Transgenic Research, **1**, 3–13 (1991)

Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice

C. BRUCE A. WHITELAW*, ALAN L. ARCHIBALD, STEPHEN HARRIS‡, MARGARET McCLENAGHAN, J. PAUL SIMONS and A. JOHN CLARK

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

Received 13 May 1991; revised 27 August 1991; accepted 4 September 1991

We are studying the tissue-specific expression of the sheep milk-whey protein gene, β-lactoglobulin. We have used sequences derived from this gene to target the expression of biomedical proteins into milk with the intention to exploit this technology in transgenic sheep as a means of protein production. In the present study, a series of β-lactoglobulin hybrid genes and β-lactoglobulin minigenes were evaluated for expression in the mammary gland of transgenic mice. In particular, we have assessed whether there is a requirement for introns for efficient transgene expression in the mammary gland, since the coding sequences of many candidate proteins are available only as cDNAs. The results suggest that the inclusion of natural introns in constructs can enhance the efficiency of transgene expression. Thus, a hybrid construct comprising 4.3 kb of the immediate 5′ flanking sequences of β-lactoglobulin fused to a genomic minigene encoding human α-antitrypsin (α_1AT) was expressed much more efficiently than an α_1AT-cDNA construct containing the same β-lactoglobulin segment. Similarly, the intact β-lactoglobulin gene was expressed more efficiently than the corresponding intronless β-lactoglobulin minigene. This effect was not seen in transient expression experiments in baby hamster kidney cells when β-lactoglobulin-α_1AT constructs were driven by SV40 enhancer sequences. The effect cannot be explained by a simple requirement for splicing, since the inclusion of the first β-lactoglobulin intron into cDNA constructs encoding human α_1AT or β-lactoglobulin itself failed to enhance the efficiency of transgene expression. It is concluded that sequence elements within introns may interact with the upstream 5′ flanking sequences of β-lactoglobulin and enable the latter to function efficiently in the mammary gland of transgenic mice.

Keywords: β-lactoglobulin, tissue specificity, hybrid genes, RNA splicing, intervening sequences

Introduction

Transgenic animals are a powerful tool for the study of eukaryotic gene expression. DNA sequences introduced into the germline are exposed to the range of developmental signals that endogenous genes experience and transgenic animals can reveal mechanisms of gene control that are not evident in cultured cells (compare Ott *et al.* (1984) with Pinkert *et al.* (1987) and Brinster *et al.* (1988)). In particular, the properties of differentiated cell types may be altered significantly by immortalization during *in vitro* culture, making them an unreliable system to study tissue-specific expression. Indeed, many cell types are intractable to *in vitro* culture and transgenic animals are the only way to intervene at the molecular-genetic level.

Transgenic animals also offer a number of biotechnological applications. The prospects for the direct genetic improvement of livestock using this technology have been reviewed by a number of authors (e.g. Simons and Land, 1987; Pursel *et al.*, 1989). In addition to improving existing traits, animals with entirely new properties can be produced. At present there is considerable interest in transgenic animals for the production of recombinant

*To whom correspondence should be addressed.

‡Present address: Glaxo Group Research, Greenford Road, Greenford, Middlesex UB6 0HE, UK.

0962-8819/91 $03.00 + .12 © 1991 Chapman & Hall

proteins, in particular human biomedical proteins, as an alternative to large scale mammalian cell culture. The approach generally taken has been to target expression to the mammary gland so that the protein product can be harvested from milk (Lathe *et al.* 1986; Clark *et al.* 1987; Whitelaw and Clark, 1989).

A number of milk protein genes have been isolated and shown to function appropriately in the mammary gland of transgenic mice (Simons *et al.*, 1987; Vilotte *et al.*, 1989; Bayna and Rosen, 1990). Regulatory sequences from these and other milk protein genes have been used to target the expression of human proteins to the mammary gland. Indeed, several human proteins with therapeutic potential have been produced in the milk of transgenic mice (Gordon *et al.*, 1987; Yu *et al.*, 1989; Archibald *et al.*, 1990; Meade *et al.*, 1990), rabbits (Buhler *et al.*, 1990) and sheep (Clark *et al.*, 1989). The proteins produced exhibited biological activity showing that the mammary gland is capable of carrying out appropriate post-translational modifications. However, a number of these studies reported both low frequencies and/or low levels of transgene expression, indicating that the configuration of regulatory sequences and protein coding sequences was not optimal. Poor transgene performance is a major constraint on the adoption of this technology, because in most instances transgenic dairy animals such as sheep or goats will be necessary for the required volume of production. Yet the efficiency of generating transgenic livestock is low and the procedure is both expensive and time-consuming (Pursel *et al.*, 1989; Clark *et al.*, 1991). Therefore, it is important to gain a better understanding of the mechanisms that control gene expression in the mammary gland so that these can be manipulated to efficiently target the expression of heterologous proteins.

We have previously shown that the sheep gene encoding the major milk whey protein β-lactoglobulin (BLG; Ali and Clark, 1988) is expressed efficiently and specifically in the mammary gland of transgenic mice (Simons *et al.* 1987; Harris *et al.* 1990). We have also described fusion genes derived from the BLG gene that express low levels of factor IX (fIX) in the milk of transgenic sheep (Clark *et al.*, 1989) and high levels of human α_1-antitrypsin (α_1AT) in transgenic mouse milk (Archibald *et al.*, 1990). The present studies were undertaken to compare the performance of a series of BLG hybrid genes and BLG minigenes designed to target expression to the mammary gland. In particular, we wished to assess the requirement for introns in obtaining efficient expression in the mammary gland of transgenic mice, since clones encoding many candidate proteins are available only as cDNA sequences and many of the corresponding genes are very large.

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 (Maniatis *et al.*, 1982). The constructs were derived by the manipulation of the genomic ovine BLG clone pSS1tgXS, which comprises of 4.3 kb of 5′ flanking sequences, the 4.7 kb transcription unit and 1.9 kb of 3′ flanking sequences (Simons *et al.*, 1987; Ali and Clark, 1988; Harris *et al.*, 1988), the ovine BLG cDNA clone pBlg931 (Gaye *et al.*, 1986), the genomic human α_1AT clone pATp7 (Kelsey *et al.*, 1987), the human α_1AT cDNA clone p8α1ppg (Ciliberto *et al.*, 1985) and the human fIX cDNA clone p5′ G3′ cVI (Anson *et al.*, 1985). The region defined by the *Pvu* II site within exon 1 of the BLG gene to the 5′ (*Sal* I) of the pSS1tgXS clone (comprising 4.3 kb of 5′ flanking sequences and 32 bp of the 5′ untranslated sequences of the BLG transcription unit) is common to all the constructs used in these studies.

BLG and AAT-B

The entire BLG gene construct (Simons *et al.*, 1987) and the AAT-B hybrid gene which comprises BLG 5′ flanking sequences fused to a α_1AT minigene, pIII-15′BLGgAAT (Archibald *et al.*, 1990), have been described previously.

