A STUDY OF OVINE β-LACTOGLOBULIN TRANSGENE EXPRESSION IN THE MOUSE MAMMARY GLAND

Kenneth W. Dobie

Ph.D.
University of Edinburgh
1996

"An organism's astonishing gift of concentrating a 'stream of order' on itself and thus escaping the decay into atomic chaos - seems to be connected with the presence of the 'aperiodic solids', the chromosome molecules, which doubtless represent the highest degree of well-ordered atomic association we know of - much higher than the ordinary periodic crystal - in virtue of the individual role every atom and every radical is playing here".

Schrödinger, E. (1944). What is life? Cambridge University Press.

DECLARATION

I declare that I have written this thesis based on my own work, and the contribution of others has been clearly indicated.

Kenneth W. Dobie

ABSTRACT

Lines of mice carrying an ovine β-lactoglobulin (BLG) transgene secrete large quantities of BLG protein into their milk. To explore the stability of transgene expression, a systematic study of expression levels in three BLG transgenic mouse lines was performed. Unexpectedly, two of these lines exhibited variable levels of transgene expression. Copy number within lines appeared to be stable and there was no evidence of transgene rearrangement. Studies on the most variable line showed that BLG production levels were stable within individual mice in two successive lactations. Backcross experiments demonstrated that the genetic background did not contribute significantly to the variation of expression levels. Tissue *in situ* hybridisation experiments revealed mosaic patterns of transgene expression within individual mammary glands from the two variable lines; in low expressing animals, discrete patches of cells expressing the transgene were observed. The concentration of transgene protein in milk reflected the proportion of mammary epithelial cells expressing BLG mRNA. Furthermore, *in situ* hybridisation to metaphase chromosomes indicated that the transgene arrays in both these lines are situated close to the centromere. These experiments suggest that variable mosaicism of transgene expression among individuals within a transgenic line is a consequence of the chromosomal location and/or the nature of the primary transgene integration event.

ACKNOWLEDGEMENTS

I am very grateful to Dr Margaret McClenaghan (Roslin Institute) for many helpful discussions, continual supervision and encouragement. I also thank Professor Richard Lathe (Centre for Genome Research) and Dr John Clark (Roslin Institute) for their supervision and many helpful discussions during my studies.

I thank Frances Thompson and Roberta Wallace for meticulous assistance with mouse husbandry and Anthea Springbett for statistical advice. During my studies at the Roslin Institute I have greatly benefited from discussions with Bert Binas, Ed Hitchin, Satish Kumar, Paul Simons, John Webster and Bruce Whitelaw. The members of the Division of Molecular Biology (Roslin Institute) provided me with the best possible environment (in many ways) for my studies. Thanks also to Andrew Dawson (Moredun Research Institute) for assistance with the micro-kjeldahl analysis, Bob Scougall (Roslin Institute) for advice on SDS-PAGE and protein quantitation and Norrie Russell, Roddy Field and Elliot Armstrong for photography and graphics.

I am grateful for a number of collaborations at the MRC Human Genetics Unit. I thank Muriel Lee and Judith Fantes for teaching me how to perform chromosome *in situ* hybridisations and the rest of the blood lab for providing such a pleasant work environment, Elizabeth Graham and Duncan Davidson for teaching me how to perform tissue *in situ* hybridisations, Linda Nicol and Peter Jeppesen for providing the CREST antibody and assistance with native chromosome analysis, Arthur Mitchell for providing the major and minor mouse satellite DNA probes, Eric Thompson for all his assistance and Dr Konrad Huppi (National Institutes of Health, U.S.A.) for providing the *c-myc* and *Pvt-1* clones.

I was financially supported by the Centre for Genome Research, The University of Edinburgh.

To family and friends for encouragement and support.

CONTENTS

Declaration....iii

Abstract....iv

Acknowledgements....v

Contents....vii

List of Figures....xiv

List of Tables....xvii

Abbreviations and Conventions....xix

CHAPTER 1 : **INTRODUCTION**

1.1. DEVELOPMENT OF THE MOUSE MAMMARY GLAND 2

1.1.1. Foetal and postnatal development 3

1.1.2. Development during gestation 4

1.1.3. Development during lactation 5

1.1.4. Involution 6

1.2. FACTORS INTRINSIC FOR MAMMARY GLAND DEVELOPMENT 6

1.2.1. Hormonal regulation 7

1.2.2. Extracellular matrix components 8

1.3. MILK COMPOSITION 10

1.3.1. Milk fat 11

1.3.2. Milk carbohydrate 11

1.3.3. Milk protein 11

1.4. FACTORS INTRINSIC FOR MILK PROTEIN GENE EXPRESSION 13

1.5. TRANSGENICS AS A METHOD OF PRODUCING BIOMEDICAL PROTEINS 16

1.5.1. Choice of biomedical proteins 16

1.5.2. Choice of expression system 17

1.6. OVINE β-LACTOGLOBULIN (BLG) 20

1.7. GENERATION OF TRANSGENIC MOUSE LINES 7, 14 & 45 22

1.8. UNSTABLE EXPRESSION OF TRANSGENES 23

1.9. AIMS OF THE PROJECT 26

***CHAPTER 2*: MATERIALS AND METHODS**

2.1 TRANSGENIC MOUSE LINES 27

2.2 MILKING OF MICE 28

2.3 PURIFICATION OF OVINE β-LACTOGLOBULIN (BLG) 29

2.4 ESTIMATION OF BLG CONCENTRATION 29

2.5 MILK PROTEIN ANALYSIS

2.5.1. Preparation of SDS-polyacrylamide gels 30

2.5.2. SDS-PAGE 31

2.5.3. Western blotting 32

2.6 DNA MANIPULATION AND ANALYSIS

2.6.1. Digesting plasmid DNA with restriction enzymes 33

2.6.2. Phenol/chloroform extraction of DNA 34

2.6.3. DNA precipitation with ethanol 35

2.6.4. Recovery of DNA fragments from low melting point (LMP) agarose gels 36

2.6.5. DNA ligation 37

2.6.6. DH5α cells 37

2.6.7. Transformation of DH5α cells with plasmid DNA 38

2.6.8. Preparation of plasmid DNA

2.6.8.1. Small-scale preparation 38

2.6.8.2. Large-scale preparation

2.6.8.2.1. Maxiprep method 39

2.6.8.2.2. CsCl-EtBr gradient method 40

2.6.9. Preparation and analysis of mouse mammary genomic DNA 41

2.6.10. Southern blotting

2.6.10.1. Transfer onto Hybond-N membrane 43

2.6.10.2. Transfer onto Zeta-Probe GT membrane 44

2.6.11. Preparation of radio-labelled DNA probes and hybridisation to Southern blot membranes 44

2.6.11.1. Multiprime DNA labelling method 45

2.6.11.2. High Prime DNA labelling method 46

2.6.11.3. Quantifying ^{32}P-labelled DNA probes 46

2.6.11.4. Hybridisation and washes 47

2.6.12. Polymerase chain reaction (PCR) analysis of tail biopsy DNA 48

2.7. RNA ANALYSIS

2.7.1. Preparation of total RNA from mouse mammary tissue 50

2.7.2. Electrophoresis of RNA, Northern blotting and hybridisation 51

2.8. PREPARATION OF AGAR PLATES

2.8.1. Plates containing antibiotic 51

2.8.2. Plates for blue/white screening 52

2.9. PREPARATION OF LAMBDA PHAGE 52

2.9.1. XL-1 Blue MRA cells 53

2.9.2. Estimation of titre of phage stocks 53

2.9.3. Preparation of a mini-stock of high titre phage 54

2.9.4. Large-scale preparation of phage DNA 54

2.10. TISSUE *IN SITU* HYBRIDISATION 56

2.10.1. Construction of BLG expression vectors 56

2.10.2. Construction of β-casein expression vectors 58

2.10.3. Coating slides with TESPA 58

2.10.4. Preparation of tissue 59

2.10.5. Tissue sectioning 60

2.10.6. Pre-hybridisation 60

2.10.7. Generation of ^{35}S rUTP labelled riboprobes 61

2.10.8. Quantifying ^{35}S rUTP labelled riboprobes 62

2.10.9. Hybridisation 63

2.10.10. Post-hybridisation 64

2.10.11. Coating slides with emulsion 64

2.10.12. Development of emulsion-coated slides 65

2.10.12. Photography of tissue *in situ* slides 65

2.11. CHROMOSOME FLUORESCENCE *IN SITU* HYBRIDISATION (FISH) 65

2.11.1. Preparation of mouse lymphocyte cells 66

2.11.2. Preparation of slides 67

2.11.3. Giemsa banding (G bands) 67

2.11.4. Karyotyping 68

2.11.5. Generation of biotin and digoxigenin labelled DNA probes 68

2.11.6. Quantitation of label incorporation 70

2.11.7. Pre-hybridisation 71

2.11.8. Preparation of probe for hybridisation 71

2.11.9. Hybridisation 72

2.11.10. Post-hybridisation 72

2.11.11. Preparation and hybridisation of antibodies 72

2.11.12. Analysis of fluorescence images

2.11.12.1. The confocal microscope 73

2.11.12.2. The Zeiss microscope with CCD 74

2.12. PREPARATION AND ANALYSIS OF "NATIVE" CHROMOSOMES 74

CHAPTER 3: **TRANSGENE EXPRESSION LEVELS AND RELATIVE COPY NUMBER IN THREE TRANSGENIC LINES**

3.1. INTRODUCTION 76

3.2. RESULTS

3.2.1. Preparation of a BLG protein standard 78

3.2.2. A quantitative assay for BLG expression levels in transgenic mouse milk 81

3.2.3. BLG protein expression levels in lines 7, 14 and 45 84

3.2.4. Analysis of BLG mRNA levels in lines 7, 14 and 45 87

3.2.5. Analysis of the BLG transgene arrays in lines 7, 14 and 45 91

3.2.6. BLG expression levels across two lactations in line 7 mice 99

3.3. CONCLUSIONS 104

CHAPTER 4: **EFFECT OF GENETIC BACKGROUND ON TRANSGENE EXPRESSION**

4.1. INTRODUCTION 106

4.2. RESULTS

4.2.1. A backcross from line 7 mice into founder strains 110

4.2.2. BLG expression levels in sub-lines through the backcross 118

4.3. CONCLUSIONS 119

CHAPTER 5: **BLG AND β–CASEIN EXPRESSION IN MID-LACTATION MOUSE MAMMARY GLAND**

5.1. INTRODUCTION 120

5.2. RESULTS

5.2.1. Tissue *in situ* hybridisation analysis of mouse mammary tissue sections. 122

5.3. CONCLUSIONS 130

CHAPTER 6: **CHROMOSOMAL LOCALISATION AND MOSAICISM OF TRANSGENE EXPRESSION**

6.1. INTRODUCTION 131

6.2. RESULTS

6.2.1. G-band and FISH analysis of lines 7, 14 & 45 135

6.2.2. Position of the transgene arrays relative to satellite DNA 137

6.2.3. Immunofluorescent / FISH analysis of line 7 chromosomes using CREST antibodies and DNA probes for BLG, *c-myc* and *Pvt-1* 150

6.3. CONCLUSIONS 161

CHAPTER 7. **DISCUSSION** 163

7.1. Variable levels of BLG protein expression within lines 7 and 45 164

7.2. Transcription versus translational mechanism for unstable expression 167

7.3. Variable transcription of the BLG transgene 168

7.4. Mosaic expression of BLG in the mammary gland 172

7.5. PEV : mechanisms of mosaic silencing 173

7.5.1. Somatic instability 173

7.5.2. Chromatin assembly model 174

7.5.3. DNA looping model 176

7.6. A model for transgene silencing in lines 7 and 45 178

7.7. PEV : a ubiquitous phenomenon? 181

7.8. Implications for mosaic transgene silencing 182

REFERENCES 184

APPENDIX I: *Statistical Terms* 225

APPENDIX II: *Publication* 226

LIST OF FIGURES

Figure 1.	Diagrammatic representation of a cluster of mammary alveoli.	5
Figure 2.	Schematic representation of mammary growth in mice.	8
Figure 3.	ECM-signalling and mammary epithelial cell differentiation.	9
Figure 4.	Mammary epithelial cell differentiation, ECM-signalling and milk protein gene expression.	14
Figure 5.	Model for the regulation of differentiation of mammary alveolar cells during puberty, pregnancy and lactation.	15
Figure 6.	A map of BLG clone SS-1.	21
Figure 7.	pGEM-4Z vector circle map.	57
Figure 8.	SDS-PAGE and wastern blot analysis of purified ovine BLG	79
Figure 9.	Quantitation of BLG standards in five gels	81
Figure 10.	SDS-PAGE analysis of mouse milk	83
Figure 11.	BLG expression level profiles in milk collected from line 7, 14 and 45 hemizygous mice.	86
Figure 12.	BLG protein and RNA expression in lines 7, 14 and 45 mice, shown as SDS-PAGE and northern blots	89
Figure 13.	BLG protein expression levels plotted against mRNA levels	90
Figure 14.	A map of the BLG transgene	92
Figure 15.	Southern blot analysis of the transgene array in lines 7, 14 and 45	94
Figure 16.	Southern blot analysis of the transgene array in line 7	95
Figure 17.	BLG expression levels in milk of line 7 animals over two consecutive lactations.	103
Figure 18.	Possible mechanisms for transgene position effects	107
Figure 19.	Breeding line 7 mice to CBA and C57BL/6 founder stains	111
Figure 20.	BLG expression levels in milk of line 7 animals; effect of genetic	

background. 115

Figure 21. Tissue morphology of the mammary gland. 123

Figure 22. *In situ* hybridisation analysis of mRNA expression patterns in non-transgenic mouse mammary tissue. 124

Figure 23. *In situ* hybridisation analysis of mRNA expression patterns in line 7 (low). 126

Figure 24. *In situ* hybridisation analysis of mRNA expression patterns in line 7 (high). 127

Figure 25. *In situ* hybridisation analysis of mRNA expression patterns in line 14. 128

Figure 26. *In situ* hybridisation analysis of mRNA expression patterns in line 45. 129

Figure 27. FISH analysis of non-transgenic mouse metaphase chromosomes. 136

Figure 28. FISH/G-band analysis of metaphase chromosomes from line 7. 139

Figure 29. FISH/G-band analysis of metaphase chromosomes from line 14. 141

Figure 30. FISH/G-band analysis of metaphase chromosomes from line 45. 143

Figure 31. Localisation of BLG and major satellite (196) DNA sequences in line 7 by chromosomal FISH to metaphase chromosomes. 144

Figure 32. Localisation of the BLG locus in line 7 relative to the centromere 147

Figure 33. Mouse chromosome 15 and the BLG transgene locus in line 7 149

Figure 34. Localisation of BLG and major satellite (196) DNA sequences in line 45 by chromosomal FISH to line 45 metaphase chromosomes. 151

Figure 35. Localisation of BLG and minor satellite DNA sequences in line 7 by chromosomal FISH to metaphase chromosomes. 153

Figure 36. Native FISH using CREST antibodies and DNA probes for BLG, *c-myc* and *Pvt-1* on line 7 chromosomes 157

Figure 37. Native FISH using CREST antibodies and DNA probes for BLG,

c-myc and *Pvt-1* on line 7 chromosomes 159

Figure 38. Mouse chromosome 15 and the BLG, *c-myc* and *Pvt-1* loci. 160

Figure 39. Models for BLG silencing (heterochromatinisation) 180

LIST OF TABLES

Table 1.	The major components of milk	10
Table 2.	Milk protein composition of sheep and mouse milk	12
Table 3.	Comparison of different farm livestock	18
Table 4.	Mammary gland-specific transgenic animals	19
Table 5.	Summary of examples of unstable transgene expression	24
Table 6.	Restriction enzymes and conditions for digestion	34
Table 7.	Range of separation for agarose gels	49
Table 8.	Calculation of the concentration of the BLG protein standard	80
Table 9.	Reproducibility study of BLG values	84
Table 10.	BLG protein expression levels in lines 7, 14 and 45	85
Table 11.	BLG protein expression levels within lines 7, 14 and 45	87
Table 12.	BLG protein and mRNA expression levels in lines 7, 14 and 45	88
Table 13.	The coefficient of variations for BLG mRNA expression within lines 7, 14 and 45	91
Table 14.	BLG expression levels and copy numbers within lines 7, 14 and 45	97
Table 15.	Sorces of variability in lactational performance	99
Table 16.	BLG expression levels in line 7 mice across two lactations	101
Table 17.	Coefficient of variation values for BLG protein expression in milk of line 7 animals	102
Table 18.	Imprinted endogenous genes	110
Table 19.	BLG expression levels in line 7 animals in a CBA backcross	113
Table 20.	BLG expression levels in line 7 animals in a C57BL/6 backcross	114
Table 21.	BLG expression levels within eight backcross populations	117
Table 22.	Comparison of the variance and mean values between the mixed C57BL/6 x CBA genetic background and successive populations	118

Table 23. Measurement of the distance between the BLG transgene locus and major satellite DNA 146

ABBREVIATIONS AND CONVENTIONS

The following abbreviations are used throughout this work:

AAT	human alpha-1-antitrypsin
BLG	β-lactoglobulin
bp	base pair(s)
cDNA	complementary DNA
CsCl	caesium chloride
DAPI	4, 6-diamidino-2-phenylindole
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	2'-deoxyguanosine 5'-triphosphate
DNase	deoxyribonuclease
DTT	dithiothreitol
dUTP	2'-deoxyuridine 5'-triphosphate
ECM	extracellular matrix
EDTA	ethylenediamine-*tetra*-acetic acid
EtBr	ethidium bromide
FCS	foetal calf serum
FISH	fluorescence *in situ* hybridisation
FITC	fluorescein isothiocyanate
FIX	human factor IX
g	gram(s)
Gb	gigabase(s)

hr	hour(s)
IPTG	isopropyl-β-D-thiogalactopyranosid
IU	international unit(s)
kb	kilobase (10^3 nt)
kD	kilodaltons
l	litre(s)
LPS	lipopolysaccharide
LMP	low melting point
M	Molar (moles/litre)
Mb	megabase(s)
MGF	mammary gland factor
µg	microgram(s)
µl	microlitre(s)
mg	milligram(s)
min	minute(s)
ml	millilitre(s)
mM	millimolar
MOPS	3-N-(morpholino) propanesulfonic acid
mRNA	messenger RNA
nt	nucleotide(s)
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
pcp	post-coitum plug
PCR	polymerase chain reaction
PEV	position-effect variegation
PI	propidium iodide

pl	picolitre(s)
RNase	ribonuclease
rpm	revolutions per minute
r.t.	room temperature
SDS	sodium dodecyl sulphate
sec	second(s)
SSC	standard saline citrate (0.15 M NaCl, 15 mM tri-sodium citrate, pH 7.0)
STAT	signal transducer and activator of transcription
TCA	trichloroacetic acid
TE	10 mM Tris-HCl pH 7.5, 1 mM EDTA
TEA	triethanolamine
TEB	terminal end buds
TR	Texas Red
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol ($C_4H_{11}NO_3$)
TTP	thymidine 5'-triphosphate
UV	ultraviolet
V	volt(s)
WAP	whey acidic protein
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER 1

INTRODUCTION

Advances in recombinant DNA technology and embryo manipulation has enabled the introduction of genes into the germline and expression of the foreign DNA (Gordon *et al.*, 1980; Wagner *et al.*, 1981; Brinster *et al.*, 1981; Gordon & Ruddle, 1981). The introduced DNA sequences were termed "transgenes" and the mice generated from viable embryos that contained the transgenes within their genome termed "transgenic" mice. The advent of transgenic technology has spawned a vast and ever-growing list of investigators utilising transgenic animals. The possibility of altering the genome of animals by introducing new gene variants has provided the opportunity for addressing fundamental questions of development, immunology, neurobiology, endocrinology, cell biology, oncogenesis, chromosome architecture, and the genetic engineering of farm animals. However, as the list of applications for transgenic technology has grown, it has become apparent that there are many examples of unstable transgene expression. Such expression has implications for investigators using this technology and it is clear that a better understanding of the regulation of transgene expression is required.

This project dealt primarily with factors involved in the expression of an ovine β-lactoglobulin (BLG) transgene in mouse mammary gland. The following section reviews mammary gland development and function, applications of mammary-specific transgene expression and the stability of transgene expression in general.

1.1. DEVELOPMENT OF THE MOUSE MAMMARY GLAND

The mammary gland undergoes morphological changes in size, structure and activity during the lifetime of the animal. These changes start during foetal life and continue even after the gland has reached maturity, as it goes through cycles of cellular differentiation / apoptosis during successive lactations. This tissue has been utilised to address numerous fundamental questions in different disciplines because of the complex developmental cycle (see 1.1.1. to 1.1.4.), the complex hormonal regulation of growth and function (1.2.1.), the interaction of growth factors and extracellular matrix (1.2.2.), the prevalence of human mammary tumours (182,000 new cases and 46,000 deaths each year in the USA alone [Boring *et al.*, 1992]) and the function of the mammary gland (i.e., milk production [1.3.]). Systems utilised to study the mammary gland range from complex three-dimensional *in vitro* tissue culture systems to *in vivo* transgenic experiments. There are many differences (e.g., hormonal regulation) in mammary development between species (Cowie *et al.*, 1980; Topper & Freeman, 1980; Knight & Peaker, 1982). However most work examining mammary development and its regulation has been performed in rodents (reviewed by Imagawa *et al.*, 1994; Medina, 1996). Whenever possible, this review will concentrate on mouse mammary gland development. Even so, there are some differences in development between mouse strains.

The mammary gland comprises the parenchyma (or epithelial portion) and the adipose stroma. The parenchyma forms a system of branching ducts from which secretory structures develop and the adipose stroma provides a substrate within which the parenchyma develops and functions. The development of the mammary gland can be divided into five main stages: foetal, postnatal, gestation, lactation and involution.

1.1.1. Foetal and postnatal development

The earliest sign of mammary-specific progenitor cells or "mammary buds" (spheres of ectodermal cells sunken into the dermis) is at day 10-11 (vaginal plug = 0 day) of foetal development (Sakakura, 1987; review). Bilateral ectodermal thickenings form the mammary lines within the underlying stroma and establish the position of the glands relative to one another between days 11 to 16 (i.e., inguinal and thoracic [pectoral] groups; Cook, 1965). From days 16 to 21 rapid proliferation of mammary epithelial cells occurs in females. Canalisation (cell death) within the core of the mammary lines results in a primary open duct in each gland. The ducts are lined by a two-cell thick epithelial layer. The primary ducts branch to form secondary ducts lined by a single layer of epithelial cells. In male mice, mesenchymal condensation occurs around days 13 to 15, due to secretion of testosterone from the foetal testis (Kratochwill, 1971; Durnberger & Kratochwill, 1980), resulting in a vestigial gland. X-ray-induced destruction of the testis in 13-day male embryos results in female-like development of the mammary gland and conversely, injection of testosterone into pregnant mothers or embryos results in the regression of mammary development in female offspring (reviewed by Imagawa *et al.*, 1994).

In newborn female mice (18-21 days after fertilisation), the mammary gland consists of only 2-4 orders of branching, connected to the outside via the nipple (one duct per nipple). These ducts serve as a framework from which the mammary gland proper develops along with all the ancillary structures (i.e., blood vessels, nerves, lymphatics, connective tissue and myoepithelial cells) under the influence of pituitary and gonadal hormones during puberty and adolescence (see 1.2.1.). This differentiation establishes the mammary fat pad which provides the support for lobulo-alveolar proliferation. Bresciani, (1971) produced autoradiographs showing

that DNA-synthesising cells were present only at the terminal end buds (TEB) at the tips of the ducts, illustrating that these are the major sites of growth in the virgin mouse. Elongation of the ducts occurs by mitotic activity in the TEB and results in a highly organised system of ducts within the fat pad in the juvenile mouse (Williams & Daniel, 1983; Daniel & Silberstein, 1987). Mitotic activity remains high until the ducts have reached the periphery of the mammary fat pads, at which time the TEB regress to blunt-ended structures containing a single layer of luminal epithelial cells with low mitotic activity. Sexual maturity is reached by 6-8 weeks by which time ductal development is complete. With each oestrous cycle, oestrogens cause the lateral buds to differentiate progressively, thus giving rise to small alveolar buds but without lobulo-alveolar formation (see 1.2.1.).

1.1.2. Development during gestation

Further development and differentiation of the gland requires fertilisation and is induced by the hormones of pregnancy (Topper & Freeman, 1980). It is estimated that 78% of mouse mammary growth takes place during gestation (Brookreson & Turner, 1959) with the number of cells doubling (measured as DNA content) every 6 days (Munford, 1963). The highest rates of epithelial cell proliferation occur between day 4 and day 12 of gestation (Traurig, 1967). Between day 8 and day 10 of gestation, the mammary alveoli start to form true lobulo-alveolar tissue with growth occurring from the ducts as well as at the end buds. A rapid increase in the number and size of alveoli occurs during the second half of gestation and can continue into lactation, resulting in the development of fully differentiated secretory lobules which are the functional secretory unit of the mammary gland (Pitelka *et al.*, 1973; Cowie *et al.*, 1980). The lobule comprises several alveoli around a single ductule which collects the secreted milk. Towards

the end of gestation the secretory epithelial cells enlarge, perhaps due to increased secretory activity (Foster, 1977). The pups are usually born at day 19 of gestation / day 1 of lactation.

1.1.3. Development during lactation

Lactogenesis begins by the 17-19th day of pregnancy. Between the last day of pregnancy and the 5th day of lactation the cell population doubles and the milk yield increases correspondingly (Knight & Peaker, 1982). Lactational performance (measured as milk yield and percentage composition with respect to total solids, lipid, and protein) increases linearly from day 8 to day 12 of lactation (Hanrahan & Eisen, 1969). The mature gland comprises specialised secretory epithelial cells that are surrounded by a layer of contractile myoepithelial cells (Fig. 1) and supporting connective and adipose tissue.

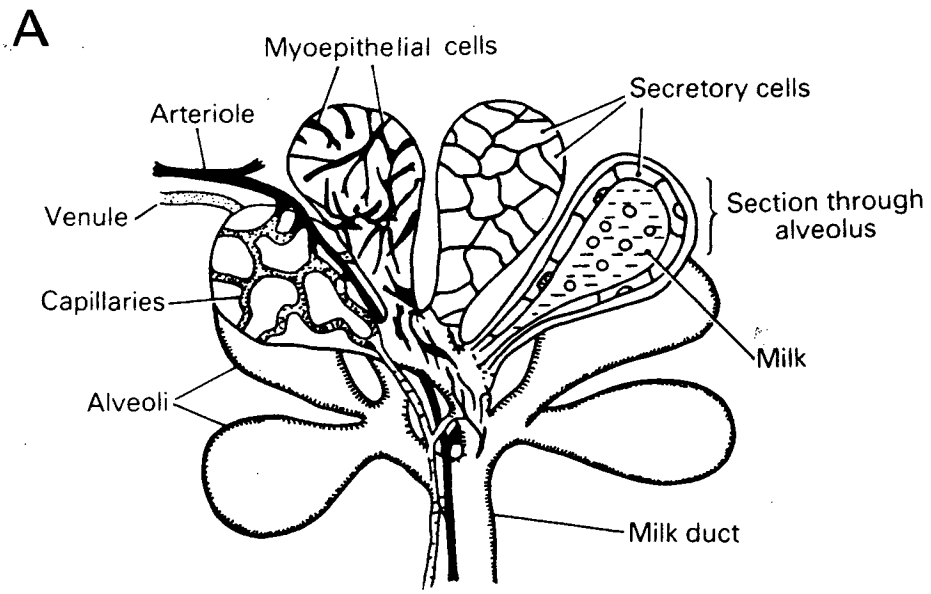


Figure 1. Diagrammatic representation of a cluster of mammary alveoli.

Taken from Mepham, (1987).

It is the myoepithelial cells that contract and force the milk out of the alveolar spaces into the ducts and eventually to the nipple.

1.1.4. Involution

The gland begins to involute following the peak milk yield at mid-lactation (day 11-12 of lactation) and the milk yield correspondingly declines until the pups are weaned at about day 21 of lactation (40 day pcp). Involution is essentially an autolytic process, with lysosomes playing a major role, while an influx of phagocytic cells aided by disruption of the tight junctions completes the degenerative process (Strange *et al.*, 1992; Boudreau *et al.*, 1995). Progressively, parenchymatous tissue is replaced by connective and adipose tissue. Little is known about whether cells "carry-over" from one lactation to the next. Involution does not return the gland to a virgin state as administration of [^{3}H] thymidine to rats suggests that there is appreciable persistence of mammary cells from one lactation to the next. This finding is consistent with the increased lactational performance in cows (e.g., the lactational performance in cows increases until the fourth lactation [Oldham & Friggens, 1989]).

1.2. FACTORS INTRINSIC FOR MAMMARY GLAND DEVELOPMENT

Development, differentiation and function of the mammary gland is influenced by the developmental lineage and the microenvironment. The microenvironment comprises hormonal influences and the extracellular matrix (ECM).

1.2.1. Hormonal regulation

The development of the gland is controlled to a large extent by the influence of several different hormones including progesterone, oestrogen, prolactin and placental lactogen (Cowie *et al.*, 1980; Topper & Freeman, 1980; Forsyth, 1986). The varied hormonal control is necessary because of the complexity of mammary gland development throughout the lifespan of the animal and the various constituents within milk (Wiens *et al.*, 1987).

Early embryonic development of the mammary tissue does not appear to be influenced by hormones, but is highly dependent on stromal factors (e.g., transforming growth factor beta [TGFβ]; Imagawa *et al.*, 1994). Later stages of foetal development (days 13 to 15) are influenced by testosterone, which suppresses mammary gland development in males (see 1.1.1.). During puberty the gland is influenced by the cyclic release of oestrogen which results in a cyclic extension of the ductal system but without lobular-alveolar development (Fig. 2). During pregnancy, oestrogen in association with pituitary hormones (e.g., prolactin), promotes ductal development, while progesterone is additionally required for lobulo-alveolar development (mammogenesis [1.1.2.]). Later in development, the foetal/placental unit may take over the role of the pituitary, placental lactogen replacing or augmenting the action of prolactin.

The mammary growth in early lactation is associated with a quite different hormonal *milieu*. Parturition is associated with an abrupt fall in progesterone and placental lactogen levels. Instead, the mammary growth may be maintained by blood prolactin, ovarian steroids (e.g. oestradiol), prostoglandins and/or local growth factors (e.g., epidermal growth factor, TGFα) in the ECM, all of which increase during this period (Imagawa *et al.*, 1994). The actions of some of these factors are mediated via intercellular signalling systems (Fig. 3); for example, prolactin acts

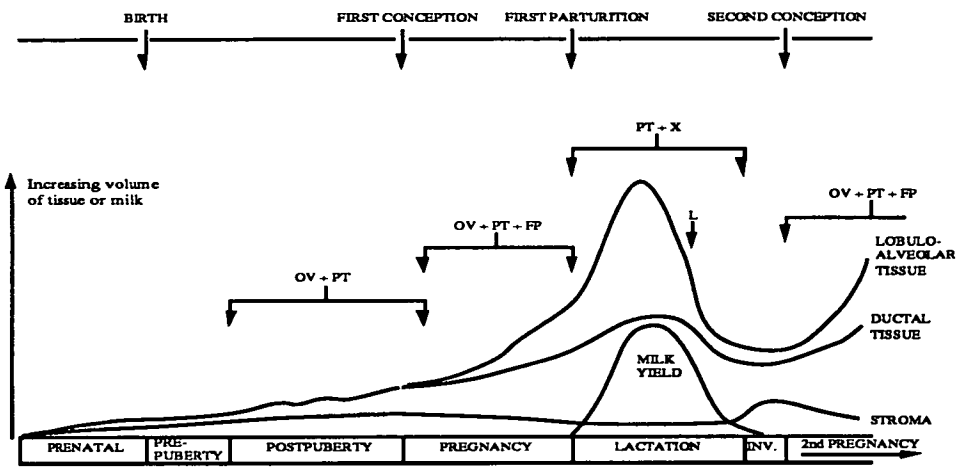


Figure 2. Schematic representation of mammary growth in mice.

OV, ovarian hormones (e.g., oestrogen); PT, pituitary hormones (e.g., prolatin); FP, foetal placental hormones (e.g., placental lactogen); X, other factors (e.g., TGFα); L, local factors (e.g., metalloproteases); INV, involution. Adapted from Mepham, (1987).

via the protein tyrosine kinase Jak2 (Campbell *et al.*, 1994; David *et al.*, 1994; Dusanter-fourt *et al.*, 1994; Gilmour & Reich, 1994; Rui *et al.*, 1994) and one of its substrates, the transcription factor MGF / STAT5 (Wakao *et al.*, 1994; Gouilleux *et al.*, 1994; DaSilva *et al.*, 1996).

1.2.2. Extracellular matrix components

Epithelial cell growth and development is dependent on the ECM which comprises the basement membrane and a variety of polysaccharides and proteins secreted by the cells (Bissell & Hall, 1987; Imagawa *et al.*, 1994). ECM molecules interact with each other and with their specific receptors on the cell surface (Fig. 3). Such interactions now are believed to play an important role in the regulation of functional differentiation (Streuli *et al.*, 1991). In the resting gland,

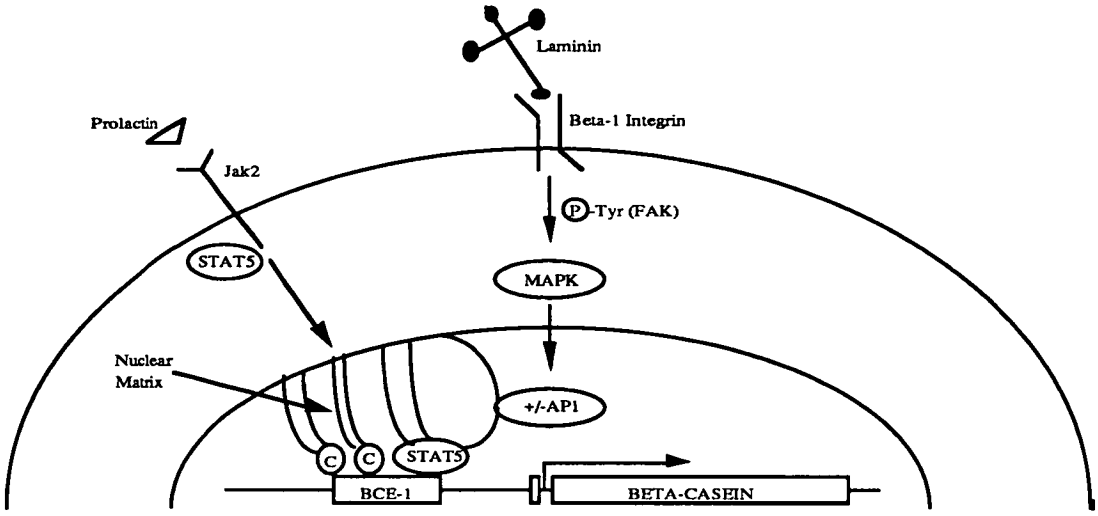


Figure 3. ECM-signalling and mammary epithelial cell differentiation.

C, CCAATT/enhancer binding protein; FAK, focal adhesion kinase; MAPK, mitogen-activated protein kinase; AP1, transcription factor; BCE-1, a 160 bp enhancer element 5' of the bovine β-casein gene; STAT5, signal transducer and activator of transcription (5). Adapted from Roskelley *et al.*, (1995).

modifications of the ECM influence ductal branching, end-bud development and epithelial proliferation (Silbertein *et al.*, 1992). During pregnancy and lactation, an intact basement membrane is required for the emergence of differentiated function (Sympson *et al.*, 1994).

The general model for the ECM-mediated differentiation is as follows: Extracellular signals are transmitted across the cell membrane via the transmembrane receptors (e.g., integrins) which recognise ECM molecules (e.g., laminin). The changes in these receptors, triggered by the ligand binding, cause an intracellular cascade of signal transduction leading to changes in gene expression, and therefore the growth and differentiation state of the cells.

1.3. MILK COMPOSITION

The function of the mammary gland is to synthesise and secrete the milk that is essential for the nourishment of mammalian young. Milk is an extremely complex biological fluid composed predominantly of water, fat, protein, carbohydrate (mainly lactose), calcium and salts as well as minor constituents such as several vitamins, enzymes, cellular metabolites, dissolved gases and trace elements (Davies *et al.*, 1983). The major constituents are the same in all species examined (Jenness, 1986) although the relative quantities vary considerably between species (Table 1).

Table 1. The major components of milk

Species	Fat (g/l)	Protein (g/l)	Lactose (mM)	Calcium (mM)
Sheep	74	55	133	58
Pig	68	48	153	104
Cow	37	34	133	30
Mouse	NA	97*	NA	NA
Human	38	10	192	7

Adapted from Davies *et al.*, (1983); NA: not available; * from the thesis of Kumar, S. (1993), University of Edinburgh

1.3.1. Milk fat.

Milk fat is composed of triacylglycerols (the major fraction), di- and monoacylglycerols, cholesterol and its esters, phospholipids and free fatty acids (long, medium and short-chain) and is of nutritional value (Davies *et al.*, 1983; Dils, 1986). The presence of these in the milk is a result of *de novo* synthesis in the mammary gland itself and from blood lipids (from the diet or from other body tissues).

1.3.2. Milk carbohydrate

Lactose is the predominant carbohydrate in the milk of most species (Davies *et al.*, 1983) but glucose, galactose and oligosaccharides are also present at much lower concentrations. Lactose maintains the osmotic balance of the milk and facilitates the absorption of calcium in the intestine.

1.3.3. Milk protein

Milk comprises two major classes of protein, namely the caseins and the milk serum or whey proteins (Table 2). The caseins and whey proteins are defined by acid precipitation and the relative abundance of each varies considerably between species (Jenness, 1986). All of the caseins and four of the whey proteins (α-lactalbumin, BLG, WAP and lactoferrin) are expressed and synthesised in the mammary gland whereas serum albumin, lysozyme and the immunoglobulins are present in milk due to transfer from the blood.

Table 2. Milk protein composition of sheep and mouse milk

Milk Protein	Concentration in milk (g/L)	
	Sheep	Mouse
Caseins		
α_S1-casein	12.0	28.0
α_S2-casein	3.8	NA
β-casein	16.0	21.0
κ-casein	4.6	2.4
Whey Proteins		
α-lactalbumin	0.8	trace
BLG	2.8	none
WAP	none	2.0
serum albumin	NA	NA
lysozyme	NA	NA
lactoferrin	NA	NA
immunoglobulins	NA	NA

Adapted from Lathe *et al.*, (1986) and the thesis of Kumar, S. (1993), University of Edinburgh; NA: not available

The caseins are thought to have a nutritional role as well as serving to transfer calcium and phosphorous from the mother to the neonate (Schmidt, 1982; Mepham, 1987) although pups suckling from mice lacking β-casein develop normally (Kumar *et al.*, 1994).

The whey proteins also appear to have functional and nutritional roles. α-lactalbumin acts as a cofactor in the synthesis of lactose by the enzyme, galactosyltransferase (Kuhn, 1983). Lactoferrin binds iron (Masson & Heremans, 1971) and is thought to transfer iron from the mother to the offspring (Saarinen & Siimes, 1979) and perhaps also inhibit the growth of serveral bacteria (Stuart *et al.*, 1984). The function of WAP (Hennighausen & Sippel, 1982) and BLG is unknown although the 3-dimensional crystal structure of bovine BLG is consistent with a role

in vitamin A transport (Papiz *et al.*, 1986; see 1.6.). WAP expression reaches levels as high as 15 mg/ml in the rabbit (Grabowski *et al.*, 1991) due to a high rate of transcription and a high degree of mRNA stability (Chen & Bissell, 1989). Overexpression of WAP in transgenic mice can result in incomplete lobulo-alveolar development and loss of milk production, implying a role in mammary development (Burdon *et al.*, 1991). The immunoglobulins confer passive immunity to the offspring (Mepham, 1987).

Several milk protein genes have been isolated and characterised (Rosen, 1987; review) and many of the *cis*-acting sequences that are important for milk protein gene expression have been identified (see 1.4. & 1.5.2.).

1.4. FACTORS INTRINSIC FOR MILK PROTEIN GENE EXPRESSION

The milk protein genes are expressed in secretory epithelial cells. Epithelial cell differentiation has been extensively studied in tissue culture by following the expression of the milk protein genes lactoferrin, β-casein and WAP (reviewed by Roskelley *et al.*, 1994). These experiments have produced evidence for a hierarchy of ECM-dependent signals that regulate mammary gland development (Fig. 4).

Firstly, architectural changes (tier 1) in cell shape (i.e., "rounding" of the epithelial cells) are necessary but not sufficient for β-casein expression (Roskelley *et al.*, 1994). This alteration in nuclear architecture appears to "prime" the cells for ligand-induced regulation of differentiation, implying that the nuclear architecture is altered via the cytoskeleton due to the rounding process. Secondly, laminin or prolactin-specific signals (tier 2) can activate an ECM-responsive element (BCE-1) and induce endogenous β-casein expression. Thirdly, morphogenesis (tier 3) results in the formation of "alveoli" in tissue culture models and the expression of WAP.

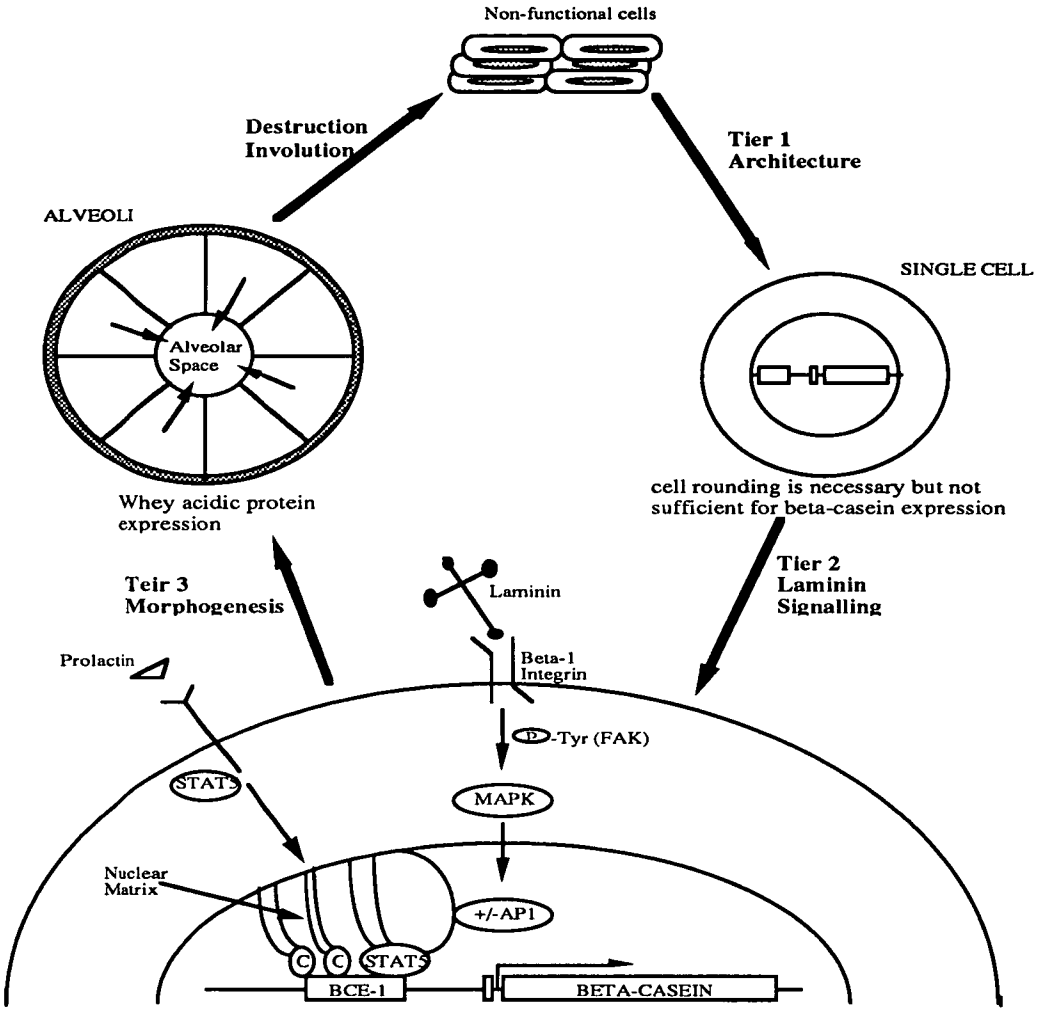


Figure 4. Mammary epithelial cell differentiation, ECM-signalling and milk protein gene expression.

C, CCAATT/enhancer binding protein; FAK, focal adhesion kinase; MAPK, mitogen-activated protein kinase; AP1, transcription factor; BCE-1, a 160 bp enhancer element 5' of the bovine β-casein gene; STAT5, signal transducer and activator of transcription (5). Taken from Roskelley *et al.*, (1995).

Finally, after weaning, degradation of the ECM by metalloproteases (Talhouk *et al.*, 1992) triggers the massive programmed cell death that occurs during involution (Strange *et al.*, 1992; Boudreau *et al.*, 1995).

Low levels of milk protein genes are expressed already in virgin mice, but their synthesis increases dramatically during pregnancy following a characteristic time course that is different for each protein. For example, β-casein and WAP expression are detectable at low levels in the virgin whereas α-lactalbumin is not (Fig. 5). β-casein expression is detectable in only a few alveolar cells by day 9 of gestation and in 30% of the cells by day 11 (Robinson *et al.*, 1995). Mosaic expression of WAP is detected by day 14 of gestation. It takes until day 18 of gestation before uniform expression of β-casein and WAP is achieved.

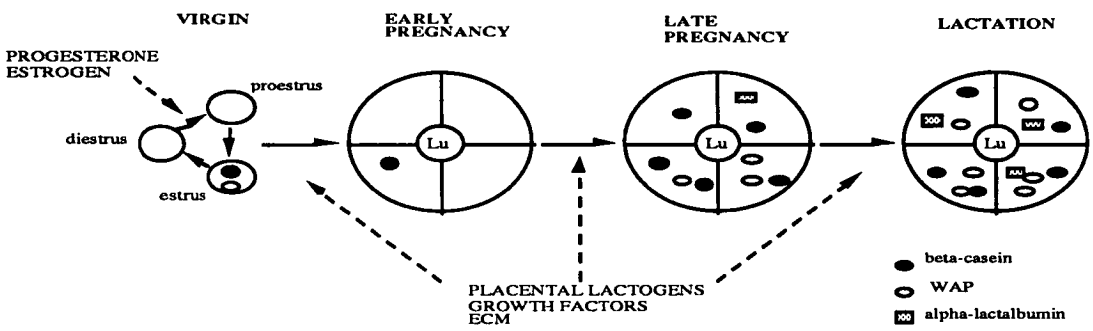


Figure 5. Model for the regulation of differentiation of mammary alveolar cells during puberty, pregnancy and lactation.

Lu, lumen; taken from Robinson *et al.*, 1995.

Mammary gland development and milk protein gene expression is in response to placental lactogens, growth factors and the ECM. The signals from these extracellular influences are transmitted to the regulatory elements of the genes via intracellular signalling. Evidence from cell and organ culture studies and

transgenic animals illustrate that the differences in milk protein gene expression may be attributed to different regulatory elements in the milk protein genes and their interactions with different hormones and/or different ECM factors. For example, WAP expression requires the synergistic action of prolactin, insulin and glucocorticoids whereas β-casein expression only requires the presence of prolactin (Hennighausen *et al.*, 1988; Pittius *et al.*, 1988; Burdon *et al.*, 1991; McKnight *et al.*, 1991; Shamay *et al.*, 1992; Wakao *et al.*, 1994). Response elements within the promoters of the two genes have been identified but the mode of the differential regulation is not understood (Li & Rosen, 1994 and 1995).

1.5. TRANSGENICS AS A METHOD OF PRODUCING BIOMEDICAL PROTEINS

The advent of transgenic technology and the cloning of several "biomedical" genes (genes that are not expressed or aberrantly expressed in a sub-population of humans and associated with various genetic disorders) has led to the enticing prospect of producing large quantities of these proteins in the milk of farm livestock (Lathe, 1985) using recombinant DNA and embryo microinjection technology.

1.5.1. Choice of biomedical proteins

Several genetically heritable disorders are caused by a deficiency in circulating levels of particular proteins. For example, haemophilia B (Christmas disease), which afflicts 1/30,000 males (Brownlee, 1987), is caused by a deficiency in human factor IX (FIX), an essential component of the blood coagulation cascade. Deficiency in human alpha-1-antitrypsin (AAT), an elastase

inhibitor, predisposes to emphysema and related disorders. AAT deficiency is one of the most common lethal hereditary disorders to affect Caucasian males of European decent (Crystal, 1989; Carlson *et al.*, 1988).

Conventionally these proteins are purified from the blood of blood donors and used in replacement therapy. This is an expensive and time consuming process because large quantities are required (e.g., AAT is present at 2 g/l in plasma and has a half-life of only 6 days) and the blood has to be screened for the presence of viruses such as those that cause hepatitis and AIDS. In the U.S.A. alone there are more than 20,000 sufferers of emphysema who require some 200 g of AAT per patient per year (Casolaro *et al.*, 1987) making a strong case for an alternative method of production. Human sequences encoding FIX and AAT have been cloned and characterised (Kurachi & Davie, 1982; Choo *et al.*, 1982; Jaye *et al.*, 1983; Ciliberto *et al.*, 1985; Kelsey *et al.*, 1987).

1.5.2. Choice of expression system

Large animal production of biomedical proteins has several advantages over bacterial (Courtney *et al.*, 1984), yeast (Sleep *et al.*, 1991) and mammalian tissue culture systems (Garver *et al.*, 1987). For example, farm animals are more likely to supply the correct post-translational modifications required for a functional human protein (Casolaro *et al.*, 1987) and they are relatively cheap to maintain. Of the three most common farm animals (pig, sheep, and cattle), Lathe *et al.*, (1986) argued that sheep would offer the most favourable compromise between several important parameters including the technical feasibility for the collection of eggs, injection of DNA into embryos and the overall cost of the procedures (Table 3).

Table 3. Comparison of different farm livestock.

	Pig	Sheep	Cattle
Availability of techniques for determining time of follicle maturation	Yes	(Yes)	(No)
Number of ovulations per animal without superovulation	10	1-3	1
Number of ovulations per animal with superovulation	15-20	4-10	possibly 6
Visualisation of pronuclei	(Yes)	(Yes)	(No)
In vitro culture of early embryo	(Yes)	(Yes)	No
Seasonal breeding	No	Yes	No
Relative cost of embryo transfer	1	2	110

Taken from Lathe *et al.*, (1986).

The mammary gland was chosen as the tissue for expression of biomedical proteins because (i) the tissue is well adapted to high levels of protein secretion into milk, (ii) milk production is strictly "inducible" by pregnancy and parturition, (iii) there are several tissue-specific milk protein promoters, (iv) collection of milk is relatively non-invasive to the animal, (v) milk is renewable.

Due to the high costs and long time-scales involved in large animal transgenics DNA constructs have generally been designed and tested in transgenic mice (Table 4). Tissue-specific expression is achieved by fusing the 5' regulatory regions from milk protein genes with the sequences encoding the biomedical protein (either genomic or cDNA). Expression is generally limited to the mammary gland and the protein products secreted into the milk. Expression levels are highest using BLG or WAP promoters whereas levels achieved with casein gene promoters have been surprisingly low, possibly due to limited regulatory sequences within the constructs. Constructs designed and tested in transgenic mice have been used to

Table 4. Mammary gland-specific transgenic animals

Source of promoter	Amount of promoter	To express	Animal	Protein levels	Reference
WAP-rat	Entire gene	Rat WAP	Mice	27% of endogenous	Bayna, 1990
WAP-mouse	2.6 kb 5', SV40 Poly A	cDNA t-PA	Mice	300 ng/ml	Gorden, 1987
WAP-mouse	2.6 kb 5', SV40 Poly A	cDNA t-PA	Goats	3 μg/ml	Ebert, 1991
WAP-mouse	Entire gene	Mouse WAP	Pigs	1 g/L	Wall, 1991
WAP-mouse	2.3 kb 5', 1.6 kb 3'	cDNA Human SOD	Mice	0.7 mg/ml	Hansson, 1994
WAP-mouse	2.3 kb 5', 1.6 kb 3'	cDNA Human Protein C	Mice	3 μg/ml	Velander, 1992a
WAP-mouse	2.3 kb 5', 1.6 kb 3'	cDNA Human Protein C	Pigs	0.001-1 mg/ml	Velander, 1992b
WAP-mouse	4.2 kb 5'	gDNA Human Protein C	Mice	0.1-0.7 mg/ml	Drohan, 1994
WAP-rabbit	17.6 kb 5'	gDNA Human AAT	Mice	10 mg/ml	Bischoff, 1992
WAP-rabbit	6.3 kb	gDNA Human GH	Mice	4-22 mg/ml	Devinoy, 1994
α-lactalbumin-bovine	Entire gene	Bovine α-lactalbumin	Mice	1.5 mg/ml	Bleck, 1993
α-lactalbumin-bovine	750 bp 5', 336 bp 3'	cDNA Bovine α-lactalbumin	Mice	0.0025-0.45 mg/ml	Vilotte, 1989
α-lactalbumin-bovine	477 bp 5', 336 bp 3'	cDNA Bovine α-lactalbumin	Mice	0.1 mg/ml	Soulier, 1992
α-lactalbumin-goat	Entire gene	Goat α-lactalbumin	Mice	1.2-3.7 mg/ml	Soulier, 1992
BLG-sheep	4 kb 5', 1.9 or 7.3 kb 3'	BLG	Mice	23 mg/ml	Simons, 1987
BLG-sheep	4 kb 5'	Human AAT	Mice	0.4-7.3 mg/ml	Archibald, 1990
BLG-sheep	4 kb 5'	Human AAT	Sheep	35 g/L	Wright, 1991
BLG-sheep	3, 5.5 or 10.8 kb 5', 8 kb 3'	BLG	Mice	1-8.5 mg/ml	Shani, 1992
BLG-sheep	3 kb 5'	gDNA Human SA	Mice	2.5 mg/ml	Shani, 1992
BLG-sheep	1.8 kb 5', 4.6 kb WAP 3'	cDNA SOD	Mice	10 ng/ml	Hansson, 1994
β-casein-rat	3.5 kb 5', 3 kb 3'	Rat β-casein	Mice	0.01-1% (mRNA)	Lee, 1988
β-casein-rabbit	2 kb 5'	Human interleukin-2	Rabbits	430 ng/ml	Buhler, 1990
β-casein-goat	3 kb 5', 6 kb 3'	Goat β-casein	Mice	12-24 mg/ml	Persuy, 1992
β-casein-goat	4.2 kb 5', 5.3 kb 3'	Goat β-casein	Mice	1 mg/ml	Roberts, 1992
β-casein-goat	6.2 kb 5', 7.1 kb 3'	cDNA bovine κ-casein	Mice	0.94-3.85 μg/μl	Guterrez, 1996
α_{S1}-casein-bovine	1.35 kb 5', 1.5 kb 3'	cDNA bovine α_{S1}-casein	Mice	0.1% (mRNA)	Clarke, 1994
α_{S1}-casein-bovine	1.35 kb 5', SV40 3'	Bacterial CAT	Mice	3 ng/ml	Clarke, 1994
α_{S1}-casein-bovine	2.9 kb 5', 3.5 kb 3'	cDNA Human IGF-1	Rabbits	1 g/L	Brem, 1994
α_{S1}-casein-bovine	21 kb 5', 2 kb 3'	gDNA Human urokinase	Mice	1-2 mg/ml	Meade, 1990
α_{S1}-casein-bovine	21 kb 5', 2 kb 3'	cDNA Human lysozyme	Mice	0.25-0.71 μg/μl	Maga, 1994
α_{S1}-casein-bovine	6.2 or 14.2 kb 5', 6.5 kb 3'	cDNA Human lactoferrin	Mice	0.1-36 μg/ml	Platenberg, 1994

t-PA, tissue-plasminogen activator; SOD, superoxide dismutase; GH, growth hormone; SA, serum albumin; CAT, chloramphenicol acetyltransferase; IGF-1, insulin-like growth factor-1; gDNA, genomic DNA; cDNA, complementary DNA; WAP & BLG, see text. Adapted from Maga and Murray, 1995.

generate transgenic rabbits, goats, pigs, and sheep. The BLG promoter has produced levels of AAT as high as 35 g/l (50% of total milk protein) in the milk of transgenic sheep (Wright *et al.*, 1991). These levels are the highest yet recorded for any biomedical protein in transgenic sheep and exceed all other pharmaceutical methods of production. Purification of AAT from sheep milk and analysis of the pharmacological activity is ongoing (Clark, A.J. pers. comm.).

