Biochem. J. (1992) 286, 31–39 (Printed in Great Britain)

Position-independent expression of the ovine β-lactoglobulin gene in transgenic mice

C. Bruce A. WHITELAW,* Stephen HARRIS,† Margaret McCLENAGHAN, J. Paul SIMONS and A. John CLARK

AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS, Scotland, U.K.

The major milk whey protein of sheep, β-lactoglobulin (BLG), is expressed specifically in the mammary gland in a developmentally regulated pattern. To identify the *cis*-acting DNA regions involved in the regulation of BLG expression, resected gene constructs were analysed in transgenic mice. BLG transgenes which contain at least the proximal 406 bp of the 5′ flanking region were expressed in all mice analysed, at levels related to transgene copy number, and thus were expressed in a position-independent manner. Expression was restricted to the mammary gland, except in a few lines where low-level expression was also detected in the salivary gland. In these mice, BLG transgenes were expressed during pregnancy and lactation in the appropriate temporal pattern. Further resection of the 5′ proximal region to −146 bp resulted in a dramatically reduced frequency of expression, without affecting tissue specificity, while a construct which retained only 79 bp of 5′ flanking region was not expressed. Chromatin analysis of isolated sheep nuclei showed that the promoter resides within a DNAaseI-hypersensitive region in the mammary gland but not in the liver. A BLG transgene displayed a similar tissue-specific pattern of DNAaseI hypersensitivity in mice. These data demonstrate an essential role of the proximal DNAaseI-hypersensitive sequences for position-independent expression of the BLG gene.

INTRODUCTION

The major milk proteins (Jenness, 1982) are expressed specifically and abundantly in the mammary gland during pregnancy and lactation (Forsyth, 1986; Gaye *et al.*, 1986; Mercier *et al.*, 1990; Harris *et al.*, 1991). In the rat, milk protein gene transcripts comprise as much as 60% of the total poly(A)$^+$ RNA present in the mid-lactation mammary gland (Mercier *et al.*, 1990). This expression pattern is regulated by a complex interaction of hormones (Topper & Freeman, 1980; Forsyth, 1986), in conjunction with intercellular and cell–extracellular-matrix interactions which are as yet poorly defined (Levine & Stockdale, 1985; Li *et al.*, 1987; Eisenstein & Rosen, 1988). Although the general trend of expression, gradually increasing through pregnancy to a maximum during lactation, is common to the major milk proteins of different species, differences both in milk composition and in hormonal induction requirements occur. It has been suggested that these differences may reflect evolutionary pressures to integrate the endocrine control of reproduction and lactation (Forsyth, 1986).

The analysis of milk protein gene expression has been restricted by the lack of a suitable *in vitro* cell culture system. Studies involving mammary epithelial cell lines have generated results, but their interpretation is limited by the inability of these cell models to mimic the entire repertoire of gene expression seen in the mammary gland (Ball *et al.*, 1988; Eisenstein & Rosen, 1988; Doppler *et al.*, 1989). An alternative approach is the use of transgenic animals, where the DNA fragment is exposed to the full range of developmental signals that endogenous genes experience. Indeed, transgenic animals can reveal mechanisms of gene control that are not evident *in vitro* (Brinster *et al.*, 1988; cf. Ott *et al.*, 1984 with Pinkert *et al.*, 1987). Although the analysis of various milk protein gene constructs in transgenic mice has begun to define the regulatory regions required for expression in the mammary gland (Simons *et al.*, 1987; Lee *et al.*, 1988; Vilotte *et al.*, 1989; Bayna & Rosen, 1990), the locations of the individual control elements have not yet been identified. Nevertheless, sequence comparison of the promoter regions of some of these genes has identified conserved regions (Hall *et al.*, 1987; Mercier *et al.*, 1990), some of which have been shown to interact with nuclear factors *in vitro* (Lubon & Hennighausen, 1987, 1988). Whether these or other regions have a functional role *in vivo* in the regulation of milk protein gene expression remains to be determined.

In ruminants, the major whey protein is β-lactoglobulin (BLG). The ovine BLG gene has been cloned and sequenced (Ali & Clark, 1988; Harris *et al.*, 1988; Ali *et al.*, 1990). Initial experiments have demonstrated that transgenic mice which carry genomic BLG gene fragments including 4.3 kb of 5′ flanking sequences express BLG specifically and abundantly in the mammary gland (Simons *et al.*, 1987). Indeed, despite the fact that rodents do not have an endogenous BLG gene, the BLG transgene is expressed in an appropriate temporal pattern during mouse mammary gland development (Harris *et al.*, 1991). In addition, hybrid constructs containing BLG 5′ sequences can target expression of heterologous proteins to the mammary gland of transgenic animals (Clark *et al.*, 1989; Archibald *et al.*, 1990). To define functionally the BLG promoter, we have generated 5′ resected gene constructs and analysed their expression pattern in transgenic mice. BLG transgenes which contain only the proximal 406 bp of 5′ flanking sequences were efficiently expressed in the mouse mammary gland in an appropriate developmental pattern. This promoter region is characterized by a strong tissue-specific DNAaseI-hypersensitive site, both in sheep and in transgenic mice. BLG transgenes which contain this compact promoter region, the transcription unit and

Abbreviations used: BLG, β-lactoglobulin; LCR, locus control region; DTT, dithiothreitol; PMSF, phenylmethanesulphonyl fluoride; HS, hypersensitive site; NF-1, nuclear factor 1.

* To whom correspondence should be addressed.

† Present address: Glaxo Group Research, Greenford Road, Greenford, Middlesex UB6 0HE, U.K.

1.9 kb of 3′ flanking sequences were expressed in all mice analysed, and showed a strong trend towards copy-number-related levels of expression.

MATERIALS AND METHODS

Generation of constructs

All commercial kits and enzymes were used as recommended by the supplier, with all constructions performed according to standard recombinant DNA procedures (Sambrook *et al.*, 1989). The BLG transgene (BLG.SS) present in mouse line BLG/14 has been previously described (Simons *et al.*, 1987). The 5′ resected constructs were derived from the genomic ovine BLG clone pSS1tgXS (Ali & Clark, 1988; Harris *et al.*, 1988), termed BLG.SX in these studies. To facilitate 5′ resection manipulations, the 4.2 kb *Sal*I–*Sph*I fragment from pSS1tgXS was subcloned into *Sal*I–*Sph*I-cut pPolyI (Lathe *et al.*, 1987), to generate pSS1tgSpS. Subsequently, pSS1tgSpS was digested either to completion with *Hpa*I, or partially with *Ava*I, and both samples to completion with *Sal*I. After blunting the cohesive ends, the various internally deleted molecules were self-ligated. The three resulting pSS1tgSpS deletion constructs were digested with *Sph*I and *Bgl*II, and the BLG-derived fragments were isolated and ligated with the 6.8 kb *Sph*I–*Xba*I fragment from pSS1tgXS to pPolyI digested with *Bgl*II and *Xba*I, to generate pSS1tgXΔH (BLGΔH), pSS1tgXΔA1 (BLGΔA1) and pSS1tgXΔA2 (BLGΔA2). The 800 bp *Bgl*II–*Sph*I fragment from pSS1tgXΔA2 was isolated and digested with *Dpn*I, *Stu*I and *Tha*I and the resulting fragments were ligated with the 6.8 kb *Sph*I–*Xba*I fragment from pSS1tgXS to pUC19 digested with *Sma*I and *Xba*I, to generate pUCXSpΔDp, pUCXSpΔSt and pUCXSpΔTh. To ease isolation of fragment for injection, these three plasmids were digested fully with *Xba*I and partially with *Eco*RI, and the relevant *Eco*RI–*Xba*I fragments were ligated to pPolyIII-I (Lathe *et al.*, 1987) digested with *Eco*RI–*Xba*I to generate pSS1tgXΔDp (BLGΔDp), pSS1tgXΔSt (BLGΔSt) and pSS1tgXΔTh (BLGΔTh). Prior to microinjection, plasmid sequences were removed by digestion with the appropriate restriction enzymes: *Sal*I–*Sal*I for BLG.SS, *Sal*I–*Xba*I for BLG.SX, BLGΔH, BLGΔA1 and BLGΔA2, and *Sfi*I–*Xba*I for BLGΔDp, BLGΔSt and BLGΔTh.