AAT-C

The internal *Sty* I (exon 2) to *Eco* RI (intron 2) fragment of pSS1tgSE was replaced with the partially complementary oligonucleotides CTTGTGATATCG and CACTATAGCTTA, inserting an inframe termination codon (TGA) in the BLG reading frame, and generating a novel *Eco* RV site within exon 2. The *Eco* RI site was blunted by filling in and ligated to the *Pvu* II site of exon 5; introns 5 and 6 were removed by replacing the region between this *Pvu* II site and a *Sma* I in exon 7 with the corresponding cDNA sequence from pBlg931. Subsequently, pSS1pUCXSAAT.TGA was generated by inserting, the 1293 bp *Taq* I–*Bst* NI α_1AT-coding fragment from p8α1ppg at the *Eco* RV site. The 8.8 kb AAT-C fragment was released from pSS1pUCXSAAT. TGA by *Sal* I plus *Xba* I digestion. This construct has the potential to encode a bi-cistronic mRNA, since it contains two open reading frames. The upstream open reading frame potentially encodes a truncated BLG protein while the second open reading frame encodes human α_1AT.

AAT-D

The *Taq* I–*Bst* NI α_1AT cDNA fragment was inserted between the *Pvu* II sites of exons 1 and 5 of pSS1tgXS and introns 5 and 6 were removed, as described above, to generate the plasmid pBJ16. *Sal* I plus *Xba* I digestion of pBJ16 released the 8.0 kb AAT-D fragment for microinjection.

FIX-D

The 1552 bp *Nhe* I–*Hind* III fragment from p5′G3′cVI, containing the human fIX coding region, was inserted between the *Pvu* II sites within exons 1 and 5 of pSS1tgXS, with introns 5 and 6 removed, as described above, to generate pBJ17. Digestion of pBJ17 with *Sal* I and *Xba* I released the 8.2 kb fIX-D fragment for microinjection.

BLG.0I

The intronless BLG-minigene pUCβlacD was generated by fusing a pair of 44 bp oligonucleotides, comprising the cDNA sequence between the *Stu* I site in exon 1 and the *Sty* I site in exon 2 of the BLG gene, with the 457 bp *Sty* I–*Sma* I fragment from pBlg931, and inserting the resulting 501 bp fragment between the corresponding sites in exon 1 and exon 7 of pSS1tgXS. *Sal* I–*Xba* I digestion of pUCβlacD released the 6.8 kb fragment BLG.0I for microinjection.

BLG.1I

The 424 bp internal *Pst* I fragment from pBlg931, comprising BLG cDNA sequences was used to replace the equivalent genomic segment between exons 2 and 7 of pSS1tgXS generating pSS1tgXSint1 containing only the first of the six native BLG introns. *Sal* I–*Xba* I digestion of pSS1tgXSint1 released the 7.5 kb fragment BLG.1I for microinjection.

SVE constructs

Constructs destined for cell transfection experiments were modified by the insertion of the SV-40 enhancer upstream of the BLG cap-site. The 179 bp *Pvu* II–*Fok* I fragment (Benoist and Chambon, 1981) fragment isolated from pUC.SVE (supplied by J.O. Bishop) was inserted in a 5′ to 3′ orientation at the *Sph* I site of pSS1tgXS, pIII-I5BLGgAAT, pSS1pUCXSAAT.TGA and pBJ16 to generate SVE.BLG, SVE.AAT-B, SVE.AAT-C and SVE.AAT-D, respectively.

GENERATION OF TRANSGENIC MICE

Production of transgenic mice by microinjection of DNA into the pronuclei of fertilized eggs from superovulated (C57BL/6 × CBA) F_1 female mice, which had been mated with F_1 stud males, was performed as described previously (Hogan *et al.*, 1986; Simons *et al.*, 1987). Transgenic lines were propagated by mating to F_1 mice. DNA, from tail biopsy material taken at weaning, was prepared by extraction with phenol/chloroform and isopropanol precipitation, digested with the appropriate restriction enzymes, subjected to agarose gel electrophoresis and blotted to nylon membranes (Hybond N, Amersham International, Amersham, UK). Southern blots were hybridized (Church and Gilbert, 1984) with random oligo-primed probes (Multiprime, Amersham International, Amersham, UK; Feinburg and Vogelstein, 1983; 1984). For rapid screening of established lines, a polymerase chain reaction (PCR) assay was performed on the crude tail digests. Segments of tail were digested overnight with 200 mg ml^{-1} proteinase K in 0.3 M sodium acetate, 10 mM Tris-Cl, pH 7.9, 1 mM EDTA, 1% SDS. The crude tail digests were chilled and centrifuged for 10 min in a microfuge at 4° C to pellet SDS and debris. A 1 μl aliquot of the supernatant was incubated at 95° C for 10 min followed by the addition of 50 ml of PCR reaction mix, and temperature cycling. PCR was performed using primers GCTTCTGGGGTCTACCAGGAAC and TCGTGCT-TCTGAGCTCTGCAG which define a 246 bp region of the 5′ end of the BLG gene, and control primers GAGTTCCGGAACTGCCTTTGGTG and CTGTGC-CACCGGGCGCATGG which define a 332 bp region of the mouse hypoxanthine–guanine phosphoribosyltransferase gene. The inclusion of the internal control primers enables non transgenic mice to be positively identified. The reaction mix contained 10 mM TrisCl pH8.3 at 25° C, 50 mM KCl, 1.5 mM $MgCL_2$, 0.01% gelatin, 0.2 mM each dNTP, 10% DMSO, each primer at a concentration of 0.1 mM and 1.25 units of *Taq* polymerase per reaction. Thirty cycles of PCR were performed incubating at 91.5° C for 30 sec and 65° C for 5 min (temperature in tube, and time at temperature). Aliquots were analysed in ethidium bromide-containing 2% agarose gels.

RNA ANALYSIS

RNA was prepared from mammary, liver, kidney, spleen, salivary and heart tissue isolated from mid-lactation females. The tissue sample was homogenized in 2 ml RNAzol (Biogenesis; Chomezynski and Sacchi, 1987) and analysed by Northern blotting. Total RNA (10 μg unless stated otherwise) was electrophoresed in 1% denaturing MOPS-formaldehyde agarose gels and then transferred and probed as described for Southern blots. RNA levels were quantified by probing dot blots loaded with serial RNA dilutions, using sheep mammary and human liver RNA as controls and corrected for loading by reprobing with ribosomal RNA sequences. After autoradiography films were scanned on a Schimadzu CS-9000 densitometer. S1 mapping was carried out as described by Davis *et al.* (1986). RNA samples were hybridized with a single-stranded 185 nucleotide *Sph* I–*Taq* I probe which encompasses the 5′ end of the BLG transcription unit and digested with S1 nuclease (Gibco BRL). Protected fragments were analysed on a 8% urea-acrylamide gel and autoradiographed.