1.6. OVINE β-LACTOGLOBULIN (BLG)

Ovine β-lactoglobulin (BLG) exists as two major alleles (Bell & McKenzie, 1967). These encode variants of BLG, termed BLG-A and BLG-B, which differ in only one amino acid; Tyr^{20} in BLG-A is replaced by His in BLG-B (Kolde & Braunitzer, 1983). Four genomic clones were isolated (Ali & Clark, 1988) from a sheep genomic DNA library. Two of these clones, termed SS1 and SS12 have been characterised by restriction enzyme and DNA sequencing analysis. A comparison of the two sequences illustrated that they differ by only 1 bp in the coding region which accounts for the Tyr/His difference (Ali & Clark, 1988; Harris *et al.*, 1988; Ali *et al.*, 1990). Clone SS-1 is a 16.2 kb *Sal*I-*Sal*I fragment that comprises 4 kb of 5' sequence, a 4.9 kb transcription unit and 7.3 kb of 3' flanking sequence (Fig. 6).

The transcription unit is transcribed into a 0.8 kb mRNA that comprises about 5% of poly(A) RNA (Mercier *et al.*, 1985). This is translated into a 36 kD dimeric protein (the dimer is composed of two identical subunits of 162 amino acids each) that is secreted into the milk. BLG is found in the milk of a diverse range of species (e.g., dogs to dolphins [Jenness, 1982; Pervaiz & Brew, 1985]) and is the major whey protein present in the milk of ruminants. However, BLG is not normally present in mouse or human milk. In other species BLG expression is

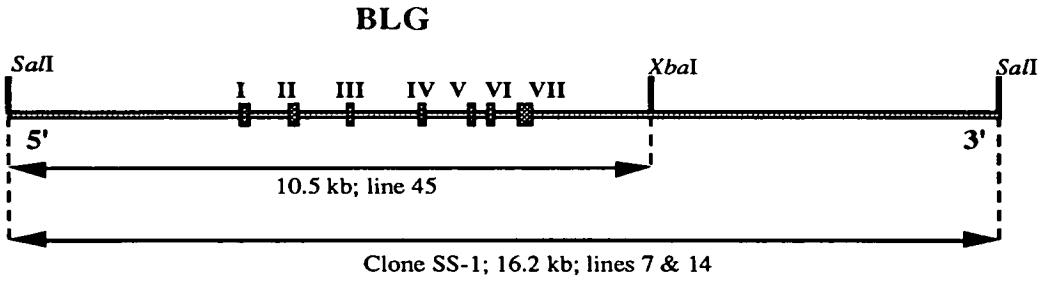


Figure 6. A map of BLG clone SS-1

BLG exons are indicated (I-VII)

limited to the mammary gland where it exhibits tight temporal regulation; in sheep, expression starts during mid-pregnancy (Gaye *et al.*, 1986), reaches a peak in lactation and tails off as involution commences.

Despite this information the function of BLG is still unknown. DNA sequence comparisons have revealed similar exon and intron arrangements with a variety of secretory proteins (mouse urinary protein, alpha-1-acid glycoprotein, rat retinol binding protein and apolipoprotein: Ali & Clark, 1988). The 3-dimensional crystal structure of bovine BLG (Papiz *et al.*, 1986) is very similar to that of human retinol binding protein (Newcomer *et al.*, 1984), suggesting a role in vitamin A transport. Indeed, BLG binds a variety of hydrophobic molecules (Lovrien & Anderson, 1969; Futterman & Heller, 1972), including retinol (Cogan *et al.*, 1976; Fugate & Song, 1980); [^{125}I]-labelled BLG-retinol complex binds to purified microvilli from the small intestine from one-week-old calf intestine (Sawyer *et al.*, 1985), giving further credence to the idea that BLG is involved in vitamin transport.

1.7. GENERATION OF TRANSGENIC MOUSE LINES 7, 14 & 45

Since 1980 several methods have been established for introducing transgenes into the germline including retroviral vectors and electroporation into embryonic stem (ES) cells. The relative merits of each technique for the generation of transgenic animals have been reviewed (Lathe *et al.*, 1986; Simons & Land, 1987; Clark *et al.*, 1987). The generation of transgenic animals by microinjection of DNA into one pronucleus of fertilised eggs is the most favoured method.

Fertilised eggs are flushed from the oviduct of superovulating female mice and immobilised by gentle aspiration using a blunt holding pipette (visualised using x400 magnification). The male pronucleus (usually the larger of the two) is injected with 2 pl of DNA solution (containing 40-400 copies of the DNA construct) using a microfine pipette and a micromanipulator. Surviving eggs (10-20%) are reimplanted into the oviduct of pseudo-pregnant mothers (see chapter 2) and allowed to develop. The proportion of offspring with the transgene is laboratory dependent; usually some 15-20% of the offspring carry the injected DNA. Integration usually takes place at a single site in the genome (for reasons that are not understood) and the locus can contain from one to several hundred copies of the transgene, usually in a head to tail array.

Clone SS1 was microinjected into mouse embryos as the full 16.2 kb fragment or as a 3' truncated 10.5 kb fragment (Simons *et al.*, 1987; Fig. 6). This resulted in three transgenic mouse lines that expressed high levels of BLG mRNA in mammary tissue and secreted the protein into the milk. Lines 7 and 14 contain ~25 and <5 copies of the 16.2 kb fragment respectively; line 45 carries ~20 copies of the 10.5 kb fragment (Whitelaw *et al.*, 1992; Dr Bruce Whitelaw, pers. comm.). The developmental regulation of BLG transgenes matches that of the endogenous β-casein gene in mice (Harris *et al.*, 1991). For example, endogenous β-casein and

BLG mRNA expression are observed at day 10 and day 12 of gestation respectively in transgenic mice and the expression of both increases until mid-lactation before falling to their original levels. During the analysis of BLG transgenic mouse lines it was noticed that one line, line 7, exhibited variable levels of mRNA and protein expression within the line (see chapter 3). This is an important phenomenon in the context of producing biomedical proteins in the milk of farm livestock as well as in the broader context of transgenic research in general.

1.8. UNSTABLE EXPRESSION OF TRANSGENES

Unstable transgene expression has been described previously; Palmiter *et al.*, (1984) reported that the level of HSV thymidine kinase (TK) expression could vary by more than an order of magnitude among transgenic progeny of the same founder. More recently a number of different transgene insertions have been described that express to variable degrees within individual cell lines or transgenic mouse lines (Table 5). In these cases, however, there has been no common explanation for the instability of expression.

Unstable expression patterns may be due to a strong selection against transgene expression. This had been demonstrated in the failure of sperm fertility engendered by testicular TK expression (Wilkie *et al.*, 1991) and in the case of high-level hepatic expression of plasminogen activator (Sandgren *et al.*, 1991). In both these cases there appears to be selection for transgene deletion. A transgene inserted into the X chromosome (Tan *et al.*, 1994) or an X-autosome translocation (Cattanach, 1974) generates a mosaic pattern of expression due to stochastic inactivation of the X chromosome. More commonly, silencing has been observed when the transgene integrates into repeat sequence or satellite DNA (Butner & Lo,

Table 5. Summary of examples of unstable transgene expression

Transgene	Cause of Unstable Expression	Reference
I-FABP-hGH	differences in clonally derived cells	Sweetser *et al.*, (1988a).
L-FABP-hGH	"	Sweetser *et al.*, (1988b).
Met-HSV *tk*	selection for transgene deletion	Wilkie *et al.*, (1991).
Alb-uPA	"	Sandgren *et al.*, (1991).
hsp68-lacZ	influence of strain-specific modifier gene(s)	McGowan *et al.*, (1989).
HSV *tk-lacZ*	"	Allen *et al.*, (1990).
Met-HRD	"	Engler *et al.*, (1991).
HSV-*tk*	integration into heterochromatin	Butner & Lo, (1986).
HSV-*tk*	"	Talarico *et al.*, (1988).
HMG CoA-*lacZ*	"	Tan *et al.*, (1993).
hCD2	"	Festenstein *et al.*, (1996).
tyrosinase-SV40E	"phenoclones"	Bradl *et al.*, (1991).
Met-tyr	"	Mintz & Bradl, (1991).
tyrosinase-tyr	"	Porter & Meyer, (1994).
mini-*white*	site of integration and/or the transgene *per se*	Dorer & Henikoff, (1994).
α-globin/*LacZ*	"	Robertson *et al.*, (1995).
brown	"	Sabl & Henikoff, (1996).
α-globin/*LacZ*	"	Robertson *et al.*, (1996).
γ/β-geo/HS2-FRT	"	Walters *et al.*, (1996).
γ/β-geo 3MRE	"	Walters *et al.*, (1996).
MBP mini-gene	unknown	Katsuki *et al.*, 1988.
mLCR5'HS2-huβ	"	Enver *et al.*, 1994.
μLCR -730 $^{A}\gamma$	*cis*-acting regulatory element	Stamatoyannopoulos *et al.*, (1993).

The transgene abbreviations are as follows: I-FABP, intestinal fatty acid binding protein gene promoter; L-FABP, liver fatty acid binding protein gene promoter, hGH, human growth hormone gene; HSV *tk*, herpes simlex virus thymidine kinase gene; Alb, mouse albumin promoter; uPA, mouse urokinase plasminogen activator gene; *hsp68*, mouse heat-shock protein 68 gene promoter; *lacZ*, *Escherichia coli* β-galactosidase gene; tyrosinase, mouse tyrosinase (monophenol, L-dopa:oxygen reductase) gene promoter; SV40E, simian virus 40 early-region transforming sequences; Met,

metallothionein gene promoter; tyr, tyrosinase cDNA; γ/β-geo/HS2-FRT, β-geo expression is driven by the human γ-globin promoter and a downstream 1 kb fragment containing the 5'HS2 element from the human β-globin locus control region (LCR) is flanked between two flp recognition target (FRT) elements (note: variegated expression is observed when the 5'HS2 element is removed); γ/β-geo 3MRE, as before except the HS2-FRT element was exchanged for three copies of the upstream region of the mouse metallothionein (mMT-1) promoter; MBP, mouse myelin basic protein promoter followed by the rabbit β-globin gene carrying the mouse MBP cDNA in the antisense orientation and by the polyadenylation sites for rabbit β-globin and SV40; mLCR5'HS2-huβ, murine β-globin locus control region DNase hypersensitive site 2 linked to a 4.8 kb human β-globin gene-containing fragment; μLCR, 2.5 kb locus control region cassette containing DNase hypersensitive sites 1-4 of the human β-globin locus; -730 $^{A}\gamma$, -730 nucleotides 5' of the human $^{A}\gamma$-globin cap site plus the human $^{A}\gamma$-globin gene.

1986; Talarico *et al.*, 1988; Festenstein *et al.*, 1996). In other cases different levels of transgene expression between animals of the same transgenic lineage have been attributed to the influence of strain-specific modifier genes (McGowan *et al.*, 1989; Allen *et al.*, 1990; Engler *et al.*, 1991).

Mosaic patterns of expression were also observed in transgenic animals bearing I-FABP-hGH or L-FABP-hGH fusion transgenes (Sweetser *et al.*, 1988a; Sweetser *et al.*, 1988b). Here the mosaicism was initially attributed to a deficit of *cis*-acting elements in the transgene. However one of the authors (J.I.G.) and colleagues later suggest that the mosaic expression implied subtle differences in clonally derived epithelial cells based on evidence that endogenous I-FABP, L-FABP and cellular retinol binding protein II (CRBP II) were expressed in a transient mosaic fashion (Rubin *et al.*, 1989). This transient mosaic expression of endogenous villi genes is reminicent of the transient mosaic expression of endogenous milk protein genes described by Robinson *et al.*, 1995 (see section 1.4.).

Mosaic expression patterns were also observed in mice carrying different tyrosinase fusion transgenes (Bradl *et al.*, 1991; Mintz & Bradl, 1991; Porter & Meyer, 1994) as well as in non-transgenic mice bearing a *chinchilla-mottled* mutation at the endogenous tyrosinase gene (Porter *et al.*, 1991). In these reports all transgenic animals examined, of different lineages, exhibited a similar pattern of light and dark transverse stripes of coat colouration, suggesting that the striated coat colour is an inherent property of the transgene.

Recently, mosaic expression of transgenes has been described in *Drosophila* (Dorer & Henikoff, 1994; Sabl & Henikoff, 1996) and mice (Robertson *et al.*, 1995; Walters *et al.*, 1996; Robertson *et al.*, 1996) where stochastic transgene silencing has been attributed to the site of transgene integration and/or the repetitive nature of the transgene *per se*.

1.9. AIMS OF THE PROJECT

It was with some of these examples in mind that we entered into a systematic study of expression levels in three BLG transgenic mouse lines (lines 7, 14 and 45) in the milk and mammary tissue. The aims of this project were to quantitate the variance of BLG in line 7 and to compare this with two other BLG expressing lines and subsequently establish the cause of the variable expression. It was anticipated that this study would enhance our knowledge about the efficiency of transgene expression and therefore contribute to the future avoidance of variable transgene expression in mice and farm livestock.

CHAPTER 2

MATERIALS AND METHODS

2.1. TRANSGENIC MOUSE LINES

Transgenic lines employed in this study were lines 7, 14 and 45: lines 7 and 14 harbour an identical transgene (16.2 kb *Sal*I-*Sal*I fragment) while line 45 carries a 3' truncated transgene (10.5 kb *Sal*I-*Xba*I) [see Fig. 6.; chapter 1).

The generation of these lines is described by Simons *et al.*, (1987). Briefly, the BLG clone SS-1 was isolated from a genomic DNA library (Ali & Clark, 1988). SS-1 comprises 4 kb 5' of the promoter, a 4.9 kb transcription unit and 7.3 kb of 3' flanking sequence. Pronuclear stage eggs were obtained from superovulated C57BL/6 x CBA F_1 females, after mating with C57BL/6 x CBA F_1 males (to induce superovulation: pregnant mare's serum is used to mimic follicle-stimulating hormone and human chorionic gonadotropin to mimic luteinizing hormone [Hogan *et al.*, 1986]). The cloned BLG gene was microinjected into either pronucleus of fertilised eggs as the entire 16.2 kb *Sal*I-*Sal*I fragment or as the 10.5 kb *Sal*I-*Xba*I fragment (5.7 kb removed from the 3' end) [Fig. 6]. Cleaved embryos (two-cell stage) were transferred to pseudopregnant (females that have been mated with vasectomised males) MF1 recipients (generation zero; G_0). Transgenic lines were established from these and maintained by systematic crossing to C57BL/6 x CBA F_1 hybrid mice (a line is a group of male and female mice decended from the same G_0 female containing the same transgenic insertion of foreign DNA).

2.2. MILKING OF MICE

All milk samples were collected from 6-8 week old transgenic mice at day 11 of lactation. Litters were standardised to 5 pups per mother at birth. Pups were separated from their mother 3 hr before milking. Mothers were injected with 0.3 International Units of oxytocin (Sigma) in 0.1 ml of distilled water into the intraperitoneal space (see Hogan *et al.*, [1986] for a description of this procedure). After 10 min an injection with anaesthetic was given (Hypnorm/Hypnovel; 10 µl/g). The mice were anaesthetised for 10-15 min before gentle massage was applied to the mammary glands and milk collected in 50 µl capillary tubes. Milk was dispensed into 50 µl amounts using a Hamilton syringe and stored at -20°C.

Defatted milk was prepared by adding 200 µl of distilled water to 50 µl of thawed mouse milk, brief vortexing to mix, followed by a 12 sec pulse spin in a microcentrifuge. Care was taken to resuspend any protein pellet that formed, without dislodging the fat around the side of the tube. The supernatant (defatted milk) was transferred into a clean microcentrifuge tube, diluted further (1:200 final dilution) in reducing buffer (0.5 M Tris-HCl pH 6.8, 10% v/v glycerol, 10% w/v SDS, 5% v/v β-mercaptoethanol [Sigma], 0.1% w/v bromophenol blue) and boiled for 4 min to denature the protein. The SDS in the buffer dissociates all proteins in the sample into their individual polypeptide subunits (i.e., the 32 kD BLG dimeric protein is dissociated into two 18 kD dimers) and the β-mercaptoethanol cleaves any disulphide bonds (there are two in BLG). Samples were stored at -20°C until ready for electrophoresis.

2.3. PURIFICATION OF OVINE β-LACTOGLOBULIN (BLG)

Sheep milk was pooled from 12 ewes (4 Black Face, 4 Grey Face and 4 Black Welsh) and BLG protein purified by the method described by Aschaffenburg & Drewry (1957).

500 ml of pooled sheep milk was warmed to 40°C, 100 g of Na_2SO_4 added, stirred until dissolved, allowed to cool to 25°C and filtered (Whatman No 4). The filtrate contained α-lactalbumin, BLG and serum albumin. All the proteins were precipitated with the exception of BLG by the addition of 1 ml of concentrated HCl (10.3 M) [BDH: specific gravity (1.16 g/ml) x purity (32.25%) / molecular weight (36.46)] for every 100 ml of filtrate with vigorous stirring reducing the pH to almost 2.0. The precipitate was separated from BLG by centrifugation (13,200 g, 30 min) and filtration. BLG was aliquoted into 1 ml amounts and stored at -70°C. The purity of the BLG preparation was checked by SDS-PAGE and Western blotting.

2.4. ESTIMATION OF BLG PROTEIN CONCENTRATION

Calculation of BLG concentration was performed at The Moredun Research Institute, Gilmerton Road, Edinburgh with guidance from Mr Andrew Dawson using the micro-Kjeldahl method (Rowland, 1938) essentially as described by Davies & Law (1983). The "N" value was multiplied by 6.38 (multiplication factor for nitrogen content of bovine BLG).

2 ml of purified BLG was precipitated with 10% w/v TCA (final concentration), the protein pellet was washed in 100% acetone, resuspended in 500 µl of 2 M NaOH (left at r.t. overnight), made up to 10 ml with distilled water, split into several aliquots (2 x 2 ml, 2 x 1 ml, 2 x 500 µl, and 2 x 250 µl) and hydrolysed (to "digest" the protein) by boiling in 3 ml of H_2SO_4 (nitrogen-free, 98% w/w:

BDH) in the presence of 2.0 g of catalyst (100 g of potassium sulphate and 1 g of selenium powdered together) in a graduated micro-Kjeldahl flask overnight. This converts the carbon in the sample to CO_2 and the hydrogen to H_2O, both of which evaporate leaving the nitrogen as NH_4^+. The NH_4^+ was converted into NH_3 by distillation with 10 M NaOH overnight into 5 ml of boric acid indicator (alkalization of the acid/protein digest). This was titrated against HCl until the clear/yellow solution turned a brown/red colour (judged by Mr Dawson). The amount of HCl required for this colour change was noted. The two samples from each dilution were averaged, a blank was subtracted, multiplied by 0.0952 (1 ml of HCl = 0.0952 mgN), multiplied by the appropriate dilution, multiplied by 6.38 (see above) and divided by two to give the concentration of BLG in mg/ml (see chapter 3).

2.5. MILK PROTEIN ANALYSIS

2.5.1. Preparation of SDS-polyacrylamide gels

Quantitative milk protein analysis was performed using SDS polyacrylamide gel electrophoresis (SDS-PAGE) [Laemmli, 1970].

20 x 20 cm, 20 x 18 cm glass plates and 1 mm gel spacers were cleaned using 70% v/v ethanol before being clamped into a vertical position to act as gel moulds. 20% resolving gel solution was prepared by mixing 1.66 ml of distilled water, 6.66 ml of 1.5 M Tris-HCl pH 8.8, 266 µl of 10% w/v SDS, and 18 ml of acrylamide (Sigma) : N, N'-bis methylene acrylamide (Sigma) [30% T, 2.67% C: 29.2 g acrylamide, 0.8 g N, N'-bis methylene acrylamide]. This solution was de-gassed under vacuum to aid polymerisation and reduce the chance of bubbles forming within the gels. Polymerisation was catalysed by adding 67 µl of 10% w/v ammonium persulphate (prepared fresh) and 13 µl of N, N, N', N'-

tetramethylethylenediamine (TEMED) [Sigma], mixed and poured immediately into the 1 mm space between the glass plates. A 1 cm layer of distilled water was floated on top to aid polymerisation and produce a smooth interface between the resolving gel and the stacking gel. The resolving gel was allowed to polymerise overnight.

The water layer on top of the resolving gel was removed and 4% stacking gels prepared by mixing 4 ml of distilled water, 1.66 ml of 0.5 M Tris-HCl pH 6.8, 67 µl of 10% w/v SDS, 866 µl of acrylamide : bis (see above), 34 µl of 10% w/v ammonium persulphate, 4 µl TEMED and poured on top of the resolving gel. Before polymerisation, a comb was placed into the stacking gel to produce the wells. The stacking gel was allowed to polymerise for 1 hr. The comb was removed and the wells washed with running buffer (25 mM Tris, 0.2 M glycine, 0.4 mM SDS) before loading the samples.

2.5.2. SDS-PAGE

15 µl of each defatted mouse milk sample in reducing buffer was loaded into polyacrylamide gels using a Hamilton syringe. After loading the gels were assembled into Bio-Rad Protean II running apparatus (usually two gels at a time) with 4.5 l of gel running buffer surrounding the glass plates (lower reservoir) and ~250 ml in the upper reservoir. Samples were electrophoresed at 32 mA (constant current) while in the stacking gel (1-2 hr) and at 48 mA while in the resolving gel (~5 hr). Gels were stained and fixed overnight using a solution of Coomassie Blue stain (50% v/v methanol, 7.5% v/v acetic acid, 2% w/v TCA, 0.04% w/v Coomassie Blue G-250 [BDH]). The mouse milk proteins were fixed and the background stain removed by washing the gels for 2 x 30 min in a solution of 23% v/v ethanol, 7% v/v acetic acid followed by 2 x 30 min in 30% v/v ethanol.

Gels were stored at 4°C and scanned at 585 nm using a Shimadzu CS-9000 densitometer.

2.5.3. Western Blotting

Electrophoresis was performed using 1.5 mm thick, 15% polyacrylamide gels. After electrophoresis, proteins were electrophoretically transferred to nitrocellulose membranes (Schleicher & Schuell) using the semi-dry method described by Khyse-Anderson, (1984). This involved constructing the following stack: two sheets of 3 MM paper (soaked in 0.3 M Tris, 20% v/v methanol, pH 10.4) onto the lower graphite anode plate, followed by one sheet of 3 MM paper (soaked in 25 mM Tris, 20% v/v methanol, pH 10.4), one sheet of nitrocellulose membrane (soaked in distilled water), the gel, two sheets of 3 MM paper (soaked in 25 mM Tris, 40 mM hexanoic acid, 20% v/v methanol, pH 9.4), and the upper graphite cathode plate. The stack was sealed with tape to prevent it from drying and transfer performed using 7 V/cm at 4°C for 45 min.

After transfer the nitrocellulose membrane was washed in 33% v/v horse serum (Gibco BRL) in washing fluid B (WFB: 0.05% v/v Tween 20 [Sigma], 0.1 M EDTA, 0.5 M NaCl in PBS, pH 7.2) at r.t. overnight, briefly washed in WFB, incubated with a rabbit antibody against BLG (a gift from Dr P. Gaye: INRA, Jouy-en-Josas, France) [1/300 dilution in 10% v/v horse serum, WFB; 20 ml total volume] at 37°C for 2 hr, WFB (two quick washes, 3 x 10 min washes, two quick washes), incubated with peroxidase-conjugated donkey anti-rabbit immunoglobulin (Scottish Antibody Production Unit [SAPU]) [1/300 dilution in 10% v/v horse serum, WFB; 20 ml total volume], WFB (as above) and washed briefly in distilled water. Immunoblots were developed with 40 mg diaminobenzidine in 40 ml of 10 mM Tris-HCl pH 7.9, 0.01% w/v hydrogen peroxide (H_2O_2) for 1-10 min until the

protein bands were clearly visible. The reaction was stopped by washing the membrane in distilled water and dried by blotting on 3 MM paper. Membranes were preserved by wrapping them in cling-film.

2.6. DNA MANIPULATION AND ANALYSIS

2.6.1. Digesting DNA with restriction enzymes

All genomic and plasmid DNA was cut using restriction enzymes and buffers (Boehringer Mannheim) according to manufacturers' instructions except that a 4 to 10-fold excess of enzyme was generally used.

e.g. 1.	10 µg	genomic DNA (in 20 µl distilled water)	
	1 µl	*Eco*RI (Boehringer Mannheim: 40 U/µl; 1 unit digests 1 µg of λDNA in 1 hr at 37°C)	
	4 µl	buffer (Boehringer Mannheim: enzyme dependent [see Table 6)	
	15 µl	distilled water	*Incubate at 37°C for 5 hr.*

e.g. 2.	1 µg	plasmid DNA (in 5 µl distilled water)	
	1 µl	*Eco*RI (10 U/µl)	
	2 µl	buffer	
	12 µl	distilled water	*Incubate at 37°C for 1 hr.*

Table 6 details the enzymes and buffers used in this study.

Table 6. Restriction enzymes and conditions for digestion

Enzyme*	Recognition sequence	Buffers‡
*Bam*HI	G/GATCC	B
*Bgl*II	A/GATCT	M
*Eco*RI	G/AATTC	H
*Eco*RV	GAT/ATC	B
*Hin*dIII	A/AGCTT	B
*Nde*I	CA/TATG	H
*Pst*I	CTGCA/G	H
*Pvu*II	CAG/CTG	M
*Sal*I	G/TCGAC	H
*Xba*I	T/CTAGA	H
*Xho*I	C/TCGAG	H

* Nomenclature (Smith & Nathans, 1973): A three-letter abbreviation for the parent organism (e.g., *Hin* for *Haemophilus influenzae* or *Bam* for *Bacillus amyloliquefaciens*), an additional letter if necessary to identify strain or serotype (*Hin*d or *Bam*H), followed by a Roman numeral to reflect the order of identification or characterisation (*Hin*dIII or *Bam*HI).

‡ Buffer B = 10 mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM 2-mercaptoethanol. Buffer H = 50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM dithioerythritol (DTE), pH 7.5. Buffer M = 10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTE, pH 7.5.

2.6.2. Phenol/chloroform extraction of DNA

Phenol/chloroform extraction is used to extract and purify nucleic acids from a variety of sources (Berger & Kimmel, 1987; Sambrook *et al.*, 1989). The aim of phenol extraction is the deproteinisation of an aqueous solution containing the DNA. Mixing phenol with the sample causes the dissociation of proteins from DNA. Centrifugation of the mixture yields two phases; a lower organic phenol phase carrying the protein (much of which is found in the white flocculent interphase) and the less dense aqueous phase containing the DNA. Used

in conjunction with phenol, chloroform improves the efficiency of extractions. This is principally due to its ability to denature protein, thereby assisting the dissociation of DNA from protein. The high density of the chloroform also enhances the separation of the phases, facilitating the removal of the aqueous phase with little cross-contamination with organic material.

An equal volume of phenol (equilibrated to pH >7.6 by repeated extraction with Tris buffer; Fisher Scientific UK) was added to each sample in microcentrifuge tubes (<500 µl samples) or glass Corex centrifuge tubes (up to 15 ml per sample), mixed by inverting the tubes for 5 min, centrifuged (12,000 g, 5 min, r.t.), the top aqueous phase was transferred into a clean tube taking care not to disturb the interphase and the process repeated using a 1:1 mix of phenol/chloroform followed by just chloroform (to remove any residual phenol). The DNA in the aqueous phase was precipitated using ethanol.

2.6.3. DNA precipitation with ethanol

For recovery of DNA from a typical reaction (e.g., for 1 µg DNA in 20 µl): added a 1/10 th volume of 3 M sodium acetate pH 5.5 and 2 volumes of 100% ethanol. This was mixed and chilled to -70°C for 15 min (gives at least 80% recovery for volumes <1 ml containing at least 10 µg/ml DNA) or overnight at -20°C for very low DNA concentrations (1 µg or less in 1 ml). For larger volumes precipitations were performed in 15-30 ml glass Corex tubes and the period at -70°C was increased to 30 min. The DNA was pelleted in a microcentrifuge (12, 000 g, 10 min 4°C). The DNA was visible as a "whitish" pellet at the bottom of the tube. In general, pellets of 10 µg were visible whereas 2 µg pellets were invisible. The supernatant was removed and the DNA pellet washed twice (centrifuged for 2 min between washes) with chilled 70% v/v ethanol to remove any solute trapped in the

precipitate. Pellets were dried in a Gyrovap (Howe) connected to an Edwards Freeze Dryer Modulyo for 1-2 min and resuspended in TE (pH 8.0) buffer (10 mM Tris-HCl, 1 mM EDTA). The EDTA chelates heavy metal ions which are commonly required for DNase activity, while the use of pH 8.0 minimises deamination. Samples were stored at -20°C.

2.6.4. Recovery of DNA fragments from low melting point (LMP) agarose gels

This procedure utilised the activity of β-agarase I (NEB), an agarose-digesting enzyme from *Pseudomonas atlantica*. β-agarase I acts by cleaving the agarose subunit, unsubstituted neoagarobiose [3,6-anhydro-α-L-galactopyranosyl-(1-3)-D-galactose] to neoagaro-oligosaccharides (Yaphe, 1957), producing carbohydrate molecules that can no longer gel and therefore releasing trapped DNA.

DNA fragments used as probes in Southern and Northern blotting procedures and for the construction of the tissue *in situ* expression vectors were recovered from 1% LMP agarose (Gibco BRL) gels. LMP agarose gels were allowed to set at 4°C. Cooling the gels made them more solid and therefore reduced the chances of the gel tearing at the wells when the comb was removed. Samples were electrophoresed and the appropriate fragments cut from the gels (visualised using a long-wavelength hand-held U.V. lamp) and transferred to microcentrifuge tubes. The weight of agarose was determined and a 1/10 th volume of 10 x β-agarase buffer (10 mM Tris-HCl, 1 mM EDTA, pH 6.5) [NEB] added. The agarose was melted at 65°C for 10 min, the tubes were transferred to 40°C (10 min), 1 unit of β-agarase (NEB) was added per 200 µl of molten agarose. The samples were mixed and then incubated at 40°C for at least of 2 hr, samples were cooled on ice for 15 min, microcentrifuged

(12,000 g, 15 min, 4°C), and transferred the supernatant into a clean microcentrifuge tube. 2 volumes of isopropanol were added, the samples were chilled to -20°C for 15 min, microcentrifuged (12,000 g, 10 min, 4°C) to pellet the DNA, washed twice with 70% v/v isopropanol, spin-vacuum dried for 5 min before resuspended in 50 µl TE buffer. Purity and yield of the fragment was checked on a mini-gel. Samples were stored at -20°C.

2.6.5. DNA ligation

"Sticky-end" ligation reactions were performed with 1 part cut plasmid : 3 parts insert (molar ratio) using the following:

~100 ng total DNA in 8 µl of distilled water

The solution was warmed to 45°C for 5 min to melt any cohesive termini that had reannealed. After chilling on ice (5 min) the following were added:

1 µl 10 x ligation buffer* (Boehringer Mannheim)

1 µl T4 DNA ligase (1 U/µl: Boehringer Mannheim)

* 20 mmol/l Tris-HCl, 10 mmol/l $MgCl_2$, 10 mmol/l dithioerythritol, 0.6 mmol/l ATP, pH 7.6.

Ligations were performed at 14°C for 1-4 hr. 1-2 µl of the ligation reaction were used to transform competent cells (see 2.6.7.).

2.6.6. DH5α cells

E. coli DH5α competent cells (Genotype: F^-, endA1, hsdR17(r_K^-, m_K^+), supE44, thi-1, λ^-, recA1, gyrA96, relA1, Δ(argF-lacZYA)U169, Ø80dlacZ[ΔM15]) were purchased from GIBCO BRL and stored at -70°C. The

Ø80dlacZ[ΔM15] marker provides α–complementation of the β-galactosidase gene, therefore enabling blue/white screening of colonies on bacterial plates containing X-gal and IPTG (see 2.8.2.).

2.6.7. Transformation of DH5α cells with plasmid DNA

DH5α cells were thawed on ice and the cells gently mixed using a blue tip. 1-2 µl (0.1-0.5 ng/µl) of ligated plasmid DNA was added to 100 µl of cells in an microcentrifuge tube, placed on ice for 30 min, heat shocked (42°C for 90 sec) and returned to ice for 2 min. 900 µl LB medium (without antibiotic) were added and incubated at 37°C for 1 hr (tape microcentrifuge tube to the top of a 15 ml glass tube) with vigorous shaking. Spread 50-100 µl of each sample onto agar plates (see 2.8.1.) using a sterile glass spreader. Incubated the plates overnight at 37°C.

2.6.8. Preparation of plasmid DNA

2.6.8.1. Small-scale preparation

Single white colonies were transferred into 15 ml glass tubes containing 2 ml of LB medium (Sambrook *et al.,* 1989) and 50 µg/ml of ampicillin (Sigma). Tubes were incubated at 37°C overnight with vigorous shaking. 1.5 ml of each culture was transferred into microcentrifuge tubes and microcentrifuged (12,000 g, 30 sec, r.t.) to pellet the cells. The supernatant was poured off and any excess liquid was blotted and removed using a pipette. The cells were resuspended in 100 µl of chilled TGE solution (50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA). 200 µl of freshly prepared 0.2 N NaOH, 1% w/v SDS was added, mixed gently by inverting the tubes 5 times, placed on ice, 150 µl of

chilled 3 M potassium/ 5 M acetate solution was added and vortexed inverted tubes for 10 sec. Cell debris was pelleted by centrifugation (12,000 g, 5 min, 4°C), the supernatant was transferred to fresh tubes and DNA isolated by phenol/chloroform, ethanol precipitation. The DNA was resuspended in 50 µl TE solution.

2.6.8.2. Large scale preparation

2.6.8.2.1. Maxiprep Method

Plasmid DNA from large-scale bacterial culture (250-500 ml) was prepared using a Magic™ Maxipreps DNA purification system (Promega) following suppliers instructions. A typical yield was 400-500 µg of plasmid DNA. This was the method employed to generate plasmids for routine analysis of recombinant DNA sequences.

Overnight bacterial cultures were decanted into 3 x 150 ml vessels, cells were pelleted by centrifugation (14,000 g, 15 min, 4°C), cells were resuspended in 1 ml of cell resuspension solution (CRS: 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 100 µg/ml RNase) using a wide bore pastette, another 15 ml of CRS and 15 ml of cell lysis solution (0.2 M NaOH, 1% SDS) was added and the solution was mixed gently but thoroughly by inverting several times. The cell suspension should become clear and viscous (can take up to 20 min). 15 ml of neutralising solution (2.55 M KAc, pH 4.8) was added and mixed immediately by inverting several times. The solution was centrifuged (14,000 g, 15 min, 4°C) and the clear supernatant carefully decanted into clean centrifuge bottles. 0.6 volumes of isopropanol was added, the solution was mixed by inversion and centrifuged (14,000 g, 15 min, 4°C) before the supernatant was discarded. The DNA pellet was resuspended in 2 ml of

TE buffer. 1 µl of the DNA solution was checked by electrophoresis in a 1% w/v agarose gel.

2 ml of DNA/TE solution was transferred into a 10 ml universal. 10 ml of Magic Maxiprep DNA Purification Resin (mix well; if necessary, warm the resin to 25-37°C [5-10 min] to dissolve any crystals. N.B. It is important not to use the resin until it is below 30°C) was added and mixed by inverting several times. A Maxicolumn tip was inserted into a vacuum manifold and the DNA/resin mix was transferred into the column. A vacuum was applied to suck the liquid through the column. 13 ml of column wash solution (200 mM NaCl, 20 mM Tris-HCl pH 7.5, 5 mM EDTA) was added, the solution was allowed to pass through the column, another 12 ml of column wash solution was added, the solution was allowed to pass through the column, 5 ml of 80% v/v ethanol was added and the resin was allowed to dry (10 min). The column was removed from the vacuum manifold and placed into a 50 ml universal (reservoir). 1.5 ml of preheated TE buffer was added to the column, after 1 min the column/reservoir was centrifuged (2.5 K, 5 min), the DNA solution was retained in the reservoir and the column was discarded. The DNA solution usually contained some resin. The resin was removed by transferring the solution into an microcentrifuge tube and microcentrifugation (12,000 g, 5 min). The supernatant which contained the dissolved plasmid DNA was removed carefully. The DNA concentration was estimated by OD_{260} or in a gel. The DNA was stored at -20°C.

2.6.8.2.2. CsCl-EtBr gradient method

The caesium chloride (CsCl)-ethidium bromide (EtBr) gradient centrifugation method described by Sambrook *et al.*, (1989) was used to

purify plasmid DNA. A high degree of purity was necessary when generating the riboprobe expression vectors for the tissue *in situ* experiments (see 2.10.1.).

Plasmid DNA was produced using the Maxiprep method described above. Each gradient contained ~100 µg of plasmid DNA, 4.75 g of CsCl, 250 µl of EtBr (stock = 10 mg/ml) in a total volume of 5.6 ml of TE buffer. The solution was mixed gently to dissolve the CsCl and transferred into Beckman Quick-Seal tubes. Enough liquid was added to make each tube weigh 9.7-9.8 g. The tubes were heat-sealed, placed into a NVT 65.2 vertical rotor (Beckmann) and ultracentrifuged (50,000 rpm, 16-18 hr, 20°C, zero deceleration) in an L8-M ultracentrifuge (Beckmann). The lower plasmid band was removed (see 2.9.4) and made up to 5 ml with distilled water. The following steps were performed 2-4 times (EtBr extraction) until the top phase was completely clear [5 ml of butan-1-ol was added, the tubes were inverted several times and centrifuged (2,500 rpm, 3 min, r.t.), the top phase was carefully removed]. The clear solution was transferred into Corex tubes and the DNA was precipitated out of solution (see 2.6.3.) except that 2 volumes of 70% v/v ethanol was used instead of 2 volumes of 100% ethanol (the water in the ethanol prevents the CsCl precipitating). The DNA pellet was resuspended in TE buffer and stored at -20°C.

2.6.9. Preparation and analysis of mouse mammary genomic DNA

High molecular weight genomic DNA was prepared by avoiding vigorous pipetting or vortexing.

Mammary tissue (0.5-1.5 g) was frozen in liquid nitrogen immediately after dissection from the animal. Frozen tissue was pulverised and ground to a fine powder under liquid nitrogen in a pestle and mortar (precooled), the powder was transferred into a 50 ml universal containing 6 ml of lysis buffer (10 mM Tris-HCl

pH 8.5, 400 mM NaCl, 2 mM EDTA, 0.5% w/v SDS), 0.5% w/v SDS and 2 ml of pronase solution (1 mg/ml pronase [Sigma], 1% w/v SDS, 2 mM EDTA), the solution was inverted several times to disaggregate any lumps of powdered tissue and incubated at 55°C overnight with gentle agitation. The solution became clear due to proteinase digestion. 2 ml of 6 M NaCl was added, the solution was mixed well for 15 sec and centrifuged (3,000 rpm, 15 min, r.t.). The supernatant was decanted into a clean tube, phenol/chloroform extractions were performed (see 2.6.2.), 5 ml of isopropanol was added and mixed gently by inverting to precipitate the DNA. DNA is seen as fine "stringy" white material floating in the solution. DNA was hooked out using a pasteur pipette (melted and shaped into a hook), air dried for 5 min, washed twice in 70% v/v ethanol, air dried for 5 min and dissolved in microcentrifuge tubes containing 500 µl of 1/10th TE buffer. The samples were stored at -20°C once the DNA was dissolved (the DNA can take up to 24 hr to go into solution). The DNA concentration was determined by measuring the absorbance at 260 nm (1/100 th dilution of each sample in TE buffer).

10 µg of mouse mammary DNA was restriction enzyme digested for 5 hr at 37°C, a 1/4 volume of loading buffer (40 mM EDTA pH 7.9, 0.1% w/v SDS, 30% w/v ficoll 400, 0.1 mg/ml bromophenol blue) was added and the solution was heated to 60°C for 5 min. The samples were cooled to r.t. and loaded into 0.8% w/v agarose gels. Gels were usually run overnight at 20 V or for 6 hr at 100 V (constant voltage).

2.6.10. Southern blotting

2.6.10.1. Transfer onto Hybond-N membrane

The aim is to transfer the DNA from the gel onto a membrane so that the DNA can be examined for the presence of specific sequences (Southern, 1975). This is achieved by the flow of liquid through the gel and membrane and UV crosslinkage of DNA with a membrane as described by Sambrook *et al.*, (1989).

Gel was soaked in 500 ml of 0.25 M HCl for 2 x 15 min (termed depurination: fragmentation of the DNA to ensure transfer of all DNA, independently of initial size [essential if fragments >8 kb are to be transferred]), briefly in distilled water (to remove excess acid), 500 ml of 0.5 M NaOH/1.5 M NaCl for 2 x 15 min (hydrolysis of the phosphodiester backbone at the site of depurination), briefly in distilled water, 500 ml of 0.5 M Tris-HCl pH 7.4/3 M NaCl for 2 x 15 min (neutralise and hence terminate hydrolysis), briefly in distilled water.

The gel was placed onto a sheet of 3 MM paper (moistened with 20 x SSC) with the ends of the paper resting in a reservoir of 20 x SSC, Saran wrap (Dow) was placed around the edges of the gel so that capillary action could only occur through the gel, a sheet of Hybond-N (Amersham) [cut to the size of the gel and moistened with 2 x SSC] was placed onto the surface of the gel taking care to avoid trapped air bubbles (bubbles block transfer; if they occurred a glass pipette was used to "roll" them out), 2 sheets of 3 MM paper (moistened with 2 x SSC) were layered on top of the Hybond-N, a stack of paper towels was placed on top of the 3 MM paper and finally an ~500 g weight was placed on top of the stack to keep all layers compressed. Transfer was allowed to occur overnight.

After transfer the stack was disassembled down to the Hybond-N membrane, the wells were marked using a soft pencil, the membrane was washed briefly in 2 x SSC and the DNA cross-linked with the membrane by UV irradiation (Stratalinker). The membrane was air-dried and wrapped in cling-film until ready for hybridisation (blots can be stored indefinitely like this but care must be taken not to chip or crack the now fragile membrane).

2.6.10.2. Transfer onto Zeta-Probe GT membrane

The procedure was identical to that described for transfer onto Hybond-N membranes except for the following steps. The acid depurination step was performed using 0.25 M HCl for 10 min. After depurination the gel was rinsed briefly in distilled water (no further treatment of the gel was necessary) before being incorporated into the gel stack as described above except that Zeta-Probe membranes (Bio-Rad) were used instead of Hybond-N membranes and 0.4 M NaOH was used instead of 20 x SSC. Transfer was complete after 4-6 hours.

2.6.11. Preparation of radio-labelled DNA probes and hybridisation to Southern blot membranes

Both methods described below are based on the method described by Feinberg & Vogelstein (1983, 1984). The cDNA strand is synthesised by the Klenow fragment of *E. coli* DNA polymerase I using the 3'-hydroxyl terminus of random oligonucleotides as primers. Modified deooxyribonucleoside-triphosphates present in the reaction are incorporated into the newly synthesised cDNA strand. The first method utilises the Multiprime DNA Labelling Kit

(Amersham); the second method utilises a pre-mixed High Prime solution (Boehringer Mannheim).

2.6.11.1. Multiprime DNA labelling method

10-50 ng of probe DNA was made up to 27 µl with distilled water. The DNA was denatured by placing the microcentrifuge tube in boiling water for 2 min followed by placing the tube on ice for at least 5 min. To this was added:

38.0 µl	distilled water
10.0 µl	Multiprime buffer solution*
5.0 µl	Primer solution‡
2.0 µl	Enzyme solution†
5.0 µl	[α^{32}P]dCTP (Amersham)

* 100 µM of each dATP, dGTP, and dTTP in 0.3 M Tris-HCl pH 7.8, magnesium chloride and 2-mercaptoethanol [Amersham]

‡ random hexanucleotides in an aqueous solution containing nuclease-free BSA [Amersham]

† 1 U/µl DNA polymerase I "Klenow" fragment in 50 mM potassium phosphate pH 6.5, 10 mM 2-mercaptoethanol, and 50% v/v glycerol [Amersham]

This was gently vortexed and microcentrifuged (5 sec pulse) to bring all the liquid to the bottom of the microcentrifuge tube. The reactions were incubated in a 37°C water bath for 30 min.

2.6.11.2. High Prime DNA labelling method

"High Prime" operates on the same principle as the Multiprime labelling system except that all the constituents are within a single reaction mixture. One vial containing 200 µl of random primer mixture contains: 1 U/µl Klenow polymerase, 0.125 mM dATP, 0.125 mM dGTP, 0.125 mM dTTP, 5 x reaction buffer and random oligonucleotides in 50% v/v glycerol (Boehringer Mannheim). The High Prime DNA labelling system therefore reduces pipetting steps which is advantageous in enzymatic reactions.

25 ng of probe DNA was made up to 11 µl with distilled water and denatured (see above) before 4 µl of High Prime solution was added followed by 5 µl of [α^{32}P]dCTP. The samples were incubated at 37°C for 10 min.

2.6.11.3. Quantifying ^{32}P-labelled DNA probes

The following method utilises a trichloroacetic acid (TCA)/filter assay to separate unincorporated [α^{32}P]dCTP's from labelled DNA probes. See Berger & Kimmel (1987) and references therein for a description of Cerenkov radiation detection.

1 µl of the reaction mix was spotted onto a filter (Watmann). Another 1 µl of the reaction mix was added to 1 ml of chilled 10% w/v TCA (to precipitate the labelled macromolecules [probe]), the tube was chilled on ice for 10 min, a clean filter was placed in a vacuum clamp, 1 ml of TCA/probe DNA was pipetted onto the filter within the vacuum clamp, the filter was washed with 10 ml of chilled 10% w/v TCA followed by 5 ml of 100% ethanol (to dry the filter). Each filter was placed in a 20 ml scintillation vial (without scintillation fluid) and the specific activity was measured using a Minaxi Tri-Carb 4000 series (United Technologies Packard)

scintillation counter (programme seven and a 60 sec sample time). Generally >60% incorporation was achieved using this method. If the incorporation was satisfactory 13 µl of 2 M NaOH was added to denature the probe DNA. If incorporation was unsatisfactory another 1 µl of Klenow was added to the sample and the reaction allowed to proceed for a further 15 min before the incorporation was re-checked. Generally the addition of an extra 1 µl of Klenow brought the incorporation up to >60%. However, if this failed, the whole process was repeated. If incorporation was still poor, fresh DNA probe was prepared.

2.6.11.4. Hybridisation and washes

Pre-hybridisation of membranes was performed in revolving glass cylinders using 50 ml of hybridisation solution (0.5 M disodium hydrogen orthophosphate [Na_2HPO_4], 7% w/v SDS, 1 mM EDTA) in a Techne Hybridiser HB-1D oven set at 65°C for 30 min. Hybridisation of probe to membrane-bound material was performed in the same cylinders using all of the denatured probe and 20 ml of hybridisation solution at 65°C overnight.

Membranes were washed to remove unhybridised probe and background signal. First wash was 2 x 20 min in revolving glass cylinders using 50 ml of low stringency wash (2 x SSC, 0.1% SDS) followed by 2 x 20 min using 50 ml of high stringency wash (0.2 x SSC, 0.01% SDS), briefly in 2 x SSC and air dried before being exposed to X-ray film (AGFA; CURIX RP1) between intensifying screens in autoradiography cassettes at -70°C (2 hr - 2 weeks) or in Storage Phosphor Screens (Molecular Dynamics) at r.t. (30 min - 72 hr). X-ray film was developed automatically using an X-ograph Compact x2 and Phosphor Screens were processed using a PhosphorImager (Molecular Dynamics). If necessary, membranes were

stripped by placing them in boiling 0.1% w/v SDS and allowed to cool to r.t. before the membrane was removed. It was then ready for re-probing.

2.6.12. Polymerase chain reaction (PCR) analysis of tail biopsy DNA

A PCR assay was used for rapid screening of mice for the presence of the transgene during the backcross experiments. Primers and methods were essentially as described by Whitelaw *et al.*, (1991). The bench and pipettes were wiped with 100% ethanol and a separate bottle of distilled water was used for PCR reactions to minimise the risk of DNA contamination leading to false positives.

Tail-tip biopsies were taken from 6-8 week old mice and digested overnight at 37°C with proteinase K (200 mg/ml) in 0.3 M sodium acetate, 10 mM Tris-HCl pH 7.9, 1 mM EDTA, 1% w/v SDS. Samples were vortexed before being frozen at -20°C. Tissue debris and SDS were pelleted by microcentrifugation (12,000 g, 15 min, 4°C), 1 µl of the supernatant was pipetted under a drop of mineral oil (Sigma) in PCR microcentrifuge tubes and the DNA denatured by heating to 95°C for 15 min before being cooled to 65°C in a Hybaid Omnigene PCR thermocycler. 50 µl of PCR reaction mix was added to each sample under the oil. The PCR reaction mix comprises the following:

29.75 µl	distilled water
5.0 µl	10 x buffer* (Boehringer Mannheim)
8.0 µl	1.25 mM dNTP mix (Pharmacia)
5.0 µl	dimethyl sulphoxide (DMSO)
0.5 µl	10 mM BLG‡ primer
0.5 µl	10 mM HPRT§ primer
0.25 µl	*Taq* DNA polymerase (Boehringer Mannheim)

* 10 x buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin)

‡ BLG primers : 5'-GCT TCT GGG GTC TAC CAG GAA-3' and 5'-TCG TGC TTC TGA GCT CTG CAG-3' amplify a 246 bp segment of the 5' end of the BLG gene

§ hypoxanthine-guanine phosphoribosytransferase (HPRT) primers : 5'-GAG TTC CGG AAC TGC CTT TGG TG-3' and 5'-CTG TGC CAC CGG GCG CAT GG-3' amplify a 332 bp segment of the HPRT gene

Samples were placed in the PCR thermocycler and given 30 alternating cycles at 92°C for 30 sec and 65°C for 5 min. Once the reaction was complete (3-4 hr), 10 µl of PCR loading dye (15% w/v Ficoll 400, 0.05% w/v SDS, 20 mM EDTA pH 8.0, 0.125% w/v orange G) was added to each sample, mixed and 30 µl electrophoresed in 2% w/v agarose gels† (20 x 20 cm) containing 0.5 µg/ml EtBr at 100 V for 30 min. Control positive, negative and blank samples were used in every gel.

Table 7. Range of separation for agarose gels

Agarose (%)	†Optimal range of separation linear DNA (kb)
0.3	60 - 5.0
0.6	20 - 1.0
0.7	10 - 0.8
0.9	7 - 0.5
1.2	6 - 0.4
1.5	4 - 0.2
2.0	3 - 0.1

Taken from Sambrook *et al.*, 1989.

2.7. RNA ANALYSIS

2.7.1. Preparation of total RNA from mouse mammary tissue

All glassware, pipettes, tips and solutions used for RNA analysis were treated with diethyl pyrocarbonate (DEPC) to destroy ribonucleases and gloves were changed frequently (Berger & Kimmel, 1987; Sambrook *et al.*, 1989).

Mammary tissue (0.5-1.5 g) was frozen in liquid nitrogen immediately after dissection from the animal. Frozen tissue was homogenised in 2 ml RNAzol B (Biogenesis Ltd) for 30 sec using an ultra-homogeniser (Janke and Kunkel). 2 x 1 ml aliquots were added to 2 ml screw-capped tubes containing 100 µl chloroform. These were inverted repeatedly for 15 sec and put in ice for 15 min, microcentrifuged (12,000 g, 15 min, 4°C), the aqueous phase transferred into a clean microcentrifuge and the RNA precipitated using an equal volume of isopropanol (4°C for 15 min). The samples were microcentrifuged (12,000 g, 15 min, 4°C) to pellet the RNA. The pellets were washed twice with 75% v/v ethanol (microcentrifuge [12,000 g, 8 min, 4°C] between washes). The RNA pellet was air dried for 10 min and resuspended in 200 µl of DEPC-treated distilled water (stored at -70°C) or in 200 µl of 100% formamide (stored at -20°C; Chomczynski, 1992). The RNA concentration was determined by measuring the absorbance at 260 nm of a 1/100 th dilution of each sample in TE buffer.

2.7.2. Electrophoresis of RNA, Northern blotting and hybridisation

Messenger-RNAs have secondary structures. Therefore denaturing gels are used for RNA electrophoresis (Berger & Kimmel 1987, Sambrook *et al.*, 1989).

10 µg of total RNA per sample was made up to 8 µl with distilled water and added to 38 µl of sample buffer (25 µl formamide, 8 µl formaldehyde and 5 µl 10 x MOPS). Samples were incubated at 65°C for 5 min and on ice for 5 min before 5 µl of running dye (50% v/v glycerol, 0.1 mg/ml bromophenol blue) and 0.5 µl of EtBr (10 mg/ml) was added to each. The samples were electrophoresed in denaturing 1.5% w/v agarose gels containing 1 x MOPS buffer and 6.8% v/v formaldehyde using 1 x MOPS as the gel tank buffer. Northern blots were performed as described in section 2.6.10. except that there was no pre-treatment of the gel before blotting onto Hybond-N (Amersham) membranes. Preparation of probes and hybridisation conditions were identical to those described in section 2.6.11..

2.8. PREPARATION OF AGAR PLATES

2.8.1. Plates containing antibiotic

LB bottom medium (1% w/v tryptone, 0.5% w/v yeast extract, 1.5% w/v agar, 0.1 M NaCl) was melted in a microwave and allowed to cool to "hand-hot". 1 ml of the appropriate antibiotic (e.g., ampicillin; stock = 50 mg/ml) was added to 500 ml LB bottom (final concentration of 100 µg/ml), swirl mixed and poured immediately into petri dishes with a bunsen flame close by to create a convection current. Once the plates had set they were placed in a 37°C oven, upside down, with the lids half off, for 1 hr to allow excess moisture to evaporate off.

Plates were stored at 4°C until ready for use (shelf life of ~2 weeks) or used immediately.

2.8.2. Plates for blue/white screening

Blue/white screening is a method for selecting recombinant plasmids that contain a sequence coding for the *lac*Z α-peptide, interrupted by a multiple cloning site. Nonrecombinant plasmids produce a functional α-peptide which, by complementing the product of the host cell *lac*Z[ΔM15] gene, leads to the production of functional β-galactosidase. Bacterial colonies harbouring the *lac*Z[ΔM15] gene on an F' factor and also containing a pGEM-4Z vector (see 2.10.1.) develop a blue colour when plated on indicator media containing IPTG and X-gal. However, when the *lac* α-peptide is disrupted by cloning into the pGEM-4Z multiple cloning region (Fig. 7), the *lac* α-peptide is usually not produced and no β-galactosidase activity is produced. Therefore, bacterial colonies harbouring recombinant pGEM-4Z vector constructs are white.

Agar plates were prepared as in section 2.8.1.. 70 µl of X-gal (stock = 20 mg/ml in dimethyl formamide), 20 µl of IPTG (stock = 0.1 M) and 50-100 µl of transformed cells were pipetted onto the surface of the plates and the liquid was spread immediately over the surface of the agar using a sterile glass rod. The plates were incubated in an oven (37°C) overnight.