Generation and identification of transgenic mice

Gel-purified insert DNA was microinjected, using standard techniques, into either pronucleus of eggs collected from superovulated (C57BL/6 × CBA) F1 female mice, which had been mated with F1 stud males, to generate transgenic mice (Hogan *et al.*, 1986; Simons *et al.*, 1987). Transgenic lines were propagated by mating with F1 mice. DNA (for Southern blot analysis) was prepared from tail biopsy material taken at weaning, digested with the appropriate restriction enzyme, subjected to agarose gel electrophoresis and blotted to nylon membranes (Hybond N; Amersham). Southern blots were hybridized (Church & Gilbert, 1984) with random oligo-primed probes (Multiprime, Amersham; Feinburg & Vogelstein, 1983, 1984). After X-ray autoradiography, a range of exposures were scanned on a Shimadzu CS-9000 densitometer, and by comparison with restricted sheep DNA and known amounts of restricted pSS1tgXS on each blot, the transgene copy number was determined. DNA loadings were corrected by reprobing for the whey acidic protein gene (a single-copy rodent gene; Gupta *et al.*, 1982). Rapid screening of established lines was by polymerase chain reaction assay on crude tail extracts (Whitelaw *et al.*, 1991).

RNA analysis

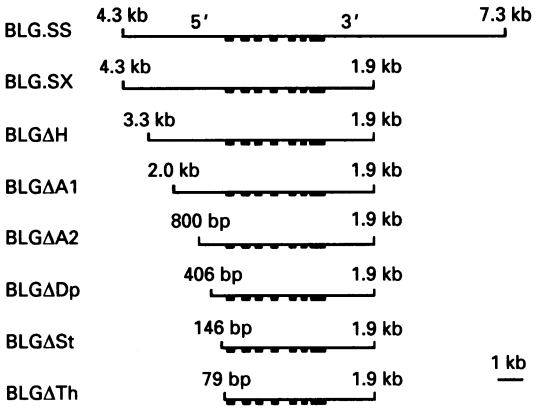
RNA was prepared from tissues isolated from mid-lactation females. The tissue sample was homogenized in 2 ml of RNAzol (Chomczynski & Sacchi, 1987), as suggested by the supplier (Biogenesis), and analysed by Northern blot. Total RNA was electrophoresed on 1% denaturing Mops/formaldehyde agarose gels, transferred and probed as for Southern blots. Blots were probed with the 424 bp *Pvu*II ovine BLG cDNA subfragment isolated from pBlg931 (Gaye *et al.*, 1986). RNA levels were quantified by probing serial RNA dilutions, dot-blotted and probed as described above, corrected for loadings and reprobing for ribosomal RNA with pXlr101 (Sollner-Webb & Reeder, 1979), and standardized with linear dilutions of a standard lactating sheep mammary RNA. A range of exposures, within a linear portion of a calibration curve, were scanned using a Shimadzu CS-9000 densitometer.

Isolation of nuclei and digestion with DNAaseI

Frozen tissue samples were ground in the presence of liquid nitrogen, homogenized in a Dounce homogenizer (Wheaton) in A/NT/L (2:3:5 by vol.; see below for composition) and filtered through several layers of Miracloth (Cambridge Bioscience). Nuclei were pelleted by centrifugation at 250 *g* for 10 min (Jouan CR3000). The nuclei were resuspended in A/NT (1:1, v/v; see below) and then repelleted as above. Storage of nuclei (10^8/ml) was at −70 °C in storage buffer [75 mM-NaCl, 0.5 mM-EDTA, 0.85 mM-dithiothreitol (DTT), 0.125 mM-phenylmethanesulphonyl fluoride (PMSF), 50% glycerol, 20 mM-Tris/HCl, pH 7.9]. All steps were carried out at 4 °C. Solution A contained 0.6 M-sucrose, 120 mM-KCl, 15 mM-NaCl, 0.3 mM-spermine, 2 mM-spermidine, 28 mM-β-mercaptoethanol, 4 mM-EDTA, 2 mM-EGTA, 0.2 mM-PMSF, 2 mM-DTT, 0.2% Triton X-100 and 10 mM-Tris/HCl, pH 7.9. Solution NT was 15 mM-NaCl/10 mM-Tris/HCl, pH 7.9, and solution L was composed of 10 mM-NaCl, 0.1% (v/v) Nonidet P-40 and 10 mM-Tris/HCl, pH 7.9.

Prior to digestion, nuclei were thawed and pelleted at 12000 *g* (Eppendorf Microfuge) for 5 min at 4 °C. The nuclei pellet was resuspended in 500 μl of digestion buffer (0.3 M-sucrose, 60 mM-KCl, 2 mM-EDTA, 2 mM-EGTA, 15 mM-NaCl, 5 mM-$MgCl_2$ and 15 mM-Tris/HCl, pH 7.5). Digestion was initiated by the addition of DNAaseI (Promega) in 5 mM-$CaCl_2$/1 mM-$MgCl_2$, and was allowed to continue for 10 min at 37 °C. Reactions were stopped by the addition of EDTA to 20 mM, and nuclei were lysed with SDS (final concentration 1%). Samples were incubated at 22 °C for 3–16 h with 2 mg of proteinase K/ml. DNA was isolated by sequential phenol and chloroform extractions, and was finally precipitated with ethanol (Sambrook *et al.*, 1989). Subsequent digestion with restriction enzyme and Southern blotting was as described above. In test reactions, without DNAaseI, digestion due to endogenous nuclear nucleases was detectable after 60 min and 20 min of incubation at 37 °C for mammary gland and liver samples respectively. That DNAaseI digestion had occurred was demonstrated by a decrease in the size average of bulk non-restricted DNA, and no hypersensitive sites were revealed by DNAaseI digestion of naked DNA from these nuclei.

Blots were probed as described above with either a 522 bp *Pvu*II–*Hin*dIII (nt 2575–3097) fragment or a 446 bp *Hin*dIII–*Bam*HI (nt 3097–3543) fragment from pSS1tgXS (Harris *et al.*, 1988). The sizes of the hybridizing bands appearing in the autoradiograms were determined by using marker DNA fragments comprising pooled samples of pSS1tgXS digested separately with *Acc*I, *Bam*HI, *Eco*RI, *Hin*dIII, *Pst*I, *Pvu*II and *Sma*I.



