TCT GAC ATC CCT AAT CCC ATT GGC TCT GAG AAC AGT GAA AAG
      TTC TCT GAC ATC CCT AAT CCT ATT GGC TCT GAG AAC AGT GAA AAG
                                   693
      ACT ACT ATG CCA CTG TGG TGA
      ACT ACT ATG CCA CTG TGG TGA taagtt
```

Figure 5. Comparison of the partial sequences between pBo 13 2 (lower) and bovine α_{s1}-casein gene cDNA reported by McKnight et al. (1989) (upper). Numbers indicate the nucleotide order in the cDNA sequence; the different base T→C is underlined. Exon and intron sequences are described by majuscules and minuscules, respectively.

in 17th exon contrary to the report by McKnight et al. (1989) and the stop code TGA located on the last code of the 17th exon. Some exon sequences are not identified, because they have not been sequenced or are not included in the pBo 13-2. As a result, the bovine α_{s1}-casein gene probably contains a minimum of 18 exons, more than the 8 exons of mouse, rat and bovine β-casein gene (Yoshimura and Oka, 1989; Jones et al., 1985; Gorodetsky et al., 1988). There are more short exons shown to be the result of an exon shuffling event and intragenic duplication of one or more ancestral exons (Bonsing and Mackinlay, 1987).

Sequence analysis for 12.6 kb of the pBo 13-2 revealed a high AT/GC ratio, 1.64. Similar results were reported for the mouse (Yoshimura and Oka, 1989) and rat (Jones et al., 1985) β-casein genes. In these cases the AT/GC ratios are 1.56 and 1.67, respectively. All of these identified exon/intron splice junction sequences are ag]G/C·····G[gt which confirms to the consensus sequence ag]·····[gt for the 5′ exon/intron splice junctions of many eukaryotic genes (Jones et al., 1985).

The genomic sequence of bovine α_{s1}-casein gene has been partial determined. However, the published information includes only the regions of 5′-upstream, exon 1 and intron 1 (Yu-Lee et al., 1986). In our laboratory, we have sequenced about 12.6 kb of bovine α_{s1}-casein genomic sequences that include eight exon/intron organizations and most sequences of introns; we intend to identify the full length sequence in the future.

Acknowledgement

This work was supported by the grants from the National Science Council of ROC (NSC 78-0409 B005 02 and 79-0409 B005 59). We are grateful to Dr. J. C. Mercier for supplying the ovine casein partial cDNA sequences used as probes.

Literature Cited

Boisnard, M. and G. Petrissant. 1985. Complete sequence of ovine α_{s2}-casein messenger RNA. Biochimie. 67:1043-1051.

Bonsing, J. and A. G. Mackinlay. 1987. Recent studies on nucleotide sequences encoding the caseins. J. Dairy Res. 54:447-461.

Chen, E. Y. and P. H. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA. 4:165-170.

Eigel, W. N., J. E. Butler, C. A. Ernstrom, H. M. Farrell, Jr., V. R. Harwalkar, R. Jenness and R. M. Whitney. 1984. Nomenclature of proteins of cow's milk: Fifth revision. J. Dairy Sci. 67:1599-1631.

Eisenstein, R. S. and J. M. Rosen. 1988. Both cell substratum regulation and hormonal regulation of milk protein gene expression are exerted primarily at the posttranscriptional level. Mol. Cell. Biol. 8:3183-3190.

Furet, J. P., J. C. Mercier, S. Soulier, P. Gaye, D. Hue-Delahaie and J. L. Vilotte. 1990. Nucleotide sequence of ovine κ-casein cDNA. Nucleic Acids Res. 18:5286.

Gorodetsky, S. I., T. M. Tkach and T. V. Kapelinskaya. 1988. Isolation and characterization of the *Bos taurus* β-casein gene. Gene. 66:87-96.

Guyette, W. A., R. J. Matuski and J. M. Rosen. 1979. Prolactin-mediated transcriptional and post-transcription control of casein gene expression. Cell. 17:1013-1023.

Helms, C., M. Y. Graham, J. E. Dutchik and M. V. Olson. 1985. Laboratory methods. A new method for purifying lambda DNA from phage lysates. DNA. 4:39-49.

Jimenez-Flores, R., Y. C. Kang and T. Richardson. 1987. Cloning and sequence analysis of bovine β-casein cDNA. Biochem. Biophys. Res. Comm. 142: 617-621.

Jones, W. K., L. Y. Yu-Lee, S. M. Clift, T. L. Brown and J. M. Rosen. 1985. The rat casein multigene family: Fine structure and evolution of the β-casein gene. J. Biol. Chem. 260:7042-7050.

Kang, Y. C., R. Jimenez-Flores and T. Richardson. 1985. Casein genes and genetic engineering of the caseins. In Evans, J. W. and A. Hollaender (Ed) Genetic Engineering of Animals. pp 95-111. Plenum Press, New York.

Maniatis, T., E. Fritsch and J. Sanbrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Lab., New York.

McKnight, R. A., R. Jimenez-Flores, Y. Kang, L. K. Creamer and T. Richardson. 1989. Cloning and sequencing of a complementary deoxyribonucleic acid coding for a bovine α_{s1}-casein A from mammary tissue of homozygous B variant cow. J. Dairy Sci. 72:2464-2473.

Meade, H., L. Gates, E. Lacy and N. Lonberg. 1990. Bovine α_{s1}-casein gene sequence direct high level expression of active human urokinase in mouse milk. Bio/Technology.8:443-446.

Mercier, J. C., P. Gaye, S. Soulier, D. Hue-Delahaie and J. L. Vilotte. 1987. Construction and identification of recombinant plasmids carrying cDNAs coding for ovine α_{s1}, α_{s2}, β-, κ-casein and β-lactoglobulin: Nucleotide sequence of α_{s1}-casein cDNA. Biochimie. 67:959-971.

Nagao, M., M. Maki, R. Sasaki and H. Chiba. 1984.

Isolation and sequence analysis of bovine α_{s1}-casein cDNA clone. Agric. Boil. Chem. 48:1663-1667.

Provot, C., M. A. Persuy and J. C. Mercier. 1989. Complete nucleotide sequence of ovine β-casein cDNA: Inter-species comparison. Biochimie. 71: 827-832.

Stewart, A. F., I. M. Willis and A. G. Mackinlay. 1984. Nucleotide sequence of bovine α_{s1} and κ-casein cDNAs. Nucleic Acids Res. 12:3895-3907.

Stewart, A. F., J. Bonsing, C. W. Beattie, F. Shah, I. M. Willis and A. G. Mackinlay. 1987. Complete nucleotide sequences of bovine α_{s2} and β-casein cDANs: Comparisons with related sequence in other species. Mol. Biol. Evol. 4:231-241.

Xie, W. Q. and M. Potts. 1989. Quick screening of *plasmid* deletion clones carring inserts of desired sizes for DNA sequencing. Gene. Anal. Techn. 6: 17-20.

Yoshimura, M. and T. Oka. 1989. Isolation and structural analysis of the mouse β-casein gene. Gene. 78:267-275.

Yu-Lee, L. Y., L. Richer-Mann, C. H. Couch, A. F. Stewart, A. S. Mackinlay and J. M. Rosen. 1986. Evolution of the casein multigene family: Conserved sequences in the 5' flanking and exon region. Nucleic Acids Res. 14:1883-1902.