2.9. PREPARATION OF LAMBDA PHAGE

Two high titre lysates containing genomic phage clones of the mouse *c-myc* and *Pvt-1* loci were a gift from Dr Konrad E. Huppi, Molecular Genetics

Section, Laboratory of Genetics, National Cancer Institute, National Institutes of Health, Bethesda, U.S.A..

2.9.1. XL-1 Blue MRA cells

E.coli XL-1 Blue MRA cells [Genotype: Δ(*mcrA*)183, Δ(*mcr*CB-hsdSMR-*mrr*)173, endA1, supE44, thi-1, gyrA96, recA1,lac] were purchased from Stratagene. 16 OD_{600} units (16 OD_{600} units = an OD_{600} of 0.5 in 32 ml of inoculated medium at 600 nm wavelength) of XL-1 Blue MRA cells were used in LBM medium (1% w/v tryptone, 0.5% w/v yeast-extract, 0.5 M NaCl, 10 mM $MgCl_2$) supplemented with 0.2% maltose (3-4 hr). Cells were harvested by centrifugation (5,000 g, 10 min, 4°C), the supernatant was collected, 16 ml of chilled 10 mM $MgSO_4$ was added, the tube was flicked (not vortexed) until all the cells were resuspended, the sample was centrifuged (5,000 g, 10 min, 4°C), the supernatant was collected, 8 ml of chilled 10 mM $MgSO_4$ (two washes with $MgSO_4$ because LBM medium is inhibitory to phage infection) were added, the cells were resuspended and the sample was stored at 4°C overnight. The cells were then ready to be used to estimate the titre of phage stocks, preparation of mini-stocks and large scale preparation of phage DNA (see below)

2.9.2. Estimation of titre of phage stocks

1 µl of phage stock was diluted 1,000-fold in SM phage buffer (0.1 M NaCl, 10 mM $MgSO_4$, 20 mM Tris-HCl pH 7.5, 2% w/v gelatin), 1 µl of the dilution was added to 200 µl of XL-1 Blue MRA cells which were prepared the previous day (see 2.9.1.), the cells were incubated at 37°C for 20 min (to allow phage time to attach to cells), 3 ml of molten (42°C) LBM-top agar (LBM with 7%

w/v agar) was added and the mixture was poured immediately onto LBM bottom-agar (LBM with 15% w/v agar) plates (15 ml of molten agar poured into 90 mm plates). The plates were incubated at 37°C overnight. Lysis of cells causes plaque formation in a lawn of non-lysed cells. The number of plaques were counted and this number was multiplied by 10^6 to give the number of plaque forming units (pfu) in 1 ml of phage stock. A high titre is 10^8-10^{10} pfu/ml.

2.9.3. Preparation of a mini-stock of high titre phage

One phage plaque was removed from a plate (as a plug, using a blue tip with the end removed). The plug was placed into 500 µl of SM buffer for 2 hr at r.t. (allowed the phage particles to diffuse into the buffer), 200 µl of the SM buffer was added to an overnight preparation of cells (see 2.9.1.) which were incubated at 37°C for 20 min and poured onto agar plates (see 2.8. [without antibiotic, IPTG or X-gal]). Confluent lysis occured after an overnight incubation at 37°C due to the high concentration of phage used. The plate was chilled to 4°C, 5 ml of chilled SM buffer added to the surface of the plate, phage particles were allowed to diffuse into the buffer for 4 hr at 4°C with gentle shaking. The SM buffer was collected and mixed with 100 µl of chloroform and centrifuged (4,000 g, 10 min, 4°C). The supernatant was removed and mixed with chloroform (0.3% final concentration) and stored at 4°C until required. The titre was usually around 10^{10} pfu/ml.

2.9.4. Large-scale preparation of phage DNA

8 ml of cells in SM buffer (prepared previous day [see 2.9.1.]) were inoculated with 1 µl (1 x 10^7-10^{10} pfu/µl; this value is dependent on the type of

phage) of high titre phage lysate. The cells and phage were incubated at 37°C for 20 min. This was added to 500 ml of LMB medium (supplemented with 10 mM $MgSO_4$ but without maltose and already warmed to 37°C) and incubated at 37°C in a shaking incubator until lysis occured. The aim was to achieve an OD_{600} = 1.3 (5-6 hr) before it crashed down to an OD_{600} = 0.3. The progress of cell growth was followed by measuring the OD_{600} of a 1 ml sample from the culture every half hour. Lysis occurred when large lumps of bacterial debris were visible in the medium.

5 ml of chloroform were added to the flask for 30 min at 37°C (ruptured any cells that had not already lysed), 3 mg of DNase I and 3 mg of RNase were added and the solution was allowed to stand at r.t. for 40 min. The samples were centrifuged to remove cell debris (5,500 g, 10 min, 4°C), the supernatant was decanted into a 1 l flask, 50 g of PEG_{6000} and 29.22 g NaCl (to precipitate the phage) was added, the solution was swirled to dissolve the PEG and salt. The mix was left at 4°C overnight. The phage were pelleted by centrifugation (5,500 g, 15 min, 4°C), the supernatant was discarded, 8 ml of SM buffer was used to resuspend the phage pellet, 4 ml of chloroform (to extract any protein) was added, the solution was inverted to mix and centrifuged (4,000 rpm, 5 min). The supernatant (contains phage) was transferred into a fresh tube.

CsCl gradients (1.5 ml of 56% CsCl solution [25.6 g of CsCl dissolved in 20 ml SM buffer and filtered through a 0.45 μm filter], 1.5 ml of 45%, and 2.5 ml of 31%) were prepared in a 14 x 95 mm clear centrifuge tubes, the interface between the 56% and 45% bands was marked with an indelible marker (this is where the phage band appears), gently added the phage solution on top of the CsCl, the tubes were placed into a Beckman SW 28 rotor (balanced) and centrifuged (23,000 rpm, 135 min, 4°C, with zero deceleration) in a Beckman ultracentrifuge to "band" the phage particles (appear blue). Once centrifugation was complete the tubes were carefully removed and clamped in front of a black background. The blue band was

removed (ignored upper bands which are partially degraded phage heads etc.) using a 19 gauge needle and syringe. The volume removed was diluted with an equal volume of SM buffer, the phage were dialysed against a large excess (at least a 1000-fold volume) of 50 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$, 10 mM NaCl for 1 hr with several changes of dialysis buffer (to remove the CsCl), the phage heads were disrupted in 20 μg/ml proteinase K, 0.5% w/v SDS, 20 mM EDTA (final concentration) at 37°C for 45 min, the phage DNA was gently extracted with 2 x phenol/chloroform and 2 x chloroform (see 2.6.2.), and precipitated with 1 volume of isopropanol and a 1/10 th volume of 3 M NaAc, pH 4.5 at -20°C for 1 hr. After centrifugation, the DNA pellet was air dried and resuspended in TE buffer.

2.10. TISSUE *IN SITU* HYBRIDISATION

Preparation of mouse mammary tissue (fixing and sectioning) was performed at The Roslin Institute and the tissue *in situ* hybridisation experiments were performed at the MRC Human Genetics Unit (HGU), Western General Hospital, Crewe Road, Edinburgh, U.K. in collaboration with Dr Duncan Davidson and Ms Elizabeth Graham using methods described by Monaghan *et al.*, (1991) and Angerer & Angerer (1992), and references therein.

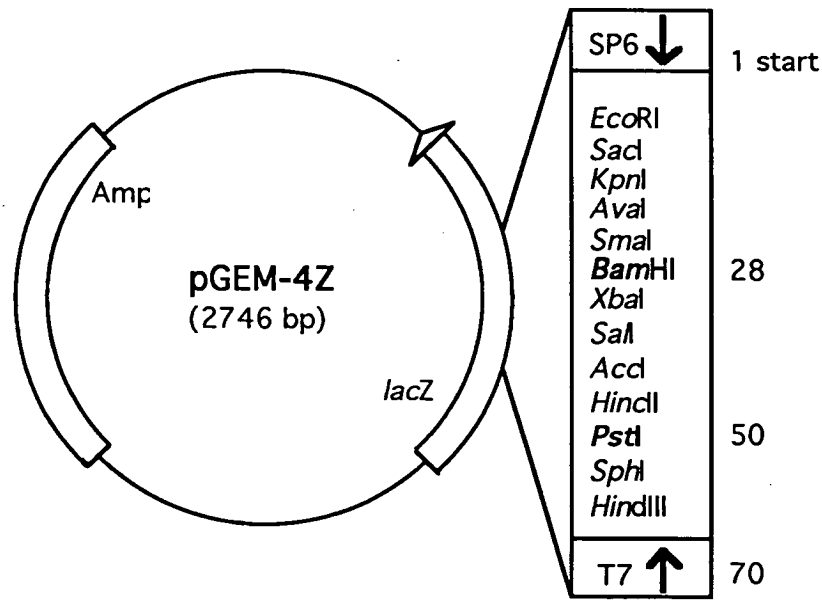
2.10.1. Construction of BLG expression vectors

Riboprobes of between 200 bp and 1 kb are recommended for tissue *in situ* hybridisation experiments utilising ^{35}S-labelled probes and paraffin embedded tissue. Probes <200 bp give a reduced hybridisation signal and probes >1 kb penetrate less effectively into fixed tissue. Riboprobes of ~400 bp were designed because of convenient *Pst*I restriction sites in both the DNA probe templates and as

a compromise between probe penetration and signal. This strategy also negated alkaline hydrolysis to alter the length of the probes.

A 424 bp *Pst*I BLG cDNA fragment (originally from pβLac-931: Gaye *et al.*, 1986) comprising all of exons 3, 4, 5, 6 and 6 bp of exon 2 and 81 bp of exon 7 was subcloned from pGEM-3 into a *Pst*I site in pGEM-4Z (Fig. 7: Promega) in both orientations.

Figure 7. pGEM-4Z vector circle map



Cloning the inserts in both orientations enabled synthesis of anti-sense (probe) and sense (control for non-specific background) riboprobes from the same promoter. The pGEM-4Z vector enabled ampicillin selection of transformed cells, blue/white screening for recombinants and a convenient cloning site close to a T7 RNA polymerase promoter that enabled highly efficient synthesis of mRNA *in vitro*. An off-centre internal *Pvu*II site was used to check the orientation of the 424 bp *Pst*I fragment. It is important to have very clean expression vector preparations for *in*

vitro mRNA synthesis (aids precursor incorporation). Therefore large quantities of plasmid DNA was prepared by the method described in section 2.6.8.1.1., followed by CsCl-EtBr gradient centrifugation (see 2.6.8.1.2.). Expression vectors were linearized using the *Bam*HI site (a site just downstream of the inserts and the T7 promoter) so that the transcripts were free from vector sequences (wastes isotope and increases the chances of background hybridisation), phenol chloroform extracted, ethanol precipitated and stored in DEPC-treated distilled water at -20°C.

2.10.2. Construction of β-casein expression vectors

A 417 bp *Pst*I β-casein cDNA fragment (originally from pCMβ13; Rosen, J., unpublished; Gupta *et al.*, 1982) comprising 5 bp of exon 6 and 412 bp of exon 7 was subcloned from pBR322 (a gift from Dr Bert Binas; current address : Max-Delbruck Centre, Berlin, Germany.) into a *Pst*I site in pGEM-4Z in both orientations. An off-centre internal *Nde*I site was used to check the orientation the 417 bp *Pst*I fragment.

2.10.3. Coating slides with TESPA

The tissue sections need to be fixed onto the slides, otherwise they float off during processing. This is achieved by coating the slides with TESPA (3-aminopropyltriethoxysilane [Sigma]) which acts as a tissue adhesive. The hydrophobic silane moiety of TESPA binds to glass, while its amino group, activated by aldehyde, is crosslinked to tissue.

Slides were generally processed in batches of 10 or 20 with the slides lying horizontal in glass racks and subsequent washes and incubations performed in 200 ml glass troughs. Batches of slides were washed in 10% v/v HCl, 75% v/v ethanol

for 10 sec, distilled water for 10 sec, 100% acetone for 10 sec and air-dried. The slides were coated with TESPA by immersing them in a solution of 2% v/v TESPA in acetone for 10 sec before two 10 sec washes in 100% acetone, distilled water for 10 sec and air-dried. Slides were stored in a slide box with silica gel.

2.10.4. Preparation of tissue

Tissue was fixed in 4% w/v paraformaldehyde (PFA) in PBS, pH 7.2 overnight at 4°C with gentle agitation and washed in PBS for 30 min at 4°C with gentle agitation. PFA is a cross-linking fixative used to retain mRNA in sections during hybridisation and wash procedures. This fixative proved to be adequate for tissue fixation and preservation of morphology and yet enabled sufficient penetration of probe to target mRNA (methods that retain RNA and preserve morphology are usually not the same as those that afford high hybridisation efficiency).

The following procedures were performed at room temperature with gentle agitation. The tissue was dehydrated in 30% ethanol for 15 min, 2 x 30 min in 70% ethanol, 2 x 30 min in 100% ethanol and 60 min in fresh 100% ethanol. The ethanol within the tissue was substituted with CNP 30 (Pentone) by performing a set of washes in CNP 30 for 2 x 30 min and overnight in fresh CNP 30. The CNP 30 was subsequently substituted with paraffin (freshly melted; it deteriorates when stored in the melted form) by performing two washes in molten paraffin wax for 1 hr at 60°C (i.e., CNP 30 acts an intermediate solvent). The tissue was finally embedded in a mould containing fresh molten paraffin which was allowed to set. These tissue wax blocks were stored at 4°C (tissue can be kept in this way for at least several years). Paraffin embedding was used because it is simple, cheap and generally preserves tissue morphology better than frozen sections (mouse embryos).

2.10.5. Tissue sectioning

The paraffin blocks with the embedded tissue were chilled to 4°C. All instruments that would come into contact with the tissue including the rotary microtome (Leitz, Wetzlar), steel blades (Tissue-Tek III), manipulation utensils and empty histological water bath were wiped down with 100% ethanol. The water bath was filled with DEPC-treated distilled water and heated to 55°C. 4 µm sections were cut to generate a "string" of abutting sections which were floated onto the water. The paraffin wax was allowed to melt slightly (~1 min) so that the tissue sections expanded to their original dimensions. Tissue sections were arranged onto slides in sets of abutting sections, usually four per slide. The slides were put into an oven (60°C) to dry overnight. Slides were stored in a slide box with some silica gel.

2.10.6. Pre-hybridisation

The aim of this step is to treat the tissue with proteinase K to partially remove proteins in order to increase target RNA accessibility to probes.

All prehybridisation dishes were washed with 100% ethanol and each step was performed with gentle agitation. The slides were placed in a slide rack and treated with the following solutions : 100% xylene for 2 x 5 min (in fume hood) to remove paraffin from tissue (residual paraffin traps probe), through 100%, 90%, 70%, 50% and 30% v/v ethanol solutions for 2 min each to rehydrate the tissue, PBS for 2 min, 4% w/v PFA in PBS pH 7.2 for 10 min, PBS for 2 x 2 min, proteinase K solution for 7.5 min (50 mM Tris, 5 mM EDTA, 5 µg/ml proteinase K type XXVIII [Sigma]), PBS for 1 min, 4% w/v PFA in PBS pH 7.2, distilled water for 10 sec, 0.1 M TEA pH 8.0, 200 ml of 0.1 M TEA with 625 µl acetic anhydride for 2 x 5 min

(used to decrease positive charge on the sections [reduce background] and help inactivate proteinase K), PBS for 2 min, 0.85% NaCl for 2 min, through 30%, 50%, 70% and 90% v/v ethanol solutions for 2 min each and 100% ethanol for 5 min each to dehydrate the tissue. Finally the slides were air-dried and stored in a box with silica gel.

2.10.7. Generation of ^{35}S rUTP labelled riboprobes

^{35}S rUTP-labelled single-stranded antisense RNA probes and control sense RNA probes were synthesised by *in vitro* transcription from the T7 promoter essentially as described by Melton *et al.*, (1984). This type of probe is easy to generate and because they form a 1:1 duplex with target RNA, they provide a consistent and quantifiable measure of temporal and spatial gene expression. The following mix was prepared in microcentrifuge tubes at room temperature:

3.0 µl	10 x buffer (NEB)
1.0 µl each	10 mM rATP, 10 mM rCTP, 10 mM rGTP (Pharm.)
1.0 µl	1 M DTT
3.0 µl	distilled water
12.0 µl	^{35}S rUTP (> 1 mCi / 100 ml) [Amersham]
5.0 µl	0.5 - 1 mg/5 ml DNA template
1.2 µl	RNase block (NBC)
0.8 µl	T7 polymerase (NBC)

The samples were vortexed briefly and pulse-centrifuged before incubating at 37°C for 25 min. Another 0.8 µl of T7 polymerase was added and incubated for a further 25 min at 37°C. 2 µl of tRNA (stock = 10 mg/ml, Sigma) and 1 µl of DNase

(stock = 10 mg/ml, NEB) was added and incubated at 37°C for 10 min. 2 µl of 100 mM EDTA and 164 µl of 50 mM DTT, TE buffer were added to give a final volume of 200 µl. The riboprobes were purified by phenol/chloroform extraction and precipitated overnight with ethanol (see 2.6.2. and 2.6.3.).

The precipitated riboprobes were pelleted by microcentrifugation (12,000 g, 30 min, r.t.). Meanwhile the following solution was prepared to wash the pellets with ; 100 µl of 1 M DTT, 1.9 ml TE buffer and 8 ml of 100% ethanol. RNA pellets were washed twice with 500 µl of the wash solution then once with 100% v/v ethanol, allowed to air-dry for 10 min and resuspended in 25 µl of 50 mM DTT, TE buffer.

2.10.8. Quantifying ^{35}S rUTP labelled riboprobes

1 µl of the probe was added to 19 µl of distilled water and 1 µl spotted onto two 2.5 cm glass microfibre filters (Whatman). One filter was washed three times with chilled 5% w/v TCA and once with chilled 100% ethanol. The percentage incorporation was measured from the two filters using the counts per minute (CPM)B values. Converted CPM values to disintegrations per minute (DPM) by assuming 50% counting efficiency; therefore multiplied CPM values by two to arrive at the DPM values. An incorporation of 1.0×10^5 dpm per 1 µl of probe was the target level of activity. To achieve this value the probe was diluted with hybridisation solution. The hybridisation solution was prepared at r.t. and comprises the following:

50%	formamide (deionized) : to reduce the hybridisation temperature
10%	dextran sulphate : increases hybridisation rate
1 x	Denhardt's solution : reduces non-specific binding
20 mM	Tris-HCl, pH 8.0
0.3 M	NaCl : high salt to reduce electrostatic binding of probe to tissue
5 mM	EDTA
10 mM	sodium phosphate
0.5 mg/ml	tRNA : carrier
50 mM	DTT : prevents probe from oxidising and inhibits RNase

The final solution also contained a 1/10 th volume of 50 mM DTT, TE buffer (N.B. It was necessary to calculate the total volume of hybridisation mix required for all of the probes before the solution was prepared. For example, four probes required about 5 ml of hybridisation solution given an average incorporation value). Probes were stored at -20°C in a box clearly labelled "radioactive" and used within 48 hr of preparation.

2.10.9. Hybridisation

The probes were denatured at 80°C for 2 min, cooled on ice and 10 µl of probe added directly onto each desiccated tissue section. An appropriately sized cover-slip was placed over the slides and the slides placed into a hybridisation box with the cover-slips facing upwards. Some tissue paper soaked in 50% v/v formamide, 5 x SSC was placed into a slide box which was then sealed with tape. The box was placed into two water-tight plastic bags. The water-tight slide box was submerged in a water bath (55°C) overnight.

2.10.10. Post-hybridisation

The hybridisation box was removed from the water bath and plastic bags and forceps used to transfer the slides into a jar containing 5 x SSC, 10 mM DTT at 55°C for 30 min. The cover-slips were removed at this stage. The slides were transferred into HS wash (5 ml 1 M DTT and 100 ml 50% v/v formamide, 2 x SSC, 0.1 M DTT) for 30 min at 65-68°C (the temperature affects the hybridisation and therefore is critical for controlling hybridised probe to background ratio), NTE wash (0.2 M NaCl, 10 mM Tris-HCl, 50 mM EDTA) for 2 x 5 min at 37°C, NTE wash with 20 μg/ml RNase for 30 min at 37°C (the RNase digests any non-hybridised probe [single-stranded], therefore reducing background), NTE wash for 5 min at 37°C, HS wash for 30 min at 65-68°C. The remaining washes were performed at room temperature. 2 x SSC for 4 x 10 min, 0.1 x SSC for 4 x 5 min, 30%, 50%, 70% and 90% v/v ethanol (each with 0.3 M ammonium acetate) for 1 min each, 100% ethanol for 2 x 5 min and air-dried.

2.10.11. Coating slides with emulsion

This step was performed in a darkroom with a safety light on. 10 ml of Ilford K5 emulsion was melted with 10 ml of distilled water at 42°C (melting takes 5-10 min). Once the "plastic" emulsion had liquefied it was poured into a vessel designed for dipping slides. Slides were dipped into the molten emulsion and arranged in a light-proof box on their ends to allow excess emulsion to run off and enable even drying of the remaining emulsion on the slides. The box was sealed from light and stored for 3 hr to allow the emulsion to dry. Once dry, the slides were arranged in a slide box, wrapped in foil and stored at 4°C for 2-6 weeks.

2.10.12. Development of emulsion-coated slides

The slides were allowed to reach room temperature before proceeding (3 hr). The slides were unwrapped in the dark room with the safe-light on and immersed in Kodak developer for 4 min, distilled water for 10 sec, AMP FIX for 5 min and distilled water for 20 min. The slides were carefully wiped to remove any remaining emulsion and dipped in 1% w/v methylene blue for 15-30 sec (to lightly stain the mammary tissue) and rinsed under a tap and air dried (do not dry in an oven as this melts the emulsion). It is important that the slides are dry before proceeding. Slides were mounted using DPX and clean cover-slips (any residual liquid on the slides crystallises in the DPX).

2.10.13. Photography of tissue *in situ* slides

Tissue *in situ* slides were examined using bright-field illumination on an Olympus BH-2 microscope fitted with neutral density filters 25 and 6, an Olympus BH2-SC swing-out condenser and an SPlan x4 objective. An Olympus U-FT focusing telescope aided sharp focusing of the images which were recorded using an Olympus C-35AD-4 camera and Kodak Extachrome 64T film.

2.11. CHROMOSOME FLUORESCENCE *IN SITU* HYBRIDISATION (FISH)

Chromosome fluorescence *in situ* hybridisation experiments were performed at the MRC HGU, Western General Hospital, Crewe Road, Edinburgh, U.K. with guidance from Ms Muriel Lee and Ms Judith Fantes using methods essentially as described in Fantes *et al.*, (1995) and references therein.

2.11.1. Preparation of mouse lymphocyte cells

Homozygous and hemizygous transgenic mice were sacrificed, their spleens removed and placed in petri dishes containing RPMI 1640 medium supplemented with 10% v/v foetal calf serum (FCS). Cells were disaggregated by puncturing the spleen using 26-gauge needles and medium forced through the tissue using a syringe, taking care not to take medium containing cells up into the syringe as this damages the cells. Cells were transferred into universals, centrifuged (600 g, 6 min), supernatants removed and cell pellets resuspended in 6-10 ml of fresh medium. Cells were stained by mixing 250 µl of the cell suspension with an equal volume of methylene blue. To calculate the cell density 7 µl of the stained cell suspension was placed in a haemocytometer (Weber Scientific International Ltd). The haemocytometer has a chamber with a height of 0.1 mm and an etched grid of 1 mm^2 (subdivided into 625 smaller squares). The volume defined by the grid was 1 x 10^{-4} ml. To calculate the number of cells per ml the total number of cells over the grid was multiplied by 2 x 10^4. Depending on the cell density the preparation was diluted or concentrated to a density of 1 x 10^7 cells/ml in RPMI medium, 10% v/v FCS, 50 µg/ml lipopolysaccaride. 5 ml aliquots were incubated at 37°C for 44-46 hr. After this period the cells were arrested at metaphase by adding 100 ng/ml colcemid (Gibco BRL), a microtubule poison, and a further incubation at 37°C for 45 min. Cells were harvested by centrigugation (600 g, 5 min). 10 ml of 75 mM KCl, a mildly hypotonic solution that swells lymphocyte cells and lyses erythrocytes, was slowly added with agitation to resuspend the cell pellet. After 10 min the cells were centriguged (600 g, 5 min), the supernatant removed and the cells fixed with a 3:1 mix of methanol:glacial acetic acid prepared fresh each time. Cells were left overnight in fixative followed by three more washes in fixative. Cells could be stored at -20°C in 8 ml of fixative at this stage.

2.11.2. Preparation of slides

Good quality chromosome spreads are essential for a successful FISH experiment. This stage in the protocol is the most difficult to control.

Glass slides were soaked in a dilute solution of HCl in ethanol and wiped with muslin to ensure clean slides. Fixed cells were removed from storage at -20°C and allowed to warm to r.t. before centrifugation (600 g, 5 min). The supernatant was removed and cells resuspended in 0.5-2 ml of fresh fixative depending on the quantity of cells (one can judge the relative cell numbers by how opaque the solution appears). The cells were taken up into a narrow-mouthed pastette and dropped onto clean slides from a height of about 30 cm in a humidity of >50% (often aided by steam from a boiling kettle). The spreading of the dropped liquid on the surface of the slide was aided by breathing first on the slide to coat it with a thin layer of moisture and gently blowing on the slide immediately after the drop had landed. This process results in metaphase chromosome spreads of differing quality. The best slides were selected by examining the chromosomes using a phase contrast microscope (Leitz Orthoplan). The region of the slide containing the highest density of cells was marked on the edge of the slide using a diamond pencil. The slides were put in a vacuum dessicator for 48-72 hr to allow the chromosomes to "age" before any further work was performed with them.

2.11.3. Giemsa banding (G bands)

Slides were placed in a coplin jar containing pre-heated 2 x SSC for 1 hr at 60°C, 0.5% trypsin (DIFCO) in PBS for 5 sec, 4% w/v Giemsa's

R66 stain (BDH) in 1 M Tris-HCl, pH 6.8 for 20 min, washed with distilled water, air-dried and mounted in DPX (BDH).

2.11.4. Karyotyping

G-banded chromosomes were visualised using a Leitz Ortholux microscope fitted with a CCD camera. Chromosomes were "digitised" and partially karyotyped using the MRC HGU Fip-2 (KMR programme) computer assisted chromosome analysis system (Piper & Lundsteen 1987; Graham & Piper 1994), visualised on a Digivision screen and images printed using a Mitsubishi video copy processor. The partial karyotypes were completed using >15 good metaphase chromosome spreads (for each line) with assistance from Ms Muriel Lee.

2.11.5. Generation of biotin and digoxigenin labelled DNA probes

DNA probes were labelled using a nick translation method essentially as described by Langer-Safer *et al.*, (1982). Nick translation is used to label DNA fragments of >1 kb. The process utilises DNase I to create single-strand nicks in double-stranded DNA. The 5'-3' exonuclease and 5'-3' polymerase actions of *E. coli* DNA polymerase I remove stretches of single-stranded DNA starting at the nicks and replaces them with new strands. In doing so, labelled dNTPs are incorporated into the probe DNA (can be insert, plasmid, lambda, cosmid or DNA from a PCR reaction). The nick translation reaction comprises the following:

2.0 µl	10 x nick translation salts*
2.5 µl each	0.5 mM dATP, dCTP, dGTP
2.5 µl	0.5 mM bio-16-dUTP or (1.0 µl dig-11-dUTP + 1.5 µl 0.5 mM dTTP)
1.0 µl	DNase (stock = 2 µg/ml‡)
6.0 µl	Probe DNA§ (500 ng, diluted to 6 µl with ddH_2O)
1.0 µl	DNA polymerase I (10U/µl)

* nick translation salts (0.5 M Tris-HCl, pH 7.5, 0.1 M Mg SO_4, 1 mM DTT, 500 µg/ml BSA fraction V) (Boehringer Mannheim)

‡ DNase I (1 mg/ml in glycerol) was diluted to 2 µg/ml in activation buffer (10 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$, 1 mg/ml BSA fraction V) at 4°C

§ genomic ovine BLG clone pSS1tgXS (Simon *et al.*, 1987) or mouse major satellite DNA (Pietras *et al.*, 1983) or minor mouse satellite DNA or mouse genomic probes for *Pvt-1* or *c-myc*

The mixture was gently vortexed and microcentrifuged (5 sec) to concentrate the solution at the base of the microcentrifuge tube. The samples were incubated at 15°C for 90 min. The reaction was stopped by adding 2 µl of 0.2 M EDTA and 1 µl of 5% v/v SDS.

Labelled DNA probes were separated from unincorporated dNTPs using Sephadex G50 columns (Pharmacia). Each column was inverted several times to resuspend the gel and placed upright to allow the gel to settle (10 min). The top and bottom caps were removed and the column was allowed to drain. The column was washed with 3 x 1 ml of TE buffer and centrifuged (800 g, 4 min) to drain excess buffer. The labelled DNA was added to the top of the column (column is placed in a universal to collect the eluate), centrifuged (800 g, 4 min) and the eluate collected

(75-100 µl). This contains the labelled DNA probe and is ready for use. The probes were stored at -20°C. Labelled probes can be stored for years at -20°C.

2.11.6. Quantitation of label incorporation

Labelled probes were diluted 1 x 10^{-3} and 1 x 10^{-4} in TE buffer. 1 µl of each dilution was spotted twice onto a grided nitrocellulose filters (Schleicher & Schuell) [filters were pre-washed in distilled water for 2 min followed by a 20 min wash in 20 x SSC and air dried. This preliminary wash in high salt helps to keep the dots small. The filters can be stored at r.t. for several months], each spot was allowed to dry and a further 1 µl of each dilution was spotted onto just one of the previous spots to give labelled probe ratios of 20:10:2:1. On the same membrane 20, 10, 2, and 1 pg of labelled lambda DNA (Boehringer Mannheim) was spotted as a comparison. The spots were air-dried and the probes were cross-linked to the membrane by exposure to 30 mJ of UV irradiation.

Filters were soaked with gentle shaking at r.t. for 5 min in 100 ml of buffer 1 (0.1 M Tris, 0.15 M NaCl, pH 7.5), 60°C for 40 min in 100 ml of buffer 2 (0.1 M Tris, 0.15 M NaCl, 3 % w/v BSA fraction V, pH 7.5), r.t. for 15 min in 10 ml of buffer 1 plus 10 µl of streptavidin alkaline phosphatase (Gibco BRL) for biotin detection or with 10 µl of anti-digoxigenin alkaline phosphatase (Boehringer Mannheim) for digoxigenin detection (both can be done together if there are two probes on the same filter with different labels). Filters were washed for 2 x 10 min in 200 ml of buffer 1 and in 100 ml of 0.1 M Tris-HCl, pH 9.5 for 5 min. The filters were sealed in polythene bags containing 5 ml of 0.1 M Tris-HCl, pH 9.5 and 2 drops each of solutions 1, 2, and 3 from the BCIP/NBT alkaline phosphatase substrate kit IV (Vector). The bags were stored in the dark for 1 hr. This process results in a colour reaction that is proportional to the amount of labelled probe DNA

on the filter and allows a comparison with the lambda standards on the same membrane. For example, a matching set of spots indicated a probe concentration of 10 μg/ml.

2.11.7. Pre-hybridisation

All further treatments of slides were performed in glass racks and immersion in 200 ml of the appropriate solution (with gentle agitation) in glass troughs at r.t. unless otherwise stated. Slides were immersed in 100 μg/ml RNase in 2 x SSC for 1 hr at 37°C (this step is only necessary when the slides have not already been through Geimsa staining), briefly washed in 2 x SSC, through 70%, 90%, and 100% ethanol (freshly prepared) for 2 min each to dehydrate the chromosomes and dried under vacuum. Slides were transferred into a pre-warmed metal rack and put in an oven set at 70°C for 5 min. Slides were immediately denatured in 70% v/v formamide, 2 x SSC for 3 min at 70°C (2 min if the slides have been Geimsa stained), quickly transferred into ice-cold 70% v/v ethanol, into 90% and 100% ethanol solutions for 2 min each to dehydrate the chromosomes and dried under vacuum. The slides were then ready for the addition of the probe.

2.11.8. Preparation of probe for hybridisation

25 ng of probe was added to 0.5 ng of Cot I DNA (Gibco BRL) and 5 μg of salmon sperm DNA (Sigma). Two volumes of 100% ethanol were added to this, mixed and spun down under vacuum until all the liquid had evaporated leaving the precipitated DNA. Probes were resuspended in 10 μl of hybridisation solution (50% v/v deionised formamide*, 2 x SSC, 1% v/v Tween 20,

10% v/v dextran sulphate) by leaving it on the bench for 1 hr. Pre-hybridisation and preparation of the probe was performed concurrently.

* Deionize 50 ml of formamide with 5 g of Amberlite monobed NB-1 resin (BDH) by stirring for an hour in a fume hood. Store frozen in aliquots of 0.5 ml. For hybridisation mix only.

2.11.9. Hybridisation

The probe was denatured at 70°C for 5 min followed by 15 min at 37°C. The slides were warmed to 37°C on a hot-plate. 10 µl of probe was spotted onto the region with the highest density of cells (marked earlier). A clean 20 x 20 cover-slip was placed on top of the probe and the cover-slip was sealed using rubber cement (TipTop). The slides were placed in a metal tray in a 37°C water bath overnight.

2.11.12. Post-hybridisation

The rubber seal was peeled off and the slides washed in 50% v/v formamide in 2 x SCC at 45°C (4 x 3 min), 2 x SSC at 45°C (4 x 3 min), 0.1 x SSC at 60°C (4 x 3 min) and transferred into 4 x SSC, 0.1% v/v Tween 20 at r.t..

2.11.13. Preparation and hybridisation of antibodies

Biotin-labelled probes were detected using sequential layers of avidin-FITC (AVFITC) or avidin-Texas Red (AVTR), biotinylated anti-avidin D (BAA, Vector) and another layer of fluorochrome-conjugated (the second and third layers were not always necessary). Digoxigenin-labelled probes were detected using

sequential layers of FITC anti-digoxigenin (Boehringer Mannheim) or Rodamine anti-digoxigenin, FITC anti-sheep (Vector) or TR anti-sheep (Vector). Antibodies were diluted in 5% w/v skimmed milk (Marvel) in 4 x SSC as follows: AVFITC and AVTR (1:500), BAA, FITC anti-sheep and TR anti-sheep (1:100), FITC anti-dig (1:40), Rodamine anti-dig (1:15).

The antibodies and skimmed milk were centrifuged (12,000 g, 15 min, 4°C) and the slides were "blocked" by incubating them on the bench with 40 µl of skimmed milk (5 min). Each antibody was added directly to the slide, a cover-slip placed on top and incubated at 37°C for 30 min in a moist chamber. The slides were washed in 4 x SSC, 0.1% v/v Tween 20 (3 x 2 min) at 45°C (for biotin detection) or 37°C (for digoxigenin detection) between each incubation step. After the last incubation with antibody the slides were drained of excess liquid and 20 µl of antifadent containing propidium iodide (PI) or 4', 6-diamidine-2-phenylindole dihydrochloride (DAPI) was added and a clean cover-slip placed on top. The cover-slips were sealed using a rubber cement (Pang). The slides were then ready to be analysed using fluorescence microscopy.

2.11.14. Analysis of fluorescence images

2.11.14.1. The confocal microscope

All the PI-stained / FITC-labelled slides were analysed using a BioRad MRC600 confocal laser scanning microscope. This was fitted with a dual A1 and A2 filter set which allowed simultaneous visualisation of PI and FITC images.

2.11.14.2. The Zeiss microscope with CCD

All DAPI-stained / FITC or TR-labelled slides were analysed using a Zeiss Axioplan fluorescence microscope. This was fitted with a triple band-pass filter set (Chroma) and a computer-driven excitation filter wheel which allowed sequencial visualisation of DAPI, FITC and TR images. Chromosomes were imaged using a cooled CCD camera fitted with a KAF 1400 chip (Photometrics). Separate images of the chromosomes and probes were merged using an Apple Mackintosh Quadra 900 computer with software developed by Digital Scientific. Images were stored on SyQuest disc and printed using a Kodak ColourEase dye sublimation printer.

2.12. PREPARATION AND ANALYSIS OF "NATIVE" CHROMOSOME SPREADS

Immunofluorescence experiments were performed at the MRC HGU, Western General Hospital, Crewe Road, Edinburgh, U.K. with guidance from Ms Linda Nicol and Dr Peter Jeppesen using methods essentially as described in Nicol & Jeppesen (1994) and references therein. The CREST patient serum was a gift from Dr Peter Jeppesen.

Lymphocyte cell culture and fixation procedures were performed as described in section 2.11.1.. Cells were resuspended in 4 ml of 75 mM KCl hypotonic solution and incubated for 10 min at 37°C, 20 µl were transferred into 10 ml of isotone (optimised sheath fluid for use in Flow Cytometry Instrumentation) and the cell density calculated using FACs flow (Becton Dickinson apparatus). Cells were diluted to a density of 2 x 10^5 cells/ml, 500 µl aliquots dispensed into Ames Cyto-Tek centrifuge buckets (Miles Scientific), centrifuged (2,000 rpm, 10

min, r.t.), slides were removed from holders and allowed to air-dry (5 min), immersed in KCM chromosome medium (KCM: 120 mM KCl, 20 mM NaCl, 10 mM Tris-HCl pH 8.0) for 10 min at r.t., the slides were transferred onto the bench and the area around the cells was dried. The slides were then ready to have antibodies applied.

CREST patient serum was diluted 1/100 in 10% v/v normal goat serum in KCM (prevents background), centrifuged (12,000 g, 5 min, r.t.), 40 µl of CREST antibody was applied to each slide, the chromosomes were covered with a small square of parafilm, the slides were incubated in a moist chamber at r.t. for 1 hr, slides were washed in KCM for 10 min, 40 µl of anti-human IgG (Sigma Immuno Chemicals [F-0132]: γ-chain specific) FITC conjugate (1/20 dilution in KCM) was added to each slide, the above incubation was repeated for 30 min, KCM for 10 min, 4% PFA for 10 min to fix the chromosomes, distilled water for 10 min, air dried, counter stained with DAPI (see 2.11.13.) and examined using the Zeiss microscope with CCD (see 2.11.14.2.). Images were stored and the slide co-ordinates of each image noted before proceeding to perform FISH analysis using DNA probes (see 2.11).

Slides with native chromosomes were treated with xylene for 30 sec, with DNA ligase for 1 hr at 37°C (to repair any damage induced by the earlier fluorescence analysis), 10 mM EDTA for 5 min, 0.1 M NaOH for 30 sec and 10 mM Tris-HCl, pH 8.0 for 30 sec. An uncrosslinking step was performed (30 sec in 0.1 M NaOH, 30 sec in 10 mM Tris-HCl pH 7.0, distilled water for 30 sec and a 10 sec fix 3:1 methanol/acetic acid) before performing FISH experiments with BLG, *c-myc* and *Pvt-1* probes using methods described in section 2.11..

CHAPTER 3

TRANSGENE EXPRESSION LEVELS AND RELATIVE COPY NUMBER IN THREE TRANSGENIC LINES

3. 1. INTRODUCTION

A genomic clone encoding ovine BLG was used to generate transgenic mice (Simons *et al.,* 1987). Three lines of transgenic mice were established (lines 7, 14 and 45) that express the BLG protein at high levels in the milk. A study of lines 7 and 14 demonstrated that transcription was initiated at the BLG promoter and that the protein was limited to the secretory epithelial cells of the mammary gland (Harris *et al.,* 1991). Expression of the transgene in line 14 reflected the temporal regulation of β-casein, an endogenous mouse milk protein gene. The expression profile of BLG during gestation was similar in sheep and in line 14 mice after compensation for the different lengths of pregnancy in the two species. Therefore the 5' *cis*-regulatory region of the transgene is capable of directing the correct temporal and spatial regulation of ovine BLG in mice.

During this study (Harris *et al.,* 1991) it was noticed that there were significant fluctuations in mRNA levels during a lactation cycle and significant variations in protein levels at mid-lactation (day 11) among animals of line 7 (McClenaghan, M.; pers. comm.). Despite the vast amount research in transgenic technology there were few examples of unstable transgene expression in the literature by 1991 (Table 5, chapter 1). Though hypervariable expression of an HMGCR transgene had been documented previously (Mehtali, M., Ph.D. thesis, University of Strasbourg, 1988; McClenaghan, M. and Lathe, R.; pers. comm.), this

project was initiated to investigate the cause of variable BLG transgene expression in mice.

As part of a systematic study of BLG expression levels in lines 7, 14 and 45 the following questions were addressed: (i) what are the precise levels of BLG protein in milk of each transgenic line?; (ii) do mRNA levels reflect protein levels?; (iii) are transgene arrays rearranged or deleted?; (iv) what is the nature of BLG mRNA expression in the mammary gland?

To address these questions a BLG protein standard was prepared from pooled sheep milk (see 3.2.1.). Milk samples were taken (mid-lactation) from approximately 30 mice of lines 7, 14 and 45. Some differences have been described between the thoracic and inguinal glands of the mouse (e.g., α-lactalbumin expression in response to hormonal stimulation; Bolander, 1990). Therefore all samples were collected from the upper left or right inguinal glands which are regarded as equivalent. The mouse milk was collected from the upper left inguinal gland and used in SDS-PAGE analysis (3.2.3.). The mouse was then sacrificed and the same gland removed and dissected in half. Each half (always the same region of the gland for all the animals) was frozen in liquid nitrogen and used to prepare RNA (3.2.4.) or DNA (3.2.5.). By this means it was possible to obtain milk, RNA and DNA systematically derived from the same gland of the same animal. The upper right inguinal gland was not milked so as to preserve tissue morphology. This gland was removed and dissected in half before being prepared for tissue *in situ* analysis (see chapter 5).

3.2. RESULTS

3.2.1. Preparation of a BLG protein standard

Quantitative analysis of BLG expression levels in the milk of transgenic mice required a BLG protein standard of known concentration. Although purified BLG is commercially available from a number of suppliers the purity is variable and the precise concentration is unreliable (due to variable quantities of water within the crystallised protein). Therefore it was necessary to generate a BLG protein standard and determine the exact concentration.

BLG was purified from sheep milk following a method described by Aschaffenburg & Drewry, (1957). SDS-PAGE demonstrated that the BLG preparation contained only one detectable protein that co-migrated with the BLG present in sheep and transgenic mouse milk (Fig. 8A). Western blot analysis using a rabbit antibody against BLG confirmed the identity of the protein (Fig. 8B).

An accurate method for determining the nitrogen content and hence the protein content of a small sample is the micro-Kjeldahl method (Rowland, 1938; Davies & Law, 1983). This method was employed to determine the concentration of protein in the BLG preparation (Table 8). Four measurements (1.172, 1.167, 1.159, and 1.230 mg/ml) derived from replicate assays were averaged to 1.2 mg/ml. This value was used to calculate and prepare the BLG standards for quantitative SDS-PAGE analysis of the milk samples collected from lines 7, 14 and 45.

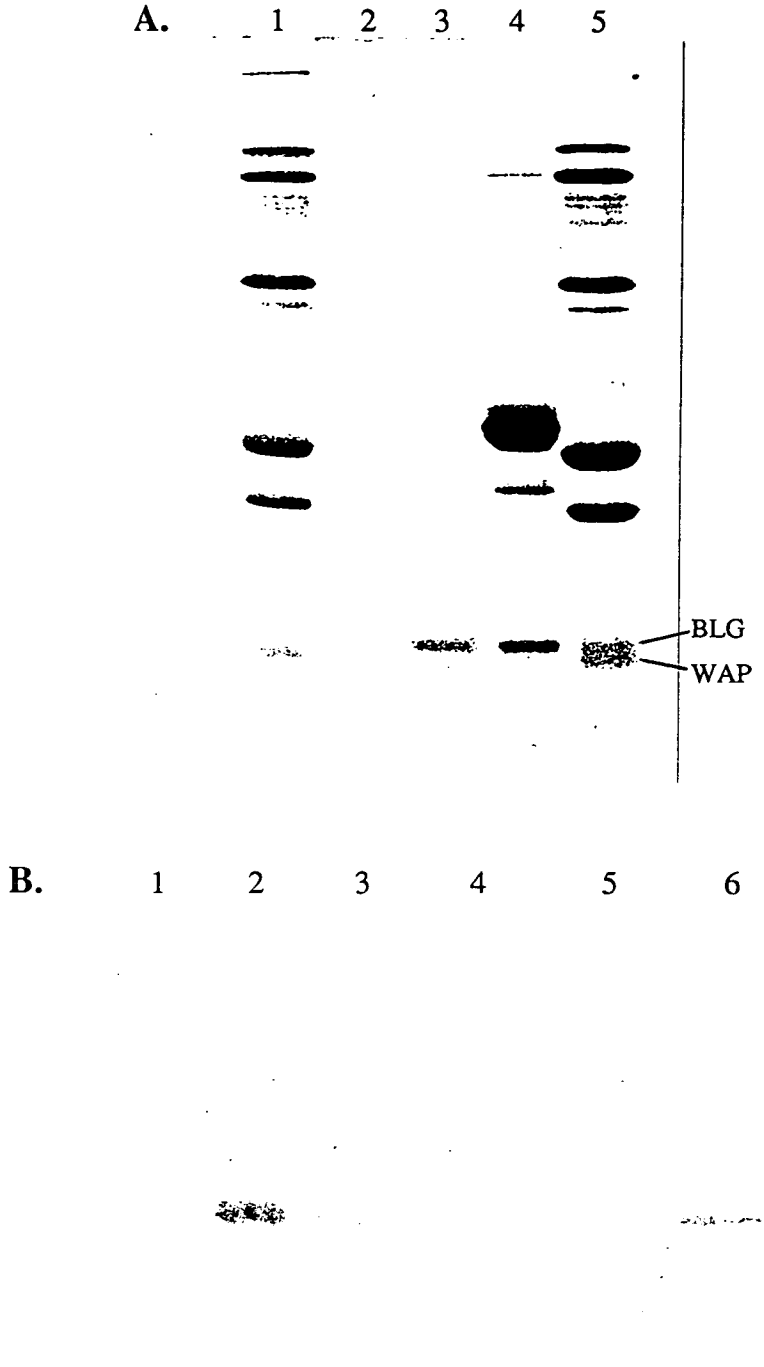


Figure 8. SDS-PAGE and western blot analysis of purified ovine BLG

(A) A 15% SDS-polyacrylamide gel stained with Coomassie Blue G-250. Equivalent amounts of defatted milk (0.16 µl) were loaded. Lane 1: 5 µl of control mouse milk; lanes 2 and 3: 1 µg and 2.5 µg of sheep BLG respectively; lane 4: 5 µl of pooled sheep milk; Lane 5: 5 µl of transgenic mouse milk (ΔSt 89.2; Whitelaw *et al.*, 1992). (B) A western blot of a 15% SDS-polyacrylamide gel. Lane 1: 5 µl of 1/125 defatted control mouse milk; lanes 2-4: 250, 100 and 20 ng of sheep BLG respectively; Lane 5: 5 µl of 1/125 defatted BLG mouse milk (ΔSt 89.2); Lane 6: 5 µl of 1/250 defatted pooled sheep milk.

Table 8. Calculation of the concentration of the BLG protein standard

Sample*	mean	mean - c†	x 0.0964‡	x Dilution¶	x 6.38§	mg/ml∞
c_1 0.227 c_2 0.210	0.218					
1_1 0.980 1_2 0.980	0.980	0.762	0.073	0.367 (x5)	2.343	1.172
2_1 0.600 2_2 0.595	0.598	0.380	0.037	0.366 (x10)	2.334	1.167
3_1 0.410 3_2 0.403	0.407	0.189	0.018	0.363 (x20)	2.319	1.159
4_1 0.307 4_2 0.329	0.318	0.100	0.010	0.386 (x40)	2.460	1.230

2 x 1 ml of ovine BLG were combined and made up to a total volume of 10 ml with distilled water. This 10 ml was divided into 2 x 2 ml (sample 1), 2 x 1 ml (sample 2), 2 x 500 µl (sample 3), and 2 x 250 µl (sample 4). *c is a distilled water control and therefore the baseline for samples 1-4; the values are the volumes of HCl added in each titration (see materials and methods), †the mean of each sample minus the mean of c, ‡1 ml of HCl = 0.0952 mgN, ¶fraction of the total 10 ml, §6.38 is the multiplication factor for the nitrogen content of bovine BLG, ∞concentration of BLG for each dilution in mg/ml (the dilutions were prepared from 2 ml of BLG stock solution)

Each quantitative SDS-polyacrylamide gel contained a 1 µg and a 2 µg BLG standard. These concentrations of BLG were within a linear range that went through the origin (Fig. 9). By comparing the BLG levels in transgenic mouse milk samples with ovine BLG standards enabled the calculation of the absolute BLG expression levels in milk samples. Each gel contained eight transgenic mouse milk samples and two BLG standards (1 µg and 2 µg). Therefore the BLG expression levels in the transgenic samples were comparable between gels.

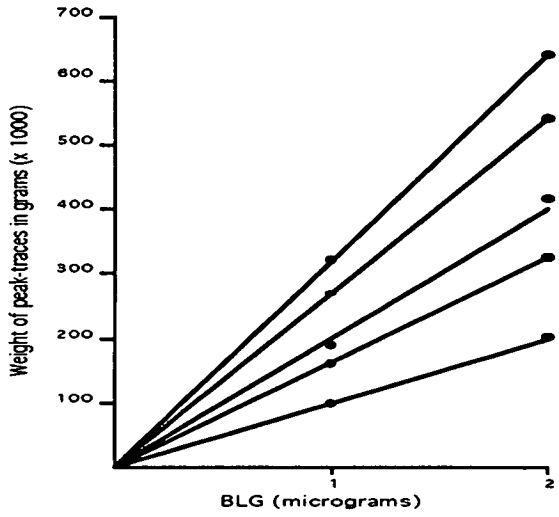


Figure 9. Quantitation of BLG standards in five gels

The area of the Coomassie blue-stained BLG peaks recorded by absorbance were calculated by weight. Measurements for 1 μg and 2 μg BLG standards in five SDS-polyacrylamide gels were generated at different stages of destaining. The BLG standards with the lowest peak measurements had been destained for the longest period (several hours).

3.2.2. A quantitative assay for BLG expression levels in transgenic mouse milk

SDS-PAGE analysis was employed to measure the BLG protein expression levels in >300 mouse milk samples. The results from this analysis were the basis of every experiment in this project and therefore exhaustive efforts were made to determine the most accurate and reproducible method for measuring BLG protein levels. The reproducibility of the assay was tested by repeatedly loading a pooled transgenic mouse milk sample into gels of differing thickness and polyacrylamide concentration, at different dilutions and under different gel scanning conditions. The following protocol was necessary to obtain

reproducible BLG values for the same sample loaded in an identical fashion within a gel. Optimal results were obtained with 1.0 mm thick gels. Mouse milk samples were loaded as 1/200 dilutions (1/250 for some of the highest expressors in line 45) of the original sample. These dilutions brought the BLG values in the transgenic mouse milk samples into the linear range of the BLG standards. Resolving gels were 20% polyacrylamide and the samples were run at 28 mA for 1 hr (stacking gel) and at 36 mA for 5-6 hr (resolving gel). This enabled the resolution of BLG and WAP (Fig. 10) so that BLG could be scanned as a discrete band. The BLG concentrations were determined by scanning densitometry. Each lane was scanned separately with the scanning beam set to the exact width of each lane; usually 9 mm for the milk samples and up to 12 mm for the BLG standards. For each scan the baseline and verticals for each peak were drawn in by hand because the automatic settings yielded variable results. From resulting traces, peaks were cut out and weighed using a fine balance. The weight of the peak-trace was proportional to the size of the peak and hence the concentration of BLG.

The data presented in table 9 demonstrates that the weight of any one peak (and therefore the amount of BLG) within a gel could be expected to lie within ±4% of the true value 95% of the time. This was felt to be an acceptable margin of error.

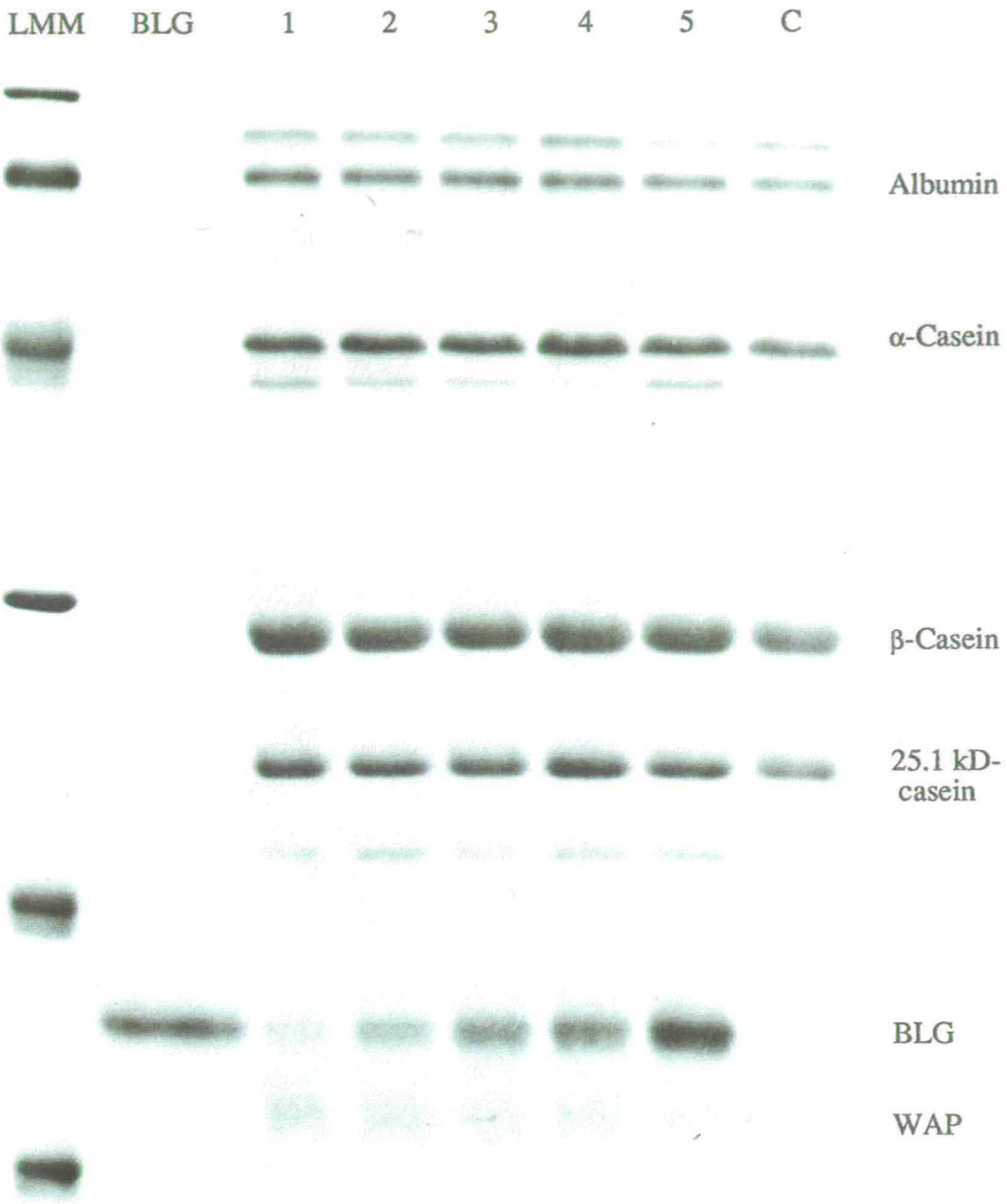


Figure 10. SDS-PAGE analysis of mouse milk

Mouse milk samples were electrophoresed in 20% polyacrylamide gels (1 mm thick) under reducing conditions and stained with Coomassie Blue. The lanes were as follows: LMM, low-molecular-mass standards (phosphorylase *b* [97.4 kDa], BSA [68 kDa], ovalbumin [43 kDa], carbonic anhydrase [29 kDa], soybean trypsin inhibitor [20.1 kDa], lysozyme [14.3 kDa]); BLG, 2 μg of ovine BLG; lanes 1-5, 1/200 dilutions of defatted transgenic mouse milk samples from line 7; C, 1/200 dilution of a defatted control non-transgenic mouse milk sample.