	Mice/lines		
	Identified	Analysed	Expressed
BLG.SS	2	2	2
BLG.SX	5	2	2
BLGΔH	7	6	6
BLGΔA1	3	3	3
BLGΔA2	7	5	5
BLGΔDp	11	10	10
BLGΔSt	10	7	1
BLGΔTh	12	8	0

Fig. 1. BLG gene constructs and expression frequency

The structure of the ovine BLG gene fragments (identified on the left) injected into fertilized mouse eggs is shown. The sizes of the 5′ and 3′ flanking regions are given in kb or bp, and exons are presented as boxes. All constructs contain the entire 4.7 kb BLG structural gene. BLG.SS and BLG.SX have been described previously as SS1 (*SalI* fragment) and a *SalI–XbaI* subfragment respectively(Simons *et al.*, 1987). The numbers of mice/lines identified by Southern blot, analysed for expression, and whether they expressed the BLG in the mammary gland, is shown on the right. Some G_0 mice either failed to transmit the transgene to their progeny, were sterile or carried the transgene on the Y chromosome.

RESULTS

Generation of resected BLG transgenic mice

The eight BLG gene constructs used in this study are illustrated in Fig. 1. All constructs contain the entire 4.7 kb BLG transcription unit (Ali & Clark, 1988; Harris *et al.*, 1988). BLG.SS is the complete insert from the phage clone SS1 (Ali & Clark, 1988), which contains 4.3 kb of 5′ flanking and 7.3 kb of 3′ flanking DNA. Each of the other constructs contains only 1.9 kb of 3′ flanking sequences, with various amounts of 5′ flanking sequences, from 4.3 kb (BLG.SX) to 79 bp (BLGΔTh).

A total of 61 G_0 transgenic mice were obtained, with < 1–30 transgene copies per cell. Animals with less than one copy per cell were presumed to be mosaic (Wilkie *et al.*, 1986; Whitelaw *et al.*, 1990) and were not analysed. In addition, three transgenic lines were obtained which carried the transgene on the Y-chromosome, precluding expression analysis. In total, 45 separate integration events were analysed in this study.

The proximal 5′ flanking sequences are required for BLG expression

Expression was analysed in G_0 or G_1 mid-lactation transgenic females. RNA was isolated from the mammary glands of 11-day post-partum females and analysed by hybridization of Northern and dot blots. Milk was also collected from these females and analysed by SDS/PAGE and Western blotting.

The presence of distal 3′ sequences in BLG.SS had no obvious effect on expression when compared with that seen for BLG.SX, which has an identical 5′ region (Simons *et al.*, 1987), or with the 5′ resected BLG transgenes lacking these distal 3′ sequences (Fig. 1 and Table 1). All BLG.SS, BLG.SX, BLGΔH, BLGΔA1, BLGΔA2 and BLGΔDp mice expressed BLG mRNA in the mammary gland and secreted BLG into milk (Fig. 1 and Table 1). Expression levels ranged from about 5% to 300% that of lactating sheep. This variation in expression was not related to the different constructs and the four lowest expressing mice/lines did not carry the same construct. Gross transgene rearrangement was not observed by Southern blot analysis (results not shown); however, the presence of localized mutations remains possible. The amount of BLG detected in mouse milk reflected, in general, the observed mRNA levels.

Further 5′ resection of this region, to leave 146 bp of the 5′ flanking DNA, resulted in a dramatic decrease in the frequency of transgene expression, with only one out of eight mice/lines expressing the BLGΔSt transgene (Fig. 1). This low expression frequency is significantly different from that observed for the BLG transgenes containing more extensive 5′ regions ($P < 0.001$; Fisher's Exact Test). In the one BLGΔSt mouse expressing BLG, the observed mRNA levels were at the lower end of the range observed for the transgenes comprising more extensive 5′ flanking sequences (Table 1). None of the eight transgenic mice/lines carrying the BLGΔTh construct expressed the transgene (Fig. 1). This construct contains only the proximal 79 bp of the 5′ flanking DNA, which includes the TATA consensus sequence.

BLG transgenes are expressed in a copy-related manner in the mammary gland

The range of expression levels was further analysed. Fig. 2(*a*) shows a scatter graph of RNA expression level versus transgene copy number for transgenic mice carrying constructs comprising 406 bp or more of 5′ flanking region. For the 10 individual BLGΔDp transgenic mice/lines, a significant linear relationship was demonstrated between transgene copy number and expression level (Fig. 2*b*), with the regression line $y = 0.31(\pm 0.15) + 0.037(\pm 0.01)x$. From the regression equation, the level of mRNA expression per BLGΔDp transgene copy is 34.7% of that in RNA of lactating sheep. Thus, per gene copy, we estimate that the BLGΔDp transgenes are expressed on average at nearly two-thirds the level of the endogenous sheep gene, since both ovine BLG alleles are active (Ali *et al.*, 1990). Significant regression lines could not be calculated for each of the other constructs due to the sample sizes.

Tissue distribution of BLG transgene expression

In sheep, BLG is expressed specifically in the mammary gland (Gaye *et al.*, 1986). We have previously shown that transgenic mice carrying either BLG.SS or BLG.SX express high levels of

Table 1. Expression levels of BLG transgenes in the mammary gland and milk

Detection levels were at least 10^4-fold lower than in standard lactating sheep. RNA levels were normalized to that of lactating sheep mammary gland. ND, not determined.

Construct/line	Copy no.	Expression of RNA
Sheep	2	1.00
BLG.SS/7	25	3.00
/14	2	0.70
BLG.SX/23	5	ND
/45	17	2.90
BLGΔH/8	9	0.31
/12	18	3.00
/14	4	0.61
/18	17	0.87
/19	2	0.44
/21	1	0.15
BLGΔA1/6	1	0.08
/9	1	0.05
/13	2	0.36
BLGΔA2/1	1	0.40
/28	3	0.50
/34	1	0.14
/39	1	0.05
/42	1	0.08
BLGΔDp/8	8	1.00
/11	3	0.92
/21	26	1.50
/28	2	0.27
/33	5	0.34
/39	30	1.24
/43	4	0.35
/44	10	0.26
/46	18	0.97
/52	3	0.21
BLGΔSt/86	2	0.20