PROTEIN ANALYSIS

Mid-lactation females were milked, and milk samples analysed by SDS-PAGE, Western blotting and radioimmunoassay as described previously (Simons *et al.*, 1987;

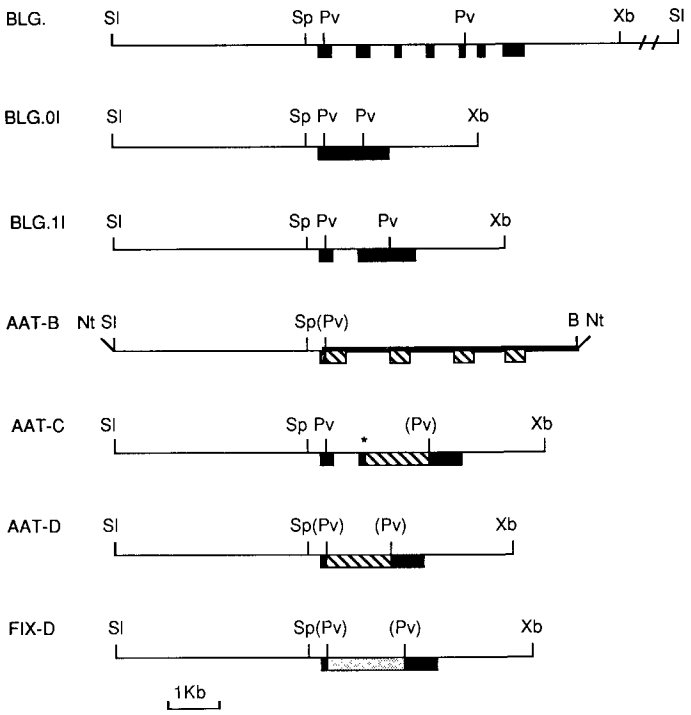

Fig. 1. Diagram of BLG, BLG minigenes and BLG hybrid gene constructs introduced into mice. Symbols: — , BLG flanking sequences and introns; ▬ , α_1AT flanking sequences and introns; ■ , BLG exons and cDNA; ▨ , α_1AT exons or cDNA; ▩ , fIX cDNA; *, position of inserted termination codon. Relevant restriction sites; B, *Bam* HI; Nt, *Not* I; Sl, *Sal* I; Sp, *Sph* I; Pv, *Pvu* II; Xb, *Xba* I; those BLG derived sites lost during ligation are shown in brackets. Note that BLG defines two constructs, containing 1.9 kb and 7.6 kb of 3′ flanking sequences, respectively (Simons *et al.*, 1987).

Archibald *et al.*, 1990). Samples from AAT-D mice were also analysed by an α1AT-specific enzyme-linked immunosorbant assay. BLG estimates were by densitometry, using a Schimadzu CS-9000 densitometer, of Coomassie-Blue stained SDS-PAGE gels, with reference to internal standards of purified BLG.

CELL CULTURE AND TRANSIENT TRANSFECTIONS

BHK (baby hamster kidney fibroblast cell line) cells were grown in Dulbecco's Modified Eagle medium (Gibco) supplemented with 10% newborn calf serum and glutamine and seeded at approximately 1×10^6 cells per 25 cm^2 flask. Plasmid DNA was introduced into cells by the calcium phosphate technique (Graham and van dr Eb, 1973) with the following modifications. The precipitate (1 ml) was added to 9 ml of culture medium. After 6 h the medium was replaced and the cells harvested 42 h later. An SVE.CAT construct, containing a 350 bp *Pvu* II–*Hind* II fragment encompassing the SV-40 enhancer/promoter (Benoist and Chambon, 1981) inserted into the *Hinc* II site of the promoterless CAT construct, pB9 (Whitelaw *et al.*, 1988), was used to standardize transfection efficiency. CAT assays were performed as previously described (Whitelaw *et al.*, 1988). RNA was harvested and the spent medium collected and analysed for the presence of secreted human α_1AT by radioimmunoassay.

Results

TARGETING THE EXPRESSION OF HUMAN α_1AT AND HUMAN FIX TO THE MAMMARY GLAND

Four BLG-hybrid genes were analysed in transgenic mice (Fig. 1; and see Materials and methods). Three of these comprise sequences encoding human α_1AT and one coding sequences of human fIX. Two of the three BLG-α_1AT hybrid genes, AAT-C and AAT-D, comprised α_1AT cDNA sequences, whereas the third, AAT-B (Archibald *et al.*, 1990) comprised α_1AT genomic sequences. FIX-D is analogous to AAT-D and comprises a 1552 bp fIX-encoding cDNA segment inserted into the BLG minigene.

Transgenic mice were produced by microinjection of the four constructs; 15 AAT-B, 7 AAT-C, 8 AAT-D and 10 FIX-D transgenic G_0 animals were identified as carrying one or more transgene copies by Southern blotting. Expression was analysed in G_0 females while transgenic lines were established from G_0 males and expression analysed in the females of subsequent generations. RNA was isolated from several tissues of mid-lactation females and analysed by Northern blotting. Milk was also collected from these animals and analysed by Western blotting and radioimmunoassay.

Construct AAT-D

AAT-D transcripts were not detected in the mammary gland, or in any of the other tissues analysed in the eight animals or lines carrying this transgene (Fig. 2). We did, however detect human α_1AT in the milk of one G_0 female carrying AAT-D in which we were unable to detect transcripts even in polyA^+ RNA selected populations. One possibility for this anomaly is mosaicism in this founder animal (Whitelaw *et al.*, 1990); a transgenic line was not established from this animal.

Construct AAT-C

In the mammary gland, 1/7 AAT-C animals or lines expressed a transcript of the expected size (1700 nucleotides) (Figs 2 and 3). In this mouse, however, human α_1AT protein was not detected in the milk. Synthesis of α_1AT directed by AAT-C mRNA is expected to involve a re-initiation of translation (Peabody and Berg, 1986; Kozak, 1987), and in cell culture this transcript is translated significantly less efficiently than AAT-B or AAT-D transcripts (data not shown).