Lane	BLG*
1	550
2	549
3	570
4	550
5	567
6	574
7	557
8	543

Table 9. Reproducibility study of BLG values

15 µl of a transgenic mouse milk sample (1/125 th dilution) was loaded into eight lanes within a gel and the BLG values determined by SDS-PAGE/ Coomassie Blue-staining and scanning densitometry.

* The weight of each BLG peak in grams (x 10^3);
n = 8, mean = 557.5, *s.d.* = 11.43, *c.v.* = 2%

3.2.3. BLG protein expression levels in lines 7, 14 and 45

Variable levels of BLG had been observed (although not quantitated) within line 7. Systematic quantitative analysis of lines 7, 14 and 45 was performed to determine the range of BLG protein expression levels within these lines. The results are presented in Table 10 and Fig. 11.

Line 14 expressed BLG in a stable fashion with only 3.5 mg/ml separating the lowest and the highest expressor (5-8.5 mg/ml). Levels of expression in the two remaining lines differed considerably between individual animals. BLG levels in milk of line 45 animals ranged from 16 to 30 mg/ml, while in line 7 there was a 8-fold difference between the lowest and highest BLG expressors (3 to 23.9 mg/ml). Table 11 shows the standard deviations (*s.d.*) and coefficients of variation (*c.v.*) for BLG protein expression in the three lines. Lines 7 and 45 had larger *s.d.* values than line 14. The *s.d.* values are a measure of the variance within a population. Therefore higher *s.d.* values in lines 7 and 45 was expected, based on the distributions presented in the histograms in Fig. 11. The *c.v.* for line 45 was similar

Table 10. BLG protein expression levels in lines 7, 14 and 45

Line 7	BLG (mg/ml)	Line 14	BLG (mg/ml)	Line 45	BLG (mg/ml)
F4	23.9	H8	8.5	E2	30.0
F10	21.5	F4	8.3	D1	29.6
F5	16.8	I10	8.3	A3	29.1
F7	14.9	I6	8.0	J2	28.7
F2	12.7	I14	7.9	G1	28.3
D5	12.3	I4	7.9	E3	28.1
E3	13.1	H3	7.4	H4	27.0
F1	10.1	F13	7.4	I1	25.7
D3	11.4	C5	7.3	A2	25.4
E3	11.3	I3	6.8	I4	25.1
F8	11.3	H10	6.7	I2	25.0
F4	10.7	G3	6.7	I3	24.9
F3	9.6	I11	6.5	E1	24.5
F9	9.4	C2	6.5	D4	23.2
E2	8.8	F9	6.5	E4	22.9
E1	8.4	B6	6.5	H2	22.6
G4	8.2	D4	6.4	D3	22.6
B2	8.1	C8	6.3	F2	22.3
G1	7.8	F10	6.1	H3	21.9
D1	7.5	C7	6.0	I5	21.8
A1	6.9	H9	5.9	F4	21.2
B5	6.8	I7	5.9	G2	21.1
G3	6.4	B2	5.8	D2	20.9
A5	5.9	G7	5.4	A4	20.4
A4	5.7	G6	5.3	F3	18.8
B1	4.4	F11	5.0	A1	18.0
D2	4.2	L14	4.7	J4	17.0
B3	4.0	L12	4.7	D5	16.0
F6	3.8				
A3	3.0				

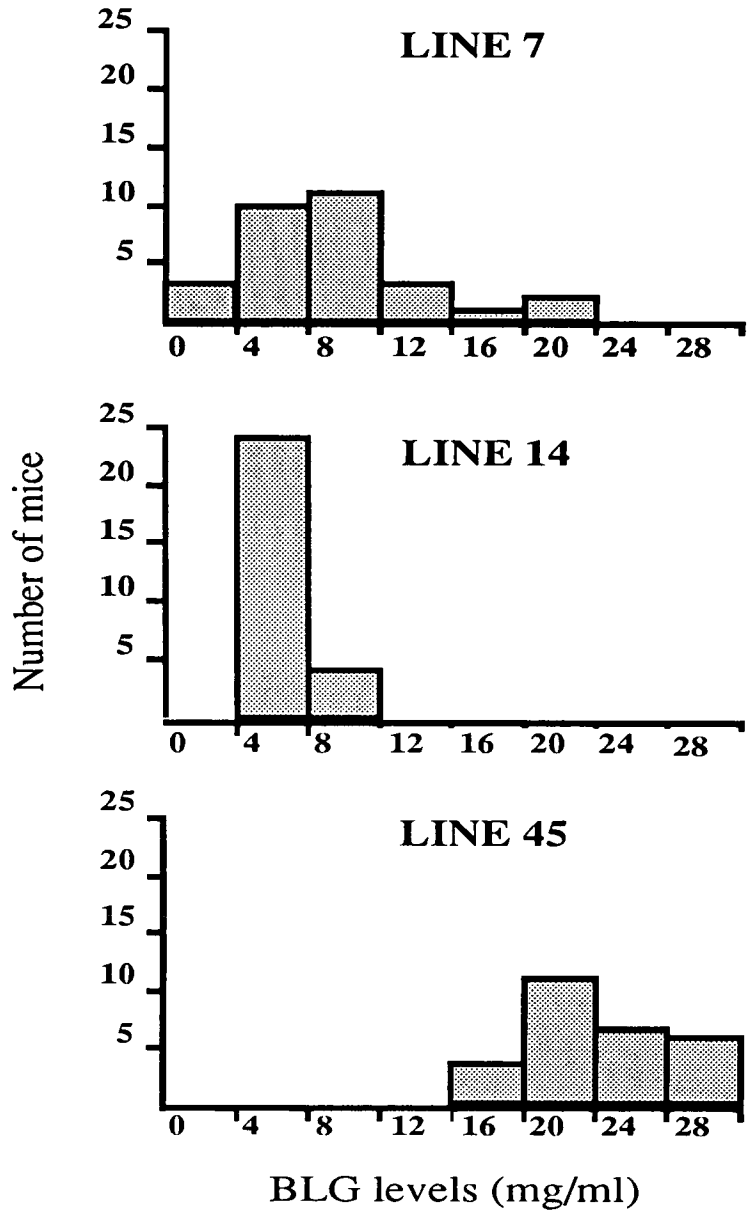


Figure 11. BLG expression level profiles in milk collected from line 7, 14 and 45 hemizygous mice

to that of line 14; this probably reflects its higher average expression level rather than comparable degrees of variation (see table 11 legend).

These results demonstrate that line 7 expresses variable levels of BLG within the line, line 14 expresses BLG in a stable fashion and that line 45 is an intermediate between these two extremes.

Table 11. BLG protein expression levels within lines 7, 14 and 45

Line	mean ± *s.d.*	*c.v.*
7	9.5 ± 4.9	52% (n = 30)
14	6.6 ± 1.1	17% (n = 28)
45	23.7 ± 3.9	16% (n = 28)

BLG protein mean values are absolute values (mg/ml); *c.v.* = 100 x *s.d.* / mean; n = number of mice

3.2.4 Analysis of BLG mRNA levels in lines 7, 14 and 45

Variable levels of BLG protein in the milk could be due to variable mRNA stability, variable mRNA translation, variable protein turnover or even unusual transcription of the transgene. To determine whether levels of BLG mRNA vary in parallel with BLG protein expression, quantitative Northern blotting using BLG and control (GAPDH) hybridisation probes was performed on total RNA from lactating tissue.

Table 12 and Figure 12 shows BLG mRNA and protein expression in individuals from lines 7, 14 and 45. The correlation coefficients for protein and RNA levels were 0.44 and 0.54 for line 14 and 45 respectively (n = 9), 0.54 for line 7 (n = 8) after the exclusion of F4 and -0.19 for line 7 including F4. These correlation values did not indicate strong linear relationships.

Table 12. BLG protein and mRNA expression levels in lines 7, 14 and 45

Line 7		Line 14		Line 45	
Protein*	RNA‡	Protein	RNA	Protein	RNA
F4 23.9	235530	H8 8.5	82245	D1 29.6	82813
F7 14.9	409459	I14 7.9	97614	J2 28.7	46936
B2 8.1	503488	F13 7.4	90737	E3 28.1	79438
B5 6.8	337695	F9 6.5	83523	I1 25.7	58881
G3 6.4	235627	I7 5.9	78155	A2 25.4	72744
A5 5.9	206387	H10 6.7	65304	H2 22.6	55551
A4 5.7	406947	B6 6.5	70542	A1 18.0	36595
B3 4.0	75340	C2 6.5	88377	F3 18.8	47694
A3 3.0	227625	H9 5.9	69808	D5 16.0	70849

* individual mouse identification and corresponding BLG protein expression level (mg/ml); ‡ BLG mRNA levels were normalised to sheep GAPDH; mRNA values are arbitrary and not comparable between lines.

Fig. 13 plots protein levels against RNA values. The outliers in each line (e.g., F4 in line 7) made a significant impact on the correlation coefficients due to the small RNA sample size. However, BLG mRNA levels did appear to be more variable in mammary tissue from animals of lines 7 and 45 than in line 14. Quantitative analysis confirmed this, as summarised in Table 13.

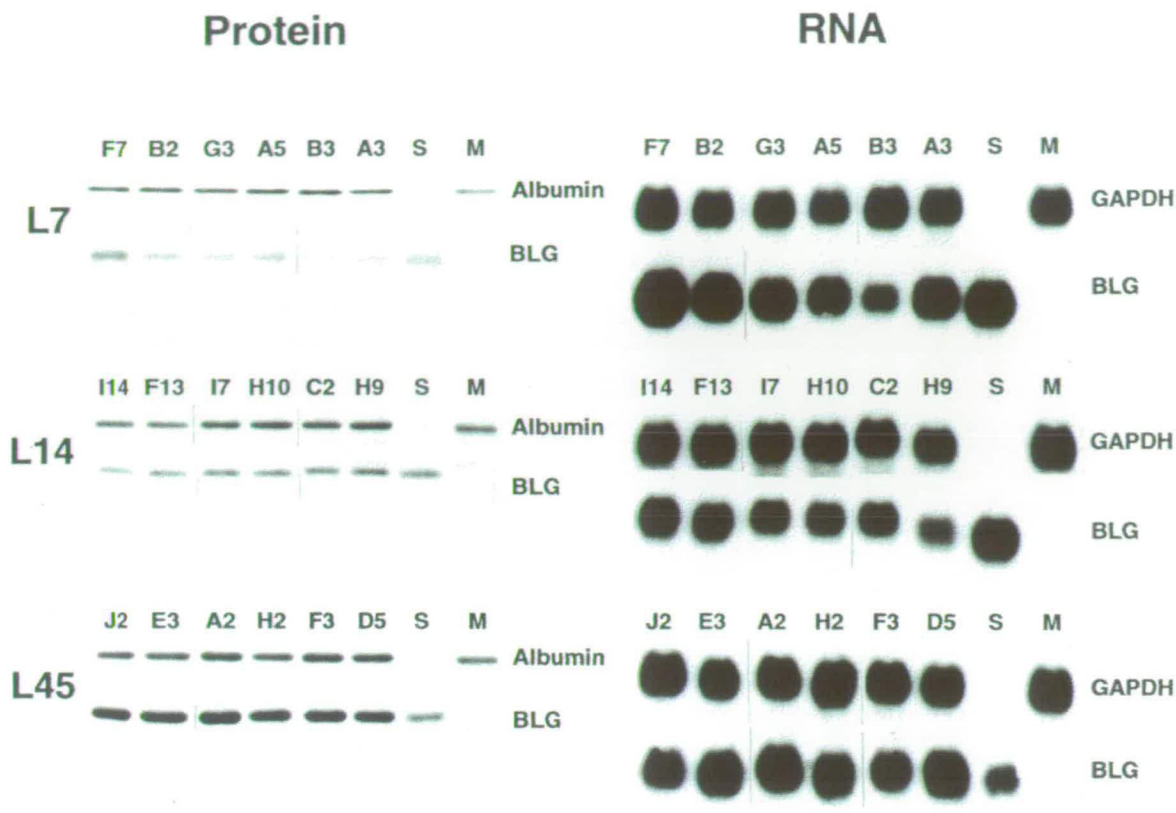


Figure 12. BLG protein and RNA expression in lines 7, 14 and 45 mice, shown as SDS-polyacrylamide gels and northern blots

The protein and RNA samples were from the same individual mice. The samples are organised from high BLG protein expression (*left*) to low BLG expression (*right*). Protein controls, 1/200 dilution of defatted sheep milk (S) and control mouse milk (M). RNA controls, 10 μg of sheep mammary RNA (S) and non-transgenic mouse mammary RNA (M). RNA blots were probed with BLG and GAPDH respectively.

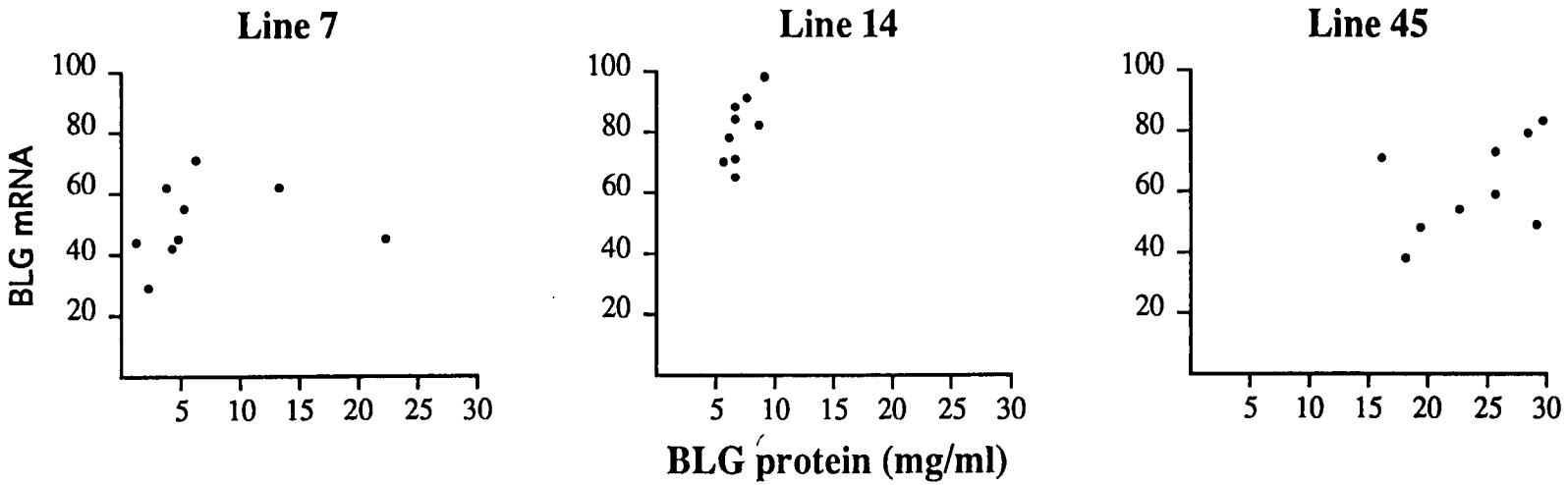


Figure 13. BLG protein expression levels plotted against BLG mRNA levels
RNA values are from Table 12, rounded up to the first two digits; Y-axis are not comparable between graphs; BLG values are absolute values in mg/ml (see Table 12.).

Table 13. The coefficient of variations for BLG mRNA expression within lines 7, 14 and 45

Line	*c.v.*
7	47% ($n = 9$)
14	13% ($n = 9$)
45	26% ($n = 9$)

RNA values were measured in arbitrary units following normalisation (within a line) to endogenous GAPDH mRNA determined by separate hybridisation to a GAPDH probe; for this reason mean values for RNA are not comparable between different lines and are not shown.

The results indicate that variable expression is present at the mRNA level in lines 7 and 45, and the relationship between protein and RNA is similar in all three lines (see also chapter 5). These experiments do not rule out variable mRNA stability within lines 7 and 45. However results presented in chapter 5 make this unlikely. Therefore variable BLG expression is probably not due to variable mRNA stability, translation or protein turnover. Instead, the most likely explanation for variable expression is that aberrant transcription of the transgene locus is occuring within lines 7 and 45.

3.2.5. Analysis of BLG transgene arrays in lines 7, 14 and 45

The formation of a transgene array in the genome is the result of a variable number of DNA fragments joining together and integrating into a single (usually) chromosomal site. Variable expression of the transgene within lines 7 and 45 could be due to two classes of effects. Firstly, variable transgene expression could be caused by gross or minor modifications to the structure of the transgene locus. For example, the transgene array may be rearranged due to breakage and rejoining before integration, the array may be partially or fully deleted

due to intrachromosomal recombination (Sandgren *et al.*, 1991) or it could include point mutations. Secondly, variable transgene expression could be imposed upon the transgene locus due to the chromosomal environment. This second class of effects are termed "position-effects" and will be dicussed in chapter 4 while this chapter will be concerned with the first class of effects.

Lines 7 and 14 harbour an identical transgene (a 16.2 kb *SalI-SalI* fragment encompassing the entire ovine BLG transcription unit; Fig. 14) while line 45 carries a 3' truncated transgene (10.5 kb *SalI-XbaI* fragment); the 3' truncation does not affect the transcription unit nor transgene function (Simons *et al.*, 1987).

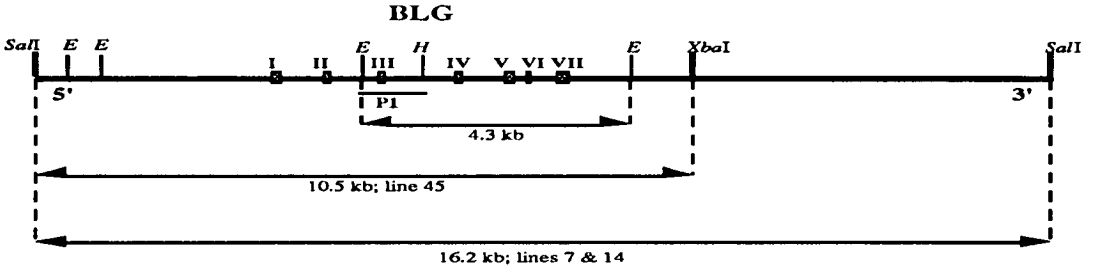


Figure 14. A map of the BLG transgene

BLG exons, I-VII; restriction enzyme sites, *Sal*I, *Eco*RI (*E*), *Hin*dIII (*H*); 1.1 kb BLG DNA hybridisation probe, P1.

To address the possibility of transgene deletion and/or rearrangement, Southern blot analysis was performed on DNA prepared from mammary tissue from lines 7, 14 and 45. The genomic DNA was restriction enzyme digested using *Eco*RI. The BLG transgene has two internal *Eco*RI restriction sites that are 4.3 kb apart (Fig. 14). An internal 1.1 kb *Eco*RI-*Hind*III probe (P1; Fig. 14) was isolated from plasmid pBJ14 (a gift from Dr Bruce Whitelaw; Roslin Institute) and used as a hybridisation probe to detect the 4.3 kb internal transgene fragment in *Eco*RI-digested mammary DNA samples. A whey acidic protein (WAP) hybridisation

probe (pBS-WAP; a gift from Dr Thomas Burdon; Centre for Genome Research, Edinburgh) or a total genomic DNA hybridisation probe was used as a control for DNA loading between lanes.

In all samples from lines 7, 14 and 45 the BLG probe identified a 4.3 kb band as predicted; there was no evidence of variations in the banding pattern that might suggest rearrangement of the transgene array in lines 14 and 45 (Fig. 15). There were 2 smaller bands (2.7 kb and 3.4 kb) present in all the line 7 individuals and one larger band (10 kb) present in five (F4, F7, F8, D2 and A3) out of nine line 7 individuals (Fig. 16A and B). The 10 kb band was equivalent to ~2 copies and the 2.7 kb and 3.4 kb bands were equivalent to <1 copy (compare with the 2 copy signal from the endogenous sheep BLG sample) indicating that they were probably due to hybridisation to fragments of the transgene at the edge of the transgene array and/or a small internal rearrangement. Nothing is known about the internal structure of the transgene array in line 7 although it is thought that lines 14 and 45 contain both head-to-head and head-to-tail arrangements of the transgene within the array (Dr Bruce Whitelaw, pers. comm.). The 10 kb band is probably present in some samples because of incomplete digestion (see the top of the lanes containing F9, B1 and B3) or partial degradation (see just below the 10 kb region in lane F1 and B1) of the DNA (Fig. 16). The extra bands were present in both high and low expressors, consistant with the interpretation that transgene rearrangement is not responsible for variable expression.

A previous estimate of copy number indicated lines 7 and 45 carried 25 and 17 transgene copies respectively (Whitelaw *et al.*, 1992), while line 14 harboured <5 copies (Dr Bruce Whitelaw, pers. comm.). Fig. 15 appears to confirm these copy numbers. Certainly line 45 has a higher copy number than line 7 and these two lines have a higher copy number than line 14. Hybridisation using BLG and control

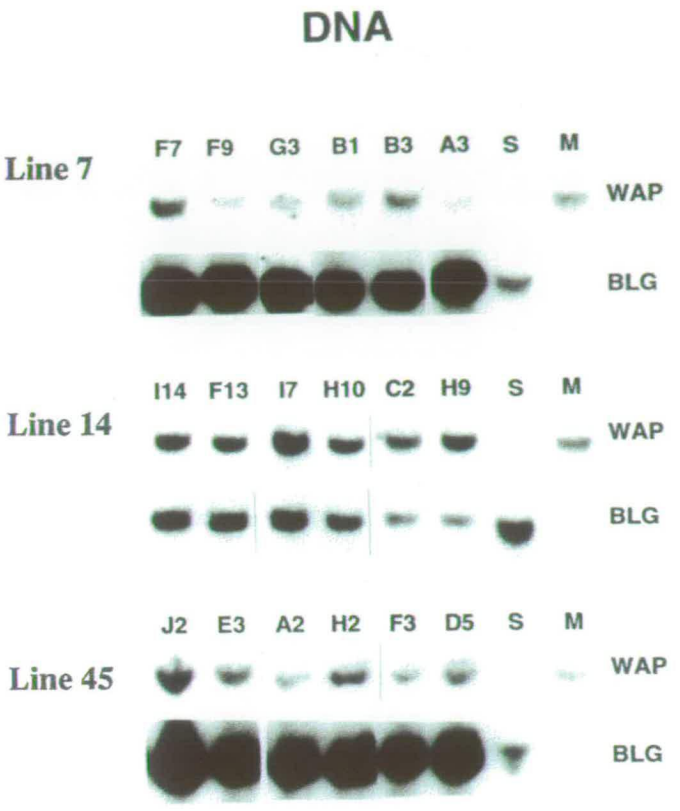


Figure 15. Southern blot analysis of the transgene array in lines 7, 14 and 45
Total genomic mouse mammary DNA from lines 7, 14 & 45 was hybridised with a WAP and a BLG probe. The samples are organised from high BLG protein expression (*left*) to low BLG protein expression (*right*). S, sheep mammary DNA; M, control mouse mammary DNA. See Table 10 for the protein expression levels.

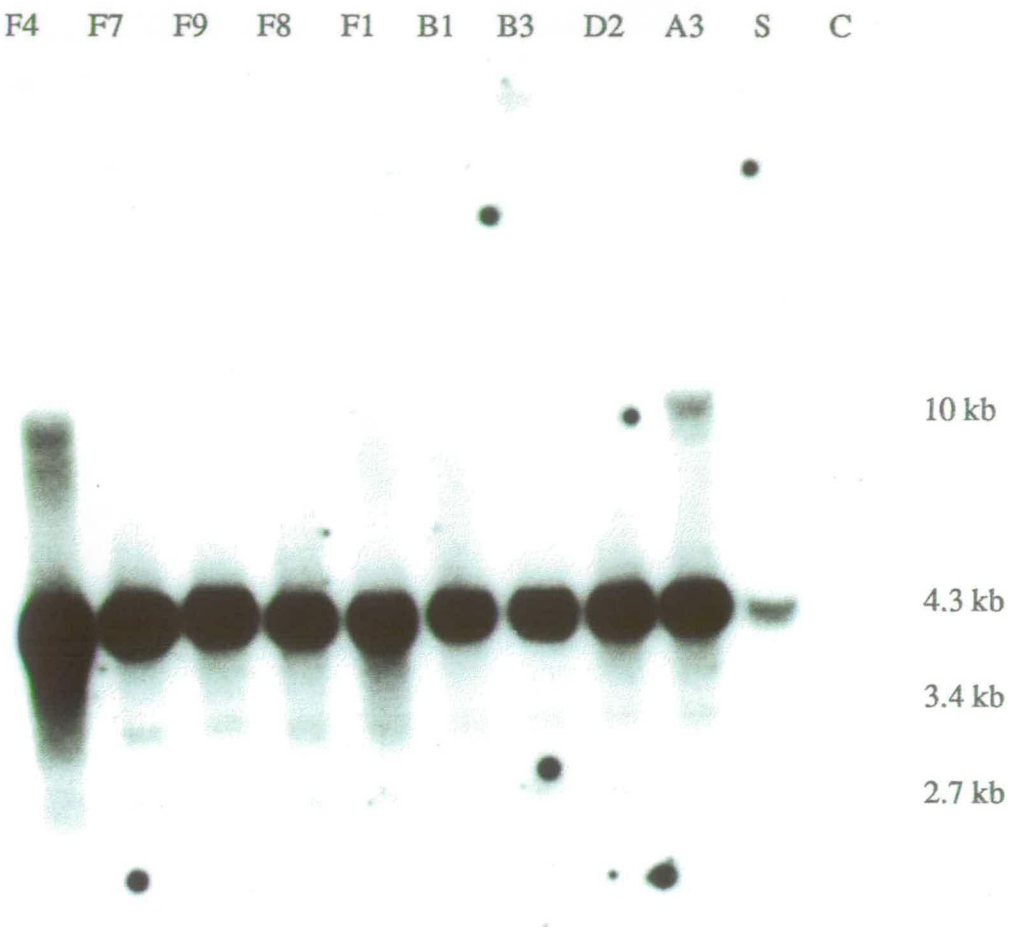


Figure 16. Southern blot analysis of the transgene array in line 7

(A) Total genomic mouse mammary DNA from line 7 was hybridised with the P1 probe. The samples are organised from high BLG protein expression (*left*) to low BLG protein expression (*right*). S, sheep mammary DNA; C, control mouse mammary DNA. See Table 10 for the protein expression levels.

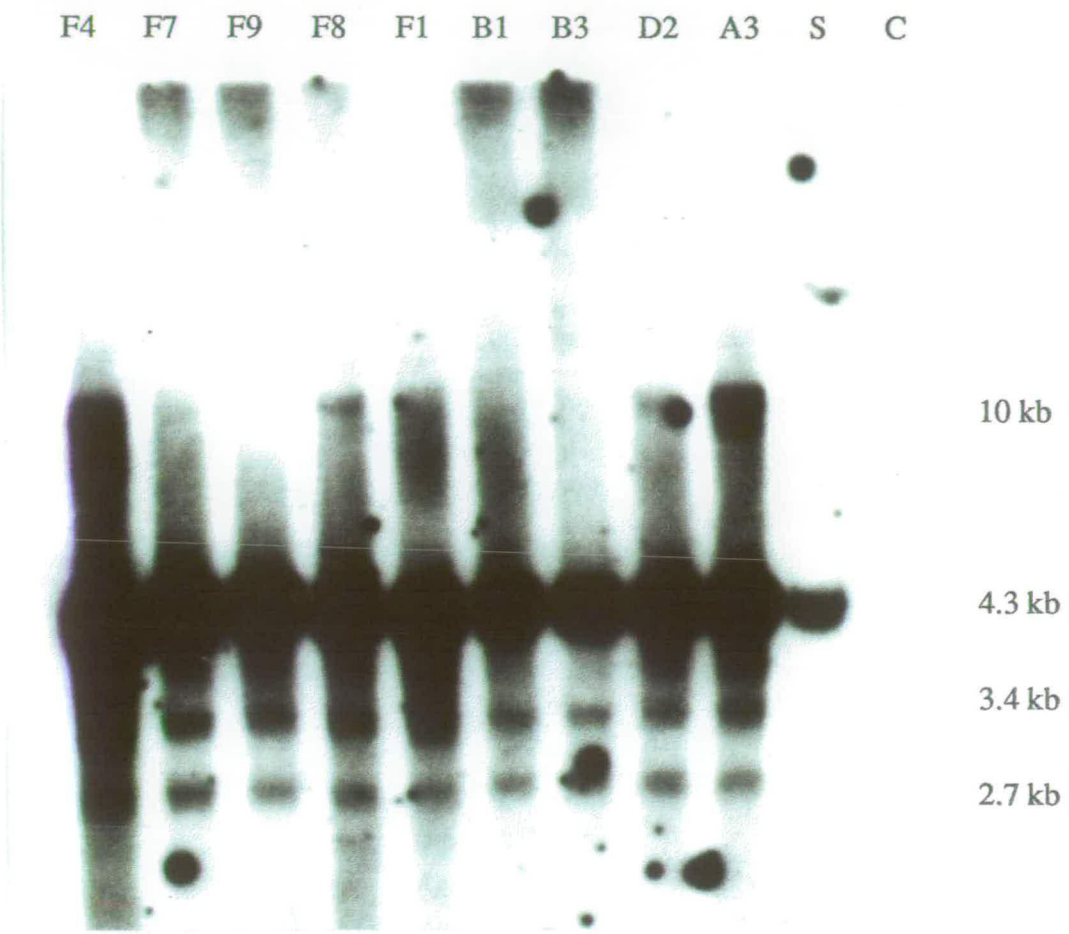


Figure 16. Southern blot analysis of the transgene array in line 7
(B) The same membrane as in **(A)**, over-exposed to film (24 hr).

probes revealed that the transgene copy number was indistinguishable within lines, even among mice from line 7 exhibiting an 8-fold range of BLG levels (Fig. 16). Despite efforts to provide a precise quantitation of relative BLG copy number in mammary gland DNA, the *c.v.* values for the BLG transgene were consistently high for all three lines. The high *c.v.* values reflects the inherent variability in the procedure (Table 14). Even line 14, which has only <5 copies of the transgene exhibited a high DNA *c.v.* value. Normalising the BLG signals to WAP (L7 *c.v.* = 26%, $n = 9$; L14 *c.v.* = 45%, $n = 9$; L45 *c.v.* = 38%, $n = 9$) or total genomic DNA (table 14) gave similar *c.v.* values, ruling out a possible transfer problem due to the size of the *Eco*RI WAP fragment (7.3 kb) in the Southern blotting procedure. All lines had similar *c.v.* values for DNA, even in line 14 which is a stable BLG expressing line. This contrasts with the protein and RNA data where the *s.d.* and *c.v.* values differed considerably for the three lines (Table 14). Furthermore the line 7 protein and RNA *c.v.* values are higher than the DNA *c.v.* value.

Table 14. BLG expression levels and copy numbers within lines 7, 14 and 45

	Protein		RNA	DNA
Line	mean ± *s.d.*	*c.v.*	*c.v.*	*c.v.*
7	9.5 ± 4.9	52% (*n*=30)	47% (*n*=9)	36% (*n*=18)
14	6.6 ± 1.1	17% (*n*=28)	13% (*n*=9)	41% (*n*=13)
45	23.7 ± 3.9	16% (*n*=28)	26% (*n*=9)	42% (*n*=13)

DNA values are measured in arbitrary units following normalisation (within a line) to endogenous total DNA determined by separate hybridisations to a total genomic probe (see materials and methods); for this reason mean values for DNA are not comparable between different lines and are not shown.

The mammary gland comprises many cell types (e.g., stromal tissue, blood vessels, myoepithelial cells) that do not express BLG. However these tissues presumably harbour the transgene array. Though the secretory epithelial cells are the most abundant cell type within the day 11 lactating mammary gland, the precise proportion of secretory epithelial cells within the mammary gland has not been determined. Therefore measurements of DNA copy number from the whole gland will include the transgene copies in the other, non-expressing cells. Given the inherent variability in the technique, secretory epithelial cell-specific deletion of the transgene array may not be detected using Southern blotting. Therefore it was not possible to rigorously exclude the possibility that the variable expression in lines 7 and 45 could be due to epithelial cell-specific transgene deletion. This point will be expanded in chapter 7.

The Southern blot analysis described in this chapter would not reveal subtle changes (e.g., a point mutation[s]) within the transgene array. To investigate this possibility it would probably have been necessary to sequence the entire transgene array in lines 7 (~400 kb), 14 (<90 kb) and 45 (~275 kb). Because of the repetitive nature of the insertions, even this would probably not have detected key point mutations. Despite extensive literature searches I have found no evidence for spontaneous point mutations within transgene arrays. Therefore although still formally possible, a point mutation that drastically affects transgene transcription is an unlikely mechanism; a more plausible mechanism that is based on past and current literature as well as the collective results presented in this study is presented in chapter 6.

These results argue that variable BLG expression is probably not due to stochastic loss or rearrangement of the transgene array in different individual line 7 and 45 mice and is instead probably due to variable mRNA transcription between individuals within a line.

3.2.6. BLG expression levels across two lactations in line 7 mice

Variable lactational performance can be due to numerous factors; none of which can explain the variable levels of BLG protein in lines 7 and 45 (Table 15).

Table 15. Sources of variability in lactational performance

Factor	Relevance to variable BLG expression
Intrinsic factors	
Genetic background	Standardised to (C57BL/6 x CBA*)
Litter size	Litters were standardised to 5 pups per mother
Reproductive state	Samples were collected at day 11 of lactation
Environmental factors	
Thermal environment	Same for all the individuals in the population
Nutrition	"
Work, exercise	"
Health status	All animals appeared healthy
Milk withdrawal	The litter sizes were standardised and milk was collected as close to 1 pm as possible for all the samples and by identical procedures
Time related factors	
Age	Standardised to 6-8 weeks
Season	Breeding is not seasonal in mice
Stage of lactation	Day 11 of lactation for all the samples

The sources of variability were adapted from Oldham & Friggens, 1989; * for an investigation of the contribution of genetic background see chapter 4

However, variable expression could be due to fluctuating expression levels within lactations, perhaps due to fluctuating levels of mammotrophic hormones. Ideally this possibility could be addressed by performing serial milking within one lactation cycle. However this approach would be unduly invasive to the mother and

the suckling offspring. Furthermore Zippelius and Schleidt, (1956) demonstrated that young rodents emit ultrasounds when in stress conditions, to which the dam responds by exhibiting maternal care patterns (e.g., time spent in nursing position). These responses may vary between individuals and therefore make serial milking within a lactation an unpredictable experiment (e.g., numbers of mice required). Instead, to determine whether expression levels exhibit temporal variation within individual animals of transgenic line 7, milk protein levels were measured in a large population of mice at day 11 of two successive lactations. This approach was far less invasive to the mice, enabled a large sample size (29) and addressed whether the expression level was fixed within the individual at mid-lactation. The results are presented in Table 16. Commencing with a group of 40 animals, 29 milk sample pairs were obtained from two successive lactations. 25% of the sample pairs were lost because (i) the mouse did not become pregnant (ii) the litter died, and (iii) the mother did not produce enough milk from the upper left inguinal gland. Milking a large population across two lactations enabled a comparison of BLG expression levels within an individual and between individuals at the same stage of lactation.

The average difference between the two lactations (L2-L1) was 2.4 mg/ml (±0.6) or 2.0 mg/ml (±0.4) if sample 37 is omitted. These differences were significantly greater than zero (paired t-test, $p < 0.05$). Thus the mice produced significantly more BLG on their second lactation (24 mice produced more, two the same and three produced less). This result might be expected because mouse milk yield and pup weight gain increase with successive lactations (Knight *et al.*, 1986) and improvement in lactational performance with successive lactations is a general phenomenon in the dairy industry (see chapter 1).

Table 16. BLG protein expression levels in line 7 mice across two lactations

Mouse	BLG expression levels (mg/ml)		
	First lactation (L1)	Second lactation (L2)	L2-L1
1	14.9	ND	/
2	4.0	2.5	-1.5
3	ND	ND	/
4	10.6	15.3	+4.7
5	17.8	22.1	+4.3
6	15.4	22.2	+6.8
7	10.3	16.1	+5.8
8	15.3	15.5	+0.2
9	ND	ND	/
10	4.8	7.4	+2.6
11	10.5	ND	/
12	12.2	12.2	0.0
13	11.4	ND	/
14	8.1	9.7	+1.6
15	12.2	ND	/
16	5.8	6.2	+0.4
17	11.4	14.2	+2.8
18	15.6	13.1	-2.5
19	13.0	18.1	+5.1
20	5.0	5.3	+0.3
21	ND	ND	/
22	ND	ND	/
23	14.5	16.6	+2.1
24	ND	ND	/
25	4.1	5.3	+1.2
26	5.9	8.7	+2.8
27	6.9	9.9	+3.0
28	4.7	4.7	0.0
29	10.5	12.6	+2.1
30	14.0	17.3	+3.3
31	10.9	12.6	+1.7
32	13.5	13.4	-0.1
33	10.1	11.1	+1.0
34	11.0	11.4	+0.4
35	9.1	10.4	+1.3
36	16.0	17.8	+1.8
37	9.8	24.2	+14.4
38	3.8	ND	/
39	5.9	9.5	+3.6
40	3.0	ND	/

ND: not determined

The BLG expression levels in the two lactations were plotted together with respect to each mouse in a bar diagram and ordered from lowest to highest expressors using the first lactation results (Fig. 17; sample 37 was omitted). This demonstrated that there was a close correlation between BLG levels measured at the first and second lactations. For example, the low expressors remained low expressors and high expressors remained high expressors; there was no evidence of the wide range of expression observed within the population being reiterated in the individual.

Albumin levels are the least affected by variations in the levels of BLG in the milk (McClenaghan *et al.*, 1995) and therefore serve as the best control for loading of samples in SDS-PAGE. The *c.v.* measurements for BLG protein were the same if calculated from absolute values or those normalised for the endogenous protein albumin (Table 17). Therefore the variations in BLG levels were not due to variations in sample loading.

Table 17. Coefficient of variation values for BLG protein expression in milk of line 7 animals

	c.v. measurements	
Sample*	Absolute‡	Albumin corrected†
1	23%	25%
2	13%	13%
3	15%	14%
4	16%	16%

* Each sample comprises seven individuals; ‡ a measure of the *c.v.* values for the weight of the BLG peak-traces in line 7 transgenic mice at day 11 of the second lactation; † absolute values were corrected against the albumin value in a pooled transgenic mouse milk control sample run on all gels.

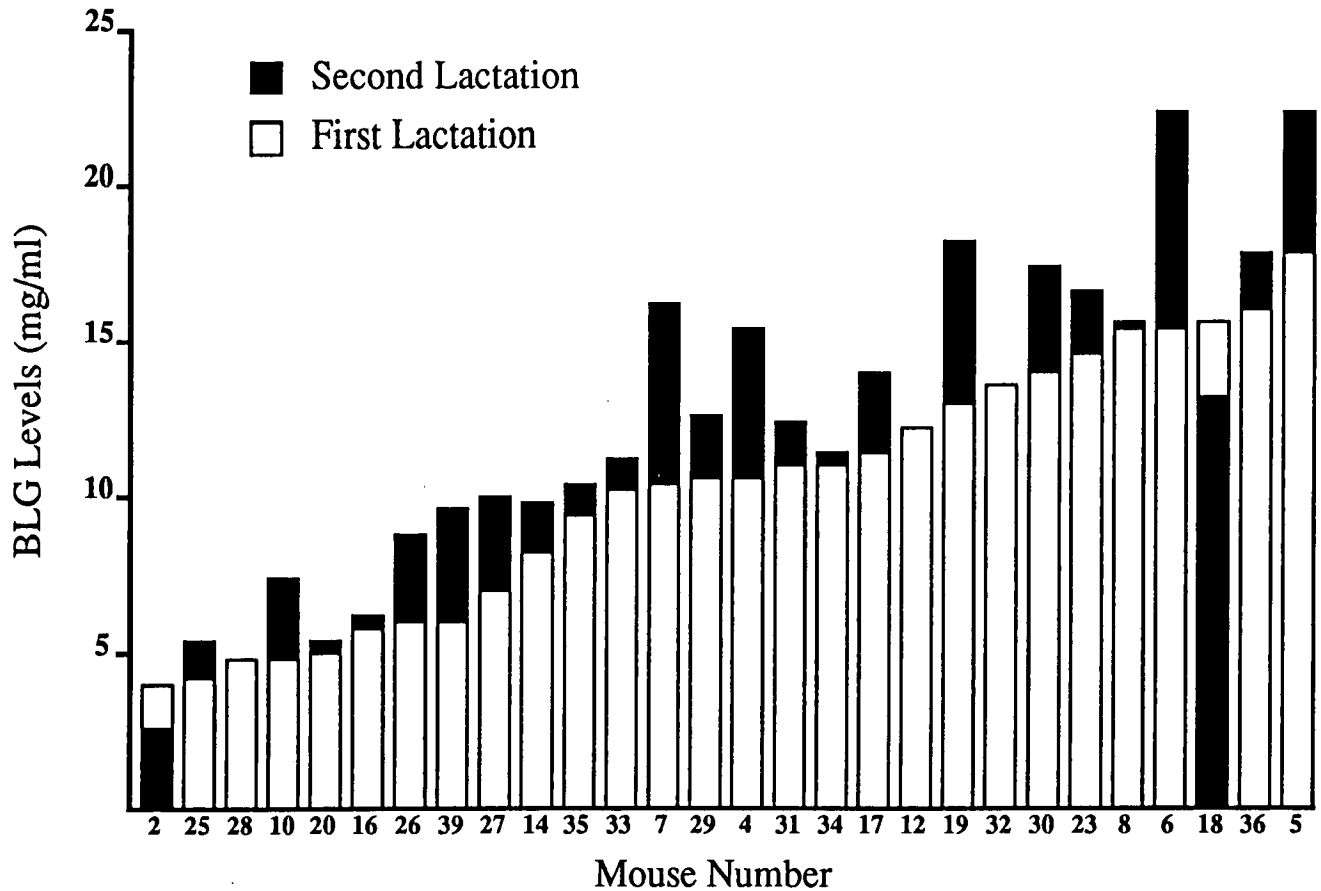


Figure 17. BLG expression levels in milk of line 7 animals over two consecutive lactations.
BLG was quantitated in milk collected at the same time point (day 11 of lactation) in two successive lactations in C57BL6 x CBA hemizygous mice. The average difference between the first and second lactations was 2 (± 0.4) mg/ml (paired t-test, $p < 0.05$).

While the line 7 population shows wide variation in the level of expression, individual levels appear to be fixed by day 11 of the first lactation.

3.3. CONCLUSIONS

A considerable amount of effort was invested into (i) purifying a pure BLG protein standard, (ii) determining the precise protein concentration of the BLG standard, and (iii) establishing a quantitative SDS-PAGE assay for the measurement of BLG concentration in transgenic mouse milk. Collectively, this enabled the accurate determination of BLG protein expression levels in several hundred samples (in this and in subsequent chapters).

Quantitative SDS-PAGE analysis of mouse milk demonstrated that the BLG protein expression levels were variable in lines 7 and 45 and stable in line 14. In line 14 levels of expression of the transgene-encoded protein are stable, while line 7 (harbouring the identical transgene) levels are highly variable; thus the DNA construct employed does not have an inherent deficiency that leads to variation in expression levels. The correlation values between BLG protein expression levels and mRNA levels were not sufficiently high to indicate strong linear relationships. However, these values did demonstrated that the relationship between protein and RNA was similar in all three lines. Furthermore, a comparison of *c.v.* values show that variation was present at the RNA level in lines 7 and 45. Though not rigorously proven, these results lead to the conclusion that the variation in protein levels is principally due to variations in transcription of the transgene insertion. Because lines 7 and 14 harbour the identicle transgene construct, it may be concluded that the particular transgene insertion locus determines variation in the levels of expression of BLG protein and mRNA levels.

While the Southern blot analysis would not reveal subtle changes in the structure of the transgene array, the data suggest that variable expression is not due to stochastic loss or rearrangement of the transgene array in different individual line 7 and 45 mice. Once established, the expression level remains relatively fixed within the individual. This result argues that, while the population shows wide variation in the level of expression, individual levels are fixed by day 11 of lactation.

CHAPTER 4

VARIABILITY IS NOT DUE TO THE GENETIC BACKGROUND

4.1. INTRODUCTION

Expression levels of randomly integrated transgenes usually differ between lines of transgenic animals (Spradling & Rubin, 1983; Palmiter & Brinster, 1986). These quantitative differences, known as "position-effects", have been attributed to the different chromosomal integration events which characterise the different lines. Transgene integration results in differing numbers of the transgene in the array (copy number) and different transgene integration sites between lines. Copy number dependent (or position-independent) levels of expression are indicative of a transgene array that is "insulated" from epigenetic influence due to the presence of a locus control region (LCR) within the construct (Grosveld *et al.*, 1987; Lang *et al.*, 1988; Talbot *et al.*, 1989; Greaves *et al.*, 1989; Bonifer *et al.*, 1990; Palmiter *et al.*, 1993; Diaz *et al.*, 1994; Krnacik *et al.*, 1995; Jones *et al.*, 1995; Aronow *et al.*, 1995; May & Enver, 1995). Position-effects can be caused by *cis-* (local chromosomal domain; e.g., promoters, enhancers and silencers) and *trans-* (epigenetic modifiers; e.g., transcription factors) acting factors that can act together or independently to produce variations in expression between lines (Kothary *et al.*, 1988; Allen *et al.*, 1988; Wilson *et al.*, 1990). For example, differences in expression levels between lines for an identical transgene can be due to integration (i) into an endogenous gene (Fig. 18), (ii) close to the enhancer of an endogenous gene, (iii) close to a CpG island (Bird, 1986) and the ubiquitous

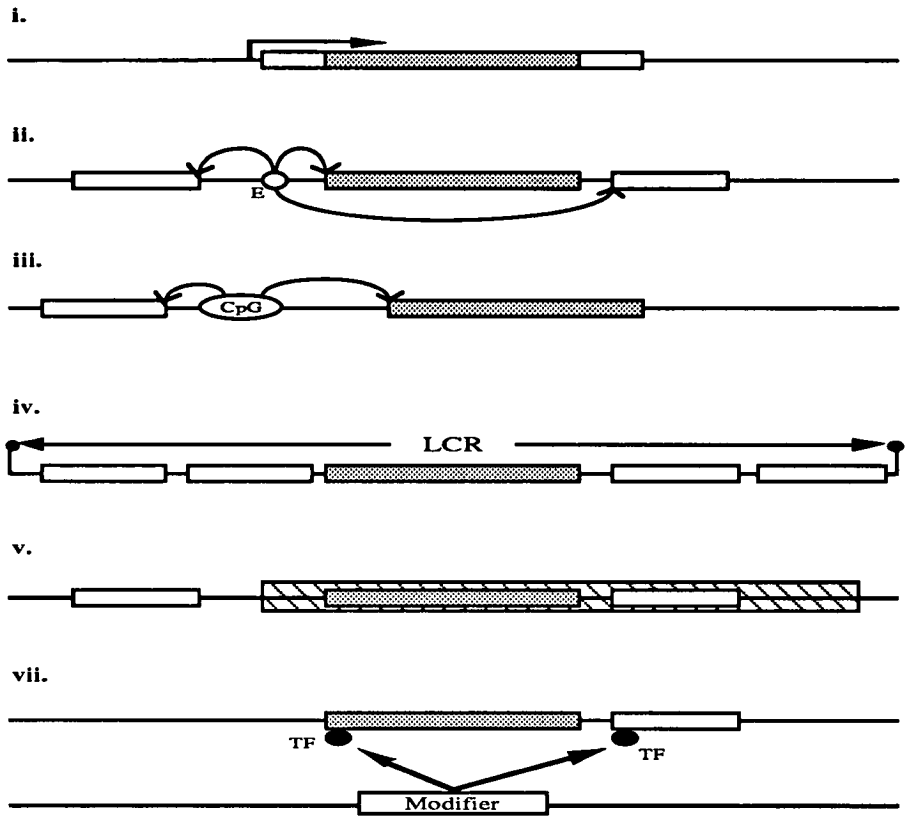


Figure 18. Possible mechanisms for transgene position effects

Open boxes (i-vii), endogenous loci; stippled boxes, transgene array; hatched box (v), imprinted region; E, enhancer; CpG, CpG island; LCR, locus control region, TF, transcription factor(s); Modifier, genotype-specific modifier locus. Taken from Kothary *et al.*, 1989.

transcription factors thought to be associated with them, (iv) within a domain that is regulated by an LCR, (v) into a region in the genome that is imprinted (Surani *et al.*, 1988), (vi) into a region in the genome that is subject to the influence of a genotype-specific modifier(s) and (vii) into or close to heterochromatin (see chapter 6).

Unstable expression occurs when the transgene is expressed in a fashion that is different from the expected temporal or spatial pattern of expression, resulting in qualitative differences concurrent with the quantitative (position effect) differences

between lines. Unstable expression comprises two categories; (1) degrees of expression that differ between lines while remaining stable within a line and (2) variable levels of expression within a line. While there are several examples of aberrant transgene expression between lines in the literature (see chapter 1, section 1.8) to my knowledge there are only three examples of variable transgene expression within a line of mice (McGowan *et al.*, 1989; Allen *et al.*, 1990; Porter & Meyer, 1994). The latter phenomenon may be rarely reported because the integration site and copy number are normally identical within a line and therefore the transgene array is under the same epigenetic influences in all individuals. Porter and Meyer (1994) did not investigate the cause of variable degrees of mosaic expression within line TgUPT9 although they did suggest that the transgene expression may be under the influence of genetic modifier loci. Allen and colleagues described variable expression of an HSV*tk-lac*Z transgene; in this study the variable expression was due to variable levels of uniform expression between individual mice rather than variable degrees of mosaic expression. Conversely, McGowan and colleagues describe variable degrees of mosaic expression of a *hsp68-lacZ* transgene within a line. In both studies the variable expression was shown to be due to the influence of a genotype-specific modifier gene(s) and the expression level was inversely related to the methylation status of the transgenes; low expressors exhibited hypermethylation whereas high expressors showed hypomethylation. The epigenetic modification (gain or loss of methylation) of the transgene locus was cumulative through successive backcrosses to a parental hybrid strain (Allen *et al.*, 1990). Furthermore, the transgene locus was susceptible to the influence of a genotype-specific modifier(s) in other non-parental genetic backgrounds. The effect of the modifier(s) was dependent upon the sex in one strain (BALB/c) therefore demonstrating a genomic imprinting effect.

Evidence for DNA methylation constituting the imprinting mark includes differential methylation of imprinted transgenes (Reik *et al.*, 1987; Sapienza *et al.*, 1987; Swain *et al.*, 1987; Chaillet *et al.*, 1991) and imprinted endogenous genes (Table 18). It is thought the silencing of a locus (containing endogenous genes or transgenes) due to genomic imprinting is an all or nothing process. Therefore integration of the BLG transgene into a region of the genome that was imprinted would result in complete silencing of expression within a line. Hence, integration of the BLG transgene into an imprinted locus can be ruled out as the mechanism for variable expression within lines 7 and 45.

Inactivation of expression as a consequence of methylation could result from the methyl adduct in the major groove interfering with the binding of transcription factors directly or indirectly through the presence of methyl-DNA binding proteins or specific repressor proteins that render the DNA inaccessible to transcription factors (Iguchi-Ariga & Schaffner, 1989; Meehan *et al.*, 1989; Boyes & Bird, 1991; Ehrlich & Ehrlich, 1993), causing heritable gene repression (Holliday & Pugh, 1975). Conversely, it is possible that DNA methylation is a consequence of gene inactivation. In the context of variable transgene expression in lines 7 and 45, a study of the methylation status of the transgene would not reveal the cause of unstable expression. Instead, a genetic backcross was performed to founder lines to investigate whether a genotype-specific modifier gene(s) was involved in variable expression.

Table 18. Imprinted endogenous genes

Maternal Allele Expression*		Paternal Allele Expression*	
Gene	Reference	Gene	Reference
Igf2r	Barlow *et al.*, 1991	*Igf2*	DeChiara *et al.*, 1991
H19	Bartolomei *et al.*, 1991	*Mas*	Villar & Pedersen, 1994
Mash2	Guillemot *et al.*, 1995	*Snrpn*	Leff *et al.*, 1992
$p57^{KIP2}$	Hatada & Mukai, 1995	*U2afbp-rs*	Hayashizaki *et al.*, 1994
		ins1/ins2	Giddings *et al.*, 1994
		IPW	Wevrick *et al.*, 1994
		Peg1/Mest	Kaneko-Ishino *et al.*, 1995

* Imprinted genes display parent-specific monoallelic expression (Surani, 1994; Barlow, 1995), although often not at all developmental stages (Szabo & Mann, 1995). *Igf2r*, insulin-like growth factor 2 receptor; *H19*, encodes an RNA with unknown function; *Mash2*, achaete-scute homology 2; $p57^{KIP2}$, a potent tight-binding inhibitor of several G_1 cyclin/Cdk complexes; *Igf2*, insulin-like growth factor 2; *Mas*, a protooncogene; *Snrpn*, small nuclear ribonucleoprotein; *U2afbp-rs*, human U2af (a snrpn) binding protein-related sequence; *ins1* and *2*, mouse insulin 1 and 2; *IPW*, imprinted gene in the Prader-Willi syndrome region; *Peg1* (also known as *Mest*), paternally expressed gene 1.

4.2. RESULTS

4.2.1. A backcross from line 7 mice into founder strains

The transgenic mice studied (lines 7, 14 and 45) were of a C57BL/6 x CBA hybrid genetic background. The differential levels of expression observed among individuals in a particular line may be due to segregation of alleles

of a modifier gene(s) that modulate transgene expression. This type of phenomenon has been and described and examined by McGowan *et al.* (1989) and Allen *et al.* (1990). Engler *et al.* (1991) also characterised strain-dependent methylation of a transgene and were able to map the position of the modifier to a specific locus *Ssm-1*.

To address the possibility of a strain effect, line 7 (the most variable line) mice were successively backcrossed to pure-bred CBA and C57BL/6 animals for four generations (Fig. 19).

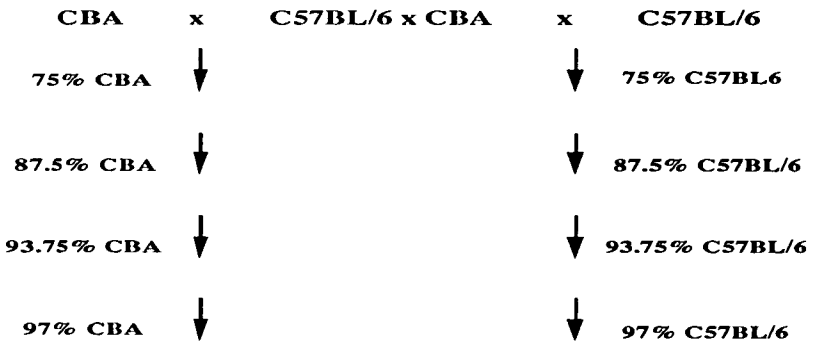


Figure 19. Breeding line 7 mice to CBA and C57BL/6 founder strains

The first cross from mixed background animals involved 8 homozygous line 7 female mice and 8 CBA male animals; similarly 8 homozygous line 7 female mice were crossed to 8 C57BL/6 male animals (16 matings in total). Hemizygous transgenic female offspring were mated once they were 6-8 weeks old and milk was collected at day 11 of lactation. This provided the first set of milk samples (backcross population: 75% CBA or 75% C57BL/6) and established 6 hemizygous sub-lines in the CBA background and 8 sub-lines in the C57 background (Tables 19 and 20). Each of the founder mothers was allocated a number (CBA 1-8 and C57 9-16) and the subsequent female offspring were designated numbers (1-5) after their

mothers number (e.g., CBA 3 produced four transgenic females: 3.1, 3.2, 3.3 and 3.4.).