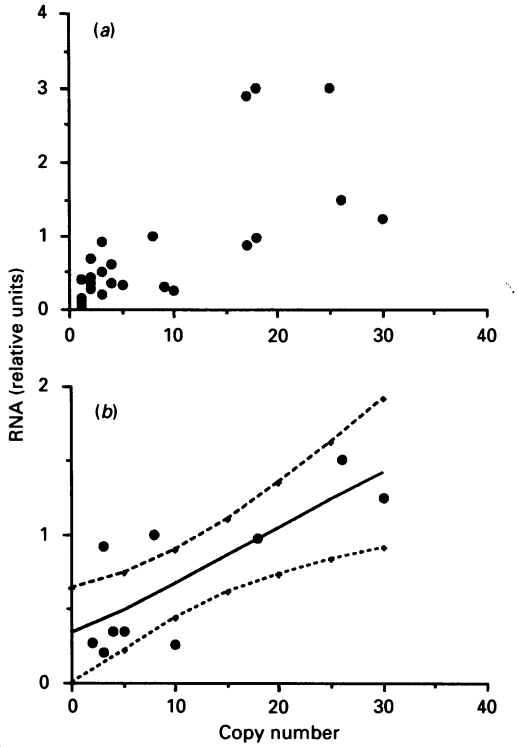


Fig. 2. Relationship between transgene copy number and expression

(*a*) Scatter graph of mRNA levels detected in the mammary gland by densitometric scanning of duplicate, serial-diluted RNA dot-blots versus transgene copy number, as estimated by densitometric scanning of Southern blots. Data are presented for BLG.SS, BLG.SX, BLGΔH, BLGΔA1, BLGΔA2 and BLGΔDp transgenic mice. Some of the single-copy transgene points overlap each other. (*b*) Linear regression (solid line) for BLGΔDp is given by $y = 0.31(\pm 0.15) + 0.037(\pm 0.01)x$, where y is the BLG mRNA expression normalized to a standard lactating sheep, and x is the transgene copy number; the 95% confidence limits are indicated (broken lines). The value for a single endogenous sheep gene is 0.5.

BLG specifically in the mammary gland (Simons *et al.*, 1987). To determine the specificity of expression of the resected transgenes, RNA was isolated from several tissues taken from mid-lactation mice and analysed for expression by Northern blot hybridization.

In transgenic mice carrying BLG gene constructs containing at least 406 bp of proximal 5′ flanking sequences, expression was detected predominantly in the mammary gland (Fig. 3*a*). By over-exposing the blots, very low levels of expression were also observed in the salivary glands in eight out of 22 of the mice/lines analysed (Fig. 3*b*). This BLG expression was not attributable to contamination with mammary gland material in these animals, as hybridization of these samples with mouse whey acidic protein or mouse β-casein cDNA probes gave no hybridization signal (results not shown). Salivary gland expression, although not restricted to a specific construct, was most prevalent in BLGΔDp mice. In all cases, the levels of salivary expression were at least 10^3-fold lower than those observed in the mammary gland. Expression of BLG was not seen in the ovine salivary gland even after long exposures of sheep RNA blots. RNA was also analysed from thymus, skeletal muscle, tongue, small intestine, lung and brain of mice of line BLGΔDp/39, which express BLG mRNA in the salivary gland, but no BLG mRNA was detected (results not shown). In the one expressing BLGΔSt line, expression of the transgene was again predominantly in the mammary gland; however, low-level expression was also seen in the liver, kidney and salivary gland (Fig. 3*c*). This non-mammary expression was also at least 10^3-fold lower than that detected in the mammary gland.

Expression of BLG transgenes during mammary gland development

During pregnancy, the mammary gland undergoes proliferation of cells and a marked differentiation of those cells (Knight & Peaker, 1982; Forsyth, 1986). In parallel, ovine BLG expression gradually increases as the mammary gland develops, to a maximum during lactation (Gaye *et al.*, 1986). We have previously demonstrated a similar developmental pattern of BLG.SS expression in mice, paralleling that seen for β-casein, an endogenous mouse milk protein gene (Harris *et al.*, 1991). To determine whether the resected BLG transgenes behaved in a similar manner, RNA was isolated from mammary tissue of age-matched siblings at selected time points during pregnancy and lactation, and analysed by Northern blotting (Fig. 4*a*). High-expressing lines were analysed to maximize the sensitivity of detection of low-level expression during early developmental time points.

All BLG transgenes which contained at least the proximal 406 bp of the 5′ flanking region demonstrated a temporal expression pattern similar to the others (Fig. 4*b*) and to that previously determined for BLG.SS (Harris *et al.*, 1991). A gradual increase was observed during pregnancy, with maximal expression occurring during lactation: by mid-pregnancy ex-

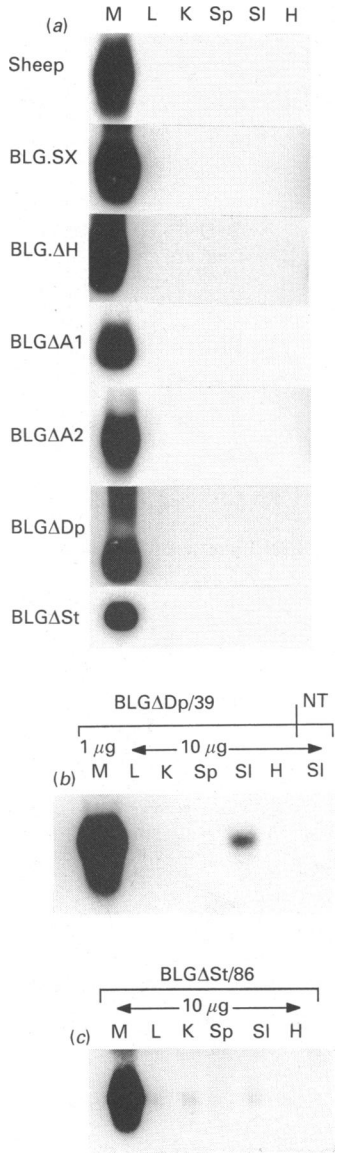


Fig. 3. Tissue specificity of BLG expression in transgenic mice

(*a*) Northern blot analysis of RNA isolated from mammary gland (M), liver (L), kidney (K), spleen (Sp), salivary gland (Sl) and heart (H) samples, isolated from individual mid-lactation transgenic mice of lines BLG.SX/45; BLGΔH/12, BLGΔA1/13, BLGΔA2/28, BLGΔDp/39 and BLGΔSt/86, and from sheep. Total RNA was resolved by denaturing gel electrophoresis, transferred to Hybond filters and hybridized to an ovine BLG-specific probe. The autoradiographs were over-exposed to maximize the sensitivity of mRNA detection. (*b*), (*c*) Low non-mammary expression in transgenic mice detected by Northern blot analysis. To allow identification of low-level expression, various amounts of total RNA were loaded as indicated above each lane. NT, non-transgenic mouse.