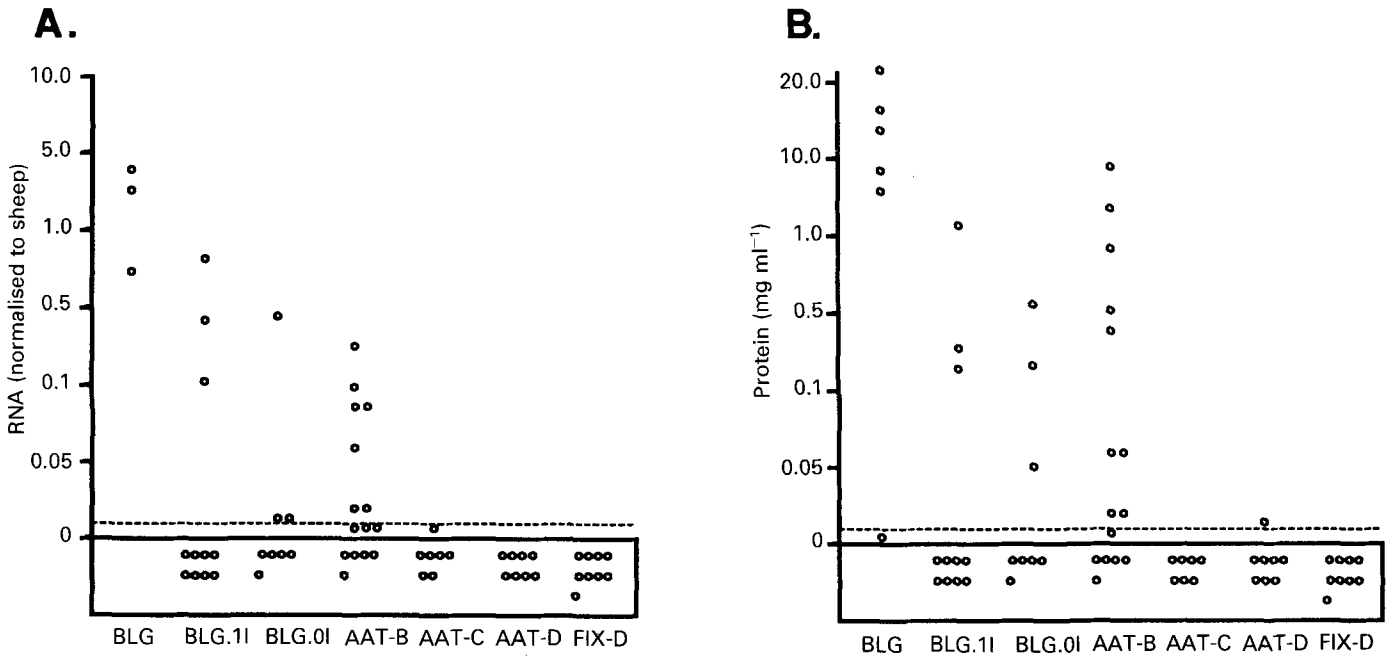

Fig. 2. Comparison of transgene expression. Steady-state BLG and α_1AT mRNA levels in the mammary gland (A) and protein levels in mouse milk (B) are compared for the various constructs described. Levels were measured as described in the Materials and methods section. Each circle corresponds to a G_0 or G_1 (when the G_0 was male) female transgenic mouse; values below the dotted line were detectable but could not be accurately determined (estimated to be $< 10^{-3}$ fold lower than standard lactating sheep for RNA; < 10 ng ml^{-1} for BLG, < 10 μg ml^{-1} for α_1AT); values below the origin represent mice in which no transgene product was detected (estimated to be at least < 10 to 4-fold lower than standard lactating sheep for RNA; < 1 ng ml^{-1} for BLG, < 5 μg ml^{-1} by radio-immunoassay and < 0.02 μg ml^{-1} by enzyme linked immunosorbant assay for α_1AT, < 1 μg ml^{-1} for fIX). The RNA levels are normalized to that of a standard lactating sheep mammary sample, while protein values are in mg ml^{-1} (lactating sheep milk contains 4.6 mg ml^{-1} BLG: Simons *et al.*, 1987). Note the change in scale at 0.1 and 1.0 units, and that RNA levels were not determined for some of the BLG mice. Statistical comparison of expression frequencies (Fisher's Exact Test) determined that BLG.0I and BLG.1I were significantly different from BLG ($p < 0.01$ and $p < 0.05$, respectively); AAT-D and fIX-D were significantly different from AAT-B ($p < 0.05$ and $p < 0.01$, respectively) while AAT-C was nearly significantly different from AAT-B ($p < 0.06$); no significant difference was detected between BLG and AAT-B.

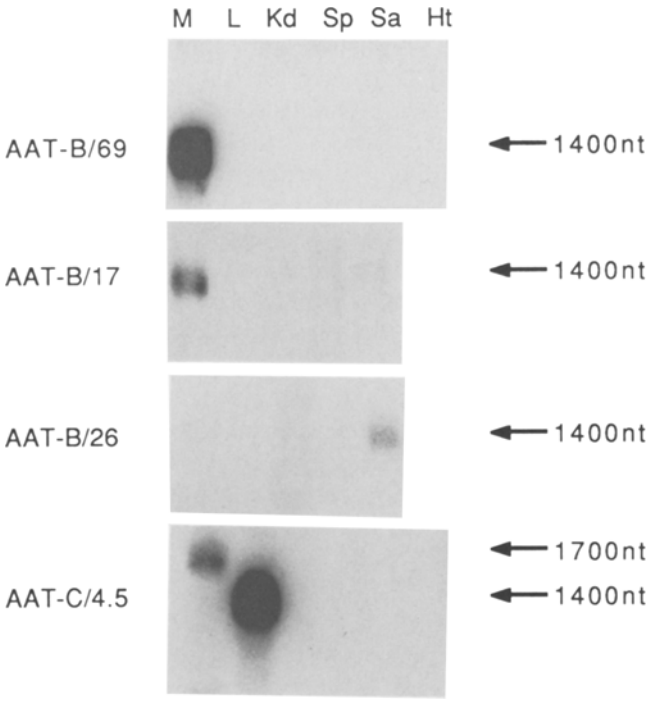

Fig. 3. Tissue-specificity of AAT-B and AAT-C transgene expression. Northern blot analysis of total RNA isolated from mid-lactation female mice; AAT-B/69; AAT-B/17; AAT-B/26 and AAT-C/4.5. Tissues analysed; M, mammary; L, Liver; K, kidney; Sp, spleen; Sa, salivary; Ht, heart. For AAT-B/17, AAT-B/26 and AAT-B/69 mice, a human α_1AT specific probe was used (Archibald *et al.*, 1990); for AAT-C/4.5, p8α1ppg was used as probe, and detects both human and mouse α_1AT. The autoradiographs were overexposed to maximise the sensitivity of mRNA detection.

Table 1. Comparison of BLG and AAT-B gene expression in the mammary gland of transgenic mice.

Mouse/line	*RNA*[a]	*Protein*[b]	*Protein: RNA ratio*	*Mean ± SE*[c]
Sheep	1.0	4.6	–	
BLG/7	3.3	23.0	7.0	
BLG/14	0.75	5.0	6.7	
BLG/45	3.45	17.9	5.2	6.3 ± 1.1
AAT-B/17	0.02	0.6	30.0	
AAT-B/35	0.25	7.3	29.2	
AAT-B/44	0.08	0.9	11.2	
AAT-B/62	0.06	3.0	50.0	
AAT-B/69	0.06	0.4	6.7	25.4 ± 15.1

[a] Determined by scanning densitometry of serial RNA dilutions, corrected for loadings by rRNA levels and normalized to standard RNA prepared from lactating sheep mammary gland.
[b] Protein levels (mg ml^{-1}) in milk were determined by RIA (Archibald *et al.*, 1990), BLG densitometry data from Simons, McClenaghan and Clark, (1987).
[c] Mean protein:RNA ratio (95% confidence limits).