The backcross experiment proceeded for another three crosses and all subsequent progeny were screened by PCR analysis to identify offspring carrying the transgene. On average 50% of progeny carried the transgene from a cross involving a transgenic (hemizygous) female with a non-transgenic inbred male. Each backcross population was maintained in sufficient numbers so as to provide approximately 20 milking females per group, with the exception of the fourth backcross populations which were deliberately scaled down in preparation for line maintenance. Due to limiting numbers and poor breeding of the inbred lines, some sub-lines were lost. However BLG protein expression levels were obtained from 11 third and 7 fourth backcross sub-lines. Therefore BLG protein expression levels for an individual mouse in the fourth backcross population could be compared with the levels from its ancestors (e.g., CBA 3.1.2.3.1. to CBA 3.1.2.3. to CBA 3.1.2. to CBA 3.1. [Table 19]).

In theory the genetic background of each backcross is 75%, 87.5%, 93.75%, and 96.9% either CBA or C57BL/6. This experiment resulted in >150 mouse milk samples which were analysed by SDS-PAGE as described in chapter 3. The results are presented in tables 19 and 20. The BLG expression levels from a C57BL/6 x CBA mixed genetic background (taken from the first lactation data in chapter 3) were plotted beside those from successive backcross populations to give the relative distribution of BLG expression ranges in successive backcross populations (Fig. 20). There appears to be a trend towards higher expression in the C57 background and lower expression in the CBA background implying a slight genetic background effect. However there is still a considerable range of expression levels in both the

Table 19. BLG expression levels in line 7 animals in a CBA backcross

B*.1	*BLG*	B*.2	*BLG*	B*.3	*BLG*	B*.4	*BLG*
1.1.	*9.6*	1.1.1.	*8.3*	1.1.1.3.	*9.4*		
1.2.	*5.4*	1.1.4.	*5.9*	1.2.1.1.	*5.0*		
1.3.	*3.0*	1.2.1.	*5.0*	1.2.1.2.	*7.5*		
1.4.	*5.7*	1.3.1.	*5.3*	1.3.3.1.	*6.6*		
		1.3.3.	*7.2*	1.4.1.3.	*5.7*		
		1.4.3.	*9.3*				
2.1.	*7.6*	2.1.1.	*6.8*	2.1.3.1.	*2.1*		
2.3.	*6.0*	2.1.3.	*4.0*	2.3.5.4.	*5.0*		
2.4.	*11.7*	2.3.5.	*6.4*	2.3.5.5.	*9.0*		
2.5.	*5.9*	2.5.4.	*6.7*	2.5.4.3.	*8.2*		
3.1.	*7.3*	3.1.2.	*14.4*	3.1.2.3.	*9.7*	3.1.2.3.1.	*12.0*
3.2.	*6.0*	3.2.3.	*5.9*	3.1.2.4.	*8.6*	3.2.3.3.1.	*9.2*
3.3.	*3.5*	3.3.3.	*4.7*	3.2.3.3.	*3.5*	3.2.3.3.2.	*5.0*
3.4.	*4.8*			3.2.3.4.	*1.7*		
		6.1.4.	*6.0*	3.3.3.1.	*6.1*		
6.1.	*11.2*	6.2.2.	*6.5*				
6.2.	*5.7*	6.3.1.	*10.6*	6.1.4.1.	*10.9*	7.2.2.1.1.	*4.1*
6.3.	*7.0*	6.3.2.	*6.7*	6.3.1.2.	*11.5*	7.2.2.1.2.	*5.4*
						7.2.2.1.3.	*12.2*
		7.1.1.	*4.8*	7.1.1.4.	*7.4*	7.2.2.1.6.	*17.5*
7.1.	*3.3*	7.1.3.	*4.2*	7.2.2.1.	*5.4*	7.2.2.1.8.	*7.6*
7.2.	*3.8*	7.2.2.	*7.0*			7.2.2.1.10.	*6.6*
7.3.	*11.3*	7.3.1.	*3.2*			7.2.2.1.11.	*4.3*
				8.1.1.1.	2.5		
8.1.	*4.7*	8.1.1.	*2.3*	8.1.2.1.	4.5	8.1.2.1.2.	*4.3*
8.2.	*6.1*	8.1.2.	*4.4*	8.1.2.2.	*4.4*	8.1.2.2.1.	*9.2*
		8.2.1.	*2.9*	8.2.1.2.	*3.4*	8.1.2.2.4.	*21.0*

*B, backcross; B1-4, mice of four successive crosses into CBA parental strain (parental genetic background contributions are respectively 75%, 87.5%, 93.75% and 97%); BLG, BLG protein expression levels in mg/ml.

Table 20. BLG expression levels in line 7 animals in a C57BL/6 backcross

B*.1	*BLG*	B*.2	*BLG*	B*.3	*BLG*	B*.4	*BLG*
9.1.	*8.3*	9.2.2.	*14.1*	10.3.1.3.	*17.4*		
9.2.	*11.3*	9.3.5.	*8.5*	10.3.2.5.	*8.2*		
9.3.	*12.7*			10.4.1.1.	*7.8*	10.4.1.1.1.	*11.0*
		10.1.1.	*5.1*	10.4.1.2.	*8.8*	10.4.1.2.2.	*9.8*
10.1.	*9.8*	10.3.1.	*14.4*	10.5.1.1.	*17.7*	14.4.1.2.3.	*11.3*
10.3.	*6.0*	10.3.2.	*15.6*	10.5.1.2.	*5.3*	10.5.1.2.2.	*14.1*
10.4.	*11.0*	10.4.1.	*12.5*	10.5.2.1.	*4.3*		
10.5.	*13.5*	10.5.1.	*15.7*	10.5.2.2.	*6.5*		
11.2.	*6.8*	10.5.2.	*8.1*				
11.3.	*9.8*			12.1.1.1.	*8.5*		
12.1.	*10.5*	12.1.1.	*5.0*				
12.2.	*6.7*			13.2.1.1.	*2.6*	13.2.1.1.3.	*6.0*
		13.2.1.	*4.7*	13.2.1.2.	*5.2*	13.2.1.1.4.	*17.5*
13.2.	*8.1*	13.2.2.	*14.9*	13.2.2.2.	*11.4*	13.2.1.2.1.	*10.3*
13.4.	*6.3*	13.4.3.	*5.7*	13.4.3.4.	*12.2*		
14.1.	*11.0*	14.1.1.	*8.0*	14.1.1.3.	*4.4*		
14.2.	*5.7*	14.2.5.	*8.7*	14.2.5.2.	*15.7*	14.2.5.2.2.	*11.3*
15.1.	*3.5*	15.1.4.	*6.0*				
15.2.	*10.0*	15.2.1.	*9.2*				
15.3.	*8.0*	15.2.2.	*4.4*			16.1.3.1.1.	*13.6*
		15.3.1.	*13.7*	16.1.2.1.	*14.0*	16.1.3.1.2.	*16.1*
16.1.	*13.7*			16.1.3.1.	*21.7*	16.1.3.2.2.	*23.1*
16.2.	*10.7*	16.1.2.	*11.7*	16.1.3.2.	*4.8*	16.1.3.2.3.	*19.5*
		16.1.3.	*11.0*	16.2.1.1.	*10.9*	16.2.1.1.1.	*17.8*
		16.2.1.	*9.4*	16.2.1.2.	*7.3*	16.2.1.2.2.	*11.7*

*B, backcross; B1-4, mice of four successive crosses into C57BL/6 parental strain (parental genetic background contributions are respectively 75%, 87.5%, 93.75% and 97%); BLG, BLG protein expression levels in mg/ml.

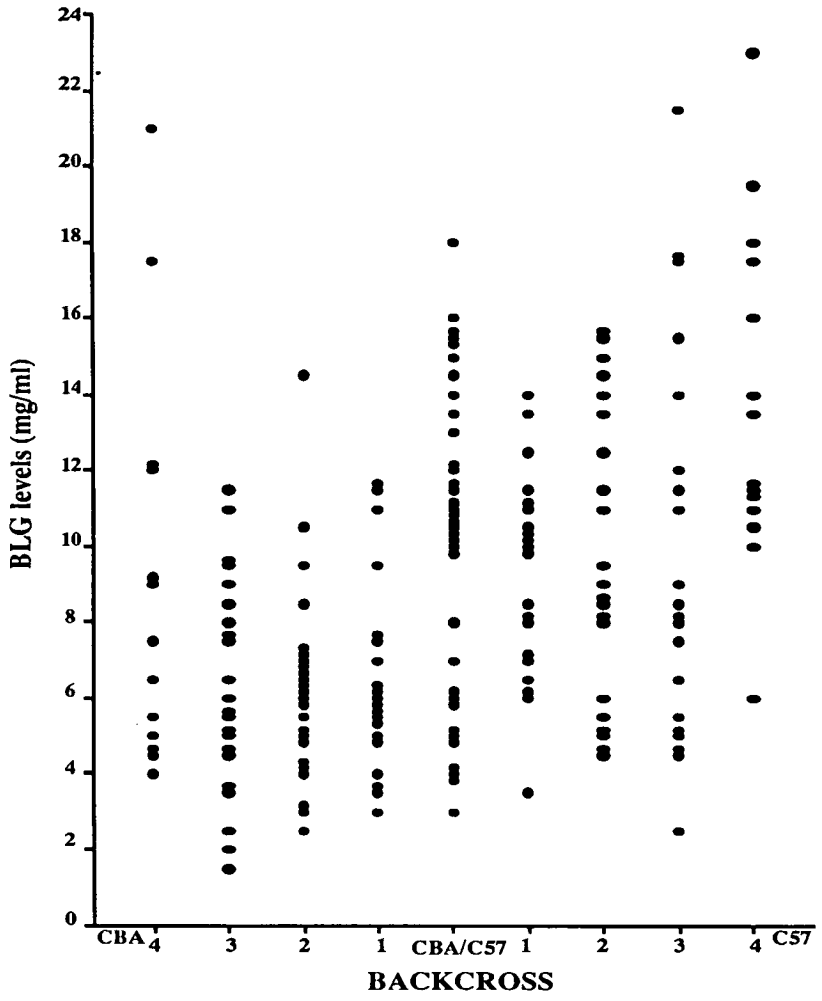


Figure 20. BLG expression levels in milk of line 7 animals; effect of genetic background.

Line 7 (C57BL/6 x CBA) animals were separately backcrossed to the two parental strains; BLG levels were recorded. CBA/C57 group (centre), individual BLG levels in mice of C57BL/6 x CBA background (50% genetic contribution from each parental strain); CBA 1-4 (left) and C57 1-4 (right) are mice of four successive crosses to each parental strain (parental genetic background contributions are respectively 75%, 87.5%, 93.75% and 97%).

C57BL/6 and CBA inbred populations, even after the fourth backcross. Others have described strain effects after only one cross into a strain exerting a modifier

influence (McGowan *et al.*, 1989; Allen *et al.*, 1990; Engler *et al.*, 1991). In this study there was no evidence of a clear segregation into high and low expressing groups, a result that would be consistent with the presence of a modifier gene(s).

The first backcross towards a C57 genetic background resulted in a significant reduction in the variance ($p < 0.05$) when compared with the mixed C57BL/6 x CBA population (Tables 21 and 22); the mean was not significantly different ($p > 0.05$). Neither the variance or the mean were significantly different in the second and third C57 backcross populations. There was still no significant difference in the variance between the fourth C57 backcross population and the mixed C57BL/6 x CBA population although there was a significant increase in the mean. In the CBA backcross experiment both the variance and the mean were significantly different ($p < 0.05$) between the first, second and third backcross CBA populations and not significantly different ($p > 0.05$) in the fourth backcross population when compared with the C57BL/6 x CBA population (Tables 21 and 22). The striking difference in distribution of expression levels between the third and fourth CBA backcross populations could be a true effect of the backcross or it could be due to a limited number of samples, i.e., some high expressors would arise in the third backcross population if more samples were measured. Because the backcrosses that preceeded the fourth cross introduced a far greater theoretical genetic difference without such a large effect on the expression level distribution it was concluded that the difference is more likely to be due to a limited number of samples.

Major alterations in transgene expression level have been documented after just one cross into an inbred line (McGowan *et al.*, 1989; Allen *et al.*, 1990; Engler *et al.*, 1991). It was predicted that the presence of a single or multiple loci would be

Table 21. BLG expression levels within eight backcross populations

Strain	First backcross (75%) mean ± *s.d.*	*c.v.*	Second backcross (87.5%) mean ± *s.d.*	*c.v.*	Third backcross (93.75%) mean ± *s.d.*	*c.v.*	Fourth backcross (97%) mean ± *s.d.*	*c.v.*
C57	9.2 ± 2.8	30% (n = 20)	9.8 ± 3.9	40% (n = 21)	9.7 ± 5.3	54% (n = 20)	13.8 ± 4.5	33% (n = 14)
CBA	6.5 ± 2.6	41% (n = 20)	5.8 ± 2.0	34% (n = 24)	6.3 ± 2.8	45% (n = 22)	9.1 ± 5.3	58% (n = 13)

s.d. = standard deviation; *c.v.* (coefficient of variation) = 100 x *s.d.* / mean; n = number of mice; BLG mean values are absolute values (mg/ml)

C57BL/6 x CBA mixed background : mean ± *s.d.* = 10.1 ± 4.2, *c.v.* = 41% (n = 35)

revealed by major alterations in both the variance and mean values for expression levels over three generations. Because no such alterations were observed by the fourth bachcross, it was concluded that the variable expression of the transgene within line 7 is not due to the influence of a genotype-specific modifier(s).

Table 22. Comparison of the variance and mean values between the mixed C57BL/6 x CBA genetic background and successive backcross populations

	C57BL/6		CBA	
Backcross	variance	mean	variance	mean
1	$p < 0.05$	$p > 0.05$	$p < 0.05$	$p < 0.05$
2	$p > 0.05$	$p > 0.05$	$p < 0.05$	$p < 0.05$
3	$p > 0.05$	$p > 0.05$	$p < 0.05$	$p < 0.05$
4	$p > 0.05$	$p < 0.05$	$p > 0.05$	$p > 0.05$

4.2.2. BLG expression levels through the backcross

The maintenance of sub-lines through the backcross enabled a comparison of BLG expression levels through sub-lines within each backcross. There were no clear trends in expression levels through the backcross. Expression levels in female progeny of high or low expressing mothers appeared to be randomly distributed and there was no evidence of an association between maternal and filial expression levels; for example, a parent expressing BLG protein at 11 mg/ml (C57 16.1.3.) gave rise to siblings with expression levels of 4.8 mg/ml (C57 16.1.3.2.) and 22 mg/ml (C57 16.1.3.1.) [Table 20]. Another striking example of the unpredictability of expression levels through the backcross is the CBA 7.2.2.1. to CBA 7.2.2.1.1-11 pedigree (Table 19). In this case the mother was expressing BLG at only 5.4 mg/ml, below the average for this backcross population (Table 21). She

produced offspring exhibiting almost the complete range of expression levels within the fourth backcross population (4.1 to 17.5 mg/ml; Table 19) and a range of intermediate expressors.

The data demonstrates that although the expression level is fixed within the individual (chapter 3) the level of expression is not predictable through the germline. A mechanism that fulfils these criteria is proposed in chapter 6.

4.3. CONCLUSIONS

A backcross into 97% CBA or 97% C57BL/6 genetic backgrounds had no significant effect on variance of BLG expression in either background; there was no clear segregation into high and low expressors. Furthermore, the expression levels were not predictable through the backcross. In conclusion, while there is a slight strain effect on the mean BLG expression level, the variable expression of the transgene within line 7 is not explained by heterogeneity of genetic background.

CHAPTER 5

BLG AND β–CASEIN EXPRESSION IN MID-LACTATION MOUSE MAMMARY GLAND

5.1. INTRODUCTION

Previous results confirmed that levels of BLG protein and mRNA varied extensively between different line 7 females. At the cellular level one can envisage two modes of variation. First, there could be a uniform level of expression in all the secretory epithelial cells of an individual mammary gland: the overall level of expression varies between individuals (Allen *et al.*, 1990). Second, the level of expression could vary from cell to cell within an individual gland giving rise to a mosaic pattern of expression, the degree of mosaicism varying between individuals (McGowan *et al.*, 1989).

In situ hybridisation to cellular mRNA with ^{35}S-labelled complementary single-stranded ("antisense") RNA probes (Angerer & Angerer, 1991, 1992; Wilkinson, 1992) was employed to discriminate between the above possibilities.

In situ hybridisation was utilised rather than immunohistochemistry for several reasons. Immunohistochemistry using antibodies for BLG has been used to localise BLG protein in the mouse mammary gland (Harris *et al.*, 1991), however the resolution was poor. Also, as milk is secreted (unlike mRNA), contamination within the sections could have been a problem using antibody detection. *In situ* hybridisation allows one to distinguish the cells in a complex tissue that express specific mRNAs. Furthermore, measuring mRNA levels provides a more accurate estimation of gene activity because mRNA accumulates before the protein it

encodes; stable proteins may persist long after the gene is repressed and the mRNA has decayed. Protein antigenicity varies after fixation with different fixatives, and detection of different antigens (i.e., BLG and β-casein) may therefore require different tissue preparations (efficiency of hybridisation of different probes is similar with the same tissue fixation procedures).

The synthesis of single-stranded RNA probes is convenient if one is *au fait* with molecular biology and the precise length of probe can be controlled (an important consideration for probe penetration into the tissue) if the appropriate restriction enzyme sites are available (see materials and methods). Single-stranded probes are more sensitive than double-stranded probes because they do not re-anneal. Most importantly, RNase digestion (see 2.10.10. in chapter 2) of any probe sticking to the tissue (specific hybridisations are not affected because RNase does not digest double-stranded complexes) reduces non-specific background signal. Labelled oligonucleotides are less sensitive because they are shorter than nucleic acid probes.

For my purposes ^{35}S-dUTP was the best labelling method. At present radiolabelled probes are more sensitive than hapten-labelled probes. The increased safety and high stability of hapten-labelled probes was not significant as the tissue *in situ* hybridisation experiment was performed twice only. Although ^{3}H-labelled probes give better resolution of signal they require longer autoradiographic exposures. ^{35}S-labelled probes give about a one cell diameter resolution which was sufficient and required an exposure of 1-2 weeks. ^{32}P-labelled probes have a shorter half-life than ^{35}S-labelled probes (14 days vs. 87 days), give poorer resolution and offer no advantage of speed of results.

In situ hybridisation studies using BLG and β-casein mRNA probes (see materials & methods) were performed on abutting sections from tissue collected at

day 11 of lactation from hemizygous individuals of lines 7, 14 and 45 (see chapter 3) and non-transgenic mice.

The structure of mid-lactation mammary gland tissue is shown in Fig. 21A. The secretory tissue is represented by the alveolus. This has a lining of secretory epithelial cells that express and secrete the milk proteins into the alveolar space. The myoepithelial cells around the outside of the alveolus then contract to squeeze the milk out into the ducts and eventually to the nipple. Therefore, when a cross section is taken through the mammary gland at mid-lactation the main features are cross sections of alveoli and the ductal systems (Fig. 21B).

5.2. RESULTS

5.2.1. Tissue *in* situ hybridisation analysis of mouse mammary tissue sections

As expected, endogenous β-casein mRNA was detected in non-transgenic tissue (Fig. 22A) whereas BLG mRNA was absent (Fig. 22B). Importantly, this demonstrated that the BLG mRNA probe did not bind in a non-specific fashion with the tissue or other mRNA's. Likewise, all sections from lines 7, 14 and 45 probed with BLG and β-casein sense probes (control for non-specific hybridisation) had no radioactive signal associated with them (Fig. 21B)

Endogenous β-casein mRNA was detected throughout the mammary epithelium in all three transgenic lines, demonstrating that essentially all secretary epithelial cells of the gland have the capacity to express a milk protein gene (Fig. 23A, Fig. 24A, Fig. 25A, Fig. 26A). Robinson and colleagues (1995) described mosaic expression (30%) of β-casein up to day 11 of gestation in mice; not until day

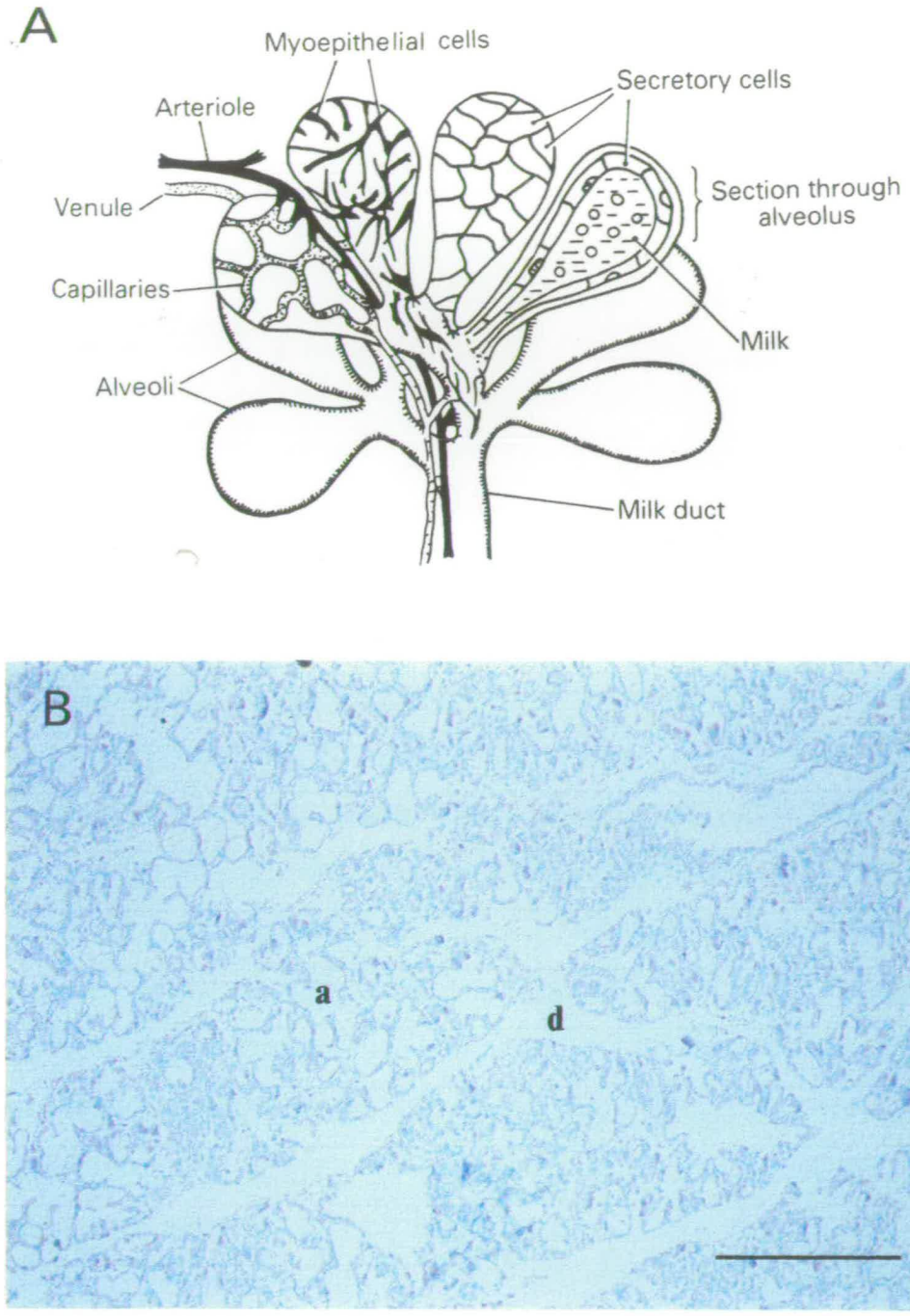


Figure 21. Tissue morphology of the mammary gland.
(**A**) Diagrammatic representation of a cluster of mammary alveoli (adapted from Mepham, 1987). (**B**) Section taken through a mid-lactation mouse mammary gland probed with a β-casein sense (control) RNA probe and stained with methylene blue. The main structural features are the rings of secretory and myoepithelial cells surrounding each alveolar space (*a*) and the ductal systems (*d*). Each image in Figs. 21 to 26 is at x4 magnification and is representative of the hybridisation pattern in the whole section. (Bar = 0.5 mm).

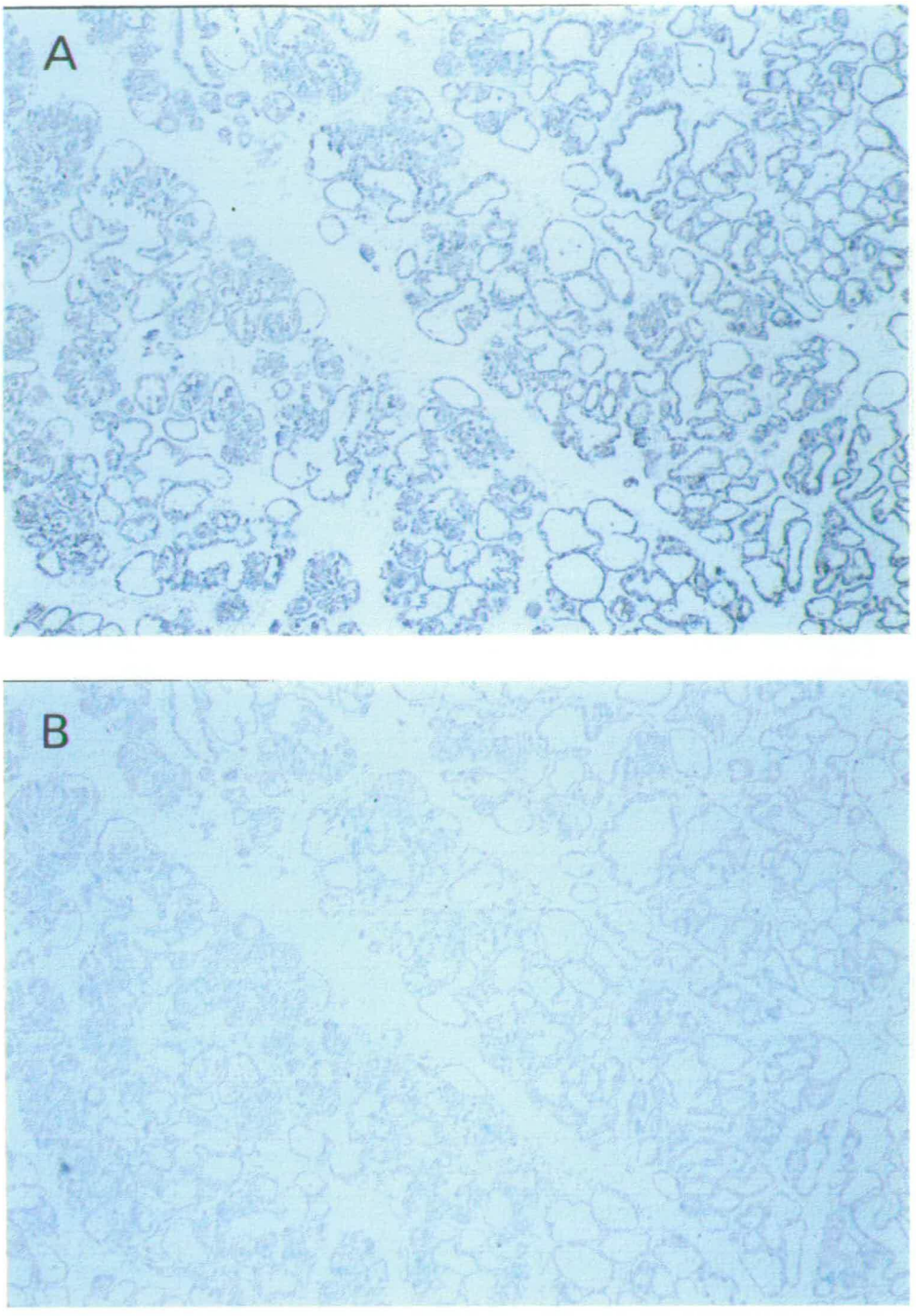


Figure 22. ***In situ*** **hybridisation analysis of mRNA expression patterns in non-transgenic mouse mammary tissue.**
(**A**) β-casein, (**B**) BLG

18 of gestation was uniform expression achieved. My study demonstrates that uniform β-casein expression is also present at day 11 of lactation.

In gland sections from individuals of line 7 expressing the BLG protein at low levels (e.g., B3: see Table 10 in chapter 3), small clusters of cells strongly expressing BLG mRNA were observed, surrounded by non-expressing cells (Fig. 23B). However some individual alveoli contained both expressing and non-expressing cells (e.g., lower left and top right region in Fig. 23B). High expressors from line 7 (e.g., F5) had the opposite pattern of expression showing small regions of cells and even individual alveoli negative for BLG mRNA, surrounded by cells abundantly expressing the transgene (Fig. 24B). Abutting sections probed for BLG and β-casein identified regions and specific alveoli that only expressed β-casein. Due to the complex 3-dimensional arrangement of the mammary epithelium no precise quantification was possible (note, Robinson *et al.*, [1995] do not mention how they quantitated the β-casein expression within the gland). Nonetheless, it did appear that the proportion of positive cells in any one section correlated with the level of BLG in the milk; highest expressors clearly had a greater proportion of cells positive for transgene mRNA. This demonstrated that variable expression was present at the mRNA level in line 7 and that it was differing degrees of mosaic mRNA expression that resulted in the high *c.v.* value (47%) measured by northern blot experiments (see chapter 3; section 3.2.4.).

Line 14 (e.g., I6) was distinctly different because BLG expression appeared to be uniform throughout all the epithelial cells within any one section, in a way that matched the pattern of β-casein expression (Fig. 25A and B). Line 45 (e.g., I2) also exhibited mosaic expression of BLG, although less so than in line 7 (Fig. 26B) which may reflect the higher average expression level in line 45.

The tissue *in situ* hybridisation results from lines 7 and 45 demonstrate that there are regions within the tissue that do not hybridise with the BLG probe.

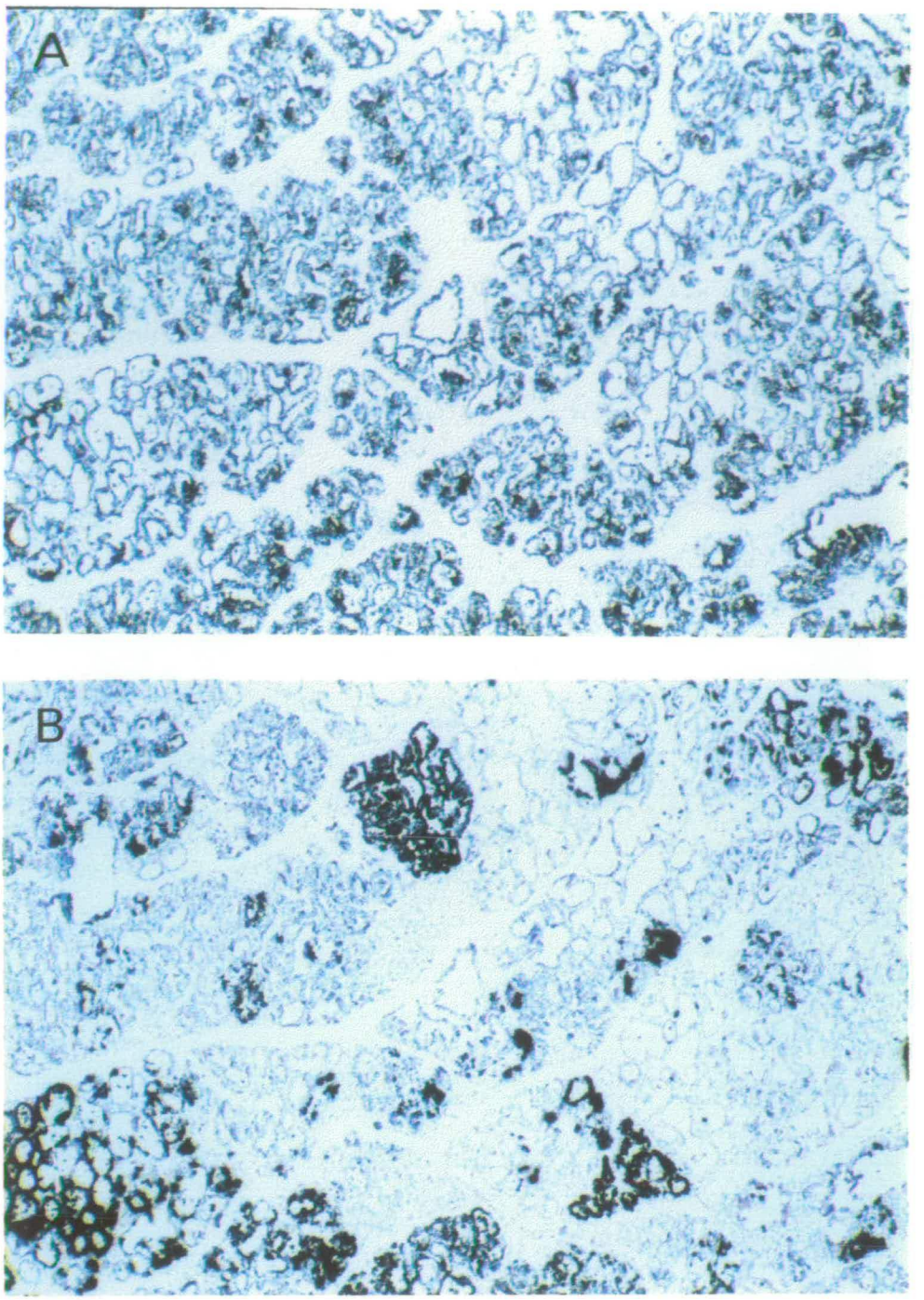


Figure 23. *In situ* hybridisation analysis of mRNA expression patterns in line 7 (low).
(**A**) β-casein, (**B**) BLG (4 mg/ml).

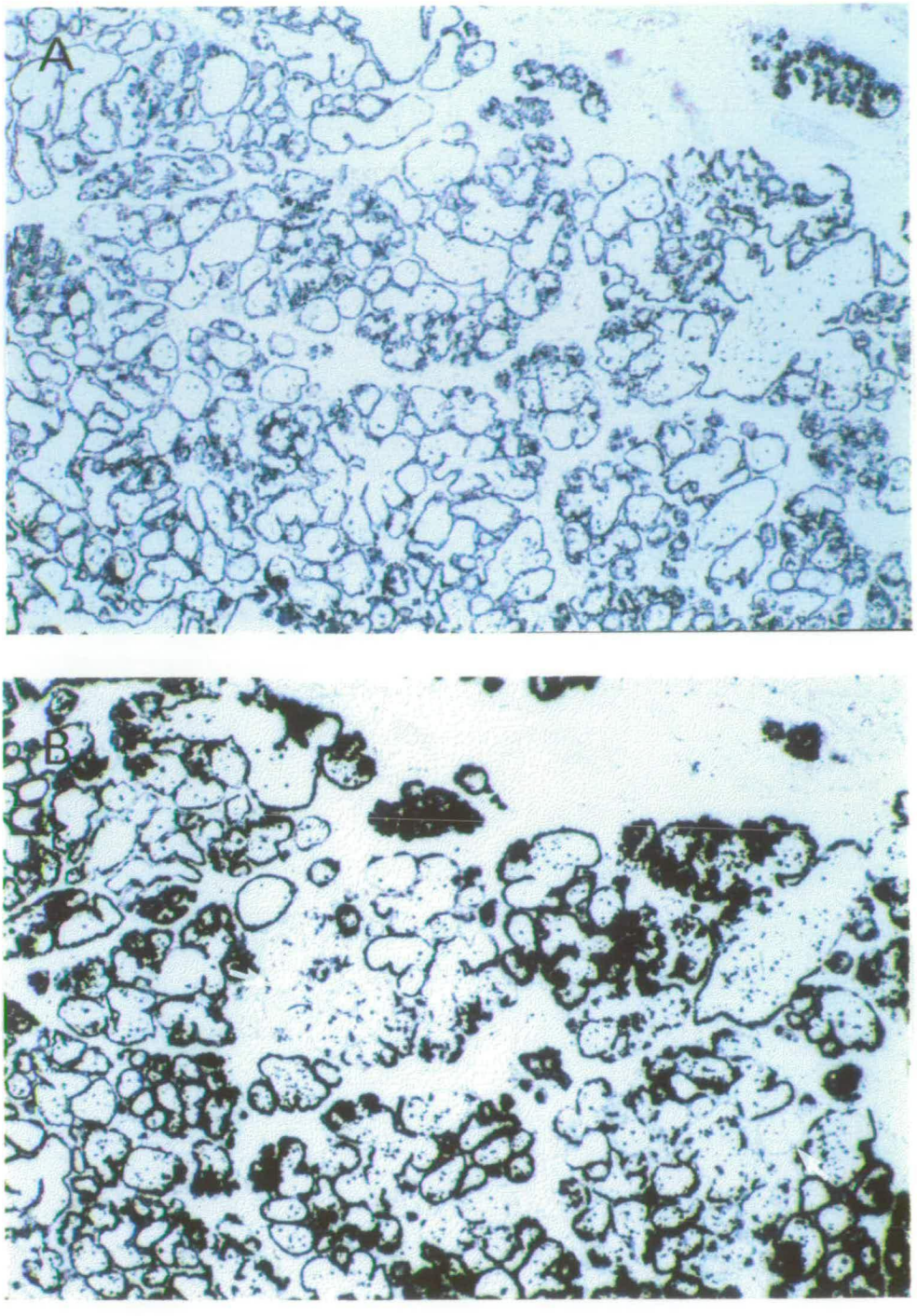


Figure 24. *In situ* hybridisation analysis of mRNA expression patterns in line 7 (high).
(**A**) β-casein, (**B**) BLG (17 mg/ml). White arrows indicate regions (*left*) and individual alveoli (*right*) that are not expressing BLG.

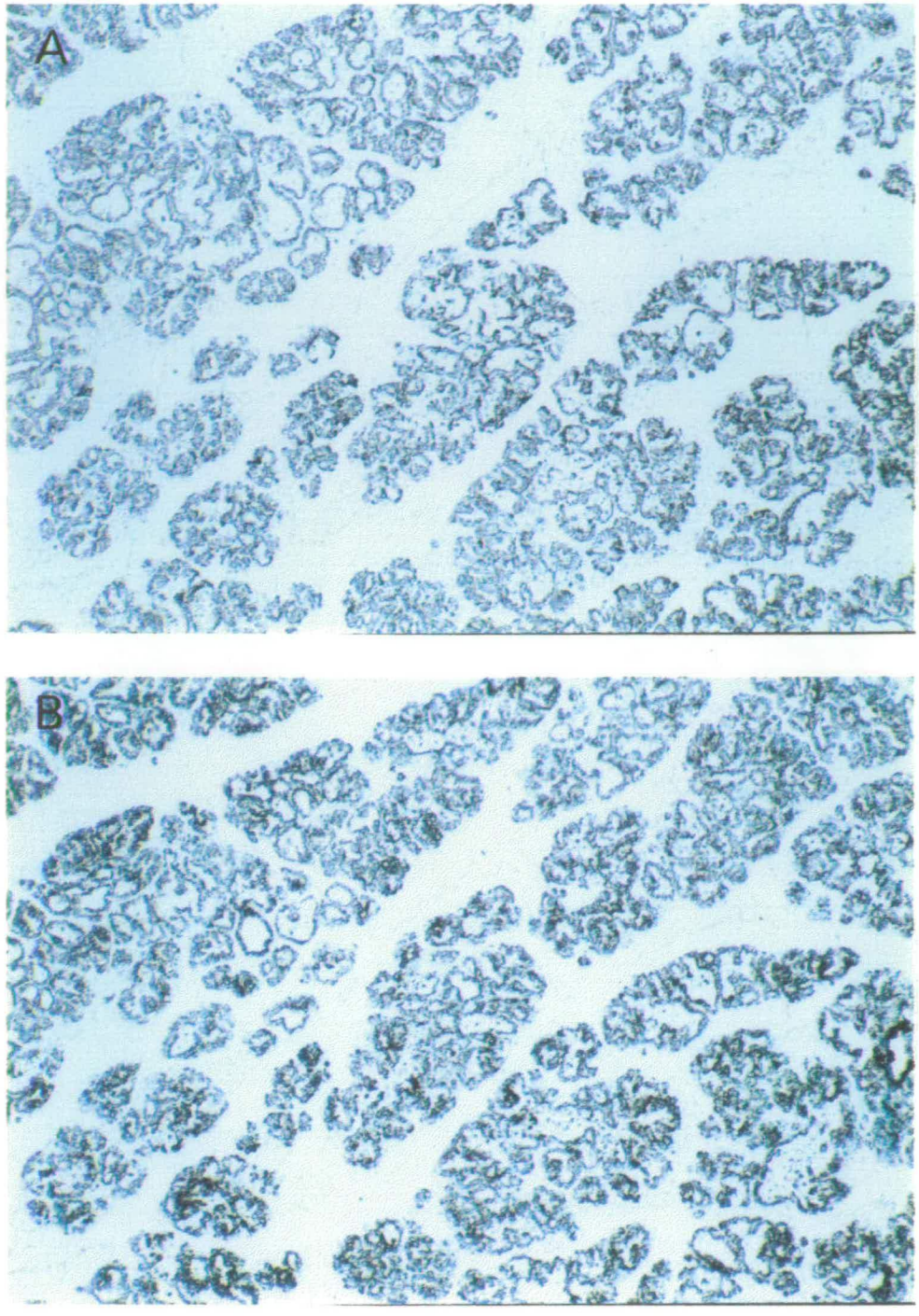


Figure 25. ***In situ*** **hybridisation analysis of mRNA expression patterns in line 14.**
(**A**) β-casein, (**B**) BLG (8 mg/ml).

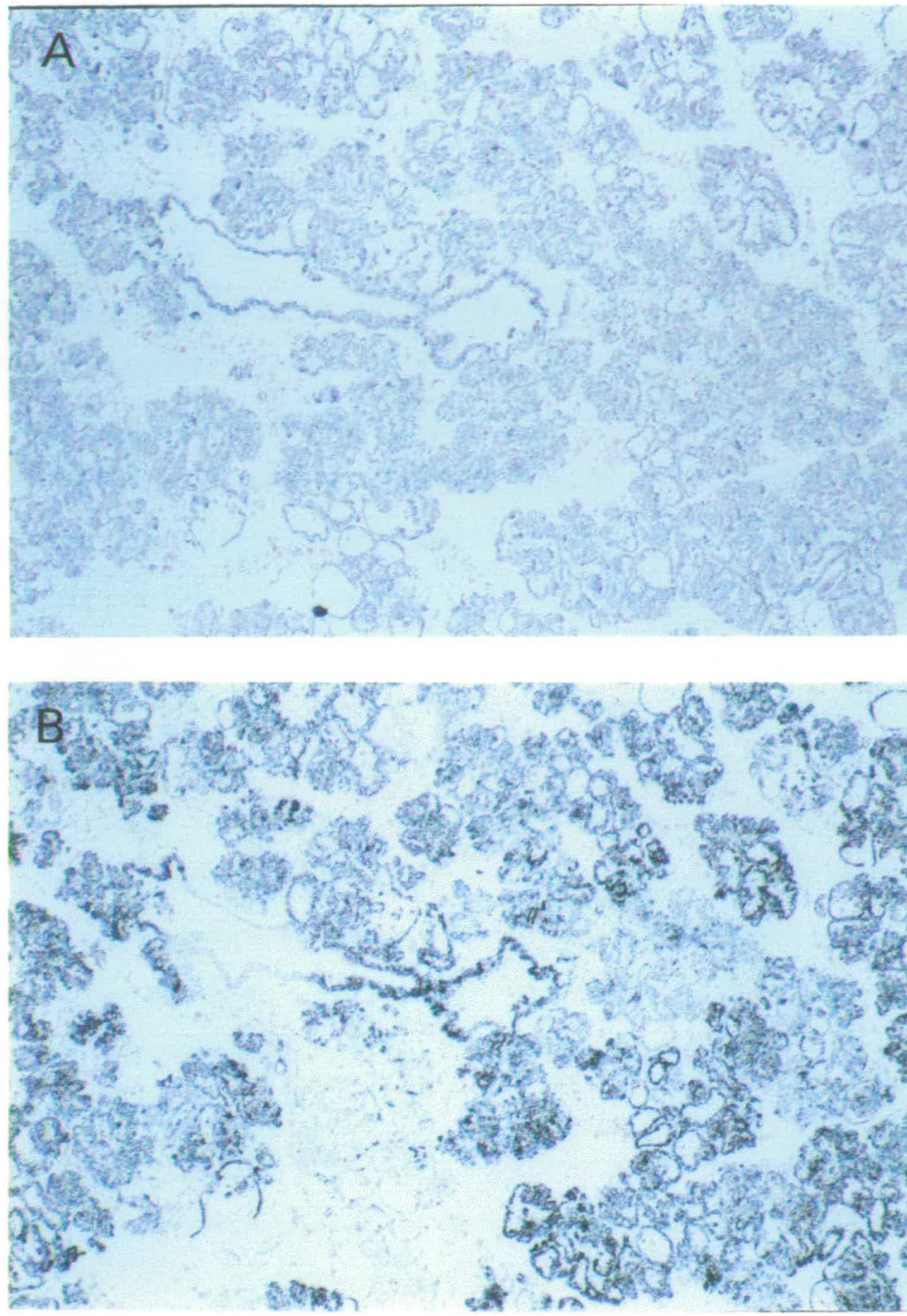


Figure 26. *In situ* hybridisation analysis of mRNA expression patterns in line 45.
(**A**) β-casein, (**B**) BLG (25 mg/ml).

5.3. CONCLUSIONS

Although the tissue *in situ* analysis did not permit precise quantitation of per-cell expression levels, the results suggest that variations in the level of BLG protein and mRNA expression in individuals of lines 7 and 45 (described in chapter 3) are due principally to variations in the proportion of mammary epithelial cells expressing the transgene.

Transgene expression was mosaic in animals of line 7 (and to a lesser extent in line 45) but uniform in all line 14 animals examined. In lines 7 and 14 the introduced transgene is identical, arguing that mosaic expression is due to the location and/or nature of the transgene array.

CHAPTER 6

CHROMOSOMAL LOCALISATION AND MOSAICISM OF TRANSGENE EXPRESSION

6.1. INTRODUCTION

The results presented in chapter 5 demonstrated that BLG mRNA was expressed in a mosaic fashion in the mammary gland in lines 7 and 45. BLG was expressed uniformly throughout the gland in line 14, as was β-casein in all three lines. There are two possible explanations for the lack of BLG expression in some mammary epithelial cells in lines 7 and 45. First, somatic inactivation of the transgene can occur due to rearrangement or deletion of the locus in some cells but not in others (Wilkie *et al.*, 1991; Sandgren *et al.*, 1991). Just such a possibility was addressed in chapter 3 and was demonstrated to be unlikely. Second, the transgene locus could be subject to transcriptional inactivation in some cells but not in others.

Transcriptional inactivation is often associated with methylation and heterochromatic condensation of a locus, best exemplified by the phenomenon of X-inactivation. However, heterochromatinisation is not restricted to the inactive X chromosome; the DNA component of autosomes can be sub-divided into euchromatic and heterochromatic DNA. Euchromatin decondenses during interphase and contains most of an organism's genes. The classical definition of heterochromatin is the region of the chromosome that remains condensed (or heterochromatinised) throughout the cell cycle (Verma, 1988). Constitutive heterochromatin is located near the centromeres, telomeres and the whole of the Y chromosome whereas facultative heterochromatin is associated with the inactive X

chromosome. The centromere is defined cytologically as the primary constriction on the chromosome at a fixed position for a given chromosome.

Transcriptional inactivation of the BLG transgene locus could be due to heterochromatic silencing, since this mechanism fulfills the following relevant criteria: (i) The inactivation process is capable of inactivating a large region of DNA; the transgene array in line 7 is >400 kb (~25 copies of a 16.2 kb fragment). Heterochromatinisation is a process thought to be involved in maintaining the inactive state of the inactive X chromosome (150 Mb) in female eukaryotic cells (Lyon, 1961; Barr & Carr, 1962; Kanda, 1973; Rastan *et al.*, 1980). (ii) The inactivation process is stable within the individual; expression levels are fixed across two lactation cycles (chapter 3). Once inactivated the X chromosome remains heterochromatinised in a permanent fashion and is heritable in a clonal fashion (Gartler & Riggs, 1983; Lyon, 1991; Brown & Willard, 1994). (iii) The expression level in a parent mouse is not transmitted to her offspring (chapter 4); therefore the fixed level of expression within the individual may be determined through the germline or during stem cell development. The state of X chromosome inactivation (heterochromatinisation), although stable within an individual, is not transmitted through the germline (Gartler *et al.*, 1975; Kratzer & Chapman, 1981; McMahon *et al.*, 1981; Monk & McLaren, 1981). (iv) The inactivation process is not 100% efficient (some degree of mosaicism was always present); abolition of expression was never observed within line 7 (chapter 5). Finally, mosaic or "variegated" expression of endogenous genes (Muller, 1930; Lewis, 1950; Baker, 1968; Spofford, 1976; Karpen, 1994) and transgenes (Dorer & Henikoff, 1994; Wallrath & Elgin, 1995; Sable & Henikoff, 1996) has been described in *Drosophila* and is termed position-effect variegation (PEV). The phenomenon involves a rearrangement of the chromosome which brings a locus (the first and best documented is the dominant allele of the *white* [w^+] gene) close to the centromere. The rearrangement is thought

to cause the heterochromatinisation of the w^+ locus in some cells but not in others resulting in a mosaic pattern of white (mutant) and red (wild type) areas on the eye. PEV has also been documented with sequences integrated close to telomeres (Gottschling *et al.*, 1990; Sandell & Zakian, 1992) or within centromeres (Allshire *et al.*, 1994) in yeast. Similarly, X-autosome chromosome translocations in mouse result in silencing of autosomal genes in proximity to the breakpoint (Cattanach, 1963; Russell, 1963; Eicher, 1970; Cattanach, 1974) implying that PEV may be a ubiquitous silencing phenomenon. The silencing process can affect genes up to 20 kb away from the heterochromatin/euchromatin breakpoint in yeast (Gottschling *et al.*, 1990) and up to 2 Mb in *Drosophila* (Weiler & Wakimoto, 1995). In summary, PEV involves a chromosomal disruption, proximity of the locus to constitutive (centromere or telomere) or facultative (inactive X chromosome) heterochromatin, heterochromatinisation of the locus and mosaic expression of the locus.

Mouse chromosomes are acrocentric (with the exception of the Y chromosome), with centromeres located at one end of the chromosome arms (Evans, 1981; Roberts *et al.*, 1988). The centromeric regions of mammalian chromosomes contain tandemly repeated satellite sequences and display characteristics of constitutive heterochromatin (e.g., late replicating and a paucity of genes). Satellite sequences in *Mus musculus* are relatively AT rich and are methylated at their CpG dinucleotides, indicative of the lack or low level of transcriptional activity. Two classes of centromeric repetitive sequences have been characterised in mouse, the minor and major (gamma) satellite sequences (Hastie, 1996). In *Mus musculus*, the minor satellite is organised as tandem repeat arrays of a 120 bp monomer sequence, may constitute up to 0.2% of its genome (the haploid copy number has been estimated to be at least 50,000) and an average of ~300 kb of satellite sequence per chromosome (Pietras *et al.*, 1983, Wong & Rattner, 1988; Kipling *et al.*, 1991). Minor satellite sequences are found at the primary constriction on the autosomes and

the X chromosome, but absent from the Y chromosome. These sequences are thought to be organised in largely uninterrupted blocks of tandem repeats (Kipling *et al.*, 1991). Cytologically, minor satellite sequences appear to be localised at the kinetochore (Wong & Rattner 1988). Closely related species of *Mus* differ in the amount and type of satellite sequence present in their centromeric heterochromatin (Wong *et al.*, 1990; Kipling *et al.*, 1994; Aker & Huang 1996). Major satellite DNA comprises a 234 bp element that is tandemly repeated 10^6 times and is localised to the region of the chromosome that immediately flanks the kinetochores (Singer, 1982; Pietras *et al.*, 1983; Lica *et al.*, 1986; Radic *et al.*, 1987; Wong & Rattner, 1988; Joseph *et al.*, 1989; Moens & Pearlman, 1990; Wong *et al.*, 1990; Matsuda & Chapman, 1991). The heterogeneity of the amount of heterochromatic DNA between closely related species (e.g., *Mus musculus, Mus caroli* and *Mus spretus*) and the repetitive nature of the sequences has led to speculation that satellite DNA has no essential function and is "junk" DNA. However studies on *Drosophila* demonstrate that many essential functions reside in heterochromatin including the sequences required for faithful chromosome pairing and segregation during mitosis and meiosis (i.e., the kinetochore/centromere; Murphy & Karpen, 1995b). Functional analysis of the *Drosophila* centromere suggests that the kinetochore is involved in spindle attachment and the flanking sequences are required for sister-chromatid cohesion and for the generation of anti-poleward forces during mitosis and meiosis (Murphy & Karpen, 1995a,b). The precise function(s) of mouse centromeric regions have yet to be dissected although regions may be structurally and/or functionally conserved between higher eukaryotes (Nicklas, 1989).

In this study there was a disruption of the normal chromosomal architecture due to transgene integration and mosaic expression of the transgene locus was observed in two out of three lines. I was therefore interested in determining whether the transgene array had integrated at or close to a centromere in lines 7 and 45 and

whether the array exhibited heterochromatinisation, akin to the observations described for PEV in *Drosophila*.

6.2. RESULTS

6.2.1. G-band and FISH analysis of lines 7, 14 and 45

Fluorescent *in situ* hybridisation (FISH) analysis was performed on chromosome spreads prepared from line 7, 14 and 45 female animals to determine the chromosomal localisation of the BLG transgene integration site.

Mouse chromosomes are acrocentric and are smaller than human chromosomes which makes them more difficult to karyotype. Furthermore, several mouse chromosomes are difficult to distinguish (e.g., chromosomes 9 and 13), even from very good chromosome preparations. Fifteen good quality G-banded metaphase chromosome spreads were analysed from each line. Karyotyping was performed partly by specialised software (see materials and methods) and partly by eye (with assistance from Muriel Lee, MRC Human Genetics Unit, Edinburgh). Digitized chromosome images were stored and the slide coordinates of each metaphase spread noted. These images were duplicated; one was used to karyotype the chromosomes and the other retained for comparison with the FISH image. The chromosome G-bands were removed by acid/alcohol hydrolysis and FISH analysis was performed using a BLG probe (see materials and methods). Fluorescent chromosome images were captured using a confocal microscope.