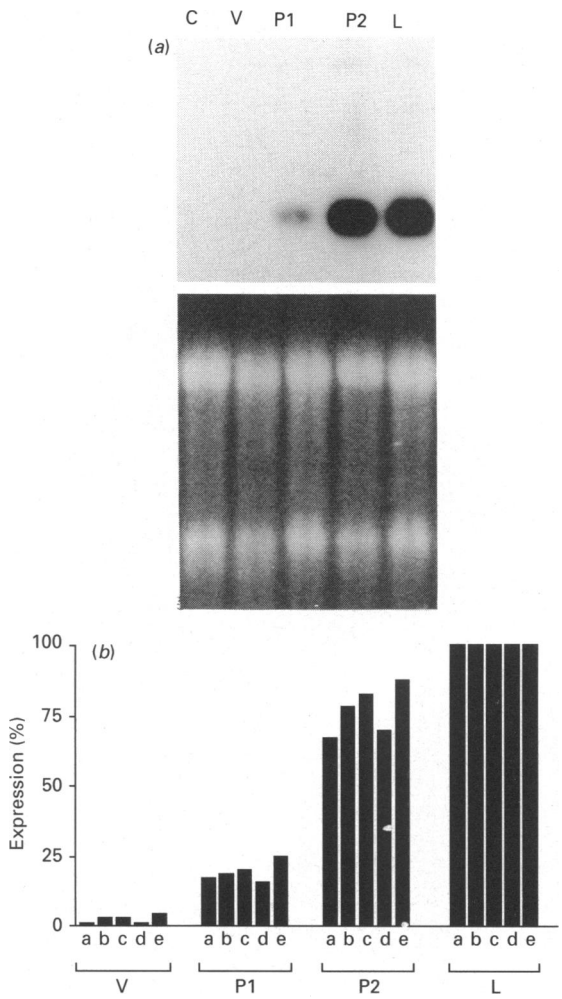


Fig. 4. Expression of BLG transgenes during mammary gland development

(*a*) Northern blot analysis of mammary RNA isolated from BLGΔDp/39 mice at various time points during pregnancy and lactation: V, virgin; P1, 10 days post-plug; P2, 16 days post-plug; L, 11 days post-partum (lactation); C, 11 days post-partum non-transgenic mouse. Total RNA was resolved by denaturing gel electrophoresis, transferred to Hybond filters and hybridized to an ovine BLG-specific probe. The ethidium bromide-stained gel is shown below. (*b*) Comparison of the temporal expression patterns seen for the various transgenes analysed; a, BLG.SS/14; b, BLGΔH/12; c, BLGΔA1/13; d, BLGΔA2/1; e, BLGΔDp/39. Values were determined by quantification of serially diluted RNA dot blots. The level of expression (average of two mice for each time point) is presented as a percentage of that detected in the mid-lactation mammary gland. All mice were age-matched siblings within each transgenic line.

pression was 6-fold higher than that detected in the virgin, a further 4-fold increase occurred by late-pregnancy, and mid-lactation levels were approximately 30% greater than that seen at the end of gestation. This pattern parallels that of endogenous β-casein transcripts (Harris *et al.*, 1991).

Mammary-specific DNAaseI-hypersensitive sites reside in the proximal 5′ region of BLG

The presence and location of DNAaseI-hypersensitive sites in the chromatin encompassing the ovine BLG gene were determined by limited DNAaseI digestion of nuclei isolated from lactating sheep mammary gland. DNA isolated from DNAaseI-treated nuclei was digested with *Hin*dIII, and the resulting fragments were identified by indirect end-labelling (Wu, 1980) of Southern blots. Two *Hin*dIII restriction fragments span the entire region of the transgenes used in this study.

The *PvHd* probe (Fig. 5*f*) detects a 7.5 kb 5′ *Hin*dIII genomic fragment (Fig. 5*a*), encompassing the first three BLG exons and approx. 5.4 kb of 5′ flanking sequences in DNA isolated from sheep nuclei which had not been treated with DNAaseI. In addition to the genomic fragment, this probe detected three other fragments of 2.4, 3.0 and 4.0 kb in DNAaseI-treated lactating-mammary chromatin (Fig. 5*a*). The 3.0 kb fragment was not detected in all preparations of nuclei analysed. These

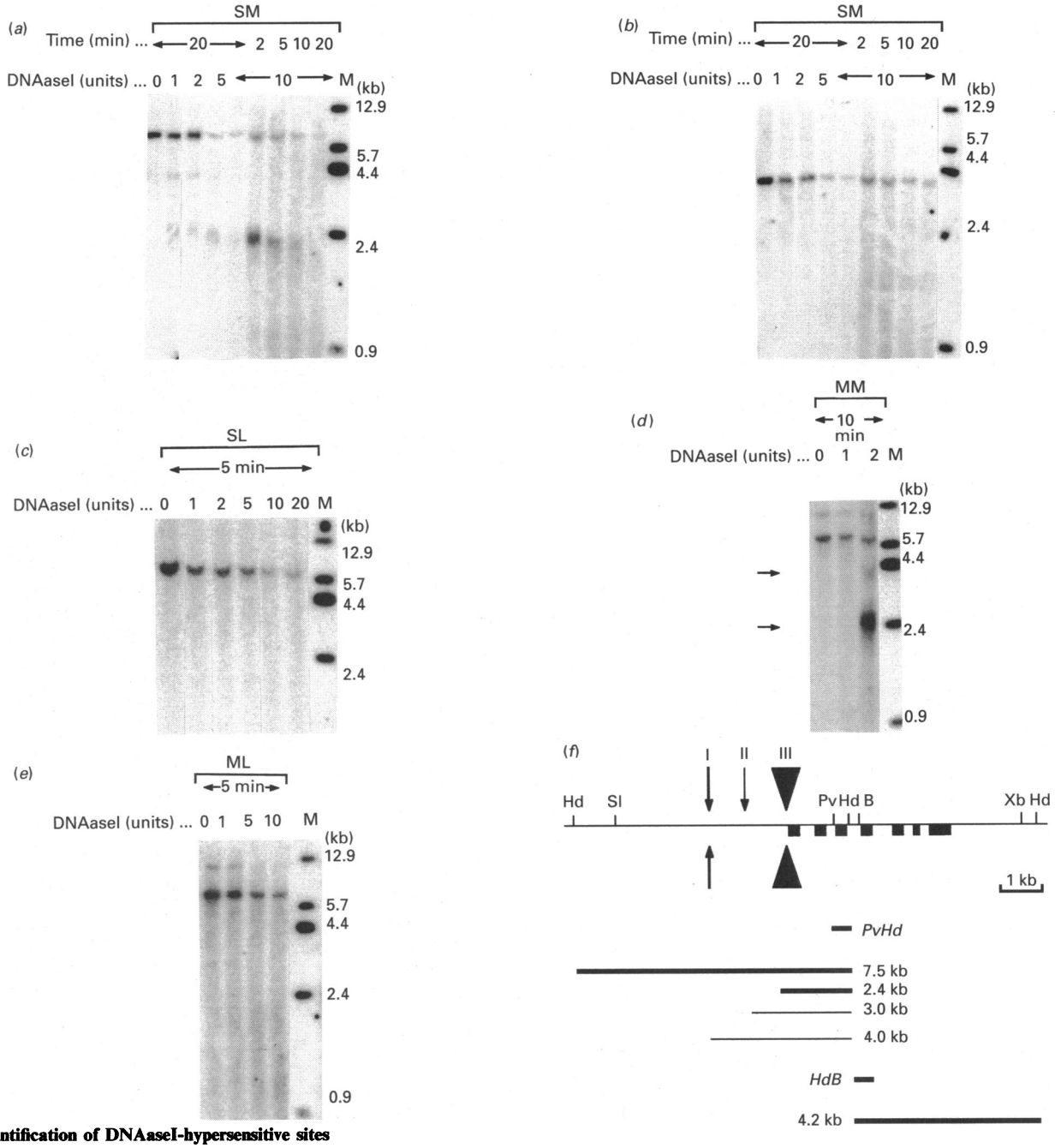


Fig. 5. Identification of DNAaseI-hypersensitive sites

Nuclei from lactating mammary glands (*a*, *b*, *d*) and liver (*c*, *e*) were incubated with increasing amounts of DNAaseI at 37 °C for various times, as indicated above each lane. The next incubation point essentially digested all bands. Purified nuclear DNA was restricted with *Hin*dIII, resolved by agarose gel electrophoresis, transferred to nylon membranes and hybridized with the *PvHd* (*a*, *c*, *d*, *e*) or the *HdB* (*b*) probes. SM, sheep mammary gland; SL, sheep liver; MM, BLG/14 transgenic mouse mammary; ML, BLG/14 transgenic mouse liver; M, marker lane. (*f*) Diagram of genomic ovine BLG showing the relevant restriction sites (Hd, *Hin*dIII; Pv, *Pvu*II; S, *Sal*I; Xb, *Xba*I). Exons are presented as boxes. Location and sizes (bp) of the *PvHd* and *HdB* probes and the observed hybridizing bands are shown diagramatically. The deduced DNAaseI-hypersensitive sites (numbered I–III; relative sensitivity indicated by arrow width) in sheep (above line) and mouse (below line) lactating mammary gland chromatin are indicated.