Construct AAT-B

AAT-B transcripts of the expected size (1400 nucleotides) were detected in the mammary gland of ten of the 15 animals (or lines) analysed (Fig. 2). AAT-B expression was restricted to the mammary gland in six of these, but in the other four expression was also detected in the salivary gland. Additionally, four animals (or lines) exhibited expression in the salivary gland but not in the mammary gland (Fig. 3). Human α_1AT was detected in the milk of all ten animals (or lines) that expressed AAT-B mRNA in the mammary gland. It is not known whether the transcripts observed in the salivary gland are also translated. In the mammary gland, there was a general correlation between mRNA levels and protein concentration in the milk (Table 1). There was, however, a wide variation in the levels of expression between the different animals (or lines) which was not correlated with transgene copy number (Fig. 2; Archibald *et al.*, 1990).

In the highest expressing AAT-B line (AAT-B/35), which produces 7–8 mg ml^{-1} α_1AT in milk, the level of α_1AT mRNA was estimated to be approximately twice that observed in human liver, the predominant site of expression of α_1AT in man. The levels of expression of AAT-B mRNA were compared directly to the levels of BLG mRNA expression in sheep and transgenic mice using a BLG-specific oligonucleotide that hybridized to the 5′ untranslated region of both transcripts. On average, the level of AAT-B mRNA was lower than the level of BLG mRNA in sheep mammary gland (Fig. 2; Table 1). It was also lower than the BLG mRNA levels in transgenic mice, most of which express at levels similar or greater than that in sheep (Fig. 2). In the mammary gland of AAT-B/35 mice the steady-state mRNA level was estimated to be a quarter that of BLG mRNA in sheep. Calculating protein or RNA ratios for BLG and AAT-B in transgenic mice and correcting for molecular weights of the two proteins suggests that the transcripts were translated and their products secreted into milk with approximately equal efficiencies (Table 1).

Construct FIX-D

No expression of RNA or protein was observed in the mammary gland or any of the other tissue analysed in the ten animals (or lines) generated (Fig. 2).

These results show that hybrid BLG genes containing cDNAs were not efficiently expressed in transgenic mice. By contrast, AAT-B which comprises genomic α_1AT sequences was expressed relatively efficiently.

HYBRID GENES ARE EXPRESSED IN FIBROBLASTS

The hybrid BLG genes analysed in transgenic mice were modified by the addition of SV40 enhancer (SVE) sequences inserted just upstream of the TATA box and tested for expression by transfection into BHK cells. This approach was taken because the available mammary-derived cell lines (Danielson *et al.*, 1984; Ball *et al.*, 1988) do not reliably mimic the differentiated state of the mammary gland (Ball *et al.*, 1988; Eisenstein and Rosen, 1988; Doppler *et al.*, 1989). The SVE–BLG constructs do not yield information on targeting expression to the mammary gland, however, they do establish whether these hybrid genes can be transcribed given the appropriate environment.

Located 10 bp upstream of the TATA box the SV40 enhancer sequences (Fig. 4a) were able to drive the expression of the BLG gene in BHK cells in transient expression experiments. The BLG transcripts were initiated correctly and were of the correct size (800 nucleotides; Fig. 4b,c). SV40-derivatives of AAT-B, AAT-C and AAT-D were expressed and RNA transcripts of the expected size were seen on Northern blots (Fig. 4d). From at least three experiments the average levels of steady-state mRNA were similar for each construct, when compared to that of control plasmids. Human α_1AT was found in the tissue-culture medium (data not shown) demonstrating that these transcripts were translated and the protein secreted. These experiments showed that, with the appropriate regulatory sequences, all three hybrid BLG-α_1AT genes were capable of being expressed at similar levels.

BLG MINIGENES ARE INEFFICIENTLY EXPRESSED IN TRANSGENIC MICE

Comparison of the performance of AAT-B to AAT-D and AAT-C suggested that natural introns could be

important in determining efficient expression of hybrid genes in transgenic mice. However, it was not clear whether this requirement was limited to targeting expression of heterologous genes, such as those encoding α_1AT or fIX, or whether it was a general requirement that would also apply to genes normally expressed in the mammary gland. To answer this question two BLG minigenes (Fig. 1), BLG.0I which is analogous to both AAT-D and FIX-D, and BLG.1I which is analogous to AAT-C, were constructed (see Materials and methods) and tested in transgenic mice.

Transgenic mice were produced by microinjection and eight BLG.0I and 11 BLG.1I transgenic G_0 animals were identified by Southern blotting. Expression was analysed in G_0 females and transgenic lines were established from G_0 males for analysis of the females of subsequent generations.

BLG.0I

As determined by both RNA and protein analyses only three out of eight mice (or lines) expressed the transgene (Fig. 2). This contrasts with transgenic mice carrying intact BLG transgenes where expression has always been detected (Fig. 2), with the exception of those animals carrying less than one copy per genome and thus presumed mosaic. Additionally, in a set of experiments involving a further 24 transgenic mice (or lines) carrying a BLG transgene comprising various amounts of 5′ flanking DNA, but otherwise identical to the 10.5 kb BLG transgene (Fig. 1), significant expression was observed in every case (C.B.A. Whitelaw, S. Harris, M. McClenaghan, J.P. Simons and A.J. Clark, unpublished data).

In the three positive BLG.0I mice (or lines) expression was tissue-specific, the BLG transcript was the expected size (800 nucleotides) and BLG was detected in the milk (Fig. 5a,b). However, from an analysis of the mRNA levels and the level of BLG protein measured in the milk the average level of expression was approximately 1/20th the level obtained with the genomic BLG transgenes.

BLG.1I

Only three out of 11 transgenic mice (or lines) expressed this transgene (Fig. 2). Again, in the three positive BLG.1I mice (or lines) expression was tissue-specific, the BLG transcripts were of the expected size and BLG was detected in the milk (Fig. 5a,b). The average level of expression in these three animals (or lines) was somewhat higher than observed for BLG.0I, i.e. about one fifth of the average level of expression observed for BLG.