The BLG probe did not hybridise with chromosomes prepared from non-transgenic mice (Fig. 27). The BLG probe did hybridise with one or two chromosomes (depending on if the mouse was hemizygous or homozygous for the

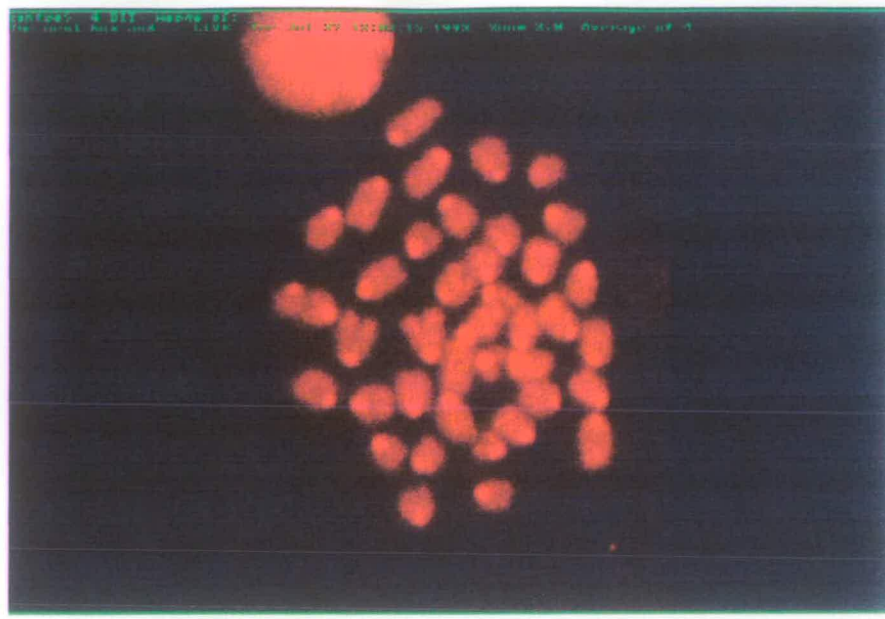


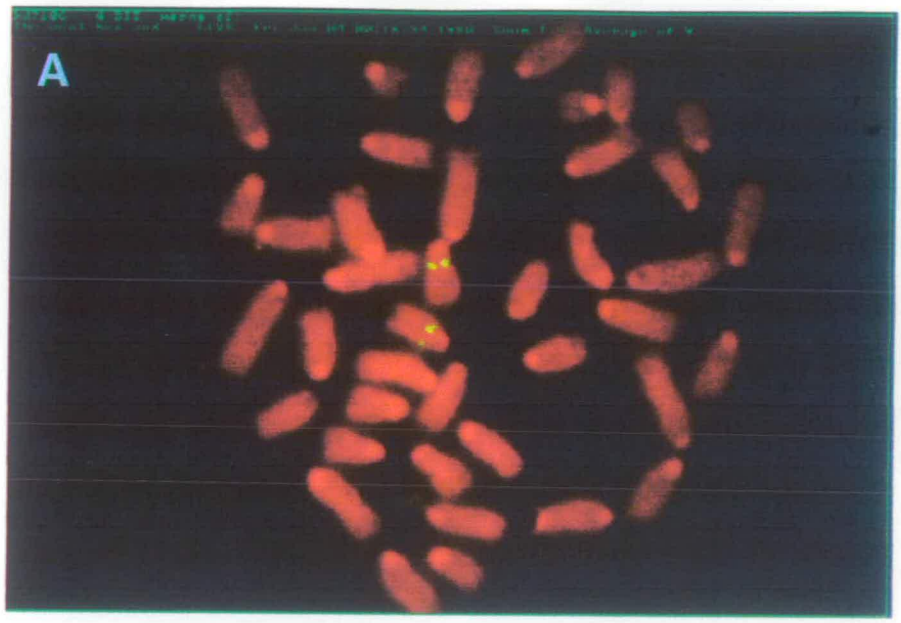
Figure 27. FISH analysis of non-transgenic mouse metaphase chromosomes. Chromosomes from non-transgenic C57BL6 x CBA mice were hybridised overnight with a biotinylated BLG probe and incubated with avidin-FITC. Chromosomes were stained red using propidium iodide.

transgene) in 100% of chromosome spreads prepared from transgenic mice. This enabled a comparison of the FISH images with the recorded G-banded images, the construction of karyotypes and hence the identification of the chromosomes carrying the transgene (Figs. 28-30). The process was repeated 15 times with the best G-banded chromosome spreads for each line. Karyotype analysis systematically identified the same chromosome harbouring the transgene within each and demonstrated that the transgene had integrated into a different chromosome in each line. The BLG transgene locus was located on chromosome 15 in line 7, chromosome 7 in line 14 and on chromosome 5 in line 45. Interestingly, the BLG locus had integrated mid-arm in line 14 but was close to the centromere in lines 7 and 45 (Figs. 28-30).

6.2.2. Position of the transgene array relative to satellite DNA

To refine the localisation of the transgene array relative to the centromere a double *in situ* hybridisation was performed using a BLG probe in conjunction with a major satellite (196) DNA probe. Major satellite DNA can comprise as much as 10% of the mouse genome and is the main DNA component of the centromeres (Singer, 1982). Therefore by combining both probes in an *in situ* hybridisation study it was possible to locate the position of the transgene integration site in these two lines relative the centromeres.

The major satellite DNA probe hybridised to the centromeric regions of metaphase chromosomes (Fig. 31). In all line 7 chromosome spreads the BLG locus was close to but outwith major satellite DNA (Fig 31A). The BLG signal was quenched to determine if any translocated satellite DNA was present at the BLG locus (Fig 31B). There was no evidence of translocated satellite DNA at the BLG transgene locus even when the satellite signal was enhanced.



B

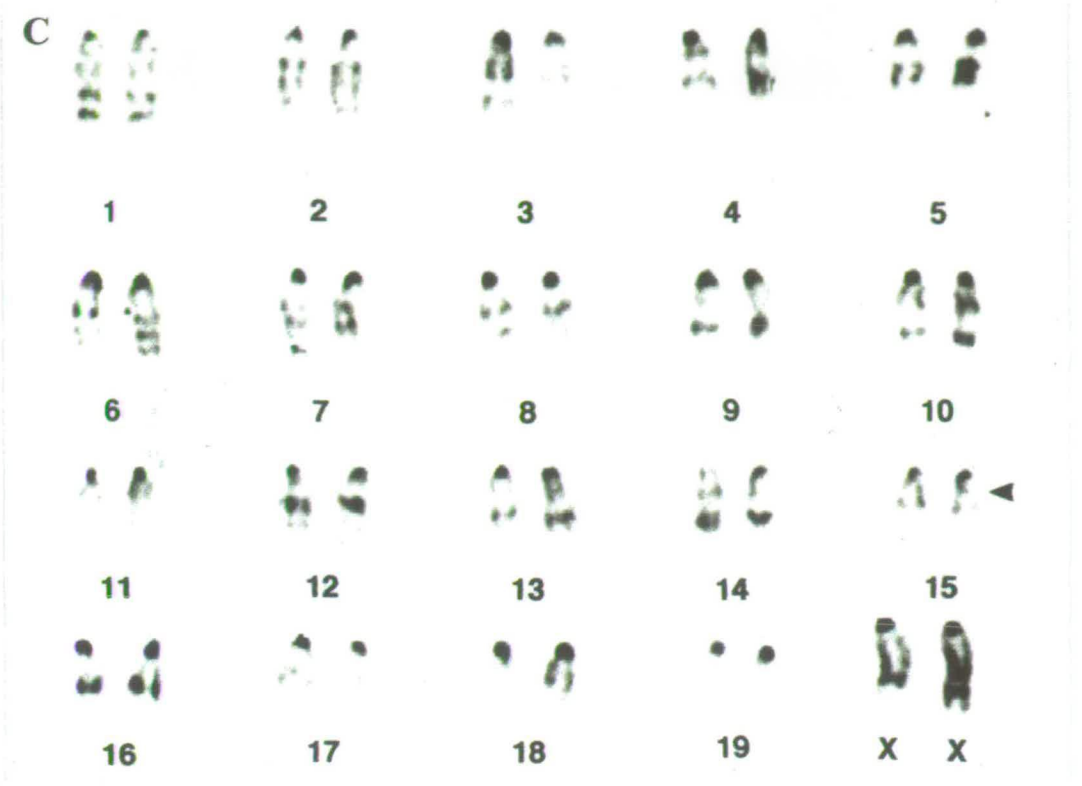
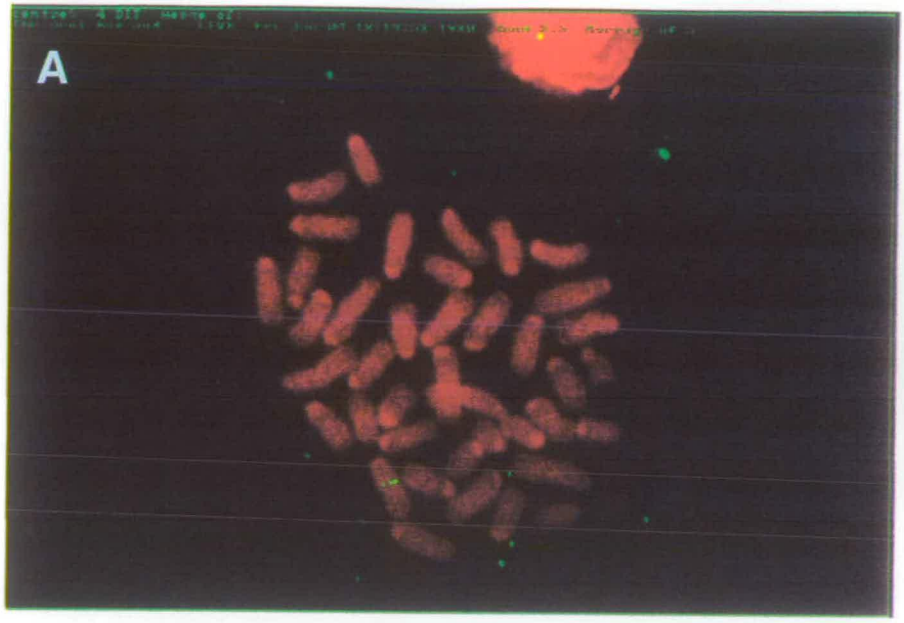


Figure 28. FISH/G-band analysis of metaphase chromosomes from line 7.

(**A**) Chromosomes from homozygous line 7 mice were stained red using propidium iodide. Biotinylated BLG probe was visualised indirectly using avidin-FITC (*green*). (**B**) G-banding in conjunction with FISH enabled the identification of the G-banded chromosomes with the integrated transgene (*arrows*). (**C**) Karyotype of line 7 G-banded chromosomes. The chromosomes harbouring the transgene (chromosome 15) are indicated with a black arrow.



B

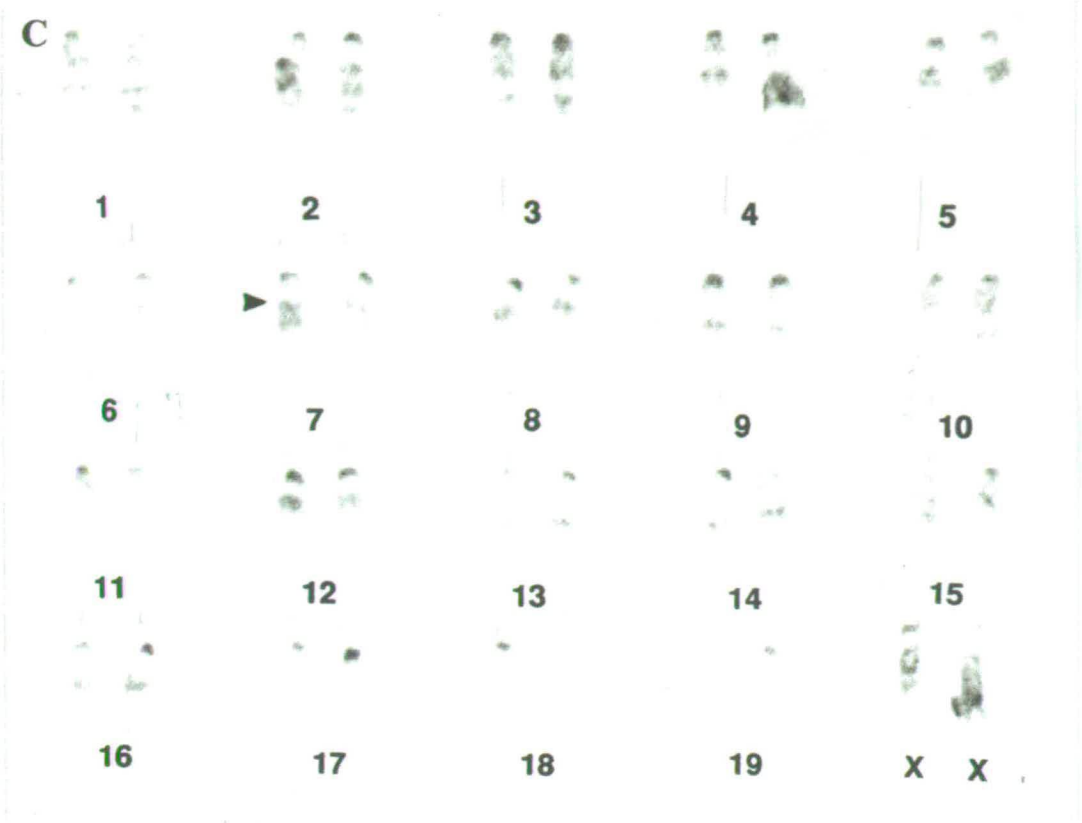
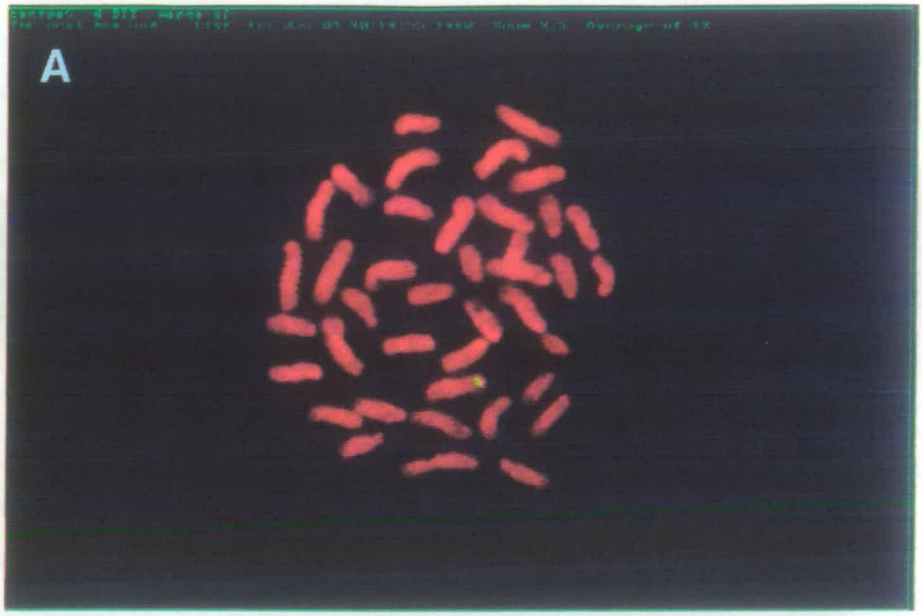
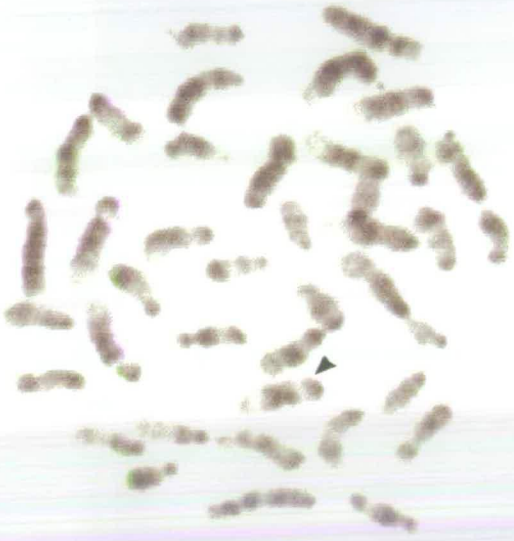


Figure 29. FISH/G-band analysis of metaphase chromosomes from line 14.

(**A**) Chromosomes from hemizygous line 14 mice were stained red using propidium iodide. Biotinylated BLG probe was visualised indirectly using avidin-FITC (*green*). (**B**) G-banding in conjunction with FISH enabled the identification of the G-banded chromosomes with the integrated transgene (*arrow*). (**C**) Karyotype of line 14 G-banded chromosomes. The chromosome harbouring the transgene (chromosome 7) is indicated with a black arrow.



B



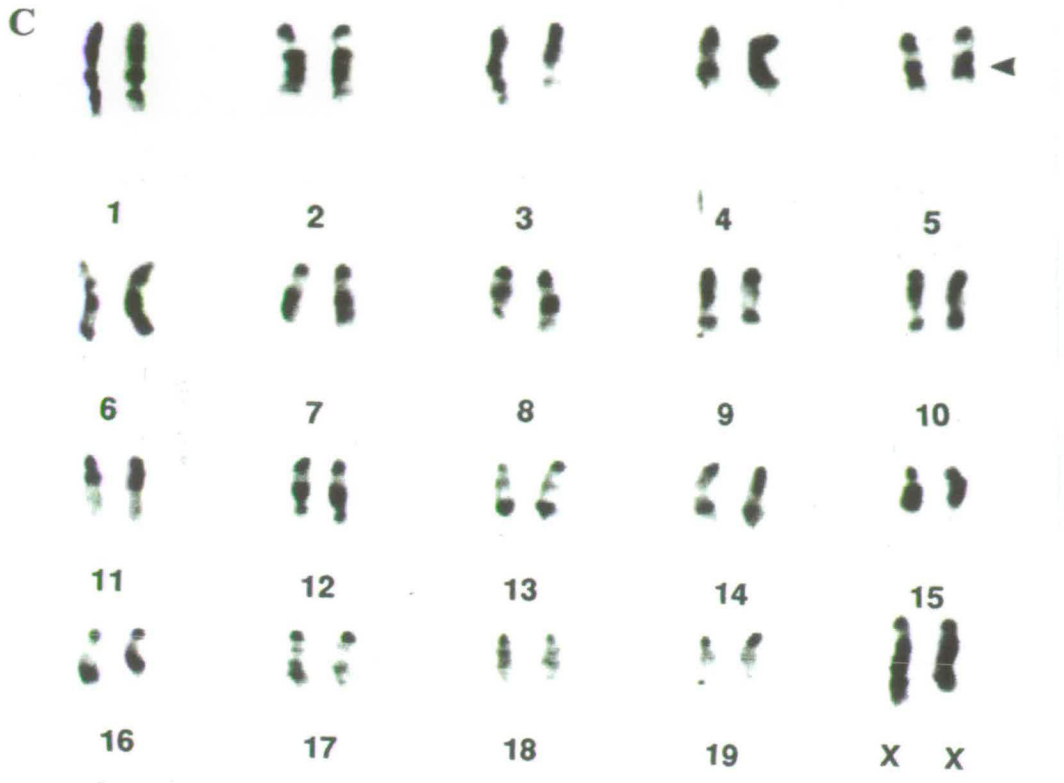
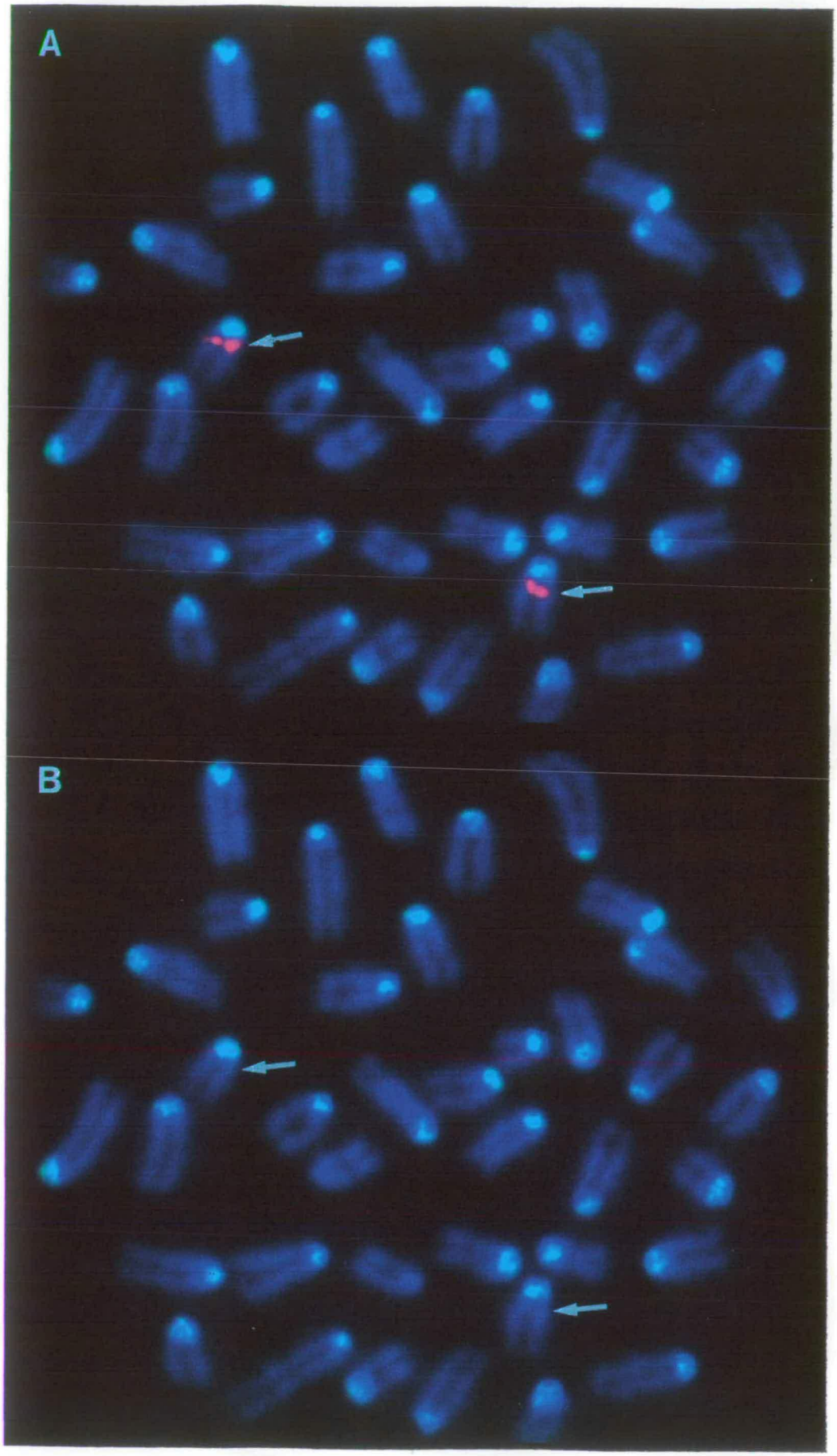


Figure 30. FISH/G-band analysis of metaphase chromosomes from line 45.

(**A**) Chromosomes from hemizygous line 45 mice were stained red using propidium iodide. Biotinylated BLG probe was visualised indirectly using avidin-FITC (*green*). (**B**) G-banding in conjunction with FISH enabled the identification of the G-banded chromosomes with the integrated transgene (*arrow*). (**C**) Karyotype of line 45 G-banded chromosomes. The chromosome harbouring the transgene (chromosome 5) is indicated with a black arrow.

Figure 31. Localisation of BLG and major satellite (196) DNA sequences in line 7 by chromosomal FISH to metaphase chromosomes.

(**A**) Chromosomes from homozygous line 7 mice were stained blue with DAPI. The BLG probe incorporated digoxigenin-11-dUTP, the major satellite DNA probe incorporated biotin-16-dUTP. The BLG probe (*arrow*) was detected with rhodamine conjugated anti-digoxigenin and Texas red conjugated anti-sheep (*red*). The major satellite probe was detected using successive layers of avidin-FITC, biotinylated anti-avidin and avidin-FITC (*green*). Each chromatid of metaphase chromosome 15 can be seen labeled with the BLG signal in these examples. (**B**) As above except the BLG signal was quenched.


A
B

Rough estimates were made of the distance between the transgene locus and the satellite DNA by taking optical readings along 56 different chromatids (Fig. 32). The edge of the satellite DNA signal was taken from two points (Fig. 32B); C1 was at the leading edge of the satellite DNA signal and C2 was half way between the edge (C1) and the middle of the satellite DNA signal. The true edge of the satellite DNA was most likely to be towards C1 because the area to which the satellite DNA probe hybridised remained essentially the same as when the signal was quenched. The BLG transgene locus signal was taken from one point (T1; Fig. 32B) which was the middle of the signal. The results are presented in Table 23.

Table 23. Measurement of the distance between the BLG transgene locus and major satellite DNA

Chromosome length* mean ±*s.d.*	T1 to C1‡ mean ±*s.d.*	T1 to C2‡ mean ±*s.d.*
41.54 ±5.02	3.35% ±5.29	12.10% ±4.76

* chromosome length was measured in microns (x10); ‡ the distance between T1 and C1 or C2 is expressed as a percentage of the total chromosome length (qtel to ptel); n = 56. This analysis was performed by Judy Fantes; MRC Human Genetic Unit, Edinburgh.

The estimated distances between the transgene locus and edge of the satellite DNA signals were used to position the transgene locus on the G-band map of mouse chromosome 15 (Fig. 33). These rough calculations position the transgene locus within the first two G-bands after the heterochromatic/euchromatic DNA boundaries.

Figure 32. Localisation of the BLG locus in line 7 relative to the centromere.

(**A**) The preparations with the longest chromosomes with sharp BLG signals were selected for this analysis. Lines were artificially drawn from one end of the chromosome to the other through the centromere and chromatids. Optical readings were taken along these lines. (**B**) Optical readings were plotted as distance vs. pixel intensity for each of the three colours, blue for chromatids, green for satellite DNA and red for the BLG locus. The ends of the chromatids are marked qtel & ptel. The edge of the centromere was measured from two points (C1 & C2). The transgene position was taken from the peak T1. This analysis was performed by Judy Fantes, MRC Human Genetics Unit, Edinburgh.

A
B
Pixel Intensity
Polygon Graph
Y-Axis
X-Axis
250.00
200.00
150.00
100.00
50.00
0.00
0.00
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
50.00
55.00
T1
c2
c1
qtel
ptel
Distance (microns [x10])

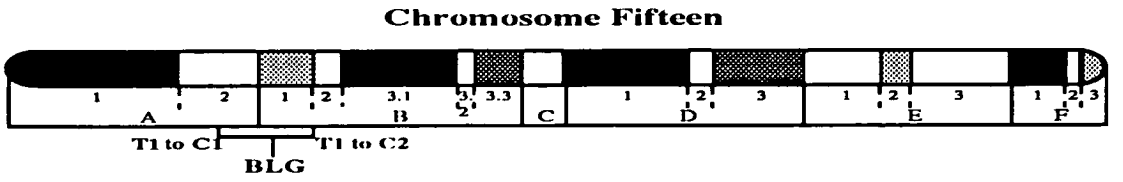


Figure 33. Mouse chromosome 15 and the BLG transgene locus

The estimated position of the BLG transgene locus on mouse chromosome 15 is indicated (band A2 or B1). The calculation of the position of the BLG locus was based on the presumption that the edge of the satellite DNA is the boundary between bands A1 and A2 (Akeson & Davisson, 1996). The distance between the A1/A2 boundary and C1 is 3.35% of the total chromosome length; the distance between the A1-A2 boundary and C2 is 12.10% of the total chromosome length (Table 23). Note that the edge of the satellite DNA was taken as the fixed point of reference this time instead of the center of the transgene signal (T1) as was the case in Table 23. The G-banded schematic diagram of mouse chromosome 15 was taken from *Mouse Genome 94* (1996), pp. 58 and is drawn to scale.

The total haploid mouse genome is estimated to be 3.3 Gb (Carden, 1994). Chromosome 15 is estimated to represent 4.17% of the total haploid genome (Evans, 1996). Therefore, based on these two presumptions, chromosome 15 is ~138 Mb in size. 3.35% of chromosome 15 therefore represents 4.6 Mb and 12.1% is equivalent to 16.7 Mb. These calculations indicate that the BLG transgene locus could be between 4.6 Mb and 16.7 Mb from major satellite DNA. However it is more likely that the lower estimate is closest to the true distance. Double *in situ* hybridisation experiments utilising a BLG probe and probes for endogenous genes (e.g., interleukin 7 receptor and growth hormone receptor) that are known to map to bands A2 or B1 (*Mouse Genome 94* [1996] pp. 58) will be necessary to map the transgene integration locus with greater precision.

In line 45 the resolution of the *in situ* mapping (estimated at around 0.5 Mb) made the BLG signal appear to abut or even overlap with the major satellite DNA

signal (Fig. 34A). However the transgene signal can be observed clearly separate from major satellite DNA in interphase nuclei where chromosomes are decondensed (Fig. 34A). Again, quenching the BLG signal indicated that there is no translocated satellite DNA at the BLG locus in line 45 (Fig. 34B).

Minor satellite DNA is found at the tip of mouse chromosomes. However, the possibility that there could be translocated minor satellite DNA at the BLG locus still existed. This was examined by performing the experiment described above but with a probe for minor satellite (Fig. 35). Results with spreads from line 7 homozygous transgenic mice demonstrated that the transgene was positioned well away from the minor satellite DNA (Fig. 35A) and there was no evidence of translocated minor satellite DNA at the BLG locus after the BLG signal was quenched (Fig. 35B).

6.2.3. Immunofluorescent / FISH analysis of line 7 chromosomes using CREST antibodies and DNA probes for BLG, *c-myc* and *Pvt-1*

Some 10-15% of Scleroderma/CREST (calcinosis, Raynaud's phenomenon, oesophageal involvement, sclerodactyly, telangiectasia) patients produce autoantibodies (known as anti-centromere antibodies [ACAs]) that recognise centromeric determinants (Moroi *et al.*, 1980; Saunders *et al.*, 1993). More specifically, the ACAs recognise a conserved 52 amino acid domain termed the chromodomain (chromo for chromatin modifier) present in a family of centromere associated proteins thought to provide structural support for other proteins involved in chromosome pairing and segregation (Kellum & Alberts, 1995). "Native" chromosomes can be prepared in a fashion that leaves the chromosome-

Figure 34. Localisation of BLG and major satellite (196) DNA sequences in line 45 by chromosomal FISH to metaphase chromosomes.

(A) Chromosomes and one interphase nucleus from hemizygous line 45 mice were stained blue with DAPI. The BLG probe incorporated biotin-11-dUTP, the major satellite DNA probe incorporated digoxigenin-11-dUTP. The BLG probe (*arrow*) was detected with avidin-FITC (*green*). The major satellite probe was detected using rhodamine conjugated anti-digoxigenin and Texas red conjugated anti-sheep (*red*). (B) As (A) except the BLG signal was quenched.

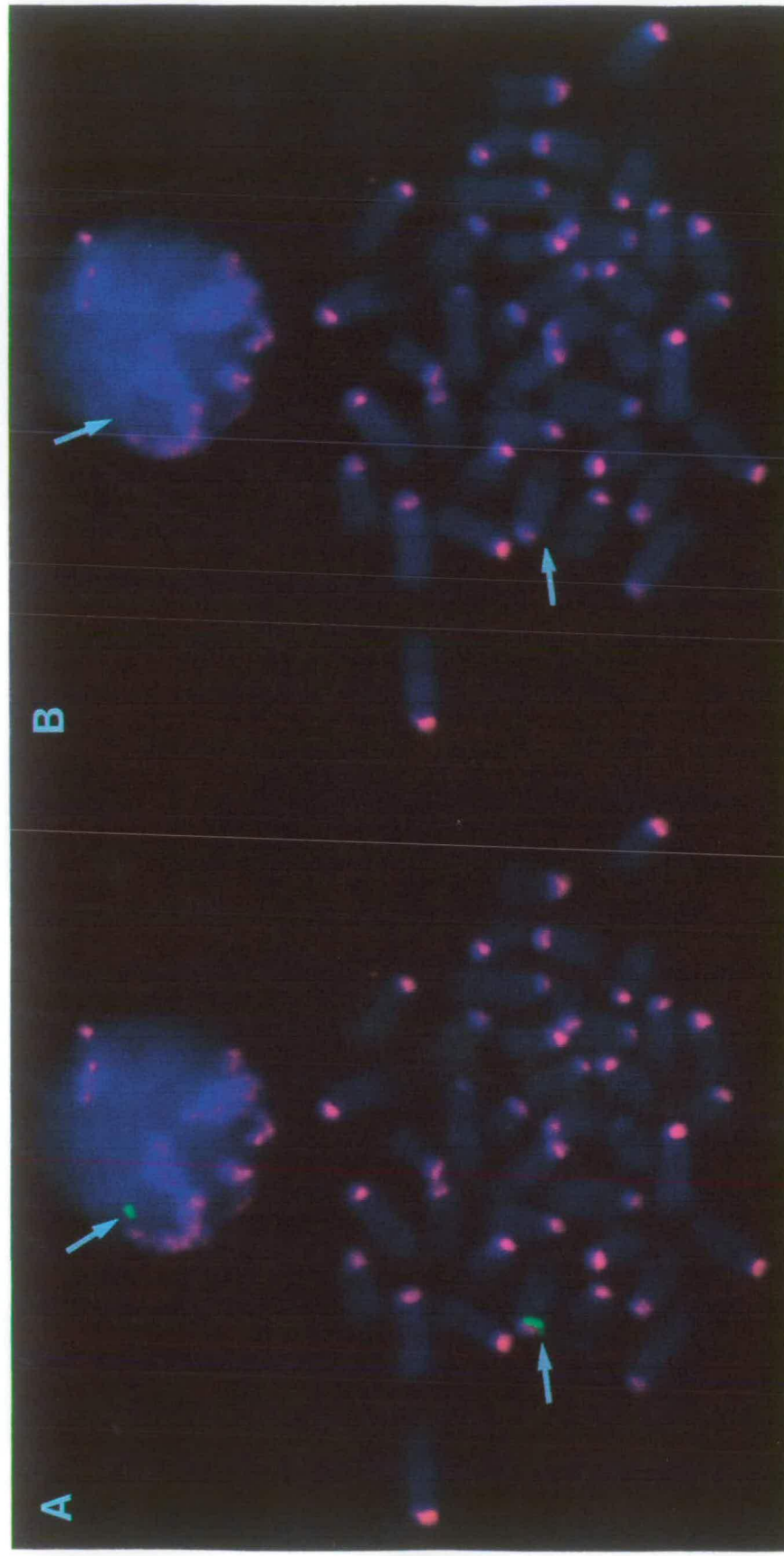
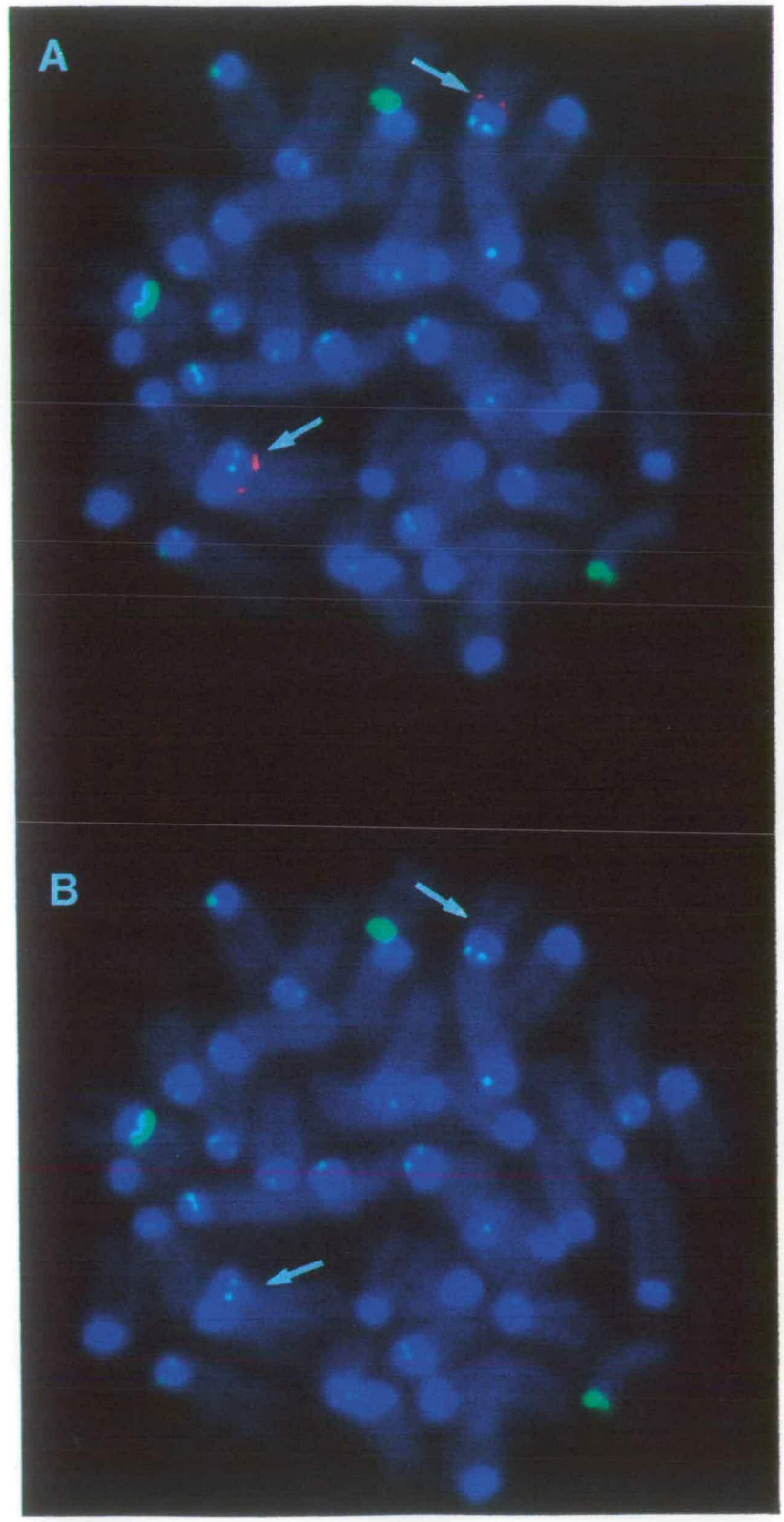

A
B

Figure 35. Localisation of BLG and minor satellite DNA sequences in line 7 by chromosomal FISH to metaphase chromosomes.

(**A**) Chromosomes from homozygous line 7 mice were stained blue with DAPI. The BLG probe incorporated digoxigenin-11-dUTP, the minor satellite DNA probe incorporated biotin-16-dUTP. The BLG probe (*arrow*) was detected with rhodamine conjugated anti-digoxigenin and Texas red conjugated anti-sheep (*red*). The minor satellite DNA probe was detected using successive layers of avidin-FITC, biotinylated anti-avidin and avidin-FITC (*green*). Each chromatid of metaphase chromosome 15 can be seen labeled with the BLG signal in these examples. (**B**) As (**A**) except the BLG signal was quenched.


A
B

associated proteins associated with the chromosomes. Immunofluorescence experiments with native chromosomes and CREST sera demonstrated that the EP fraction (Jeppesen & Nicol, 1986) recognises a mammalian homologue of the *Drosophila* heterochromatin-binding protein 1 (HP1; James & Elgin, 1986; Elgin, 1996) that is specifically associated with pericentromeric heterochromatin in *Mus musculus* and humans (Nicol & Jeppesen, 1994). HP1 shares the chromodomain or "HP1/*Pc* box" (Clark & Elgin, 1992) with the Polycomb protein (*Pc*; Paro & Hogness, 1991) which is thought to regulate certain developmental events by repressing homeotic gene expression (Moehrle & Paro, 1994). A further correlation between PEV and homeotic gene repression has been demonstrated by the fact that regulatory regions that respond to *Pc* group repression can induce variegation of an adjacent *white* gene (Fauvarque & Dura, 1993). PEV can be modified by mutations in genes that are then termed suppressors (e.g., *Suvar(2)5* [HP1]; Eissenberg *et al.*, 1990; Eissenberg *et al.*, 1992) or enhancers (e.g., the gene encoding GAGA, a protein associated with gene activation; Farkas *et al.*, 1994) of variegation. The trithorax-group of genes in *Drosophila* are general activators of homeotic genes and demonstrate structural and functional similarities with the enhancers of variegation. Therefore PEV serves as a marker in genetic screens to identify other heterochromatin or chromatin-affecting proteins (Reuter *et al.*, 1990; Grazino *et al.*, 1992). More than 50 suppressor or enhancer loci of PEV have been characterised to date and several other candidates have been identified (Weiler & Wakimoto, 1995).

Results presented in chapters 5 and 6 established that (i) there is mosaic expression of the transgene in two lines and (ii) the transgene arrays in the same two lines are close to their respective centromeres. These are the two main features associated with PEV. Immunofluorescent analysis using CREST antibodies was performed on native chromosomes prepared from line 7 mice to investigate whether there was evidence of heterochromatinisation of the BLG locus.

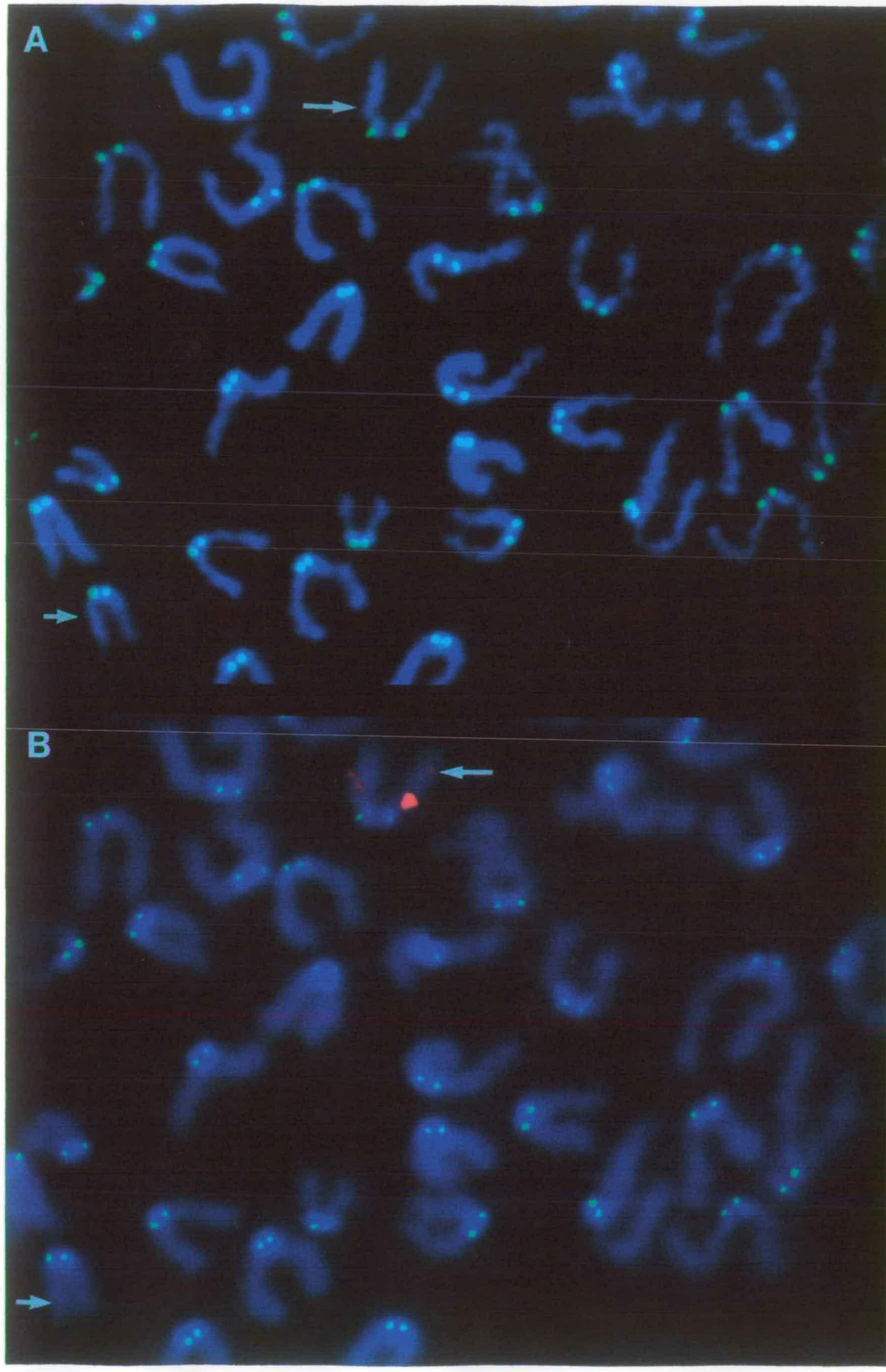
Native metaphase chromosomes were prepared from lymphocyte cell cultures from low and high BLG expressing individuals from line 7. Immunofluorescence experiments using CREST antibodies resulted in the characteristic "double-dot" binding to the kinetochores (Fig. 36A). The degree of pericentromeric (lighter labeling around the kinetochores; Fig. 37A) varied between slides in preparations from the same animal and from different animals. These differences are attributable to the inherent technical difficulty of producing chromosome preparations of consistent quality.

The position of the BLG locus was identified by following the immunofluorescent analysis with FISH using a BLG and two endogenous chromosome 15 probes, *c-myc* and *Pvt-1* (Fig. 38). The endogenous probes enabled the identification of both chromosome 15's in chromosome preparations from hemizygous line 7 individuals (Fig. 36B and Fig. 37B). This strategy enabled the comparison of the heterochromatic staining on the chromosome harbouring the transgene and its homologue without the transgene. Out of more than 60 metaphase chromosome spreads examined from low and high line 7 BLG expressors there was no evidence for co-localisation of the BLG and CREST signals (Figs. 36 and 37).

This result could suggest that the BLG locus is not epigenetically silenced by heterochromatinisation of the locus. However there are several other possible reasons why there was no evidence for co-localisation of the BLG and CREST signals. (i) Crucially, the CREST EP serum is specific for just one heterochromatic protein (Nicol & Jeppesen, 1994). The BLG locus may be heterochromatinised by a complex of proteins that does not contain this protein. Other proteins that contain a chromodomain have been identified in mouse (Singh *et al.*, 1991; Wreggett *et al.*,

Figure 36. Native FISH using CREST antibodies and DNA probes for BLG, *c-myc* and *Pvt-1*on line 7 chromosomes.

(A) Native chromosome spreads from hemizygous line 7 mice were incubated with CREST antibody. The CREST antibody was visualised indirectly using anti-human IgG-FITC (*green*). The chromosomes were stained blue with DAPI. Images were recorded and the coordinates of each spread noted before proceeding with FISH analysis using BLG and endogenous gene probes. The chromosomes harbouring the transgene (*long arrow*) and the homologue (*short arrow*) are indicated. (**B**) The BLG, *c-myc* and *Pvt-1* probes incorporated digoxigenin-11-dUTP and visualised indirectly using rhodamine conjugated anti-digoxigenin and Texas red conjugated anti-sheep (*red*). The CREST signal was retained from the previous hybridisation (*green*). The *c-myc / Pvt-1* loci are indicated on the chromosome with the transgene signal (*long arrow*) and the homologue (*short arrow*).


A
B

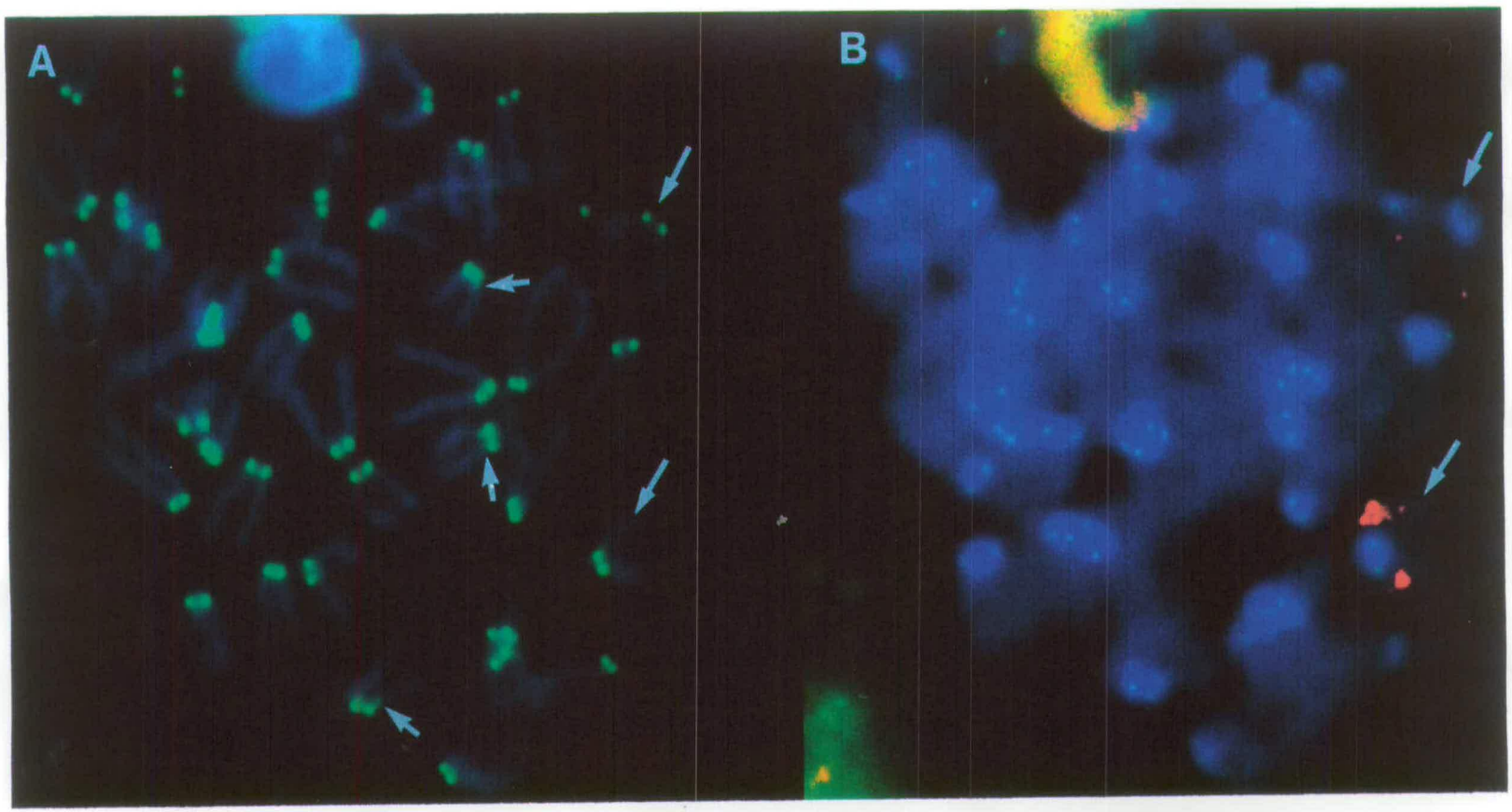


Figure 37. Native FISH using CREST antibodies and DNA probes for BLG, *c-myc* and *Pvt-1* on line 7 chromosomes.
(**A**) The chromosomes indicated (*short arrows*) exhibited distinctive pericentromeric CREST signals. The two chromosome 15's are also indicated (*long arrows*). (**B**) The chromosomes with the transgene (*lower arrow*) and its homologue (*upper arrow*) without the transgene are indicated.

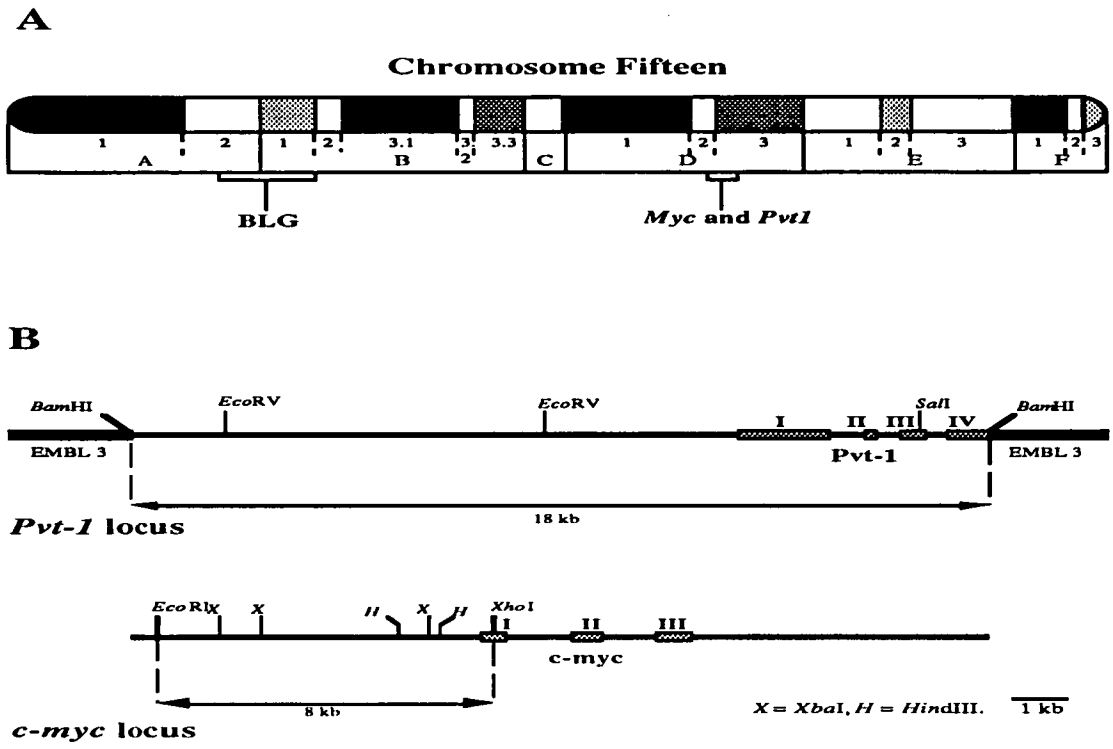


Figure 38. Mouse chromosome 15 and the BLG, *c-myc* and *Pvt-1* loci.

(**A**) Schematic representation of a G-banded chromosome 15. The *c-myc* and *Pvt-1* loci are mapped to band D2 or D3 (*Mouse Genome 94* [1996] pp. 58). Also illustrated is the estimated position of the BLG transgene locus. (**B**) Restriction maps (incomplete) of the *c-myc* and *Pvt-1* loci. The *c-myc* and *Pvt-1* clones were a gift from Dr Konrad Huppi, NIH, USA.

1994; Singh, 1994). M31 is a 24.7 kD protein that is thought to be a component of constitutive heterochromatin and is the closest homologue to HP1 with 51% homology. M32 is more distantly related to HP1 and is thought to be a component of silenced chromosomal domains. (ii) The metaphase spreads were prepared from lymphocyte cell cultures that were derived from the spleen and not from mammary tissue (at present there is no suitable protocol for mammary gland disaggregation and cell culture). Silencing of the BLG locus in the two tissues may not utilise the same processes. (iii) Finally, the CREST antibody serum may bind to the BLG locus but the signal may not be detectable because it is a relatively small target (400-500 kb).

6.3. CONCLUSIONS

The transgene arrays have integrated into chromosomes 15 (line 7), 7 (line 14), and 5 (line 45). Lines 7 and 45 exhibit variable protein and mRNA expression levels, mosaic expression in the mammary gland and the transgenes have integrated close to their respective centromeres. Line 14 exhibits stable protein and mRNA expression, uniform expression in the mammary gland and the transgene has integrated mid-arm on chromosome 7. The BLG locus does not co-localise with major or minor satellite DNA and could be as close as 4-5 Mb from major satellite DNA in line 7 and even closer in line 45. These results imply that variegated expression of the transgene may be associated with the position of the locus relative to the centromere, akin to PEV in *Drosophila*.

Immunofluorescence experiments provided no evidence for co-localisation of the CREST antibody signal with the BLG locus on native metaphase chromosome preparations from line 7 animals. However this negative result does not rule out the

possibility that the BLG locus has been silenced in some cells due to a stochastic heterochromatinisation process.

CHAPTER 7

DISCUSSION

The introduction of new gene combinations into the mammalian germline by pronuclear injection of DNA is widely used in a variety of applications. The transgene usually integrates as a tandemly repeated and predominantly head-to-tail array at a single random site within the genome and therein segregates as a single Mendelian locus (Palmiter & Brinster, 1986). Though the transgene insertion site can clearly have an influence upon the tissue-specificity and level of expression, it has often been assumed that expression of the transgene, once integrated, is stable within a transgenic line. Indeed, some detailed studies have failed to detect variations in the level of transgene expression (on average, less than 20% variation in the level of reporter gene expression among different individuals of each line examined [Al-Shawi *et al.*, 1990]).