fragments identify two weak DNAaseI-hypersensitive sites, respectively −1800 bp (HSI) and −800 bp (HSII) from the transcription start site, and a strong hypersensitive site (HSIII) which, located at −300 to +100 bp, encompasses the promoter region (Fig. 5*f*). No equivalent hypersensitive sites were detected in chromatin from lactating sheep liver nuclei (Fig. 5*c*). The *HdB* probe detects a 4.2 kb *Hin*dIII genomic fragment which comprises the 3′ BLG exons and approx. 2.0 kb of 3′ flanking sequences (Fig. 5*f*). Using this probe, no DNAaseI-hypersensitive sites were detected in mammary (Fig. 5*b*) or liver (results not shown) chromatin.

To determine whether BLG transgenes displayed a similar DNAaseI-hypersensitivity pattern, nuclei were isolated from mid-lactation mammary gland tissue from a line of BLG.SS mice (BLG/14). This line carries two copies of the BLG.SS construct and expresses BLG abundantly and specifically in the mammary

gland (Simons *et al.*, 1987) in the appropriate developmental pattern (Harris *et al.*, 1991). Fragments corresponding to two hypersensitive sites were identified using the *PvHd* probe (Fig. 5*d*), in addition to the bands representing the internal *Hin*dIII transgene fragment (6.5 kb) and the fragment comprising transgene and flanking mouse sequences (10.0 kb). These sites co-localized to HSI and HSIII that were identified in lactating sheep mammary nuclei. HSII was not detected, and may reflect the inability to perform a similarly extensive DNAaseI titration due to the limited number of nuclei available from the mouse mammary gland or the relatively weak and inconsistent nature of this site in sheep. Again, no hypersensitive sites were detected in liver chromatin (Fig. 5*e*). In unrestricted DNAaseI-digested mammary chromatin, a fragment of the approximate size of the transgene fragment was detected (results not shown), suggesting that the promoter regions of both copies of the transgene were hypersensitive.

DISCUSSION

Expression of BLG in transgenic mice: defining the promoter region

The ovine BLG gene is expressed specifically in the sheep mammary gland (Gaye *et al.*, 1986; Harris *et al.*, 1991). We have previously shown that BLG transgenes containing 4.3 kb of 5′ flanking sequences are expressed efficiently, with an appropriate developmental pattern, in the mammary gland of transgenic mice (Simons *et al.*, 1987; Harris *et al.*, 1991). We have also shown that this 5′ region is capable of targeting the expression of a heterologous gene to the mammary gland (Archibald *et al.*, 1990). In this present study we show that the proximal 406 bp of 5′ flanking sequences are sufficient to direct efficient expression of the BLG gene to the mammary gland. Within this promoter region, important regulatory elements are present between −406 and −146, since the removal of these sequences dramatically reduces the frequency of transgenic mouse lines expressing BLG. The fact that expression was observed in one out of eight cases, again predominantly in the mammary gland, with a BLG transgene containing only 146 bp of the immediate 5′ flanking sequences indicates that additional regulatory elements, capable of targeting expression to the mammary gland, are located more proximally in the promoter, or downstream of the transcription start site. The ability of these elements to function in the absence of the region between −146 and −406 would appear to be highly dependent on the site of transgene integration. An alternative, but unlikely, explanation is that the expression observed in this one case was due to the fortuitous integration of the BLGΔSt transgene in the vicinity of endogenous enhancer sequences that mediate expression in the mammary gland.

For many genes, the analysis of 5′ flanking regulatory sequences has identified both proximal and distal control elements which are necessary for appropriate tissue-specific expression (e.g. Pinkert *et al.*, 1987). The studies described here demonstrate that the 5′ flanking sequences required for appropriate tissue-specific and developmental regulation of the BLG gene are located in a relatively compact promoter region. By contrast, two other milk protein genes have been found to possess more extensive promoters. Thus transgenic mice carrying a whey acidic protein gene containing 950 bp of 5′ flanking sequences or a β-casein gene containing 500 bp of 5′ flanking region expressed the transgenes predominantly in the mammary gland, but in both cases appropriate expression was not observed. The whey acidic protein transgenes, although expressed at levels which approached that of the endogenous mouse gene, presented an inappropriate temporal expression pattern (Bayna & Rosen, 1990), while the β-casein transgenes responded appropriately during mammary gland development, but the expression levels obtained were extremely low when compared with the endogenous mouse gene (Lee *et al.*, 1988).

BLG transgenes are expressed in the salivary gland

BLG expression was detected at low levels in the salivary glands of a number of transgenic mice. Salivary gland expression of other milk protein genes and hybrid genes derived from them, including BLG, has also been reported in transgenic mice (Lee *et al.*, 1988; Pittius *et al.*, 1988; Archibald *et al.*, 1990; Gurzburg *et al.*, 1991). Since salivary expression is a consistent phenomenon, it is unlikely to be due solely to position effects at the site of transgene insertion.

The mammary gland and salivary glands have several characteristics in common. They share similar morphology and cell types, and possess an overlapping range of cell surface receptors (Abbey & Witorsch, 1984). The salivary gland also responds to at least some of the hormones which stimulate milk protein synthesis in the mammary gland (Hatakeyama *et al.*, 1988). Therefore the expression of milk protein genes in the salivary gland may reflect the presence of similar *trans*-acting factors in these two tissues. The inappropriate, low-level salivary expression seen in this study could be due to species differences in the relative affinities of the *trans*-acting factors for the ovine BLG regulatory sequences in the mouse and sheep salivary glands. Alternatively, the lack of expression of endogenous milk protein genes in salivary tissue may be due to the presence of negative regulatory elements missing from the constructs used for gene transfer.

The BLG promoter encompasses a mammary-specific DNAaseI-hypersensitive site

Expressed genes often reside within an open chromatin structure (Weintraub, 1985; Elgin, 1988). This is manifested as an increased sensitivity to nucleases of the chromatin in the vicinity of actively transcribed genes (Jantzen *et al.*, 1986). Sites which are hypersensitive to nuclease digestion are frequently found within these regions and represent a disruption in the nucleosomal array. In numerous cases, these sites have been mapped to important regulatory elements which interact specifically with transcription factors (Elgin, 1988). In sheep, the BLG promoter, as defined by the resection analysis described in this study, exhibits a strong, tissue-specific DNAaseI-hypersensitive site (HSIII), suggesting the interaction of *trans*-acting factors with sequence elements in this region. The fact that a BLG transgene also exhibits the same strong hypersensitive site shows that its formation is mediated by *cis*-DNA elements within the transgene and that its presence is associated with BLG expression in the mammary gland.