The performance of BLG.0I and BLG.1I when

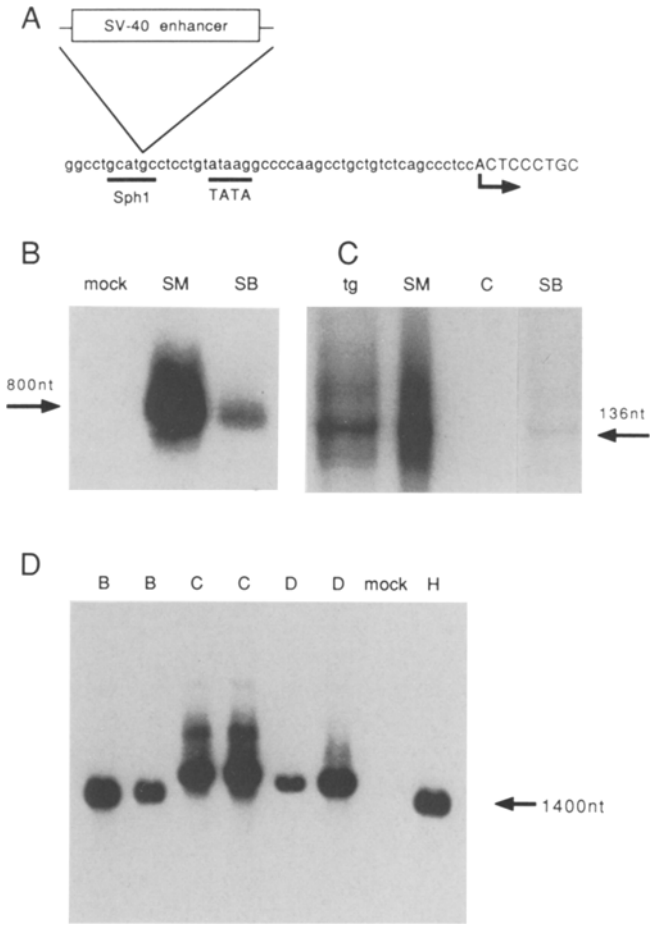

Fig. 4. Expression of SV40 driven BLG-α_1AT hybrid genes in transiently transfected BHK cells. (A) Sequence at the BLG transcription start site (Ali and Clark, 1988; Harris *et al.*, 1988) showing position of SV-40 enhancer insertion. The site of insertion at *Sph* I and the TATA box are both underlined. The arrow indicates the BLG transcription initiation site and the first 9 bp of BLG exon 1 are in upper case. (B) Northern blot of RNA isolated from cells either transiently transfected with SVE.BLG (SB) or no DNA (mock) and from lactating sheep mammary gland (SM), hybridized with a ^{32}P-labelled ovine BLG specific probe. (C) S1 protection analysis of the SVE.BLG transcript. A 185 nucleotide ^{32}P-labelled probe, complementary to nucleotides −44 and +141 with respect to the BLG transcription start site, was hybridized to RNA and after S1 nuclease digestion the protected fragments were electrophoresed on 8% urea-acrylamide gels; transgenic mouse BLG/14 mammary (tg), lactating sheep mammary (SM), control nontransgenic mouse mammary (C), and BHK cells transiently transfected with SVE.BLG (SB). A 136 nucleotide protected fragment was detected in the sheep, transgenic mouse and BHK-derived samples indicating that the transcriptional start site utilized in BHK cells is the same as that used in transgenic mice and sheep (Gaye *et al.*, 1986; Ali and Clark, 1988). (D) Northern blot analysis of RNA isolated from cells transiently transfected with SVE.AAT-B (B), SVE.AAT-C (C), SVE.AAT-D (D), no DNA (mock) and HepG2 cells (H). Total RNA (10 μg; except for HepG2 sample, 1 μg) was electrophoresed on 1% denaturing gels, transferred to hybond filters and hybridized to a ^{32}P-labelled α_1AT specific probe. The arrow indicates the size of the α_1AT transcript detected in HepG2 cells.

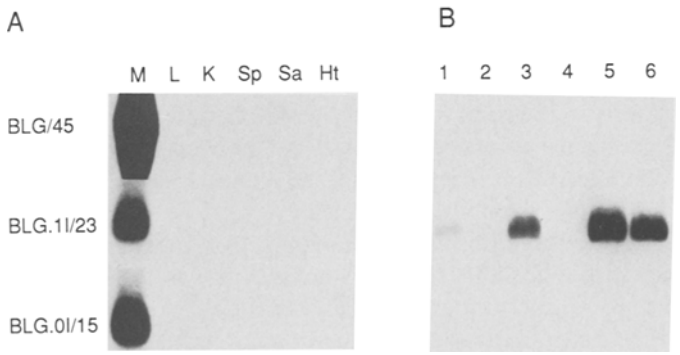

Fig. 5. Tissue-specific expression of BLG and BLG minigenes in transgenic mice. (A) Northern blot analysis of RNA isolated from mid-lactation female mice hybridized to a ^{32}P-labelled BLG specific probe. Tissues analysed; M, mammary; L, liver; K, kidney; Sp, spleen; Sa, salivary; Ht, heart. Autoradiographs were overexposed to maximize the sensitivity of mRNA detection. (B) Western blot analysis of mouse milk. Equivalent amounts of whey protein preparations were probed with a polyclonal antibody to ovine BLG (Simons *et al.*, 1987). Lanes 1, BLG.0I/15; lane 2, control mouse milk; lane 3, BLG.1I/23; lane 4, control mouse milk; lane 5, BLG/45; lane 6, standard lactating sheep.

compared to the intact BLG gene suggests that introns may be required for determining efficient expression of milk protein genes in the mammary gland.

Discussion

The BLG gene is efficiently expressed in transgenic mice (Simons *et al.*, 1987). A hybrid gene AAT-B, comprising 4.3 kb of 5′ BLG flanking sequences fused to an α_1AT minigene, is also expressed relatively efficiently in transgenic mice (Archibald *et al.*, 1990; Table 1). Thus, the mammary gland is capable of high-level expression of genes that are normally expressed in other tissues. However, the same upstream BLG sequences failed to efficiently drive several other α_1AT and flX hybrid genes. Therefore, the design of the transgene into which this segment is incorporated is crucial for efficient expression.

A common feature of all the poorly expressed constructs in this and our previous work (Clark *et al.*, 1989) is the use of cDNA sequences encoding the protein of interest. Similarly, Gordon *et al.* (1987) and Yu *et al.* (1989) have reported low level expression of cDNA-based constructs containing the mouse whey acidic protein (WAP) gene promoter. Poor expression of the constructs may also be due to transgene configuration since it is known that the intact WAP gene functions relatively efficiently in transgenic mice (Baynna and Rosen, 1990; Burdon *et al.*, 1991) and pigs (Wall *et al.*, 1991). In contrast, Buhler *et al.* (1990) have described low level expression of a construct in which genomic sequences of human interleukin 2 were fused to the 5′ region of rabbit β-casein gene. The poor performance of this construct may simply reflect the absence of essential sequences in the 5′ casein segment, since intact β-casein transgenes are poorly expressed in transgenic mice (Lee *et al.*, 1988). A bovine αs_2-casein gene fused to genomic sequences encoding the human urokinase gene was expressed at high levels in the mammary gland (Meade *et al.*, 1990). Likewise, AAT-B, which comprises native α_1AT introns and some 3′ flanking sequences, was also expressed relatively efficiently in the mammary gland, in contrast to the cDNA containing hybrid constructs AAT-C and AAT-D. In these experiments, an effect of the 3′ flanking sequences on expression cannot be ruled out. The unmodified BLG gene was also expressed substantially more efficiently than its intronless counterpart. Taken together, these results suggest a requirement for some or all the natural introns for efficient expression.

The failure of cDNA-based hybrid genes to be expressed at high frequency in the mammary gland does not appear to be due to a simple requirement for splicing, because AAT-C (which contained the first intron of BLG) was expressed as poorly as its intronless counterparts AAT-D and FIX-D. Furthermore BLG.1I, containing only the first BLG intron, was expressed no more efficiently than its intronless counterpart, BLG.0I. However, BLG minigenes were expressed more efficiently than their α_1AT or flX cDNA counterparts suggesting that sequences within these cDNAs can influence expression.