Sequence comparison of the α-lactalbumin and casein gene promoters from several species has revealed a 30 bp region which is highly conserved. This region has been termed the 'milk box' (Hall *et al.*, 1987; Mercier *et al.*, 1990). Nuclear factors, including NF-1, have been shown to interact with the murine α-lactalbumin promoter milk box *in vitro* (Lubon & Hennighausen, 1987) and with a comparable region in the murine whey acidic protein gene promoter (Lubon & Hennighausen, 1988). The BLG promoter contains a region with only a weak similarity to the milk box. It does, however, contain a number of NF-1-binding sites, as well as multiple sites that bind a mammary-gland-specific nuclear factor which may interact to drive expression in the mammary gland (Watson *et al.*, 1991).

Two minor upstream hypersensitive sites (HSI and HSII) were detected in sheep mammary chromatin. HSII was not detected in all sheep mammary samples, and was not detected in transgenic

mice. No hypersensitive sites were detected in liver chromatin. No essential role for HSI and HSII, with respect to BLG expression, is evident from these studies, since their deletion has no overt effect on the efficiency of transgene expression. They may simply reflect DNA structures which display enhanced sensitivity to nuclease digestion in tissues where the BLG gene is expressed. Alternatively, these sites may reflect a function which was masked in the expression studies by the presence of downstream sequences.

BLG transgenes are expressed in a copy-related manner in the mammary gland

The level of expression of most gene constructs introduced into transgenic mice does not correspond to the transgene copy number (Lee *et al.*, 1988; Bonnerot *et al.*, 1990). This is thought to be due to the influence of the chromosomal DNA sequences in the vicinity of the integration site (Allen *et al.*, 1990; Al-Shawi *et al.*, 1990). Indeed, this so-called 'position effect' can be so extreme that a transgene may be completely inactive in a significant proportion of the transgenic animals generated in an experiment.

The β-globin-like genes provide a notable exception to the position-dependent behaviour of most transgenes. Thus a variety of β-globin constructs have been shown to be expressed in transgenic mice independently of their site of integration, and exhibit a strong correlation between copy number and level of expression (Grosveld *et al.*, 1987; Orkin, 1990). These constructs all contain a distal 5′ flanking region, now termed the Locus Control Region (LCR; Orkin, 1990), which is responsible for this effect. Other genes for which position-independent expression has been reported include those for CD2 (Lang *et al.*, 1988), lysozyme (Bonifer *et al.*, 1990) and α-globin (Higgs *et al.*, 1990), and *fps/fes* (Greer *et al.*, 1990). The α- and β-globin LCRs are characterized by strong tissue-specific DNAaseI-hypersensitive sites, as is the 3′ flanking region of the CD2 gene responsible for position-independent expression. The β-globin LCR and the equivalent region from CD2 have both been shown to confer position-independent expression on heterologous genes (Ryan *et al.*, 1989; Greaves *et al.*, 1989). At present, it is not clear by what mechanism these sequences mediate position-independent expression.

BLG genes comprising 406 bp or more of the proximal 5′ flanking sequences showed a clear relationship between copy number and level of expression in the mammary gland, indicating that they may function in a position-independent manner. Sequences within the promoter region of BLG are clearly essential for efficient expression. Furthermore, BLG expression correlates with the presence of a strong DNAaseI-hypersensitive site in this region. By contrast to the β-globin LCR and the 3′ regulatory region identified for the CD2 gene, the BLG promoter does not function independently since it does not drive the expression of a heterologous gene, α_1-antitrypsin, in a position-independent manner (Archibald *et al.*, 1990). This conclusion is supported by the experiments utilizing two BLG minigenes, which lack BLG introns, in which expression was detected in less than half the mice analysed (Whitelaw *et al.*, 1991). Our current hypothesis is that the position-independent expression of BLG transgenes requires an interaction between the promoter region and element(s) located within and/or downstream of the gene. Intriguingly, no DNAaseI-hypersensitive sites have been detected within this region.

Targeting of efficient expression of foreign proteins to the mammary gland

There is considerable interest in transgenic animals as production systems for proteins. The approach generally taken has been to target expression of the proteins of interest to the mammary gland, with the intention of harvesting them from milk (Wilmut *et al.*, 1990). We have previously used BLG sequences to target expression of human factor IX and human α_1-antitrypsin to the mammary glands of sheep and mice (Clark *et al.*, 1989; McClenaghan *et al.*, 1991); however, with only one hybrid construct was a high level of expression detected (Archibald *et al.*, 1990; Wright *et al.*, 1991; Whitelaw *et al.*, 1991). Other workers attempting the same approach have also been hampered by low expression levels (Gordon *et al.*, 1987; Yu *et al.*, 1989; Buhler *et al.*, 1990; Denman *et al.*, 1991; but see Meade *et al.*, 1990). Defining the elements within the BLG gene that mediate efficient expression and understanding how they function will facilitate the design of hybrid genes which can reliably target the expression of heterologous genes to the mammary gland.

We are grateful to Jen Anderson, Sandra Couperwhite, Morag Robertson, Wendy Shephard, Roberta Wallace and John Webster for excellent technical assistance, Anthea Springbett for statistical analysis, Henryk Lubon for advice on nuclei isolation, and Christine Watson for helpful discussion. We thank Jean-Claude Mercier and Jeff Rosen for providing plasmids pBlg931, pMWAP-14 and pCMβ13.