The poor performance of the cDNA constructs appears to be due to attempting to express these hybrid genes in transgenic mice, because SV40-driven constructs are capable of expression in fibroblasts. Thus, given functional regulatory sequences and the right environment these transgenes are capable of being expressed. These results are similar to those of Brinster *et al.* (1988). However, in the experiments of Brinster *et al.* (1988), stable transformants selected for expression of a cotransformed marker were analysed, raising the possibility that intron effects were masked. The transient transfection experiments described here, where no selection was made, therefore emphasise the lack of a requirement for introns *in vitro*. This conclusion contrasts with that of Buchman and Berg (1988), who showed intron-dependence *in vitro* and suggests that the effect is mediated after the initiation of transcription. Whatever the mechanism(s) of intron-dependence reported by Buchmann and Berg, (1988) it is clearly different from that observed in transgenic mice.

The performance of BLG hybrid genes and BLG minigenes in transgenic mice is reminiscent of the behaviour of transgenes carrying 'weak promoters' (e.g. Lee *et al.*, 1988; Allen *et al.*, 1990). Such transgenes are particularly sensitive to the site of integration and are often expressed at a reduced level or not at all. In contrast to the entire BLG transgene (C.B.A. Whitelaw, S. Harris,

M. McClenaghan, J.P. Simons and A.J. Clark, unpublished data), there was no obvious correlation between transgene copy-number and expression level for the other constructs described in this work. Indeed, for AAT-B the highest expressor had a copy-number of one (Archibald *et al.*, 1990). All the constructs reported in this paper comprise identical 5′ flanking BLG sequences and it would seem reasonable to postulate that it is the interaction between these sequences and the downstream sequences that determines how effectively position effects will be overcome. This interaction appears to be manifest only when genes are exposed to developmental influences, as is the case in transgenic animals. One possibility is that these putative *cis*-acting sequences enable a (*trans*)gene to adopt the appropriate chromatin configuration thus allowing access of the appropriate transcription factors to the 5′ promoter/enhancer region. The precise nature of these sequences is not known, they may be discretely organized or distributed throughout the gene. One possibility is that they correspond to nuclear matrix attachment sites which have recently been described for other genes (Blasquez *et al.*, 1989; Bonifer *et al.*, 1990).

From a practical point of view, these results have several implications and when taken together with those of Brinster *et al.* (1988) and Palmiter *et al.* (1991) demonstrate that intronic sequences can enhance the efficiency of gene expression in transgenic mice. Thus, limitations are imposed on the construction of transgenes, particularly those in which the structural gene encompasses large regions of chromosome, e.g. Factor VIII, 185 kb (Gitscheir *et al.*, 1984). However, genomic minigenes such as AAT-B do function quite efficiently in transgenic mice (Hammer *et al.*, 1987). Therefore, it should be possible to efficiently express constructs comprising a mixture of cDNA and genomic sequences. Our results also suggest that the requirement for introns is not due to a requirement for splicing *per se* but may be due to the presence of *cis*-acting elements. The identity, organization and precise function of such elements is not known but it is conceivable that once identified they could be incorporated into transgene constructs to ensure efficient expression.

Acknowledgements

We are grateful to Jen Anderson, John Bowman, Sandra Couperwhite, Morag Robertson, Wendy Shephard, Roberta Wallace and John Webster for technical assistance. We would like to thank John Bishop, George Brownlee, Ricardo Cortese, Gavin Kelsy and Jean-Claude Mercier for providing plasmids pUC.SVE, p5′G3′cVI, p8α_{1ppg}, pATp7 and pBlg931, respectively. We also thank Anthea Springbett for statistical analysis, Simak Ali for the S1 mapping analysis and Gordon Wright (Pharmaceutical Proteins Ltd) for the α_1AT enzyme-linked immunosorbant assays. This work was funded, in part, by Pharmaceutical Proteins Limited.

References

Ali, S. and Clark, A.J. (1988) Characterisation of the gene encoding ovine β-lactoglobulin. *J. Mol. Biol.* **199**, 415–26.

Allen, D.N., Norris, M.L. and Surani, M.A. (1990) Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell* **61**, 853–61.

Anson, D., Austin, D.E.G. and Brownlee, G.G. (1985) Expression of active human clotting factor IX from recombinant DNA clones in mammalian cells. *Nature (London)* **315**, 683–5.

Archibald, A.L., McClenaghan, M., Hornsey, V., Simons, J.P. and Clark, A.J. (1990) High-level expression of biologically active human α_1-antitrypsin in the milk of transgenic mice *Proc. Natl Acad. Sci. USA* **87**, 5178–82.

Ball, R.K., Friis, R.R., Schoenenberger, C.A., Doppler, W. and Groner, B. (1988) Prolactin regulation of β-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line. *EMBO J.* **7**, 2089–95.

Bayna, E.M. and Rosen, J.M. (1990) Tissue-specific, high level expression of the rat whey acidic protein gene in transgenic mice. *Nucl. Acids Res.* **18**, 2877–985.

Benoist, C. and Chambon, P. (1981) *In vivo* sequence requirements of the SV40 early promoter region. *Nature (London)* **290**, 304–10.

Blasquez, V.C., Sperry, A.O., Cockerill, P.N. and Gerrard, W.T. (1989) Protein : DNA interactions at chromosomal loop attachment sites. *Genome* **31**, 503–9.

Bonifer, C., Vidal, M., Grosveld, F. and Sippel, A.E. (1990) Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. *EMBO J.* **9**, 2843–8.

Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E. and Palmiter, R.D. (1988) Introns increase transcriptional efficiency in transgenic mice. *Proc. Natl Acad. Sci. USA* **85**, 836–40.

Buchman, A.R. and Berg, P. (1988) Comparison of intron-dependent and intron-independent gene expression. *Mol. Cell. Biol.* **8**, 4395–405.

Buhler, T.A., Bruyere, T., Went, D.F., Stranzinger, G. and Burki, K. (1990) Rabbit β-casein promoter directs secretion of human interleukin-2 into milk of transgenic rabbits. *Bio/Technology* **8**, 140–3.

Burdon, T., Sankaran, L., Wall, R., Spencer, M. and Henninghausen, L. (1991). Expression of a whey acidic protein transgene during mammary development. *J. Biol. Chem.* **266**, 6909–14.

Chomezynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by guanidium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162**, 156–9.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–5.

Ciliberto, G., Dente, L. and Cortese, R. (1985) Cell-specific expression of a transfected human α1-antitrypsin gene. *Cell* **41**, 531–40.

Clark, A.J., Simons, J.P., Wilmut, I. and Lathe, R. (1987)

Pharmaceuticals from transgenic livestock. *Trends Biotech.* **5**, 20–4.

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. and Wilmut, I. (1989) Expression of human anti-hemophilic factor IX in the milk of transgenic sheep. *Bio/Technology* **7**, 487–92.

Clark, A.J., Archibald, A.L., McClenaghan, M., Simons, J.P., Whitelaw, C.B.A. and Wilmut, I. (1991) The germline manipulation of livestock: progress during the past five years. *Proc. New Zealand Soc. Anim. Prod.* **50**, 167–80.

Danielson, K.G., Oborn, C.J., Durban, E.M., Butel, J.S. and Medina, D. (1984) Epithelial mouse mammary cell line exhibiting normal morphogenesis *in vivo* and functional differentiation *in vitro*. *Proc. Natl Acad. Sci. USA* **81**, 3756–60.

Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) *Basic Methods in Molecular Biology*, pp. 1–388. Amsterdam, New York and London: Elsevier Science Publishing Co. Inc.

Doppler, W., Groner, B. and Ball, R.K. (1989) Prolactin and glucocorticoid hormones synergistically induce expression of transfected rat β-casein gene promoter constructs in a mammary epithelial cell line. *Proc. Natl Acad. Sci. USA* **86**, 104–8.

Eisenstein, R.S. and Rosen, J. (1988) Both cell substratum regulation and hormonal regulation of milk protein gene expression are exerted primarily at the post-transcriptional level. *Mol. Cell. Biol.* **8**, 3183–3190.

Fienberg, A.P. and Vogelstein, B. (1983) Technique for radiolabelling DNA restriction fragments to high specific activity. *Anal. Biochem.* **132**, 6–13.

Fienberg, A.P. and Vogelstein, B. (1984) Addendum: A technique for radiolabelling DNA restriction fragments to high specific activity. *Anal. Biochem.* **137**, 266.

Gaye, P., Huc-Delahaie, D., Mercier, J.-C., Soulier, S., Vilotte, J.-L. and Furet, J.-P. (1986) Ovine β-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland. *Biochimie* **68**, 1097–107.

Gitschier, J., Wood, W.I., Goralka, T.M., Wion, K.L., Chen, E.Y., Eaton, D.E., Vehar, G.A., Capon, D.J. and Lawn, R.M. (1984) Characterisation of the human factor VIII gene. *Nature (London)* **312**, 326–30.

Gordon, K., Lee, E., Vitale, J.A., Smith, A.E., Westphal, H. and Hennighausen, L. (1987) Production of human tissue plasminogen activator in transgenic mouse milk. *Bio/Technology* **5**, 1183–7.

Graham, F.L. and van der Eb, A.J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456–67.

Hammer, R.E., Krumlauf, R., Camper, S.A., Brinster, R.L. and Tilghman, S.M. (1987) Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. *Science* **235**, 53–8.

Harris, S., Ali, S., Anderson, S., Archibald, A.L. and Clark, A.J. (1988) Complete nucleotide sequence of the ovine β-lactoglobulin gene. *Nucl. Acids Res.* **16**, 10379.

Harris, S., McClenaghan, M., Simons, J.P., Ali, S. and Clark, A.J. (1990) Gene expression in the mammary gland. *J. Reprod. Fert.* **88**, 707–15.

Hogan, B., Constantini, F. and Lacy, E. (1986) *Manipulating the Mouse Embryo: A Laboratory Manual.* New York: Cold Spring Harbour.

Kelsey, G., Povey, S., Bygrave, A.E. and Lovell-Badge, R.H. (1987) Species- and tissue-specific expression of human α1-antitrypsin in transgenic mice. *Genes Dev.* **1**, 161–71.

Kozak, M. (1987) Effects of intercistronic length on the efficiency of reinitiation by eukaryotic ribosomes. *Mol. Cell. Biol.* **7**, 3438–45.

Lathe, R., Clark, A.J., Archibald, A.L., Bishop, J.O., Simons, J.P. and Wilmut, I. (1986) Novel products from livestock. In: Smith, King and McKay (eds) *Exploiting New Technologies in Animal Breeding*, pp. 99–102. Oxford: Oxford University Press.

Lee, K.-F., DeMayo, F.J., Atiee, S.H. and Rosen, J.M. (1988) Tissue-specific expression of the rat β-casein gene in transgenic mice. *Nucl. Acids Res.* **16**, 1027–41.

Maniatis, T., Fitsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbour Laboratory, New York: Cold Spring Harbour.

Meade, H., Gates, L., Lacy, E. and Lonberg, N. (1990) Bovine alpha s_1-casein gene sequences direct high level expression of active human urokinase in mouse milk. *Bio/Technology* **8**, 443–6.

Ott, M.-O., Sperling, L., Herbomel, P., Yaniv, M. and Weiss, M.C. (1984) Tissue-specific expression is conferred by a sequence from the 5′ end of the rat albumin gene. *EMBO J.* **3**, 2505–10.

Palmiter, R.D., Sandgren, E.P., Avarbock, M.R., Allen, J.M. and Brinster, R.L. (1991) Heterologous introns can enhance expression in transgenes in mice. *Proc. Natl Acad. Sci. USA* **88**, 478–82.

Peabody, D.S. and Berg, P. (1986) Termination-reinitiation occurs in the translation of mammalian cell mRNAs. *Mol. Cell. Biol.* **6**, 2695–703.

Pinkert, C.A., Ornitz, D.M., Brinster, R.L. and Palmiter, R.D. (1987) An albumin enhancer located 10 kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. *Genes Dev.* **1**, 268–76.

Pursel, V.G., Pinkert, C.A., Miller, K.F., Bolt, D.J., Campbell, R.G., Palmiter, R.D., Brinster, R.L. and Hammer, R.E. (1989) Genetic engineering of livestock. *Science* **244**, 1281–8.

Simons, J.P. and Land, R.B. (1987) Transgenic livestock. *J. Reprod. Fert.* **34**, 237–50.

Simons, J.P., McClenaghan, M. and Clark, A.J. (1987) Alteration of the quality of milk by expression of sheep β-lactoglobulin in transgenic mice. *Nature (London)* **328**, 530–2.

Vilotte, J.-L., Soulier, S., Stinnarke, M.-G., Massoud, M. and Mercier, J.-C. (1989) Efficient tissue-specific expression of bovine α-lactalbumin in transgenic mice. *Eur. J. Biochem.* **186**, 43–8.

Wall, R.J., Pursel, V.G., Shamay, A., McKnight, R., Pittius, C. and Hennighausen, L. (1991) High-level synthesis of a heterologous milk protein in the mammary glands of transgenic swine. *Proc. Natl Acad. Sci. USA* **88**, 1696–700.

Whitelaw, B., Wilkie, N.M., Jones, K.A., Kadonaga, J.T., Tjian, R. and Lang, J.C. (1988) Transcriptionally active domains in the 5′ flanking sequence of human c-*myc*. In: *Growth*

Factors, Tumor Promoters and Cancer Genes, pp. 337–51. New York: Alan R. Liss.

Whitelaw, C.B.A. and Clark, A.J. (1989) Animal bioreactors. *AgBiotech News Inform.* **1**, 701–5.

Whitelaw, C.B.A., Archibald, A.L., Harris S., McClenaghan, M., Simons, J.P., Springbett, A., Wallace, R. and Clark, A.J. (1990) Frequency of germline mosaicism in founder transgenic mice. *Mouse Genome* **88**, 114.

Yu, S.-H., Deen, K.C., Lee, E., Hennighausen, L., Sweet, R.W., Rosenberg, M. and Westphal, H. (1989) Functional human CD4 protein produced in milk of transgenic mice. *Mol. Biol. Med.* **6**, 255–61.