REFERENCES

Abbey, L. M. & Witorsch, R. J. (1984) Oral Surg. **58**, 682–687
Ali, S. & Clark, A. J. (1988) J. Mol. Biol. **199**, 415–426
Ali, S., McClenaghan, M., Simons, J. P. & Clark, A. J. (1990) Gene **91**, 201–207
Allen, D. N., Norris, M. L. & Surani, M. A. (1990) Cell **61**, 853–861
Al-Shawi, R., Kinnaird, J., Burke, J. & Bishop, J. O. (1990) Mol. Cell. Biol. **10**, 1192–1198
Archibald, A. L., McClenaghan, M., Hornsey, V., Simons, J. P. & Clark, A. J. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 5178–5182
Ball, R. K., Friis, R. R., Schoenenberger, C. A., Doppler, W. & Groner, B. (1988) EMBO J. **7**, 2089–2095
Bayna, E. M. & Rosen, J. M. (1990) Nucleic Acids Res. **18**, 2877–2985
Bonifer, C., Vidal, M., Grosveld, F. & Sippel, A. E. (1990) EMBO J. **9**, 2843–2848
Bonnerot, C., Grimber, G., Briand, P. & Nicolas, J.-F. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6331–6335
Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E. & Palmiter, R. D. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 836–840
Buhler, T. A., Bruyere, T., Went, D. F., Stranzinger, G. & Burki, K. (1990) Bio/Technology **8**, 140–143
Chomczvnski, P. & Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
Church, G. M. & Gilbert, W. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 1991–1995
Clark, A. J., Bessos, H., Bishop, J. O., Brown, P., Harris, S., Lathe, R., McClenaghan, M., Prowse, C., Simons, J. P., Whitelaw, C. B. A. & Wilmut, I. (1989) Bio/Technology **7**, 487–492
Denman, J., Hayes, M., O'Day, C., Edmunds, T., Barlett, C., Hirani, S., Ebert, K. M., Gordon, K. & McPherson, J. M. (1991) Bio/Technology **9**, 839–843
Doppler, W., Groner, B. & Ball, R. K. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 104–108
Eisenstein, R. S. & Rosen, J. M. (1988) Mol. Cell. Biol. **8**, 3183–3190
Elgin, S. C. R. (1988) J. Biol. Chem. **263**, 19259–19262
Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13
Feinberg, A. P. & Vogelstein, B. (1984) Anal. Biochem. **137**, 266
Forsyth, I. A. (1986) J. Dairy Sci. **69**, 886–903
Gaye, P., Hue-Delahaie, D., Mercier, J.-C., Soulier, S., Vilotte, J.-L. & Furet, J.-P. (1986) Biochimie **68**, 1097–1107
Gordon, K., Lee, E., Vitale, J. A., Smith, A. E., Westphal, H. & Hennighausen, L. (1987) Bio/Technology **5**, 1183–1187
Greaves, D. R., Wilson, F. D., Lang, G. & Kioussis, D. (1989) Cell **56**, 979–986
Greer, P., Maltby, V., Rossant, J., Bernstein, A. & Pawson, T. (1990) Mol. Cell. Biol. **10**, 2521–2527
Grosveld, F., van Assendelft, G. B., Greeves, D. R. & Kollias, G. (1987) Cell **51**, 975–985
Gupta, P., Rosen, J. M., Eustachio, P. D. & Ruddle, F. H. (1982) J. Cell Biol. **93**, 199–204

Gurzburg, W. H., Salmons, B., Zimmermann, B., Muller, M., Erfle, V. & Brem, G. (1991) Mol. Endocrinol. **5**, 123–133
Hall, L., Emery, D. C., Davis, M. S., Parker, D. & Craig, R. K. (1987) Biochem. J. **242**, 735–742
Harris, S., Ali, S., Anderson, S., Archibald, A. L. & Clark, A. J. (1988) Nucleic Acids Res. **16**, 10379
Harris, S., McClenaghan, M., Simons, J. P., Ali, S. & Clark, A. J. (1991) Dev. Genet. **12**, 299–307
Hatakeyama, S., Sahima, M., Shirasuna, K., Satoh, M. & Suzuki, A. (1988) Cancer **62**, 716–722
Higgs, D. R., Wood, W. G., Jarman, A. P., Sharpe, J., Lida, J., Pretorius, I.-M. & Ayyub, H. (1990) Genes Dev. **4**, 1588–1601
Hogan, B., Costantini, F. & Lacy, E. (1986) Manipulating The Mouse Embryo: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
Jantzen, K., Fritton, H. P. & Igo-Kemenes, T. (1986) Nucleic Acids Res. **14**, 6085–6099
Jenness, R. (1982) in Developments In Dairy Chemistry-1: Proteins (Fox, P. F., ed.), pp. 87–115, Applied Science, London
Knight, C. H. & Peaker, M. (1982) J. Reprod. Fertil. **65**, 521–536
Lang, G., Wotton, D., Owen, M. J., Sewell, W. A., Brown, M. H., Mason, D. Y., Crumpton, M. J. & Kioussis, D. (1988) EMBO J. **7**, 1675–1682
Lathe, R., Vilotte, J.-L. & Clark, A. J. (1987) Gene **57**, 193–201
Lee, K.-F., DeMayo, F. J., Atiee, S. H. & Rosen, J. M. (1988) Nucleic Acids Res. **16**, 1027–1041
Levine, J. F. & Stockdale, F. E. (1985) J. Cell Biol. **100**, 1415–1422
Li, M. L., Aggeler, J., Farson, D. A., Haiter, C., Hassel, J. & Bissel, M. J. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 136–140
Lubon, H. & Hennighausen, L. (1987) Nucleic Acids Res. **15**, 2103–2121
Lubon, H. & Hennighausen, L. (1988) Biochem. J. **256**, 391–396
McClenaghan, M., Archibald, A. L., Harris, S., Simons, J. P., Whitelaw, C. B. A., Wilmut, I. & Clark, A. J. (1991) Animal BioTech. **2**, 161–176
Meade, H., Gates, L., Lacy, E. & Lonberg, N. (1990) Bio/Technology **8**, 443–446
Mercier, J.-C., Vilotte, J.-L. & Provot, C. (1990) in Genome Analysis in Domestic Animals (Gelderman, H. & Ellendorff, F., eds.), pp. 233–258, VCH, Weinheim
Orkin, S. H. (1990) Cell **63**, 665–672
Ott, M.-O., Sperling, L., Herbomel, P., Yaniv, M. & Weiss, M. C. (1984) EMBO J. **3**, 2505–2510
Pinkert, C. A., Ornitz, D. M., Brinster, R. L. & Palmiter, R. D. (1987) Genes Dev. **1**, 268–276
Pittius, C. W., Hennighausen, L., Lee, E., Westphal, H., Nicols, E., Vitale, J. & Gordon, K. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 5874–5878
Ryan, T. M., Behringer, R. R., Townes, T. M., Palmiter, R. D. & Brinster, R. L. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 37–41
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor
Simons, J. P., McClenaghan, M. & Clark, A. J. (1987) Nature (London) **328**, 530–532
Sollner-Webb, B. & Reeder, R. H. (1979) Cell **18**, 485–499
Topper, Y. J. & Freeman, C. S. (1980) Physiol. Rev. **60**, 1049–1106
Vilotte, J.-L., Soulier, S., Stinnarke, M.-G., Massoud, M. & Mercier, J.-C. (1989) Eur. J. Biochem. **186**, 43–48
Watson, C. J., Gordon, K. E., Robertson, M. & Clark, A. J. (1991) Nucleic Acids Res. **19**, 6603–6610
Weintraub, H. (1985) Cell **42**, 705–711
Whitelaw, C. B. A., Archibald, A. L., Harris, S., McClenaghan, M., Simons, J. P., Springbett, A., Wallace, R. & Clark, A. J. (1990) Mouse Genome **88**, 114
Whitelaw, C. B. A., Archibald, A. L., Harris, S., McClenaghan, M., Simons, J. P. & Clark, A. J. (1991) Transgenic Res. **1**, 3–13
Wilkie, T. M., Brinster, R. L. & Palmiter, R. D. (1986) Dev. Biol. **118**, 9–18
Wilmut, I., Archibald, A. L., Harris, S., McClenaghan, M., Simons, J. P., Whitelaw, C. B. A. & Clark, A. J. (1990) J. Reprod. Fertil. **41**, 135–146
Wright, G., Carver, A., Cottom, D., Reeves, D., Scott, A., Simons, P., Wilmut, I., Garner, I. & Colman, A. (1991) Bio/Technology **9**, 830–834
Wu, C. (1980) Nature (London) **286**, 854–860
Yu, S.-H., Deen, K. C., Lee, E., Hennighausen, L., Sweet, R. W., Rosenberg, M. & Westphal, H. (1989) Mol. Biol. Med. **6**, 255–261

Received 11 October 1991/12 February 1992; accepted 26 February 1992