This contrasts with the situation found in *Drosophila*, yeast and plants where transgene expression can be subject to stochastic influences that lead to variegation of expression levels. We thus entered into a detailed study of the stability of transgene expression in three transgenic lines harbouring the ovine BLG gene. The transgene is expressed selectively within the lactating mammary gland and BLG protein is exported into milk. Of three lines studied, only one (line 14) expressed the transgene in a stable manner and wide variations in the individual level of expression were observed among individuals of the other two lines (lines 7 and 45). Most other examples of unstable transgene expression describe differences (e.g., degrees of mosaic expression) between different transgenic lines. This study provides the first detailed evidence for variegated expression of a transgene between individuals within transgenic lines.

DNA copy number within these lines appeared to be stable and there was no significant evidence of transgene rearrangement. Despite variation across the population, extensive analysis of the most variable line (line 7) demonstrated that the expression levels are fixed within the individual (chapter 3). The genetic background of this line, however, does not play a significant role in this variation (chapter 4). In this pedigree, as well as in a second line (line 45), mammary tissue *in situ* hybridisation experiments revealed mosaic expression of the transgene (chapter 5); in low expressing animals, discrete patches of cells expressing the transgene were observed. The concentration of the transgene protein in milk reflected the proportion of mammary epithelial cells expressing BLG mRNA. In contrast, all the secretory epithelial cells within transgenic glands expressed the gene for an endogenous milk protein (β-casein) uniformly. *In situ* hybridisation to metaphase chromosomes indicated that the transgene arrays in lines 7 and 45 are situated close to the centromere.

7.1. Variable levels of BLG protein expression within lines 7 and 45

In the experiments described in chapter 3 the ranges of BLG protein expression levels were different in lines 7, 14 and 45 (mean [mg/ml] ± *s.d.* = 9.5 ± 4.9, 6.6 ± 1.1 and 23.7 ± 3.9 respectively). BLG is thus expressed in an unstable fashion in line 7 and in a stable fashion in line 14. Line 45 is more difficult to classify in terms of stable or unstable expression. Though line 45 exhibits less variability than line 7, this line (45) has a larger mean expression level than line 7. A possible explanation is that some factor(s) restricts the range of expression in line 45, reducing the extent of variation by suppressing transgene product levels at the upper end of the range. For example, there may be an upper limit or "ceiling effect" to the amount of BLG that can be expressed within the mammary gland. The gland

appears to have a finite secretory capacity for milk proteins (McClenaghan *et al.*, 1995). In chapter 3 the mean BLG expression level was 23.7 mg/ml (n = 28) for line 45 hemizygous mice. The data from chapter 3 would predict a mean of 45-50 mg/ml in homozygous line 45 mice. However, the mean BLG expression level was only 33.2 mg/ml (n = 8) in line 45 homozygous mice in a separate study (McClenaghan *et al.*, 1995). The large difference between the predicted and measured homozygous mean expression levels in the two studies could be due to the small sample size (n = 8) in the latter study. Alternatively the homozygous mean could be limited to around 30-35 mg/ml by constraints on mammary gland synthetic capacity. Moreover, when BLG protein expression is increased (e.g., by crossing two hemizygous line 45 mice) the levels of the endogenous milk proteins were reduced (McClenaghan *et al.*, 1995). The stage/level of competition for milk protein production was not identified. An increase in expression level of one protein at the expense of another could be due to competition for transcription factors (e.g., MGF / STAT5) involved in milk protein gene expression (Wakao *et al.*, 1994; Gouilleux *et al.*, 1994; DaSilva *et al.*, 1996). Interestingly, the compensatory effect is different for each endogenous protein; WAP is proportionately reduced more than α- and β-casein (McClenaghan *et al.*, 1995). Alternatively, the gland could exert a limit to the amount of BLG expressed by a negative feedback mechanism.

Tissue *in situ* hybridisation experiments revealed mosaic expression of BLG mRNA in the mammary gland of line 7 and 45 animals. In line 7, the proportion of cells positive for transgene mRNA appeared to correlate with protein expression levels. In mammary gland sections from the lowest expressors regions of expressing cells were surrounded by non-expressing cells. In the highest expressors this situation was reversed, with small regions and even single alveoli comprising non-expressing cells surrounded by expressing cells. While the proportion of mammary gland cells expressing the transgene clearly varied between individuals of line 7,

individual differences in line 45 were less pronounced. Here the patterns of expression were similar to those observed in line 7 individuals expressing the transgene in the majority of mammary gland cells. There was little evidence of patches of expressing cells surrounded by non-expressing cells. This may be due to the higher average expression level in line 45. However, line 45 individuals did demonstrate a degree of variegation between individuals within the line that corresponded with the protein expression levels. This result suggests that the variation of BLG protein levels in line 45 animals was a true measure of the range of expression within the line and not a statistical wobble or spread due to the high average expression level.

Unfortunately, it was not possible to accurately quantitate tissue *in situ* hybridisation images. This would require serial sections through the gland (due to the complex three-dimentional structure of the gland) and exhaustive estimates of numbers of expressing and non-expressing cells in each section. A strategy similar to this was described by Munford (1963) and serves to demonstrate that such an undertaking would be a separate project in its own right.

In summary, only line 14 exhibited stable expression akin to endogenous milk protein gene expression. This served to demonstrate that the transgene construct itself is not responsible for variegated expression. However, the other two lines studied (7 and 45) showed considerable variation within the lines. These variations were manifested at the level of protein expression within the milk and the proportion and mRNA expressing cells within the mammary gland.

7.2. Transcriptional versus translational mechanism for unstable BLG expression

Eukaryotic gene expression is regulated at a variety of transcriptional and/or the post-transcriptional levels. Chapter one summarised a multitude of factors involved in transcription of the milk protein genes within the mammary gland including transcription factors (Wakao *et al.*, 1994; Li & Rosen, 1995; DaSilva *et al.*, 1996), hormones (Topper & Freeman, 1980; Hobbs *et al.*, 1982; Wiens *et al.*, 1987), cell-cell (Levine & Stockdale, 1985; Chen & Bissell, 1989) and cell-ECM interactions (Lee *et al.*, 1984; Blum *et al.*, 1987; Chen & Bissell, 1989). Some of these influences lead to differential regulation of individual milk protein mRNAs (Hobbs *et al.*, 1982). For example, studies using primary cultures of mouse mammary epithelial cells maintained on floating collagen gels, with or without addition of ECM components, demonstrated that casein mRNAs were hormonally regulated while WAP and α-lactalbumin were not expressed (Lee *et al.*, 1984, 1985). Steady-state mRNA levels are determined by a variety of parameters including initiation and elongation of transcription, 5' cap addition, polyadenylation, splicing, transfer from the nucleus to the cytoplasm, subcellular targeting of transcripts and the role of mature mRNA degradation. Although little is known about the role of extracellular influences on milk protein mRNA stability in the mammary gland, there are several well characterised examples of regulation of gene expression by control of mRNA stability in other systems (for review, see Atwater *et al.*, 1990). For example, oestrogen was reported to have a >30-fold stabilising effect on the half-life of vitellogenin mRNA in *Xenopus laevis* liver (Brock & Shapiro, 1983). Milk protein expression is also governed by protein targeting to constitutive and regulated secretory pathways that transport the mature polypeptide through and out of the secretory epithelial cells (Turner *et al.*, 1992). It

is therefore perhaps unsurprising that high correlations between individual BLG protein and mRNA expression levels were not observed within lines 7, 14 and 45. The correlation coefficients were similar for all three lines (after excluding one outlier in line 7) arguing that the relationship between RNA and protein is the same in all three lines. The variability of BLG mRNA expression levels reflected the variability of protein expression levels; line 7 was the most variable for protein and mRNA expression, line 14 was the most stable and line 45 had an intermediate range relative to lines 7 and 14. This relationship could be due to variable mRNA stability within lines 7 and 45. However the tissue *in situ* hybridisation experiments demonstrated that there was a good correlation between the BLG protein expression level and the number of mRNA expressing cells in line 7. The theoretical possibility remains that non-expressing cells transcribe the transgene but degrade the transcript to such an extent that it becomes undetectable by in situ hybridisation. However, it seems more likely that variable protein levels are more due to variable transcription of the transgene.

7.3. Variable transcription of the BLG transgene

Unstable transgene expression could be due to stochastic deletion or rearrangement of the locus. However no evidence was found for significant loss of transgene copies within individuals exhibiting the full range of expression within all three lines. Based on the relatively uniform intensity of the BLG bands on Southern blots the *c.v.* values appeared surprisingly high for all three lines. This was initially attributed to the preparation of the DNA samples and the quantitation of the bands. Exhaustive efforts were made to reduce cumulative errors by cleaning the DNA twice (2 x 3 phenol/chloroforms), ensuring complete resuspension of the DNA, measurement of DNA concentrations in triplicate,

production of blots with low background hybridisation, careful quantitation using a phosphoimager and the use of two (total mouse mammary DNA and WAP) internal controls for loading. Despite every effort the *c.v.* values remained high in each line, implying that they reflect the inherent variability in the technique. It is doubtful that the *c.v.* values represent variations in transgene copy number that are responsible for variable mRNA and protein expression levels. For example, line 14 has *c.v.* values of 41-45% for copy number within the line without demonstrating variations in mRNA and protein expression levels. In contrast, line 45 has *c.v.* values of 38-42% for copy number and does demonstrate significant variations in mRNA and protein. Unfortunately similar studies have not been performed by others in the laboratory and therefore a comparison of "basal" variation for the procedure cannot be made. Significantly, the DNA copy number *c.v.* values were similar in all three lines while the protein (*s.d.*) and mRNA (*c.v.*) values varied considerably within the three lines. Furthermore, the line 7 protein and mRNA *c.v.* values (52% and 47% respectively) were higher than the DNA *c.v.* values (26-42%), making transgene deletion an unlikely explanation for variable expression.

These results contrast with a separate study in which unstable transgene expression was shown to be due to stochastic deletion of the locus in two lines of transgenic mice (Sandgren *et al.*, 1991). In this study expression of an albumin-urokinase-type plasminogen activator fusion construct in the liver was lethal in transgenic mice. However two lines stochastically reverted from an original 10 and 4 copies of the transgene to 2-3 and 1 copy respectively, resulting in the abolition of expression and subsequent selection for cells exhibiting the deletion. These deletions were clearly visible using Southern analysis even though the liver is a heterogenous tissue comprising several cell types. Detailed Southern analysis of transgene DNA demonstrated that the most likely mechanism responsible for the loss of cytotoxic transgene expression was homologous recombination between

copies of the transgene within the array. Transgene deletion was manifested as changes in liver phenotype because the rearrangements imparted a selective advantage on the cells exhibiting the rearrangement and the liver permitted amplification of these cells. Similarly, expression of a metallothionein/herpes simplex virus thymidine kinase fusion construct in embryonic germ cells resulted in sperm infertility (Braun *et al.*, 1990). Intrachromosomal recombination between duplicated genomic DNA flanking the transgene deleted the array and resulted in fertile males due to clones of wild-type sperm (Wilkie *et al.*, 1991). Again, the deletions were clearly visible by Southern analysis.

Line 7 exhibited variable expression and line 14 exhibited stable expression even though both lines contain an identical construct. The size of the transgene array could be the determining factor; the larger arrays in lines 7 and 45 could be more susceptible to intrachromosomal recombination. However an array of only 4 copies exhibited deletion down to just one copy in the study described by Sandgren and collegues (1991).

The total DNA content of the mouse mammary gland increases 5- to 7-fold between the virgin and day 14 of lactation (Brookreson & Turner, 1959; Nicoll & Tucker, 1965). Mammary tissue from virgin C3H/Crgl/2 mice comprises about 23% mammary parenchymal DNA and 77% stromal DNA (Nicoll & Tucker, 1965). While the DNA content of the mammary gland fat pad did not increase between the virginal state and day 10 of lactation the parenchymal tissue increased to about 89% (a 27-fold increase) of total DNA. However it is not known what percentage of the cells within the gland at mid-lactation are of the secretory epithelial type (Prof. Barry Gusterson, pers. comm.). Given the inherent variability in the Southern blot data and the heterogeneity of cell types within the mammary gland, stochastic deletion of the locus in lines 7 and 45 cannot be ruled out as the cause of variable expression. However this explanation is hard to reconcile mechanistically. The

mammary gland is similar to the liver and testis (Sandgren *et al.*, 1991 and Wilkie *et al.*, 1991 respectively) in that the gland exhibits massive cell proliferation (Traurig, 1967) and cell selection could plausibly operate. However, expression of BLG does not appear to be toxic to the mother or offspring (Simons *et al.*, 1987) and the secretory epithelial cells appear morphologically normal (Harris *et al.*, 1991). Therefore it is unclear whether mammary stem cells exhibiting intrachromosomal deletion of the BLG locus could have a selective advantage over those without transgene deletion. It is not known for certain whether the mammary gland contains a population of stem cells that divide and differentiate to generate the secretory epithelial cells. However, the cycles of cell proliferation and involution during successive lactations is consistant with the presence of stem cells (for review see Daniel & Silberstein, 1987). Autoradiographic studies of glands at early pregnancy demonstrate that proliferation occurs at the terminal end buds (TEBs) of the ductal system in rats (Dulbecco *et al.*, 1982) and from the TEBs and the ducts themselves in mice (Bresciani, 1971) suggesting the presence of pluripotent cells within the ductal branches. The high proportion of non-expressing cells in low expressors from line 7 would have had to arisen by deletion of the transgene in >50% of the mammary stem cells or within most of the cell lineages very early in mammary development. This seems very unlikely unless the transgene had integrated into a region within the genome that is a hotspot for recombination.

In summary, while these studies would not reveal subtle changes in the structure of the transgene array, variable expression is probably not due to stochastic loss or rearrangement of the transgene array in different individual line 7 and 45 mice.

7.4. Mosaic expression of BLG in the mammary gland

BLG was expressed in a mosaic fashion in lines 7 and 45 while β-casein expression was uniform. This result argues against the possibility that mosaicism is caused by cycling of active and inactive domains due to hormonal regulation, as previously suggested for endogenous milk proteins in sheep and mice (Molenaar *et al.*, 1992; Barash *et al.*, 1994) or transgenes in mice (Faerman *et al.*, 1995). Mosaic expression of four endogenous mouse milk genes and for a whey acidic protein (WAP)-reporter fusion gene has been described during gestation in mice (Robinson *et al.*, 1995). However, expression of all of the genes was observed throughout the mammary epithelium by day 1 of lactation. While the observations illustrated and described in chapter 5 do not rigorously rule out threshold effects due, for instance, to sub-optimal levels of lactogenic hormones leading to stochastic onset of milk protein gene expression, this possibility is unlikely. First, sections were taken from glands at a time (day 11 of lactation) when milk production is near-maximal and all secretory cells are producing β-casein. Second, line 14, harbouring the identical construct to that present in the highly variable line 7, shows stable expression throughout the gland.

The patching of cells in which the transgene is active or inactive is suggestive of clonal expansion. Because no evidence was found of transgene loss or rearrangement, the results presented in this study suggest that events leading to epigenetic silencing of transcription occur stochastically in individual progenitor cells; these are then transmitted through cell division to daughter cells and give rise to the observed mosaic pattern of expression. The closest precedent for these observations is afforded by the striated coat colour patterns recorded by Mintz and collegues for a mouse mutant affecting the endogenous tyrosinase gene (Porter *et al.*, 1991) or in animals harbouring tyrosinase fusion transgenes (Bradl *et al.*, 1991;

Mintz & Bradl, 1991). Robertson *et al.*, (1995) have reported differing levels of heterocellular expression for a globin promoter-αHS-40 transgene among transgenic mouse lines and shown that the degree of silencing increases with age (Robertson *et al.*, 1996). Although levels are fixed within lines, the underlying mechanism(s) may be similar, resulting in phenotypic differences that reflect the clonal origin of mosaic patches (phenoclones [Mintz, 1970]).

7.5. Position effect variegation: mechanisms of mosaic silencing

PEV was first documented by Muller in 1930 when he described a mosaic or variegated phenotype in the eye of *Drosophila melanogaster*. Variegation has fascinated and confounded investigators in the field since its discovery. PEV now applies to several phenomena (discussed below) but is most commonly observed in *Drosophila* when a chromosome rearrangement relocates a euchromatic endogenous gene locus close to a heterochromatic centromere or telomere. The translocation can result in the silencing of the locus in a variable proportion of cells, hence the variegation. Despite over 65 years of intensive study the phenomenon of PEV is still not adequately explained. However in the last few years there have been some exciting developments in the field. In attempting to dissect the cause of variegated expression research has focused on unstable gene structure or unstable gene expression.

7.5.1. Somatic instability

In *Drosophila*, surprisingly few genomic rearrangements (Rushlow *et al.*, 1984; Tartof *et al.*, 1984; Hayashi *et al* 1990; Karpen & Spradling, 1990) or transposon insertions (Spradling & Rubin, 1983;

Daniels *et al.*, 1986; Karpen & Spradling, 1992; Tower *et al.*, 1993; Zhang & Spradling, 1993) exhibiting PEV have been analysed at the DNA level. Variegation of genes due to chromosome rearrangement (Henikoff, 1981; Rushlow *et al.*, 1984; Hayashi *et al.*, 1990) or insertion into subtelomeric regions (Levis, 1989; Karpen & Spradling, 1992) sometimes show evidence of altered copy number and structure of the variegating gene. However, these alterations are generally considered as insufficient to explain the dramatic alteration in phenotype. The lack of evidence for substantial alterations to the loci could be because the examples failed to exhibit sufficiently high levels of variegation and Southern blot analysis would not reveal altered molecules representing a small fraction of the total. Other studies (Spradling & Karpen, 1990; Karpen & Spradling, 1990; Spradling, 1993) however, have demonstrated detailed examples of DNA elimination (up to 39-fold reduction) occuring at a locus (*yellow*) exhibiting strong variegation. The authors argue that some DNA alterations are due to stochastic instability of juxtaposed heterochromatic sequences and that copy number reduction (under-representation) represents a subclass of variegating regions. Spradling (1993) suggests that somatic rearrangements of heterochromatic regions may serve as a process for the regulation of genes (e.g., rRNA genes) embedded within heterochromatin or as overall organisers of nuclear chromatin through development.

7.5.2. Chromatin assembly model

Until recently, unstable gene expression of endogenous genes relocated close to centromeric or telomeric heterochromatin had been attributed to a heterochromatinisation process. The heterochromatinisation model (Zuckerkandl, 1974; Sinclair *et al.*, 1983; Locke *et al.*, 1988; Grigliatti, 1991) proposes that multimeric protein complexes normally associated with

heterochromatin slide or spread along the chromosome past the heterochromatin/euchromatin breakpoint and into adjacent relocated euchromatin (Belyaeva & Zhimulev, 1991; Martin-Morris *et al.*, 1993). In this model, mass-action or self-assembly of heterochromatin-associated proteins (see chapter 6) cause stochastic condensation of the DNA at some stage in development resulting in the variable inactivation of genes between the centromere and the vicinity of the breakpoint (Tartof *et al.*, 1984; Spradling & Karpen, 1990; Eissenberg & Elgin, 1991; Grigliatti, 1991; Henikoff, 1992; Reuter & Spierer, 1992). Evidence for this model is provided by the cytological observation that euchromatin juxtaposed with heterochromatin appears to merge with the highly compacted chromocentre on *Drosophila* polytene chromosomes (Schultz & Caspersson, 1939; Prokofyeva-Belgovskaya, 1939; Cole & Sutton, 1941; Zuckerkandl, 1974; Locke *et al.*, 1988). Furthermore, the inactivation of the endogenous loci is directional; loci closest to the breakpoint are more strongly affected than those further away, implying a spreading of the silencing process from the heterochromatic centromere. The degree of variegation of a locus can be reduced by relocation away from heterochromatin (Dubinin & Sidorov, 1935; Panshin, 1935; Kaufmann, 1942; Judd, 1955). Therefore it is proximity of the locus to heterochromatin *per se*, and not mutation of the locus due to translocation that is the cause of the variegated expression.

Evidence, however, for the mass-action model has been lacking. If the silencing process does indeed spread from the centromere then why does it not continue to spread along the whole of the chromosome as is thought to occur in X chromosome inactivation? It has been proposed that there are "insulator or blocking" elements interspersed along the chromosome that prevent the infinate spread of the silencing (Tartof *et al.*, 1984). However molecular evidence for the existance of these elements has not arisen (Kellum & Schedl, 1991) although a locus control region appears to insulate a transgene from variegated expression in mice

(Festenstein *et al.*, 1996). Close inspection of chromosome regions associated with some PEV rearrangements demonstrate discontinuous compaction. This suggests that instead of uniform spread along the chromsosome a more local heterochromatinisation process takes place (Belyaeva & Zhimulev, 1991). Furthermore, two variegating loci within the same cell can show differential variegation, condensation and binding to heterochromatin protein 1 (HP1; Bishop, 1992; Belyaeva *et al.*, 1993). This argues against the idea that the silencing is caused by the binding of proteins that are uniformly distributed throughout the nucleus. Perhaps the most striking example of variegation that cannot be explained by this model is the dominant PEV induced by certain alleles of the *brown* (*bw*) gene. The $brown^{Dominant}$ (bw^{D}) mutation results when a >2 Mb sequence of heterochromatin integrates within the *bw* gene. This can result in variegated "*trans*-inactivation" of the wild-type *bw* homologue on the unaltered chromosome (Slatis, 1955; see also 7.5.3.).

7.5.3. DNA looping model

Recently, two papers (Dorer & Henikoff, 1994; Sabl & Henikoff, 1996) reported that heterochromatinisation of a transgene in *Drosophila* is not only dependent upon proximity to the centromere but also on the number of repeats present at the transgene locus; silencing increased in step with copy number. The most recent study (Sabl & Henikoff, 1996) using a *Drosophila* brown eye colour transgene has resulted in a different model from the classical heterochromatic spreading hypothesis to explain PEV. The transgene $P(bw^{+})$ was inserted 50-75 kb from heterochromatin. Initially this showed an intermediate level of variegation. Transgene duplication or inversion resulted in different levels of variegation; reversal of orientation resulted in almost no inactivation, duplication in reversed or

the same orientation increased inactivation while triplication resulted in almost complete extinction of expression. It is hard to reconcile these observations with the model for a heterochromatinisation process spreading along the chromosome. Instead, the authors suggest a pairing-looping mechanism, whereby somatic pairing involving middle repetitive elements can lead to looping and sequestration of euchromatic genes into a heterochromatic compartment. In separate studies, homology-dependent silencing of bw^D and bw^+ is dependent on pairing of the chromosomes (Dreesen *et al.*, 1991) and the position of the bw^D allele relative to the centromere; variegation is suppressed when bw^D is translocated further from the centromere (Talbert *et al.*, 1994) and enhanced when bw^D is closer to the centromere (Henikoff *et al.*, 1995). Cytological studies demonstrate that the insertion of a large block of heterochromatin into *bw* causes the allele and its wild-type homologue (due to chromosome pairing) to be moved into a heterochromatic region of the nucleus (Csink & Henikoff, 1996; Dernburg *et al.*, 1996). These observations have led to the idea that the nucleus may be compartmentalised into regions that are permissive for gene transcription (in which euchromatin is located) and regions that are non-permissive for gene transcription (and in which heterochromatin is located) [see also Karpen, 1994]. In addition, sub-regions of euchromatin may be packaged into condensed chromatin structures or domains that prevent access to transcriptional activators (Locke *et al.*, 1988; Grigliatti, 1991; Wallrath & Elgin, 1995). Transcriptionally permissive and non-permissive domains may be regulated by macromolecular complexes that mediate changes in chromatin structure via nucleosome positioning (reviewed by Kingston *et al.*, 1996). Once established, the non-permissive transcriptional state can be maintained through successive cell divisions during mitosis, as is the case with the silent mating-type locus in yeast (Loo & Rine, 1995) and X chromosome inactivation in female eukaryotic cells

(Lyon, 1991; Brown & Willard, 1994), resulting in clones of expressing and non-expressing cells.

7.6. A model for transgene silencing in lines 7 and 45

While the mechanism responsible for BLG transgene silencing in some mammary cells but not in others remains to be elucidated, the phenomenon of PEV in *Drosophila* affords an informative parallel. The experiments described in this study argue that a similar process may occur in mouse. First, in both mouse lines exhibiting mosaic BLG expression the transgene array is integrated in the vicinity of the centromere while, in the stably-expressing line, the transgene locus was located some distance from the centromere. Second, gene variegation in *Drosophila* can extend over 50-60 polytene bands (1-2 Mb) from a rearrangement breakpoint (Demerec, 1940; Hartmann-Goldstein, 1967; Weiler & Wakimoto, 1995). While no precise measure can be provided, estimates from FISH chromosomal spreads suggest that the transgene insertion of line 7 is some 4-5 Mb from centromeric heterochromatin. This is broadly consistant with the situation found in *Drosophila*, particularly if relative genome sizes are taken into account (*Mus musculus* = 3.3 Gb, *Drosophila melanogaster* = 165 Mb; Carden, 1994). Third, the recent experiments by Henikoff and colleagues indicate that heterochromatinisation in *Drosophila* is not only dependent on proximity to the centromere but also on the number of repeats present at the transgene locus. In this study lines 7 and 45 both display mosaic expression and both have some 20 copies of the transgene, while the stably expressing line 14 contains only <5 copies. Furthermore, Mehtali *et al.*, (1990) described increased extinction of transgene activity (as judged by reporter gene expression and DNA methylation measurements) with increased copy number. Mosaic expression was reported for a

transgene integrated into the heterochromatic portion of the Y chromosome (Practcheva *et al.*, 1994) while heterochromatic features of an unusually large (1.1 Mb) β-globin transgene integrated into a peritelomeric region were also observed (Manuelidis, 1991). It is therefore plausible to suggest that both integration site and number of transgene copies predispose to the mosaic expression phenotype.

A link between heterochromatinisation and variegation in lines 7 and 45 has yet to be demonstrated. In *Drosophila* the degree of variegation is remarkably sensitive to the presence or absence of the Y chromosome (Karpen & Spradling, 1990). This is probably because *Drosophila* only has four autosomes. The heterochromatic Y chromosome is thought to act as a "sink" for heterochromatin-associated proteins; the presence of the Y chromosome drains the availability of heterochromatin-associated proteins from the variegating locus, the absence of the Y chromosome releases more heterochromatin-associated proteins. It would be interesting to test if a similar strategy would work in mice by using XO female mice harbouring the BLG transgene. Would this reduce the range of expression levels, lower the mean and increase the degree of variegation (number of non-expressing cells) within line 7? Major satellite DNA comprises up to 10% of the genome in *Mus* which is equivalent to ~330 Mb. The X chromosome represents 6.18% of the total genome (Evans, 1996) which is ~204 Mb. Therefore this experiment would remove more than one third of the heterochromatin. This may act as a significant release of heterochromatin-associated proteins that could affect BLG expression within line 7. An alternative strategy to investigate the possibility of a heterochromatic protein complex at the BLG locus would be to fractionate fragmented chromatin by adsorption to CREST, M32 or M31 antibodies that recognise heterochromatin-associated proteins. The extent of differential representation of transgene DNA in the bound and unbound fractions should correlate with expression levels.

To examine the possibility that the mosaic patterns reflect clonal expansion from committed progenitor cells (and determine whether there is a link between X chromosome inactivation and epigenetic silencing of euchromatic loci) it would be valuable to cross line 7 mice to animals harbouring an X-linked *lac*Z transgene (Tan *et al.*, 1994). The coincidence of β-gal and BLG expression would argue for clonal expansion.

Three heterochromatic silencing mechanisms could cause variegated expression of the BLG transgene locus. Firstly, as has been classically described for PEV in *Drosophila*, the BLG locus could be silenced due to heterochromatin-associated proteins spreading from the centromere (Fig. 39A).

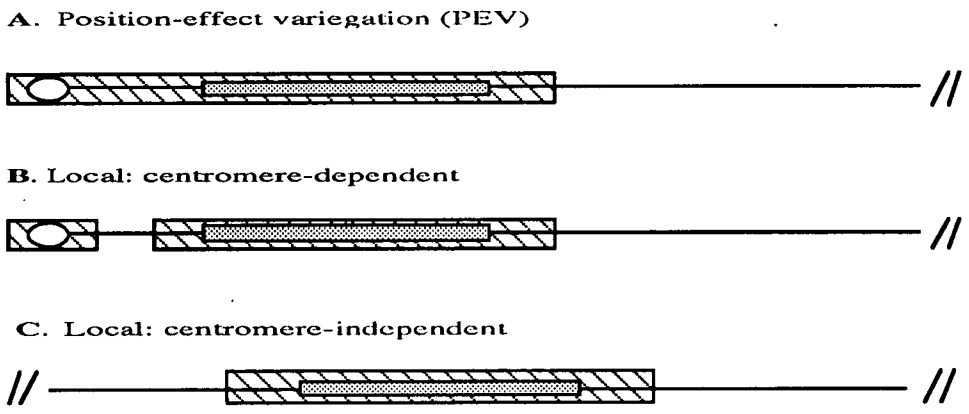


Fig. 39. Models for BLG silencing (heterochromatinisation)

Open circle, centromere; plain line, euchromatin; stipled boxes, transgene array; hatched boxes, heterochromatin

This could be tested by double chromosome *in situ* hybridisation mapping of endogenous genes in the vicinity the BLG integration locus (e.g., GHR and IL7R) and examining whether these genes also exhibit variegated expression. Secondly, the BLG locus could be silenced because of its proximity to the centromere (Fig. 39B). However, instead of heterochromatin-associated proteins spreading from the

centromere, this model predicts that pairing of the transgene locus with heterochromatic regions is facilitated by proximity, leading to transfer into a heterochromatic domain within the nucleus. Thirdly, the BLG locus could be silenced because of the repeat structure of the transgene array and the silencing process is independent of the position of the locus relative to the centromere (Fig. 39C). If heterochromatin-associated proteins are involved in silencing of the BLG locus then current literature would suggest that the second mechanism is the most likely cause of silencing in lines 7 and 45. The second and third mechanisms could be distinguished between by generating more transgenic lines with different copy numbers and integration sites and determine if there is a relationship between copy number, position relative to the centromere and transgene variegation.

7.7. PEV: A ubiquitous phenomenon?

The phenomenon of PEV has stimulated and held the interest of geneticists since its discovery. It represents one of a variety of diverse, yet related processes. These include X-chromosome inactivation in female mammals (see chapter 6), silencing of euchromatic genes in the vicinity of X-autosome translocation junctions in mice (Cattanach, 1974) or genes adjoining the Xq27 locus in human fragile-X syndrome (Laird, 1987), genetic imprinting in human disease (Scrable *et al.*, 1989; Hall, 1990; Nicholls, 1993), endogenous and transgene silencing in plants (Matzke & Matzke, 1995; Martienssen & Richards, 1995), the control of mating type in *Saccharomyces cerevisiae* (Laurenson & Rine, 1992; Rivier & Rine, 1992; Loo & Rine, 1995) and processes involved in heritable silencing of developmentally regulated genes (Eissenberg *et al.*, 1995; Orlando & Paro, 1995; Elgin, 1996). A crucial feature of PEV and related processes is that the states of silencing are stored as heritable chromatin structures instead of a feedback

loop of regulatory proteins that diffuse from place to place. The study of PEV and heterochromatin structure may provide useful models for understanding the molecular basis of general mechanisms involved in the genetic control of chromosome domains.

7.8. Implications for mosaic transgene expression

Mosaic expression of transgenes in mice is not uncommon (see chapter 1) and has implications for investigations utilising gene addition technology. For instance, in gene expression studies the pattern of transgene expression would give a less than complete picture of the pattern of expression of the endogenous gene. In medical research, gene addition is now being tested as a means of causing tissue specific cell ablation in the treatment of brain tumours (Culver *et al.*, 1992), or in the study of the consequences of the loss of thyroid hormone production without surgery in mice (Wallace *et al.*, 1994). Variegated expression in such experiments would prevent 100% ablation and therefore result in variable numbers of cells remaining. Similarily, the genetic manipulation of animal organs for transplant purposes might be prejudiced by unpredictable transgene expression (Cozzi & White, 1995). In biotechnological applications, such as the production of proteins of biomedical interest in the milk of transgenic livestock (Clark *et al.*, 1987; Archibald *et al.*, 1990; Wright *et al.*, 1991), mosaic expression will reduce the yield of product. Prescreening on the basis of chromosomal localisation and copy number could potentially avoid the need for large scale breeding experiments to verify the stability of expression, while the use of large transgene constructs (e.g., based on bacterial or yeast artificial chromosomes) would be expected to hinder DNA looping between repeat motifs (if this does indeed contribute to PEV) and yield stable expression patterns.

Since this project was completed other colleagues have reported variable expression within transgenic lines (Drs B. Binas and T. Burdon, pers. comm.; Drs S. Morley and J.J. Mullins, pers. comm.) suggesting that the phenomenon may not be uncommon. Particularly intriguing is the possibility that some endogenous genes within the mammalian genome, especially those close to heterochromatic regions (e.g., immunodeficiency, centromeric instability, facial anomalies syndrome [ICF; Gimelli *et al.*, 1993] and faciocapulohumeral muscular dystrophy [Zatz *et al.*, 1995]) or present as tandem duplications at a single locus, may be susceptible to the type of mosaic silencing described here. Recently, chromosome rearrangements located 4-400 kb (5' or 3') from genes have been associated with some human (e.g., SOX9 / campomelic dysplasia and autosomal sex reversal; PAX6 / aniridia) and mouse (*steel* locus / female sterility) genetic disorders (reviewed by Milot *et al.*, 1996; Bedell *et al.*, 1996). For example, detailed molecular analysis of two aniridia (absence of iris) human pedigrees have demonstrated that the disorder is associated with chromosome rearrangements greater than 85 kb downstream of the PAX6 gene (Fantes *et al.*, 1995). Although a link between the aberrant phenotypes and the genomic rearrangements has yet to be established, it is tempting to speculate that the rearrangements cause relocation of the affected loci into transcriptionally non-permissive chromatin domains. It would appear that our understanding of gross chromosomal architecture will be as critical in our understanding of gene regulation as our understanding of local *cis*-acting elements.

REFERENCES

Aker, M. & Huang, H.V. (1996). Extreme heterogeneity of minor satellite repeat arrays in inbred strains of mice. *Mammalian Genome 7*: 62-64.

Akeson, E.C. & Davisson, M.T. (1996). Centromeric heterochromatin variants. In: *Genetic variants and strains of the laboratory mouse.* (eds., Lyon, M.F., Rastan, S. & Brown, S.D.M.), 3rd edition, Vol. 2, Oxford University Press, U.K., pp. 1506-1508.

Ali, S. & Clark, A.J. (1988). Characterization of the gene encoding ovine beta-lactoglobulin. *J. Mol. Biol. 199*: 415-426.

Ali, S., McClenaghan, M., Simons, J.P. & Clark, A.J. (1990). Characterisation of the alleles encoding ovine β-lactoglobulins A and B. *Gene 91*: 201-207.

Allen, N.D., Cran, D.G., Barton, S.C., Hettle, S., Reik, W. and Surani, M.A.H. (1988). Transgenes as probes for active chromosomal domains in mouse development. *Nature 333*: 852-855.

Allen, N.D., Norris, M.L. & Surani, M.A. (1990). Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell 61*: 853-861.

Allshire, R.C., Javerzat, J.P., Redhead, N.J., Cranston, G. (1994). Position effect variegation at fission yeast centromeres. *Cell 76*: 157-169.

Al-Shawi, R., Burke, J., Jones, C.T., Simons, J.P. & Bishop, J.O. (1988). A *Mup* promoter-thymidine kinase reporter gene shows relaxed tissue-specific expression and confers male sterility upon transgenic mice. *Mol. Cell. Biol. 8*: 4821-4828.

Al-Shawi, R., Kinneard, J., Burke, J. & Bishop, J.O. (1990). Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect. *Mol. Cell. Biol. 10*: 1192-1198.

Andres, A.-C., Schonenberger, C.-A., Groner, B., Hennighausen, L., LeMeur, M. & Gerlinger, P. (1987). Ha-ras oncogene expression directed by a milk protein gene promoter: tissue specificity, hormonal regulation, and tumor induction in transgenic mice. *Proc. Natl. Acad. Sci. USA 84*: 1299-1303.

Angerer, L.Y. & Angerer, R.C. (1991). Localisation of mRNAs by *in situ* hybridisation. *Methods Cell Biol. 35*: 37-71.

Angerer, L.Y. & Angerer, R.C. (1992). *In situ* hybridisation to cellular RNA with radiolabelled RNA probes. In: *In Situ Hybridization; A Practical Approach*, (ed., Wilkinson, D.G.), Oxford University Press, Oxford, pp. 15-31.

Archibald, A.L., McClenaghan, M., Hornsey, V., Simons, J.P. & 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*: 5178-5182.

Aronow, B.J., Ebert, C.A., Valerius, M.T., Potter, S.S., Wiginton, D.A., Witte, D.P. & Hutton, J.J. (1995). Dissecting a locus control region: facilitation of enhancer function by extended enhancer-flanking sequences. *Mol. Cell Biol. 15*: 1123-1135.

Aschaffenburg, R. & Drewry, J. (1957). Improved method for the preparation of crystalline β-lactoglobulin and α-lactalbumin from cow's milk. *Biochem. J. 65*: 273-277.

Atwater, J.A., Wisdom, R. & Verma, I.M. (1990). Regulated mRNA stability. *Annu. Rev. Genet. 24*: 519-541.

Baker, W.K. (1968). Position-effect variegation. *Advan. Genet. 14*: 133-169.

Barash, I., Faerman, A., Ratovitsky, T., Puzis, R., Nathan, M., Hurwitz, D.R. & Shani, M. (1994). Ectopic expression of β-lactoglobulin/human serum albumin fusion genes in transgenic mice. *Transgen. Res. 3*: 141-151.

Barlow, D.P., Stoger, R., Herrmann, B.G., Saito, K., Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus. *Nature 349*: 84-87.

Barlow, D.P. (1995). Gametic imprinting in mammals. *Science 270*: 1610-1613.

Barr, M.L. & Carr, D.H. (1962). Correlation between sex chromatin and chromosomes. *Acta. Cytol. 6*: 34-45.

Bartolomei, M.S., Zemel, S. & Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. *Nature 351*: 153-155.

Bayna, E.M. & Rosen, J.M. (1990). Tissue specific expression of the rat whey acidic protein gene in transgenic mice. *Nucl. Acids Res. 18*: 2977-2985.

Bedell, M.A., Jenkins, N.A. & Copeland, N.G. (1996). Good genes in bad neighbourhoods. *Nature Genetics 12*: 229-232.

Bell, K. & McKenzie, H.A. (1967). The whey proteins of ovine milk: β-lactoglobulin. *J. Mol. Biol. 199*: 415-426.

Belyaeva, E.S. & Zhimulev, I.F. (1991). Cytogenetic and molecular aspects of position effect variegation in *Drosophila*. III. Continuous and discontinuous compaction of chromosomal material as a result of position effect variegation. *Chromosoma 100*: 453-466.

Belyaeva, E.S., Demarkova, O.V., Umbetova, L.F. & Zhimulev, I.F. (1993). Cytogenetic and molecular aspects of position effect variegation in *Drosophila*. V. Heterochromatin-associated protein HP1 appears in euchromatic chromosomal regions that are inactivated as a result of position effect variegation. *Chromosoma 102*: 583-590.

Berger, S.L. & Kimmel, A.R. (1987). Methods in enzymology: guide to molecular cloning techniques. Vol. 152, Academic Press.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. *Nature 321*: 209-213.

Bischoff, R., Degryse, E., Perraud, F., Dalemans, W., Ali-Hadji, D., Thepot, D., Devinoy, E., Houdebine, L.M. & Pavirani, A. (1992). A 17.6 kbp region located upstream of the rabbit WAP gene directs high level expression of a functional human protein variant in transgenic mouse milk. *FEBS 305*: 265-268.

Bishop, C.P. (1992). Evidence for intrinsic differences in the formation of chromatin domains in *Drosophila melanogaster*. *Genetics 132*: 1063-1069.

Bissell, M.J. & Hall, H.G. (1987). Form and function in the mammary gland: the role of extracellular matrix. In: The mammary gland. (eds., Neville, M. & Daniel, C.), Plenum Press, New York, pp. 97-146.

Bleck, G.T. & Bremel, R.D. (1993). Variation of expression of a bovine α-lactalbumin transgene in milk of transgenic mice. *J. Dairy Sci. 77*: 1897-1904.

Bolander, F.F. (1990). Differential characteristics of the thoracic abdominal mammary glands from mice. *Exp. Cell Res. 189*: 142-144.

Bonifer, C., Vidal, M., Grosveld, F. & Sippel, A.E. (1990). Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. *EMBO J. 9*: 2843-2848.

Boring, C.C., Squires, T.S. & Tong, T. (1992). Cancer statistics. *Ca. Cancer J. Clin. 42*: 19-38.

Boudreau, N., Sympson, C.J., Werb, Z. & Bissell, M.J. (1995). Suppression of ICE and apoptosis in mammary epithelial cells by extracellular matrix. *Science 267*: 891-893.

Boyes, J. & Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell 64*: 1123-1134.

Bradl, M., Larue, L. & Mintz, B. (1991). Malignant melanoma in transgenic mice. *Proc. Natl. Acad. Sci. USA 88*: 6447-6451.

Braun, R.E., Lo, D., Pinkert, C.A., Widera, G., Flavell, R.A., Palmiter, R.D. & Brinster, R.L. (1990). Infertility in male transgenic mice: disruption of sperm development by HSV-tk expression in post meiotic germ cells. *Biol. Reprod. 43*: 684-693.

Brem, G., Hartl, P., Besenfelder, U., Wolf, E., Zinovieva, N. & Pfaller, R. (1994). Expression of synthetic cDNA sequences encoding human insulin-like growth factor-1 (IGF-1) in the mammary gland of transgenic rabbits. *Gene 149*: 351-355.

Bresciani, F. (1971). Ovarian steroid control of cell proliferation in the mammary gland and cancer. In: *Basic actions of sex steroids on target organs*, (eds., Hubinont, P.O., Leroy, F. & Galand, P.), Karger, Basel, pp. 130-159.

Brinster, R.L., Chen, H.Y., Trumbauer, M., Senear, A.W., Warren, R. & Palmiter, R.D. (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene in eggs. *Cell 27*: 223-231.

Brock, M.L. & Shapiro, D.J. (1983). Estrogen stabilizes vitellogen mRNA against cytoplasmic degradation. *Cell 34*: 207-214.

Brookreson, A.D. & Turner, C.W. (1959). Normal growth of mammary gland in pregnant and lactating mice. *Proc. Soc. exp. Biol. Med. 102*: 744-745.

Brown, C.J. & Willard, H.F. (1994). The human X inactivation center is not required for the maintenance of X chromosome inactivation. *Nature 368*: 154-156.

Brownlee, G.G. (1987). The molecular pathology of haemophilia B. *Biochem. Soc. Trans. 15*: 1-8.

Buhler, T.A., Bruyere, T., Went, D.F., Stranzinger, G. & Burki, K. (1990). Rabbit β-casein promoter directs secretion of human interleukin-2 into the milk of transgenic rabbits. *Bio/Technology 8*: 140-143.

Burdon, T., Sankaran, L., Wall, R.J., Spencer, M. & Hennighausen, L. (1991). Expression of whey acidic protein during mammary development: evidence for

different mechanisms of regulation during pregnancy and lactation. *J. Biol. Chem. 266*: 6909-6914.

Burdon, T., Wall, R.J., Shamay, A., Smith, G.H. & Hennighausen, L. (1991). Over-expression of an endogenous milk protein gene in transgenic mice is associated with impaired mammary development and a *milchlos* phenotype. *Mech. Dev. 36*: 67-74.

Butner, K. & Lo, C.W. (1986). Modulation of *tk* expression in mouse percentromeric heterochromatin. *Mol. Cell. Biol. 6*: 4440-4449.

Campbell, G.S., Argetsinger, L.S., Ihle, J.N., Kelly, P.A., Rillema, J.A. & Carter-Su, C. (1994). Activation of JAK2 tyrosine kinase by prolactin receptors in Nb_2 cells and mouse mammary gland explants. *Proc Natl. Acad. Sci. 91*: 5232-5236.

Carden, M.J. (1994). Genome organisation. In: *The encyclopedia of molecular biology*. (eds., Kendrew, J. & Lawrence, E.), Blackwell, U.K., pp. 422.

Carlson, J.A., Rogers, B.B., Sifers, R.N., Hawkins, H.K., Finegold, M.J. & Woo, S.L.C. (1988). Multiple tissues express α1-antitrypsin in transgenic mice and man. *J. Clin. Invest. 82*: 26-36.

Casolaro, M.A., Fells, G., Wewers, M., Pierce, J.E., Ogushi, F., Hubbard, R., Sellers, S., Forstrom, J., Lyons, D., Kawasaki, G. & Crystal, R.G. (1987). Augmentation of lung antineutrophil elastase capacity with recombinant human α1-antitrypsin. *J. Appl. Physiol. 63*: 2015-2023.

Cattanach, B.M. (1963). The inactive X hypothesis and position effects in the mouse. *Genetics 48*: 884-885.

Cattanach, B.M. (1974). Position effect variegation in the mouse. *Genet. Res. Camb. 23*: 291-306.

Chaillet, J.R., Vogt, T.F., Beier, D.R., Leder, P. (1991). Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell 66*: 77-83.

Chen, L.H. & Bissell, M.J. (1989). A novel regulatory mechanism for whey acidic protein gene expression. *Cell Regul. 1*: 45-54.

Chomczynski, P. (1992). Solubilization in formamide protects RNA from degradation. *Nucl. Acids Res. 20*: 3791-3792.

Choo, K.H., Gould, K.G., Rees, D.J.G. & Brownlee, G.G. (1982). Molecular cloning of the gene for anti-haemophilic factor IX. *Nature 299*: 178-180.

Ciliberto, G., Dente, L. & Cortese, R. (1985). Cell-specific expression of a transfected human α_1-antitrypsin gene. *Cell 41*: 531-540.

Clark, A.J., Simons, J.P., Wilmut, I. & Lathe, R. (1987). Pharmaceuticals from transgenic livestock. *Trends Biotechnol 5*: 20-24.

Clark, R.F. & Elgin, S.C. (1992). Heterochromatin protein 1, a known suppressor of position-effect variegation, is highly conserved in *Drosophila*. *Nucleic Acids Res. 20*: 6067-6074.

Clarke, R.A., Sokol, D., Rigby, N., Ward, K., Murray, J.D. & Mackinlay, A.G. (1994). Mammary gland specific expression of bovine α_{S1}-casein derived transgenes in mice. *Transgenics 1*: 313-319.

Cogan, U., Kopelman, M., Mokady, S. & Shinitzky, M. (1976). Binding affinities of retinol and related compounds to retinol binding proteins. *Eur. J. Biochem. 6*: 71-78.

Cole, P.A. & Sutton, E. (1941). The absorption of ultraviolet radiation by bands of the salivary gland chromosomes of *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol. 9*: 66-70.

Cook, M.J. (1965). The anatomy of the laboratory mouse. Academic Press, London.

Courtney, M., Buchwalder, A., Tessier, L.-H., Jaye, M., Benavente, A., Balland, A., Kohli, V., Lathe, R., Tolstoshev, P. & Lecocoq, J-P. (1984). High-level production of biologically active human α1-antitrypsin in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA 81*: 669-673.

Cowie, A.T. & Tindal, J.S. (1971). Physiology of the mammary gland. In *The physiology of lactation.* Edwards Arnold, London, pp. 1-53.

Cowie, A.T., Forsyth, I.A. & Hart, I.C. (1980). Hormonal control of lactation. Springer-Verlag, Berlin.

Cozzi, E. & White, D.J.G. (1995). The generation of transgenic pigs as potential organ donors for humans. *Nature Medicine 1*: 964-966.

Csink, A.K. & Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in *Drosophila. Nature 381*: 529-531.

Culver, K.W., Ram, Z., Wallbridge, S., Ishii, H., Oldfield, E.H. & Blaese, R.M. (1992). In vivo transfer with retroviral vector-producer cells for treatment of experimental brain tumors. *Science 256*: 1550-1552.

Daniel, C.W. & Silberstein, G.B. (1987). Postnatal development of the rodent mammary gland. In: *The mammary gland: development, regulation and function* (eds., Neville, M.C. & Daniel, C.W.), Plenum Press: New York & London, pp. 3-36.

Daniels, S.B., McCarron, M., Love, C., Clark, S.H. & Chovnick, A. (1986). The underlying bases of gene expression in stable transformants of the *rosy* locus in *Drosophila melanogaster. Genetics 113*: 265-285.

Davies, D.T., Holt, C. & Christie, W.W. (1983). The composition of milk. In: *Biochemistry of lactation*, (ed., Mepham, T.B.), Elsevier Science Publication. B.V. Amsterdam, pp. 71-117.

Davies, T.D. & Law, A.J.R. (1983). Variation in the protein composition of bovine casein micelles and serum casein in relation to the micellar size and milk temperature. *J. Dairy Res. 50*: 67-75.

DaSilva, L., Rui, H., Erwin, R.A., Zack Howard, O.M., Kirken, R.A., Malabarba, M.G., Hackett, R.H., Larner, A.C. & Farrar, W.L. (1996). Prolactin recruits STAT1,

STAT3 and STAT5 independent of conserved receptor tyrosines TYR402, TYR479, TYR515 and TYR580. *Mol. Cell Endocin. 117*: 131-140.

David, M., Petricoin III, E.F., Igarashi, K-I., Feldman, G.M., Finbloom, D.S. & Larner, A.C. (1994). Prolactin activates the interferon-regulated p91 transcription factor and the Jak2 kinase by tyrosine phosphorylation. *Proc. Natl. Acad. Sci. USA 91*: 7174-7178.

DeChiara, T.M., Robertson, E.J. & Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell 64*: 849-859.

Demerec, M. (1940). Genetic behavior of euchromatic segments inserted into heterochromatin. *Genetics 25*: 618-627.

Dernburg, A.F., Broman, K.W., Fung, J.C., Marshall, W.F., Philips, J., Agard, D.A. & Sedat, J.W. (1996). Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell 85*: 745-759.

Devinoy, E., Thepot, D., Stinnakre, M.-G., Fontaine, M.-L., Grabowski, H., Puissant, C., Pavirani, A. & Houdebine, L.-M. (1994). High level production of human growth hormone in the milk of transgenic mice: the upstream region of the rabbit whey acidic protein (WAP) gene targets transgene expression to the mammary gland. *Transgenic Res. 3*: 79-89.

Diaz, P., Cado, D. & Winoto, A. (1994). A locus control region in the T cell receptor α/δ locus. *Immunity 1*: 207-217.

Dils, R.R. (1986). Comparitive aspects of milk fat synthesis. *J. Dairy Sci. 69*: 904-910.

Dorer, D.R. & Henikoff, S. (1994). Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. *Cell 77*: 993-1002.

Dreesen, T.D., Henikoff, S. & Loughney, K. (1991). A pairing-sensitive element that mediates *trans*-inactivation is associated with the *Drosophila brown* gene. *Genes Dev. 5*: 331-340.

Drohan, W.N., Zhang, D., Paleyanda, R.K. Chang, R., Wroble, M., Velander, W. & Lubon, H. (1994). Inefficient processing of human protein C in the mouse mammary gland. *Transgenic Res. 3*: 355-364.

Dubinin, N.P. & Sidorov, B.N. (1935). The position effect of the hairy gene. *Biol. Zh. Mosk. 4*: 555-568.

Dulbecco, R., Henahan, M. & Armstrong, B. (1982). Cell types and morphogenesis in the mammary gland. *Proc. Natl. Acad. Sci. USA 79*: 7346-7350.

Durnberger, H. & Kratochwill, K. (1980). Specificity of tissue interaction and origin of mesenchymal cells in the androgen response of the embryonic mammary gland. *Cell 19*: 465-471.

Dusanter-fourt, I., Muller, O., Ziemiecki, A., Mayeux, P., Drucker, B., Djiane, J., Wilks, A., Harpur, A.G., Fisher, S. & Gisselbrecht, S. (1994). Identification of JAK protein tyrosine kinases as signaling molecules for prolactin. Functional analysis of prolactin receptor and prolactin-erythropoietin receptor chimera expressed in lymphoid cells. *EMBO J. 13*: 2583-2591.

Ebert, K.M., Selgrath, J.P., DiTullio, P., Denman, J., Smith, T.E., Memon, M.A., Schindler, J.E., Monastersky, G.M., Vitale, J.A. & Gorden, K. (1991). Transgenic production of a variant of human tissue plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. *Bio/Technology 9*: 835-838.

Ehrlich, M. & Ehrlich, K.C. (1993). Gene methylation patterns and expression. In: *DNA methylation: molecular biology and biological significance*. (eds., Jost, J.P. & Saluz, H.P.), Birkauser Verlag, Basel, pp. 145-168.

Eicher, E.M. (1970). X-autosome translocations in the mouse: total inactivation versus partial inactivation of the X chromosome. *Adv. Genet. 15*: 175-259.

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V. & Elgin, S.C.R. (1990). Mutation in a heterochromatin-specific chromosomal protein is

associated with the suppression of position-effect variegation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA 87*: 9923-9927.

Eissenberg, J.C. & Elgin, S.C. (1991). Boundary functions in the control of gene expression. *Trends Genet. 7*: 335-340.

Eissenberg, J.C., Morris, G.D., Reuter, G. & Hartnett, T. (1992). The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics 131*: 345-352.

Eissenberg, J.C., Elgin, S.C.R. & Paro, R. (1995). Epigenetic regulation in *Drosophila*: a conspiracy of silence. In: *Chromatin structure and gene expression.* (ed., Elgin, S.C.R.), Oxford University Press, U.K., pp. 147-171.

Elgin, S.C.R. (1996). Heterochromatin and gene regulation in Drosophila. *Curr. Opin. Genet. Dev. 6*: 193-202.

Engler, P., Haasch, D., Pinkert, C.A., Doglio, L., Glymour, M., Brinster, R., and Storb, U. (1991). A strain-specific modifier on mouse chromosome 4 controls the methylation of independent transgene loci. *Cell 65*: 939-947.

Enver, T., Li, Q., Gale, K.B., Hu, M., May, G.E., Karlinsey, J.E., Jimenez, G., Papayannopoulou, T. & Costantini, F. (1994). Analysis of the developmental and transcriptional potentiation functions of 5'HS2 of the murine β-globin locus control region in transgenic mice. *Dev. Biol. 165*: 574-584.

Evans, E.P. (1981). Karyotype of the house mouse. *Symp. Zool. Soc. London 47*: 127-139.

Evans, E.P. (1996). Standard normal chromosomes. In: *Genetic variants and strains of the laboratory mouse*. (eds., Lyon, M.F., Rastan, S. & Brown, S.D.M.), 3rd edition, Vol. 2, Oxford University Press, U.K., pp. 1446.

Faerman, A., Barash, I., Puzis, R., Nathan, M., Hurwitz, D.R. & Shani, M. (1995). Dramatic heterogeneity of transgene expression in the mammary gland of lactating

mice: a model system to study the synthetic activity of mammary epithelial cells. *J. Histoch. Cytoch. 43*: 461-470.

Fantes, J.A., Oghene, K., Boyle, S., Danes, S., Fletcher, J.M., Bruford, E.A., Williamson, K., Seawright, A., Schedl, A., Hanson, I., Zehetner, G., Bhogal, R., Lehrach, H., Gregory, S., Williams, J., Little, P.F.R., Sellar, G.C., Hoovers, J., Mannens, M., Weissenbach, J., Junien, C., Heyningen, V.V. & Bickmore, W. (1995). A high-resolution integrated physical, cytogenetic, and genetic map of human chromosome 11: distal p13 to proximal p15.1. *Genomics 25*: 447-461.

Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., Mannens, M., Danes, S., van Heyningen, V. & Hanson, I. (1995). Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum. Mol. Genet. 4*: 415-422.

Farkas, G., Gausz, J., Galloni, M., Keuter, G., Gyurkovics, H. & Karch, F. (1994). The *Trithorax*-like gene encodes the *Drosophila* GAGA factor. Nature 371: 806-808.

Fauvarque, M.O. & Dura, J.M. (1993). Polyhomeotic regulatory sequences induce developmental regulator-dependent variegation and targeted P-element insertions in *Drosophila*. *Genes Dev. 7*: 1508-1520.

Feinberg, A.P. & Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem. 132*: 6-13.

Feinberg, A.P. & Vogelstein, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Addendum. Anal. Biochem. 137*: 266-267.

Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Miliou, A., Jones, M. & Kioussis, D. (1996). Locus control region function and heterochromatin-induced position effect variegation. *Science 271*: 1123-1125.

Forsyth, I.A. (1986). Variation among species in the endocrine control of mammary growth and function: the roles of prolactin, growth hormone and placental lactogen. *J. Dairy Sci. 69*: 886-903.

Foster, R.C. (1977). Changes in mouse mammary epithelial cell size during mammary gland development. *Cell Diff. 6*: 1-8.

Fugate, R.D. & Song, P.S. (1980). Spectroscopic characterization of β-lactoglobulin -retinol complex. *Biochim. Biophys. Acta. 625*: 28-42.

Futterman, S. & Heller, J. (1972). The enhancement of florescence and the decreased susceptibility to enzymatic oxidation of retinol complexed with bovine serum albumin, β-lactoglobulin, and the retinol-binding protein of human plasma. *J. Biol. Chem. 247*: 5168-5172.

Gartler, S.M., Andina, R.T. & Gant, N. (1975). Ontology of X chromosome inactivation in the female germ line. *Exp. Cell Res. 91*: 454-457.

Gartler, S.M. & Riggs, A.D. (1983). Mammalian X-chromosome inactivation. *Ann. Rev. Genet. 17*: 155-190.

Garver, R.I., Chytil, A., Karlsson, S., Fells, G.A., Brantly, M.L., Courtney, M., Kantoff, P.W., Nienhuis, A.W., French Anderson, W. & Crystal, R.G. (1987). Production of glycosylated physiologically "normal" human α1-antitrypsin by mouse fibroblasts modified by insertion of a human α1-antitrypsin cDNA using a retroviral vector. *Proc. Natl. Acad. Sci. USA 84*: 1050-1054.

Gaye, P., Hue-Delahaie, D., Mercier, J.C., Soulier, S., Vilotte, J.L. & Furet, J.P. (1986). Ovine β-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland. *Biochimie 68*: 1097-1107.

Giddings, S.J., King, C.D., Harman, K.W., Flood, J.F. & Carnaghi, L.R. (1994). Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. *Nature Genet. 6*: 310-313.

Gilmour, K.C. & Reich, N.C. (1994). Receptor to nucleus signaling by prolactin and interleukin 2 via activation of latent DNA-binding factors. *Proc. Natl. Acad. Sci. USA 91*: 6850-6854.

Gimelli, G., Varone, P., Pezzolo, A., Lerone, M. & Pistoia, V. (1993). ICF syndrome with variable expression in sibs. *Med. Genet. 30*: 429-432.

Gordon, J.W., Scangos, G.A., Plotkin, D., Barbosa, J.A. & Ruddle, F.H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci. USA 77*: 77380-77384.

Gordon, J.W. & Ruddle, F.H. (1981). Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science 214*: 1244-1246.

Gordon, K., Lee, E., Vitale, J.A., Smith, A.E., Westphal, H. & Hennighausen, L. (1987). Production of human tissue plasminogen activator in transgenic mouse milk. *Bio/Technology 5*: 1183-1187.

Gottschling, D.E., Aparico, O.M., Billington, B.L. & Zakian, V.A. (1990). Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell 63*: 751-762.

Gouilleux, F., Wakao, H., Mundt, M. & Groner, B. (1994). Prolactin induces phosphorylation of Tyr694 of Stat5 (MGF), a prerequisite for DNA binding and induction of transcription. *EMBO J. 13*: 4361-4369.

Grabowski, H., Le Bars, D., Chene, N., Attal, J., Malienou-N'Gassa, R., Puissant, C. & Houdebine, L.M. (1991). Rabbit whey acidic protein concentration in milk, serum, mammary gland extract and culture medium. *J. Dairy Sci. 74*: 4143-4154.

Graham, J. & Piper, J. (1994). Automatic karyotype analysis. In: *Methods in molecular biolgy:Chromosome analysis protocols.* (ed., Gosden, J.R.), Humana Press, Totowa, New Jersey, pp. 141-185.

Grazino, V., Pereira, A., Laurenti, P., Graba, Y., Levis, R.W., Parco Le, Y. & Pradel, J. (1992). Cell lineage-specific expression of *modulo*, a dose-dependent modifier of variegation in *Drosophila*. *EMBO J. 11*: 4471-4479.

Greaves, D.R., Wilson, F.D., Lang, G. & Kioussis, D. (1989). Human CD2 3'-flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. *Cell 56*: 979-986.

Grigliatti, T. (1991). Position-effect variegation - an assay for non-histone chromosomal proteins and chromatin assembly and modifying factors. *Methods Cell Biol. 35*: 587-627.

Grosveld, F., Blom van Assendelft, G., Greaves, D.R. & Kollias, G. (1987). Position-independent, high-level expression on the human β-globin gene in transgenic mice. *Cell 51*: 975-985.

Guillemot, F., Caspary, T., Tilghman, S.M., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Anderson, D.J., Joyner, A.L., Rossant, J. & Nagy, A. (1995). Genomic imprinting of *Mash2*, a mouse gene required for trophoblast development. *Nature Genet. 9*: 235-242.

Gupta, P., Rosen, J.M., D'Eustachio, P. & Ruddle, F.H. (1982). Localization of the casein gene family to a single mouse chromosome. *J. Cell. Biol. 93* : 199-204.

Gutierrez, A., Meade, H.M., Jimenez-Flores, R., Anderson, G.B., Murray, J.D. & Medrano, J.F. (1996). Expression of a bovine κ-casein cDNA in the mammary gland of transgenic mice utilizing a genomic milk protein gene as an expression cassette. *Transgenic Res. 5*: 271-279.

Hall, J.G. (1990). How imprinting is relevant to human disease. *Development* Supplement, 141-148.

Hanrahan, J.P. & Eisen, E.J. (1970). A lactation curve for mice. *Lab. Anim. Care 20*: 101-104.

Hansson, L., Edlund, M., Edlund, A., Johansson, T., Marklund, S.L., Fromm, S., Stromqvist, M. & Tornell, J. (1994). Expression and characterization of biologically active human extracellular superoxide dismutase in milk of transgenic mice. *J. Biol. Chem. 69*: 5358-5363.

Harris, S., Ali, S., Anderson, S., Archibald, A.L. & Clark, A.J. (1988). Complete nucleotide sequence of the genomic ovine β-lactoglobulin gene. *Nucl. Acids Res. 16*: 10379-10380.

Harris, S., McClenghan, M., Simons, J.P., Ali, S. & Clark, A.J. (1991). Developmental regulation of the sheep β-lactoglobulin gene in the mammary gland of transgenic mice. *Dev. Genet. 12*: 299-307.

Hartmann-Goldstein, I.J. (1967). On the relationship between heterochromatization and variegation in Drosophila with special reference to temperature-sensitive periods. *Genet. Res. 10*: 143-159.

Hastie, N.D. (1996). Highly repeated DNA families in the genome of *Mus musculus*. In: *Genetic variants and strains of the laboratory mouse.* (eds., Lyon, M.F., Rastan, S. & Brown, S.D.M.), 3rd edition, Vol. 2, Oxford University Press, U.K., pp. 1425-1442.

Hatada, I. & Mukai, T. (1995). Genomic imprinting of $p57^{KIP2}$, a cyclin-dependent kinase inhibitor, in mouse. *Nature Genet. 11*: 204-206.

Hayashi, S., Rudell, A., Sinclair, D. & Grigliatti, T. (1990). Chromosome structure is altered by mutations that suppress or enhance position effect variegation. *Chromosoma 99*: 391.

Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Sasaki, N., Hirose, K., Imoto, H., Okuizumi, H., Muramatsu, M., Komatsubara, H., Shiroishi, T., Moriwaki, K., Katsuki, M., Hatano, N., Sasaki, H., Ueda, T., Mise, N., Takagi, N., Plass, C. & Chapman, V.M. (1994). Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method. *Nature Genet. 6*: 33-40.

Hazelrigg, T., Levis, R. & Rubin, G.M. (1984). Transformation of *white* locus DNA in *Drosophila*: Dosage compensation, *zeste* interaction, and position effects. *Cell 36*: 396-381.

Henikoff, S. (1981). Position-effect variegation and chromosome structure of a heat shock puff in *Drosophila*. *Chromosoma 83*: 381-

Henikoff, S. (1992). Position effects and related phenomena. *Curr. Opin. Genet. Dev. 2*: 907-912.

Henikoff, S., Jackson, J.M. & Talbert, P.B. (1995). Distance and pairing effects on the *brown*^{DOMINANT} heterochromatic elements in *Drosophila*. *Genetics 140*: 1007-1017.

Hennighausen, L.G. & Sippel, A.E. (1982). Mouse whey acidic protein is a novel member of the family of 'four-sulphide core' proteins. *Nucl. Acids Res. 10*: 2677-2684.

Hennighausen, L., Westphal, C., Sankaran, L. & Pittius, C.W. (1988). Regulation of expression of milk protein genes. In: *Transgenic technology in medicine and agriculture*. (eds., First, N. & Haseltine, F.), Boston: Butterworth, pp. 61-70.

Hennighausen, L., Ruiz, L. & Wall, R. (1990). Transgenic animals - production of foreign proteins in milk. *Curr. Opin. Biotechnol. 1*: 74-78.

Hobbs, A.A., Richards, D.A., Kessler, D.J. & Rosen, J.M. (1982). Complex hormonal regulation of casein gene expression. *J. Biol. Chem. 257*: 3598-3605.

Hogan, B., Constantini, F. & Lacy, E. (1986). Manipulating the mouse embryo: A laboratory manual. *Cold spring Laboratory Press.*

Holliday, R. & Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science 187*: 226-232.

Iguchi-Ariga, S.M.M. & Schaffner, W. (1989). CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev. 3*: 612-619.

Imagawa, W., Yang, J., Guzman, R. & Nandi, S. (1994). Control of mammary gland development. In: *The physiology of reproduction.* (eds., Knobil, E. & Neill, J.D.), Vol. 2, 2nd ed., Raven Press, New York, pp. 1033-1065.

Jahner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Lohler, J., Simon, I. & Jaenisch, R. (1982). *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. *Nature 298*: 623-628.

James, T.C. & Elgin, S.C.R. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila and its gene. *Mol. Cell Biol. 6*: 3862-3872.

Jaye, M., De La Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P. & Lecocq, J.P. (1983). Isolation of a human antihaemophilic factor cDNA using a unique 52 base synthetic oligonucleotide probe deduced from the amino acid sequence of bovine factor IX. *Nucl. Acids Res. 11*: 2325-2335.

Jenness, R. (1982). Inter-species comparison of milk proteins. In: *Developments in dairy chemistry* (ed., Fox, P.F.), Vol. 1, Applied Sciences Publication, New York, pp. 87-114.

Jenness, R. (1986). Lactational performance of various mammalian species. *J. Dairy Sci. 69*: 869-885.

Jeppesen, P. & Nicol, L. (1986). Non-kinetochore directed autoantibodies in scleroderma/CREST: identification of an activity recognising a metaphase chromosome core non-histone protein. *Mol. Biol. Med. 3*: 369-384.

Jeppesen, P., Mitchell, A., Turner, B. & Perry, P. (1992). Antibodies to defined histone epitopes reveal variations in chromatin conformation and underacetylation of centric heterochromatin in human metaphase chromosomes. *Chromosoma 101*: 322-332.

Jones, B.K., Monks, B.R., Liebhaber, S.A. & Cooke, N.E. (1995). The human growth hormone gene is regulated by a multicomponent locus control region. *Mol. Cell Biol. 15*: 7010-7021.

Joseph, A., Mitchell, A.R. & Miller, O.J. (1989). The organisation of mouse satellite DNA at centromeres. *Exp. Cell Res. 183*: 494-500.

Judd, B.H. (1955). Direct proof of a variegated-type position effect at the *white* locus in *Drosophila melanogaster. Genetics 27*: 537-549.

Kafri, T.M., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. & Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes & Dev. 6*: 705-714.

Kanda, N. (1973). A new differential technique for staining the heteropycnotic X-chromosome in female mice. *Exptl. Cell Res. 80*: 463-467.

Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S.C., Ishino, F. & Surani, M.A. (1995). *Peg1/Mest* imprinted gene on chromosome 6 identified by cDNA subtraction hybridisation. *Nature Genet. 11*: 52-59.

Karpen, G.H. & Spradling, A.C. (1990). Reduced DNA polytenization of a minichromosome region undergoing position-effect variegation in *Drosophila. Cell 63*: 97-107.

Karpen, G.H. & Spradling, A.C. (1992). Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome *Dp1187* by single *P* element insertional mutagenesis. *Genetics 132*: 737-753.

Karpen, G.H. (1994). Position-effect variegation and the new biology of heterochromatin. *Curr. Opin. Genet. Dev. 4*: 281-291.

Katsuki, M., Sato, M., Kimura, M., Yokoyama, M., Kobayashi, K. & Nomura, T. (1988). Conversion of normal behavior to shiverer by myelin basic protein antisense cDNA in transgenic mice. *Science 24*: 593-595.

Kaufmann, B.P. (1942). Reversion from roughest to wild type in *Drosophila melanogaster*. *Genetics 27*: 537-549.

Kellum, R. & Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell 64*: 941-950.

Kellum, R. & Alberts, B.M. (1995). Heterochromatin protein 1 is required for correct chromosome segregation in *Drosophila* embryos. *J. Cell Sci. 108*: 1419-1431.

Kelsey, G.D., Povey, S., Bygrave, A.E. & Lovell-Badge, R.H. (1987). Species- and tissue-specific expression of human α1-antitrypsin in transgenic mice. *Genes Dev. 1*: 161-171.

Khyse-Anderson, J. (1984). Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J. biochem. biophys. Meth. 10*: 203-209.

Kingston, R.E., Bunker, C.A. & Imbalzano, A.N. (1996). Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev. 10*: 905-920.

Kipling, D., Ackford, H.E., Taylor, B.A. and Cooke, H.J. (1991). Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. *Genomics 11*: 235-241.

Kipling, D., Wilson, H.E., Mitchell, A.R., Taylor, B.A. & Cooke, H.J. (1994). Mouse centromere mapping using oligonucleotide probes that detect variants of the minor satellite. *Chromosoma 103*: 46-55.

Kipling, D., Mitchell, A.R., Masumoto, H., Wilson, H.E., Nicol, L. & Cooke, H.J. (1995). CENP-B binds a novel centromeric sequence in the Asian mouse *Mus caroli*. *Mol. Cell. Biol. 15*: 4009-4020.

Knight, C.H. & Peaker, M. (1982). Development of the mammary gland. *J. Reprod. Fert. 65*: 521-536.

Knight, C.H., Maltz, E. & Docherty, A.H. (1986). Milk yield and composition in mice: effects of litter size and lactation number. *Comp. Biochem. Physiol. 84A*: 127-133.

Kold, H.-J. & Braunitzer, G. (1983). The primary structure of ovine β-lactoglobulin, 2. Discussion and genetic aspects. *Milchwissenschaft 38*: 70-72.

Kothary, R.K., Clapoff, S., Brown, A. Campbell, R., Peterson, A. and Rossant, J. (1988). A transgene containing *lacZ* inserted into the *dystonia* locus is expressed in neural tube. *Nature 335*: 435-437.

Kothary, R.K., Allen, N.D. & Surani, M.A.H. (1989). Transgenes as molecular probes of mammalian developmental genetics. In: *Oxford surveys of eukaryotic gene expression.* (ed., McClean, N.), Vol. 6, Oxford University Press, pp. 145-178.

Kratochwill, K. (1971). In vitro analysis of the hormonal basis for the sexual dimorphism in the embryonic development of the mouse mammary gland. *J. Embryol. Exp. Morphol. 25*: 141-153.

Kratzer, P.G. & Chapman, V.M. (1981). X chromosome reactivation in oocytes of *Mus caroli. Proc. Natl. Acad. Sci. USA 78*: 3093-3097.

Krnacik, M.J., Li, S., Liao, J. & Rosen, J.M. (1995). Position-independent expression of whey acidic protein transgenes. J. Biol. Chem. 270: 11119-11129.

Kuhn, M.J. (1983). The biosynthesis of lactose. In: *Biochemistry of lactation.* (ed., Mepham, T.B.), Elsevier Science Publication. B.V. Amsterdam, pp. 159-176.

Kumar, S. (1993). Gene targeting in embryonic stem cells and the manipulation of milk composition. Ph.D Thesis, The University of Edinburgh.

Kumar, S., Clarke, A.R., Hooper, M.L., Horne, D.S., Law, A.J.R., Leaver, J., Springbett, A., Stevenson, E. & Simons, J.P. (1994). Milk composition and lactation of β-casein-deficient mice. *Proc. Natl. Acad. Sci. USA 91*: 6138-6142.

Kurachi, K. & Davie, E.W. (1982). Isolation and characterization of a cDNA coding for human factor IX. *Proc. Natl. Acad. Sci. USA 79*: 6461-6464.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*: 680-685.

Laird, C.D. (1987). Proposed mechanism of inheritance and expression of the human fragile-X syndrome of mental retardation. *Genetics 117*: 587-599.

Lang, G., Wotton, D., Owen, M.J., Sewell, W.A., Brown, M.H., Mason, D.Y., Crumpton, M.J. & Kioussis, D. (1988). The structure of the human CD2 gene and its expression in transgenic mice. *EMBO J. 7*: 1675-1682.

Langer-Safer, P.R., Levine, M. & Ward, D.C. (1982). Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proc. Natl. Acad. Sci. USA 79*: 4381-4385.

Lathe, R. (1985). Molecular tailoring of the farm animal germline. AFRC ABRO report, pp. 7-10.

Lathe, R., Clark, A.J., Archibald, A.L., Bishop, J.O., Simons, J.P. & Wilmut, I. (1986). Novel products from livestock. In *Exploiting new technologies in animal breeding: genetic developments*, (eds., Smith, C., King, J.W.B. & McKay, J.C.), Oxford University Press, Oxford, pp. 91-102.

Laurenson, P. & Rine, J. (1992). Silencers, silencing, and heritable transcription states. *Microbiol. Rev. 56*: 543-560.

Lee, E.Y.-H., Parry, G. & Bissell, M.J. (1984). Modulation of secreted proteins of mouse mammary epithelial cells by the collagenous substrata. *J. Cell Biol. 98*: 146-155.

Lee, E.Y.-H., Lee, W.-H., Kaetzel, C.S., Parry, G. & Bissell, M.J. (1984). Interaction of the mouse mammary epithelial cells with collagen substrata: regulation of casein gene expression and secretion. *Proc. Natl. Acad. Sci. USA 82*: 1419-1423.

Lee, K.-F., DeMayo, F.J., Atiee, S.H. & Rosen, J.M. (1988). Tissue specific expression of the rat β-casein gene in transgenic mice. *Nucl. Acids Res. 16*: 1027-1041.

Lee, K.-F., Atiee, S.H. & Rosen, J.M. (1989). Differential regulation of rat β-casein-chloramphenicol acetyltransferase fusion gene expression in transgenic mice. *Mol. Cell Biol. 9*: 560-565.

Leff, S.E., Brannan, C.I., Reed, M.L., Ozcelik, T., Francke, U., Copeland, N.G. & Jenkins, N.A. (1992). Maternal imprinting of the mouse *Snrpn* gene and conserved linkage homology with the human Prader-Willi syndrome region. *Nature Genet. 2*: 259-264.

Levine, J.F. & Stockdale, F.E. (1985). Cell-cell interactions promote mammary epithelial cell differentiation. *J. Cell Biol. 100*: 1415-1422.

Levis, R. (1989). Viable deletions of a telomere from a *Drosophila* chromosome. *Cell 58*: 791-801.

Lewis, E.B. (1950). The phenomenon of position effect. *Advan. Genet. 3*: 73-115.

Li, S. & Rosen, J.M. (1994). Glucocorticoid regulation of rat whey acidic protein gene expression involves hormone-induced alterations of chromatin structure in the distal promoter region. *Mol. Endocrinol. 8*: 1328-1335.

Li, S. & Rosen, J.M. (1995). CTF/NF1 and mammary gland factor (STAT5) play a critical role in regulating rat whey acidic protein gene expression in transgenic mice. *Mol. Cell Biol. 15*: 2063-2070.

Lica, L.M., Narayanswami, S. & Hamkalo, B.A. (1986). Mouse satellite DNA, centromere structure, and sister chromatid pairing. *J. Cell Biol. 103*: 1145-1151.

Lin, C.Q. & Bissell, M.J. (1993). Multi-faceted regulation of cell differentiation by extracellular matrix. *FASEB J. 7*: 737-743.

Locke, J. Kotarski, M.A. & Tartof, K.D. (1988). Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effects. *Genetics 120*: 181-198.

Loo, E. & Rine, J. (1995). Silencing and heritable domains of gene expression. *Annu. Rev. Genet. 11*: 520-548.

Lovrien, R. & Anderson, W.F. (1969). Resolution of binding sites in β-lactoglobulin. *Arch. Biochem. Biophys. 131*: 139-144.

Lyon, M.F. (1961). Gene action in the X chromosome of the mouse (*Mus musculus*). *Nature 190*: 372-373.

Lyon, M.F. (1991). The quest for the X inactivation centre. *Trends Genet. 7*: 69-70.

Maga, E.A., Anderson, G.B., Huang, M.C. & Murray, J.D. (1994). Expression of human lysozyme mRNA in the mammary gland of transgenic mice. *Transgenic Res. 3*: 36-42.

Maga, E.A. & Murray, J.D. (1995). Mammary gland expression of transgenes and the potential for altering the properties of milk. *Bio/Technology 13*: 1452-1457.

Manuelidis, L. (1991). Heterochromatic features of an 11-megabase transgene in brain cells. *Proc. Natl. Acad. Sci. USA 88*: 1049-1053.

Martienssen, R.A. & Richards, E.J. (1995). DNA methylation in eukaryotes. *Curr. Opin. Genet. Dev. 5*: 234-242.

Martin-Morris, L.E., Loughney, K., Kershisnik, E.O., Poortinga, G. & Henikoff, S. (1993). Characterization of sequences responsible for *trans*-inactivation of the *Drosophila brown* gene. *Cold Spring Harbor Symp. Quant. Biol. 58*: 577-584.

Masson, P.L. & Heremans, J.F. (1971). Lactoferrin in milk from different species. *Comp. Biochem. Physiol. 39B*: 119-129.

Matsuda, Y. & Chapman, V.M. (1991). *In situ* analysis of centromeric satellite DNA segregating in *Mus* species crosses. *Mammalian Genome 1*: 71-77.

Matzke, M.A. & Matzke, A.J.M. (1995). How and why plants inactivate homologous (trans)genes? *Plant Physiol. 107*: 679-685.

May, G. & Enver, T. (1995). Targeting gene expression to haemopoietic stem cells: a chromatin-dependent upstream element mediates cell-type specific expression of the stem cell antigen CD34. *EMBO J. 14*: 564-574.

McClenaghan, M., Springbett, A., Wallace, R.M., Wilde, C.J. & Clark, A.J. (1995). Secretory proteins compete for production in the mammary gland of transgenic mice. *Biochem J. 310*: 637-641.

McGowan, R., Campbell, R., Peterson, A. & Sapienza, C. (1989). Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse. *Genes Dev. 3*: 1669-1676.

McKnight, R.A., Burdon, T., Pursel, V.G., Shamay, A., Wall, R.J. & Hennighausen, L. (1991). The whey acidic protein. In: *Breast cancer: molecular and cellular biology*. (eds., Dickson & Lippman), Boston: Kluwer Academic Publishers, pp. 399-412.

McMahon, A., Fosten, M. & Monk, M. (1981). Random X chromosome inactivation in female primordial germ cells in the mouse. *J. Embryol. Exp. Morphol. 64*: 251-258.

Meade, H., Gates, L., Lacy, E. & Lonberg, N. (1990). Bovine α_{S1}-casein gene sequences direct high level expression of active human urokinase in mouse milk. *Bio/Technology 8*: 443-446.

Medina, D. (1996). The mammary gland: a unique organ for the study of development and tumorigenesis. *J. Mam. Gland Biol. Neop. 1*: 5-19.

Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L. & Bird, A.P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell 58*: 499-507.

Mehtali, M., LeMeur, M. & Lathe, R. (1990). The methylation-free status of a housekeeping transgene is lost at high copy number. *Gene 91*: 179-184.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. & Green, M.R. (1984). Efficient *in vitro* synthesis of biologically active RNA and DNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucl. Acids Res. 12*: 7035-7056.

Mepham, T.B. (1987). *Physiology of lactation*. Open University Press, Milton Keynes and Philadelphia.

Mercier, J.C., Gaye, P., Soulier, S., Hue-Delahaie, D. & Vilotte, J.L. (1985). Construction and identification of recombinant plasmids carrying cDNAs coding for ovine α_{s1}-, α_{s2}-, β-, κ-casein and β-lactoglobulin. Nucleotide sequence of α_{s1}-casein cDNA. *Biochimie 67*: 959-971.

Milot, E., Fraser, P. & Grosveld, F. (1996). Position effects and genetic disease. *Trends Genet. 12*: 123-126.

Mintz, B. (1970). Gene expression in allophenic mice. *Symp. Int. Soc. Cell Biol. 9*: 15-42.

Mintz, B. & Bradl, M. (1991). Mosaic expression of a tyrosinase fusion gene in albino mice yields a heritable striped coat pattern in transgenic homozygotes. *Proc. Natl. Acad. Sci. USA 88*: 9643-9647.

Moehrle, A. & Paro, R. (1994). Spreading the silence: epigenetic transcriptional regulation during *Drosophila* development. *Dev. Genet. 15*: 478-484.

Moens, P.B. & Pearlman, R.E. (1990). Telomere and centromere DNA are associated with the cores of meiotic prophase chromosomes. *Chromosoma 100*: 8-14.

Molenaar, A.J., Davis. & Wilkins, R.J. (1992). Expression of α-lactalbumin, α-S1-casein, and lactoferrin gene is heterogenous in sheep and cattle mammary tissue. *J. Histoch. Cytoch. 40*: 611-618.

Monaghan, A.P., Davidson, D.R., Sime, C., Graham, E., Baldock, R., Bhattacharya, S.S. & Hill, R.E. (1991). The *Msh*-like homeobox genes define domains in the developing vertebrate eye. *Development 112*: 1053-1061.

Monk, M. & McLaren, A. (1981). X chromosome activity in fetal germ cells of the mouse. *J. Embryol. Exp. Morphol. 63*: 75-84.

Monk, M., Boubelik, M. & Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development 99*: 371-382.

Moroi, Y., Peebles, C., Fritzler, M.J., Steigerwald, J. & Tan, E.M. (1980). Autoantibody to centromere (kinetochore) in scleroderma sera. *Proc. Natl. Acad. Sci. USA 77*: 1627-1631.

Muller, H.J. (1930). Types of visible variations induced by X-rays in *Drosophila*. *J. Genet. 22*: 299-334.

Munford, R.E. (1963). Changes in the mammary glands of rats and mice during pregnancy, lactation and involution. *J. Endocrinol. 28*: 1-44.

Murphy, T.D. & Karpen, G.H. (1995a). Interactions between the *nod*$^{+}$ kinesin-like gene and extracentromeric sequences are required for transmission of a *Drosophila* minichromosome. *Cell 81*: 139-148.

Murphy, T.D. & Karpen, G.H. (1995b). Localisation of centromere function in a *Drosophila* minichromosome. *Cell 82*: 599-609.

Newcomer, M.E., Jones, T.A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P.A. (1984). The three-dimentional structure of retinol-binding protein. *EMBO J. 3*: 1451-1454.

Nicholls, R.D. (1993). Genomic imprinting and uniparental dysomy in Angelman and Prader-Willi syndrome. *Am. J. Med. Genet. 46*: 16-25.

Nicoll, C.S. & Tucker, H.A. (1965). Estimates of parenchymal, stromal, and lymph node deoxyribonucleic acid in mammary glands of C3H/Crg1/2 mice. *Life Sci. 4*: 993-1001.

Nicklas, R.B. (1989). The motor for poleward chromosome movement in anaphase is in or near the kinetochore. *J. Cell Biol. 109*: 2245-2255.

Nicol, L. & Jeppesen, P. (1994). Human autoimmune sera recognize a conserved 26 kD protein associated with mammalian heterochromatin that is homologous to heterochromatin protein 1 of *Drosophila*. *Chromosome Res. 2*: 245-253.

Oldham, J.D. & Friggens, N.C. (1989). Sources of variability in lactational performance. *Proc. Nut. Soc. 48*: 33-43.

Orlando, V. & Paro, R. (1995). Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr. Opin. Genet. Dev. 5*: 174-179.

Palmiter, R.D., Wilkie, T.M., Chen, H.Y. & Brinster, R.L. (1984). Transmission distortion and mosaicism in an unusual mouse pedigree. *Cell 36*: 869-877.

Palmiter, R.D. & Brinster, R.L. (1986). Germ-line transformation of mice. *Annu. Rev. Genet. 20*: 465-499.

Palmiter, R.D., Sandgren, E.P., Koeller, D.M. & Brinster, R.L. (1993). Distal regulatory elements from the mouse metallothionein locus stimulate gene expression in transgenic mice. *Mol. Cell Biol. 13*: 5266-5275.

Panshin, I.B. (1935). New evidence for the position-effect hypothesis. *C.R. (Dokl) Acad. Sci. URSS 9*: 85-88.

Papiz, M.Z., Sawyer, L., Eliopoulos, E.E. North, A.C.T., Findlay, J.B. C., Sivaprasadarao, R., Jones, T.A., Newcomer, M.E. & Kraulis, P.J. (1986). The

structure of β-lactoglobulin and its similarity to plasma retinol-binding protein. *Nature 324*: 383-385.

Paro, R. & Hogness, D.S. (1991). The polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci. USA 88*: 263-267.

Pervaiz, S. & Brew, K. (1985). Homology of β-lactoglobulin, serum retinol-binding protein, and protein HC. *Science 228*: 335-337.

Persuy, M.-A., Stinnakre, M.-G., Printz, C., Mahe, M.-F. & Mercier, J.C. (1992). High expression of caprine β-casein gene in transgenic mice. *Eur. J. Biochem. 205*: 887-893.

Pietras, D.F., Bennett, K.L., Siracusa, L.D., Woodworth-Gutai, M., Chapman, V.M., Gross, K.W., Kane-Haas, C. & Hastie, N.D. (1983). Construction of a small *Mus musculus* repetitive DNA library: identification of a new satellite sequence in *Mus musculus*. *Nucl. Acids Res. 11*: 6965-6983.

Piper, J. & Lundsteen, C. (1987). Human chromosome analysis by machine. *Trends Genet. 3*: 309-313.

Pitelka, D.R., Hamamoto, S.T., Duafala, J.G. & Nemanic, M.K. (1973). Cell contacts in the mouse mammary glands. *J. Cell Biol. 56*: 797-818.

Pittius, C.W., Hennighausen, L., Lee, E., Westphal, H., Nichols, E., Vitale, J. & Gordon, K. (1988). A milk protein gene promoter directs the expression of human tissue plasminogen activator cDNA to the mammary gland in transgenic mice. *Proc. Natl. Acad. Sci. USA 85*: 5874-5878.

Platenburg, G.J., Kootwijk, E.P.A., Kooiman, P.M., Woloshuk, S.L., Nuijens, J.H., Krimpenfort, P.J.A., Pieper, F.R., de Boer, H.A. & Strijker, R. (1994). Expression of human lactoferrin in milk of transgenic mice. *Transgenic Res. 3*: 99-108.

Porter, S.D., Larue, L. & Mintz, B. (1991). Mosaicism of tyrosinase-locus transcription and chromatin structure in dark vs. light melanocyte clones of homozygous *chinchilla-mottled* mice. *Dev. Genetics 12*: 393-402.

Porter, S.D. & Meyer, C.J. (1994). A distal tyrosinase upstream element stimulates gene expression in neural-crest-derived melanocytes of transgenic mice: position-independent and mosaic expression. *Development 120*: 2103-2111.

Pravtcheva, D.D., Wise, T.L., Ensor, N.J. & Ruddle, F.H. (1994). Mosaic expression of an Hprt transgene integrated in a region of Y heterochromatin. *J. Exp. Zool. 268*: 452-468.

Prokofyeva-Belgovskaya, A.A. (1939). Cytological mechanism of mosaicism and of chromosome rearrangement. *C.R. (Dokl) Acad. Sci. U.R.S.S.N.S. 22*: 270-273.

Radic, M.Z., Lundgren, K. and Hamkalo, B.A. (1987). Curvature of mouse satellite DNA and condensation of heterochromatin. *Cell 50*: 1101-1108.

Rastan, S., Kaufman, M.H., Handyside, A. & Lyon, M.F. (1980). X chromosome inactivation in the extraembryonic membranes of diploid pathogenetic mouse embryos demonstrated by differential staining. *Nature 288*: 172-173.

Razin, A. & Shemer, R. (1995). DNA methylation in early development. *Hum. Mol. Genet. 4*: 1751-1755.

Reik, W., Collick, A., Norris, M.L., Barton, S.C. & Surani, M.A. (1987). Genomic imprinting determines methylation of paternal alleles in transgenic mice. *Nature 328*: 248-251.

Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A. & Spierer, P. (1990). Dependence of position-effect variegation in Drosophila on dose of a gene encoding an unusual zinc-finger protein. *Nature 344*: 219-223.

Reuter, G. & Spierer, P. (1992). Position effect variegation and chromatin proteins. *Bioessays 14*: 605-612.

Rivier, D.H. & Rine, J. (1992). Silencing: the establishment and inheritance of stable, repressed transcription states. *Curr. Opin. Genet. Dev. 2*: 286-292.

Roberts, B., DiTullio, P., Vitale, J., Hehir, K. & Gordon, K. (1992). Cloning of goat β-casein encoding gene and expression in transgenic mice. *Gene 121*: 255-262.

Roberts, C., Weith, A., Passage, E., Michot, J.L., Mattei, M.G. and Bishop, C.E. (1988). Molecular and cytogenetic evidence for the location of *Tdy* and *Hya* on the mouse Y chromosome short arm. *Proc. Natl Acad. Sci. USA 85*: 6446-6449.

Robertson, G., Garrick, D., Wu, W., Kearns, M., Martin, D. & Whitelaw, E. (1995). Position-dependent variegation of globin transgene expression in mice. *Proc. Natl. Acad. Sci. 92*: 5371-5375.

Robertson, G., Garrick, D., Wilson, M., Martin, D.I.K. & Whitelaw, E. (1996). Age-dependent silencing of globin transgenes in the mouse. *Nucl. Acids Res. 24*: 1465-1471.

Robinson, G.W., McKnight, R.A., Smith, G.H. & Hennighausen, L. (1995). Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. *Development 121*: 2079-2090.

Rosen, J.M. (1987). Milk protein gene structure and expression. In: *The mammary gland: development, regulation and function* (eds., Neville, M.C. & Daniel, C.W.), Plenum Press: New York & London, pp. 301-322.

Roskelley, C.D., Srebrow, A. & Bissell, M.J. (1995). A hierarchy of ECM-mediated signalling regulates tissue-specific gene expression. *Curr. Opin. Cell Biol. 7*: 736-747.

Rowland, S.J. (1938). The determination of the nitrogen distribution in milk. *J. Dairy Res. 9*: 42-46.

Rubin, D.C., Ong, D.E. & Gordon, J.I. (1989). Cellular differentiation in the emerging fetal rat small intestinal epithelium: mosaic patterns of gene expression. *Proc. Natl. Acad. Sci. USA 86*: 1278-1282.

Rui, H., Kirken, R.A. & Farrar, W.L. (1994). Activation of receptor-associated tyrosine kinase JAK2 by prolactin. *J. Biol. Chem. 269*: 5364-5368.

Rushlow, C.A., Bender, W. & Chovnick, A. (1984). Studies on the mechanism of heterochromatic position effect at the *rosy* locus of *Drosophila melanogaster*. *Genetics 108*: 603-615.

Russell, L.B. (1963). Mammalian X chromosome activation/inactivation limited in spread and in region of origin. *Science 140*: 976-978.

Saarinen, U.M. & Siimes, M.A. (1979). Iron absorption from breast milk, cow's milk, and iron-supplemented formula: an opportunistic use of changes in total body iron determined by hemoglobin, ferritin, and body weight in 132 infants. *Pediat. Res. 13*: 143-147.

Sabl, J.F. & Henikoff, S. (1996). Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in Drosophila. *Genetics 142*: 447-458.

Sakakura, T. (1987). Mammary embryogenesis. In: *The mammary gland: development, regulation and function*. (eds., Neville, M.C., Daniel, C.W.), Plenum Press, New York & London, pp 37-66.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). Molecular cloning: A Laboratory Manual, Second Edition (Cold Springs Harbor, New York: Cold Springs Harbor Laboratory Press).

Sandell, L.L. & Zakian, V.A. (1992). Telomeric position effect in yeast. *Trends Cell Biol. 2*: 10-14.

Sandgren, E.P., Palmiter, R.D., Heckel, J.L., Daugherty, C.C., Brinster, R.L. & Degen, J.L. (1991). Complete hepatic regeneration after somatic deletion of an albumin-plasminogen activator transgene. *Cell 66*: 245-256.

Sapienza, C., Peterson, A.C., Rossant, J. & Balling, R. (1987). Degree of methylation of transgenes is dependent on gamete of origin. *Nature 328*: 251-254.

Saunders, W.S., Chue, C., Goebl, M., Craig, C., Clark, R.F., Powers, J.A., Eissenberg, J.C., Elgin, S.C., Rothfield, N.F. & Earnshaw, W.C. (1993). Molecular cloning of a human homologue of *Drosophila* heterochromatin protein HP1 using anti-centromere autoantibodies with anti-chromo specificity. *J. Cell Sci. 104*: 573-582.

Sawyer, L., Papiz, M.Z., North, A.C.T. and Elipoulos, E.E. (1985). ??????? *Biochem. Soc. Trans. 13*: 265-266.

Schmdt, D.G. (1982). Association of caseins and casein micelle structure. In: *Developments in dairy chemistry* (ed., Fox, P.F.), Vol. 1, Applied Science Publication, New York, pp. 61-86.

Schultz, J. & Caspersson, T. (1939). Heterochromatic regions and the nucleic acid metabolism of the chromosomes. *Archiv. Exp. Zellforsch. 22*: 650-654.

Scrable, H., Cavenee, W., Ghavimi, F., Lovell, M., Morgan, K. & Sapienza, C. (1989). A model for embryonal rhabdomyosarcoma tumorigenesis that involves genomic imprinting. *Proc. Natl. Acad. Sci. USA 86*: 7480-7484.

Shamay, A., Pursel, V.G., Wall, R.J. & Hennighausen, L. (1992). Induction of lactogenesis in transgenic virgin pigs: evidence for gene and integration site-specific hormonal regulation. *Mol. Endocrinol. 6*: 191-197.

Shani, M., Barash, I., Nathan, M., Ricca, G., Searfoss, G.H., Dekel, I., Faerman, A., Givol, D. & Hurwitz, D.R. (1992). Expression of human serum albumin in the milk of transgenic mice. *Transgenic Res. 1*: 195-208.

Silberstein, G.B., Flanders, F.C., Roberts, A.B. & Daniel, C.W. (1992). Regulation of mammary morphogenesis: evidence for extracellular matrix-mediated inhibition of ductal budding by transforming growth factor beta-1. *Dev. Biol. 152*: 354-362.

Simons, J.P. & Land, R.B. (1987). Transgenic livestock. *J. Reprod. Fert. Suppl. 34*: 237-250.

Simons, J.P., McClenaghan, M. & Clark, A.J. (1987). Alteration of the quality of milk by expression of sheep β-lactoglobulin in transgenic mice. *Nature 328*: 530-532.

Sinclair, D.A.R., Mottus, R.C. & Grigliatti, T.A. (1983). Genes which suppress position-effect variegation in *Drosophila melanogaster* are clustered. *Mol. Gen. Genet. 191*: 326.

Singer, M.F. (1982). Highly repeated sequences in mammalian genomes. *Int. Rev. Cytol. 76*: 67-112.

Singh, P.B., Miller, J.R., Pearce, J.J., Burton, R.D., Paro, R., James, T.C. & Gaunt, S.J. (1991). A sequence motif found in a *Drosophila* heterochromatin protein is conserved in animals and plants. *Nucl. Acids Res. 19*: 789-793.

Singh, P.B. (1994). Molecular mechanisms of cellular determination: their relation to chromatin structure and parental imprinting. *J. Cell Sci. 107*: 2653-2668.

Slatis, H.M. (1955). A reconsideration of the brown-dominant position effect. *Genetics 40*: 246-251.

Sleep, D., Belfield, G.P., Ballance, D.J., Steven, J., Jones, S., Evans, L.R., Moir, P.D. & Goody, A.R. (1991). *Saccharomyces cerevisiae* strains that overexpress heterologous proteins. *Bio/Technology 9*: 183-187.

Smith, H.O. & Nathans, D. (1973). A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J. Mol. Biol. 81*: 419.

Soulier, S., Vilotte, J.L., Stinnakre, M.G. & Mercier, J.C. (1992). Expression analysis ruminant α-lactalbumin in transgenic mice: developmental regulation and general location of important *cis*-regulatory elements. *FEBS 297*: 13-18.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol. 98*: 503-517.

Spofford, J.B. (1976). Position-effect variegation in *Drosophila.* In *Genetics and biology of Drosophila.* (eds., Ashburner, M. & Novitski, E.), Vol. 1c, London Academic Press, pp. 955-1019.

Spradling, A.C. & Rubin, G.M. (1983). The effect of chromosome position on the expression of the *Drosophila* xanthine xehydrogenase gene. *Cell 34*: 48-58.

Spradling, A.C. & Karpen, G.H. (1990). Sixty years of mystery. *Genetics 126*: 779-784.

Spradling, A.C. (1993). Position effect variegation and genomic instability. *Cold Spring Harb. Symp. Quant. Biol.* 585-596.

Stamatoyannopoulos, G., Josephson, B., Zhang, J-W. & Li, Q. (1993). Developmental regulation of human γ-globin genes in transgenic mice. *Mol. Cell Biol. 13*: 7636-7644.

Strange, R., Li, F., Saurer, S., Burkhardt, A. & Friis, R.R. (1992). Apoptotic cell death and tissue remodelling during mouse mammary involution. *Development 115*: 49-58.

Streuli, C.H., Bailey, N. & Bissell, M.J. (1991). Control of mammary epithelial differentiation: basement membrane induces tissue-specific gene expression in the absence of cell-cell interaction and morphological polarity. *J. Cell Biol. 115*: 1383-1395.

Stuart, J., Norrell, S. & Harrington, J.P. (1984). Kinetic effect of human lactoferrin on the growth of *Escherichia coli* 0111. *Int. J. Biochem. 16*: 1043-1048.

Surani, M.A., Reik, W. & Allen, N.D. (1988). Transgenes as molecular probes for genomic imprinting. *Trends Genet. 4*: 59-62.

Surani, M.A. (1994). Genomic imprinting: control of gene expression by epigenetic inheritance. *Curr. Opin. Cell Biol. 6*: 390-395.

Swain, J.L., Stewart, T.A. & Leder, P. (1987). Parental legacy determines methylation and expression of an autosomal transgene: a molecular mechanism for parental imprinting. *Cell 50*: 719-727.

Sweetser, D.A., Hauft, S.M., Hoppe, P.C., Birkenmeier, E.H. & Gordon, J.I. (1988a). Transgenic mice containing intestinal fatty acid-binding protein-human growth hormone fusion genes exhibit correct regional and cell-specific expression of the reporter gene in their small intestine. *Proc. Natl. Acad. Sci. USA*: 9611-9615.

Sweetser, D.A., Birkenmeier, E.H., Hoppe, P.C., McKeel, D.W. & Gordon, J.I. (1988b). Mechanisms underlying generation of gradients in gene expression within the intestine: an analysis using transgenic mice containing fatty acid binding protein-human growth hormone fusion genes. *Genes Dev. 2*: 1318-1332.

Sympson, C.J., Talhouk, R.S., Alexander, C.M., Chin, S.K., Clift, S.M., Bissell, M.J. & Werb, Z. (1994). Targeted expression of stromelysin-1 in mammary gland provides evidence for a role of proteinases in branching morphogenesis and the requirement for an intact basement membrane for tissue-specific gene expression. *J. Cell Biol. 125*: 681-693.

Szabo, P. & Mann, J.R. (1995). Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes Dev. 9*: 1857-1868.

Talarico, D., Peverali, A.F., Ginelli, E., Meneveri, R., Mondello, C. & Valle, G.D. (1988). Satellite DNA induces unstable expression of the adjacent herpes simplex virus *tk* gene cotransfected in mouse cells. *Mol. Cell. Biol. 8*: 1336-1344.

Talbert, P.B., LeCiel, C.D. & Henikoff, S. (1994). Modification of the *Drosophila* heterochromatin mutation *brown*^{DOMINANT} by linkage alterations. *Genetics 136*: 559-571.

Talbot, D., Collis, P., Antoriou, M., Vidal, M., Grosveld, F. & Greaves, D.R. (1989). A dominant control region from the human β-globin locus conferring integration site-independent gene expression. *Nature 338*: 352-355.

Talhouk, R.S., Bissell, M.J. & Werb, Z. (1992). Coordinated expression of extracellular marix-degrading proteinases and their inhibitors regulates mammary epithelial function during involution. *J. Cell Biol. 118*: 1271-1282.

Tan, P.P., Williams, E.A. & Tam, S.S. (1994). X-chromosome inactivation occurs at different times and in different tissues of the post-implantation mouse embryo. *Dev. Genet. 15*: 491-503.

Tartof, K.D., Hobbs, C. & Jones, M. (1984). A structural basis for variegating position effects. *Cell 37*: 869-878.

Topper, Y.J. & Freeman, C.S. (1980). Multiple hormone interactions in the developmental biology of the mammary gland. *Physiol. Rev. 60*: 1049-1056.

Tower, J., Karpen, G., Craig, N. & Spradling, A. (1993). Preferential transposition of *Drosophila P* elements to nearby chromosome sites. *Genetics 133*: 347-359.

Traurig, H.U. (1967). A autoradiographic study of cell proliferation in the mammary gland of the pregnant mouse. *Anat. Rec. 159*: 239-248.

Turner, M.D., Rennison, M.E., Handel, S.E., Wilde, C.J. & Burgoyne, R.D. (1992). Proteins are secreted by both constitutive and regulated secretory pathways in lactating mouse mammary epithelial cells. *J. Cell Biol. 117*: 269-278.

Velander, W.H., Page, R.L., Morcol, T., Russell, C.G., Canseco, R., Young, J.M., Drohan, W.N., Gwazdauskas, F.C., Wilkins, T.D. & Johnson, J.L. (1992a). Production of biologically active protein C in the milk of transgenic mice. *Ann. N.Y. Acad. Sci. 665*: 391-403.

Velander, W.H., Johnson, J.L., Page, R.L., Russell, C.G., Subramanian, A., Wilkins, T.D., Gwazdauskas, F.C., Pittius, C. & Drohan, W.N. (1992b). High level of expression of a heterologous protein in the milk of transgenic swine using the cDNA encoding human protein C. *Proc. Natl. Acad. Sci. USA 89*: 12003-12007.

Verma, R.S. (1988). *Heterochromatin: molecular and structural aspects.* Cambridge University Press, Cambridge.

Villar, A.J. & Pedersen, R.A. (1994). Paternal imprinting of the *Mas* protooncogene in mouse. *Nature Genet. 8*: 373-379.

Vilotte, J.L., Soulier, S., Stinnakre, M.G., Massoud, M. & Mercier, J.C. (1989). Efficient tissue-specific expression of bovine α-lactalbumin in transgenic mice. *Eur. J. Biochem. 186*: 43-48.

Vonderhaar, B.K. & Ziska, S.E. (1989). Hormonal regulation of milk protein expression. *Annu. Rev. Physiol. 51*: 641-652.

Wagner, E.F., Stewart, T.A. & Mintz, B. (1981). The human β-globin gene and a functional viral thymidine kinase gene in developing mice. *Proc. Natl. Acad. Sci. USA 78*: 5016-5020.

Wakao, H., Gouilleux, F. & Groner, B. (1994). Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *EMBO J. 13*: 2182-2191.

Wall, R.J., Pursel, V.G., Shamay, A., McKnight, R.A., Pittius, C.W. & Hennighausen, L. (1991). High level synthesis of a heterologous milk protein in the mammary glands of transgenic swine. *Proc. Natl. Acad. Sci. USA 88*: 1696-1700.

Wallace, H., McLaren, K., Al-Shawi, R. & Bishop, J.O. (1994). Consequences of thyroid hormone deficiency induced by the specific ablation of thyroid follicle cells in adult transgenic mice. *J. Endocrinol. 143*: 107-120.

Wallrath, L.L. & Elgin, S.C.R. (1995). Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev. 9*: 1263-1277.

Walters, M.C., Magis, W., Fiering, S., Eidemiller, J., Scalzo, D., Groudine, M. & Martin, D.I.K. (1996). Transcriptional enhancers act in *cis* to suppress position-effect variegation. *Genes Dev. 10*: 185-195.

Weiler, K.S. & Wakimoto, B.T. (1995). Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genet. 29*: 577-605.

Wevrick, R., Kerns, J.A. & Francke, U. (1994). Identification of a novel paternally expressed gene in the Prader-Willi syndrome region. *Hum. Mol. Genet. 3*: 1877-1882.

Whitelaw, C.B.A., Archibald, A.L., Harris, S., McClenaghan, M., Simons, J.P. & Clark, A.J. (1991). Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice. *Transgenic Res. 1*: 3-13.

Whitelaw, C.B.A., Harris, S., McClenaghan, M., Simons, J.P. & Clark, A.J. (1992). Position-independent expression of the ovine β-lactoglobulin gene in transgenic mice. *Biochem. J. 286*: 31-39.

Wiens, D., Park, C.S. & Stockdale, F.E. (1987). Milk protein expression and ductal morphogenesis in the mammary gland *in vitro*: hormone-dependent and -independent phases of adipocyte-mammary epithelial cell interactions. *Dev. Biol. 120*: 245-258.

Wilkie, T.M., Braun, R.E., Ehrman, E.J., Palmiter, R.D. & Hammer, R.E. (1991). Germ-line intrachromosomal recombination restores fertility in transgenic MyK-103 male mice. *Genes Dev. 5*: 38-48.

Wilkinson, D.G. (1992). The theory and practice of *in situ* hybridisation. In*: In situ hybridisation; a practical approach.* (ed., Wilkinson, D.G.), Oxford University Press, Oxford, pp. 1-13.

Williams, J.M. & Daniel, C.W. (1983). Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol. 97*: 274-290.

Wilson, C., Bellen, H.J. & Gehring, W.J. (1990). Position effects on eukaryotic gene expression. *Annu. Rev. Cell Biol. 6*: 679-714.

Wong, A.K.C. & Rattner, J.B. (1988). Sequence organisation and cytological localisation of the minor satellite of mouse. *Nucleic Acids Res. 16*: 11645-11661.

Wong, A.K.C., Biddle, F.G. & Rattner, J.B. (1990). The chromsomal distribution of the major and minor satellite is not conserved in the genus *Mus*. *Chromosoma 99*: 190-195.

Wreggett, K.A., Hill, F., James, P.S., Hutchings, A., Butcher, G.W. & Singh, P.B. (1994). A mammalian homologue of *Drosophila* heterochromatin protein 1 (HP1) is a component of constitutive heterochromatin. *Cytogenet. Cell Genet. 66*: 99-103.

Wright, G., Carver, A., Cottom, D., Reeves, D., Scott, A., Simons, J.P., Wilmut, I. & Colman, A. (1991). High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Bio/Technology 9*: 830-834.

Yaphe, W. (1957). The use of agarase from *Pseudomonas atlantica* in the identification of agar in marine algae (*Rhodophyceae*). *Can. J. Microbiol. 3*: 987-993.

Yeivin, A. & Razin, A. (1993). Gene methylation patterns and expression. In: *DNA methylation: molecular biology and biological significance*. (ed., Jost, J.P. & Saluz, H.P.), Birkauser Verlag, Basel, pp. 523-568.

Zang, P. & Spradling, A. (1993). Efficient and dispersed local *P* element transposition from *Drosophila* females. *Genetics 133*: 361-373.

Zatz, M., Marie, S.K., Passos-Bueno, M.R., Vainzof, M., Campiotto, S., Cerqueira, A., Wijmenga, C., Padberg, G. & Frants, R. (1995). High proportion of new

mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *Am. J. Hum. Genet. 56*: 99-105.

Zippelius, H.M. & Schleidt, W.M. (1956). Ultraschall-Laute bei jungen Mausen. *Naturwissenchaften 43* : 509-520.

Zuckerkandl, E. (1974). A possible role of "inert" heterochromatin in cell differentiation. Action and competition for "locking" molecules. *Biochimie. 56*: 937-954.

APPENDIX I

Statistical Terms

Variance: average of the squared distance of observations (sample values e.g. BLG expression levels) from the mean

Mean: sum of all the values divided by the number of values

Standard Deviation (*s.d.*): square root of the variance (same measurement as the variance but on the same scale as the mean)

Coefficient of variation: (*s.d.* / mean) x 100 (the *s.d.* as a proportion of the mean)

Correlation coefficient: a measure of the degree to which the two variables (e.g. BLG mRNA and protein expression levels) are associated linearly with each other

APPENDIX II

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 6659–6664, June 1996
Genetics

Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus

(β-lactoglobulin/epigenetic silencing/position-effect variegation)

KENNETH W. DOBIE*†, MURIEL LEE‡, JUDITH A. FANTES‡, ELIZABETH GRAHAM‡, A. JOHN CLARK†, ANTHEA SPRINGBETT†, RICHARD LATHE*, AND MARGARET MCCLENAGHAN†§

*Centre for Genome Research, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JQ, United Kingdom; †Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, United Kingdom; and ‡Medical Research Council Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, United Kingdom

Communicated by Allan C. Spradling, Carnegie Institution of Washington, Baltimore, MD, February 9, 1996 (received for review November 2, 1995)

ABSTRACT **Mice carrying an ovine β-lactoglobulin (BLG) transgene secrete BLG protein into their milk. To explore transgene expression stability, we studied expression levels in three BLG transgenic mouse lines. Unexpectedly, two lines exhibited variable levels of transgene expression. Copy number within lines appeared to be stable and there was no evidence of transgene rearrangement. In the most variable line, BLG production levels were stable within individual mice in two successive lactations. Backcrossing demonstrated that genetic background did not contribute significantly to variable expression. Tissue *in situ* hybridization revealed mosaicism of transgene expression within individual mammary glands from the two variable lines; in low expressors, discrete patches of cells expressing the transgene were observed. Transgene protein concentrations in milk reflected the proportion of epithelial cells expressing BLG mRNA. Furthermore, chromosomal *in situ* hybridization revealed that transgene arrays in both lines are situated close to the centromere. We propose that mosaicism of transgene expression is a consequence of the chromosomal location and/or the nature of the primary transgene integration event.**

β-Lactoglobulin (BLG) is a major ovine milk protein. The function of BLG is unknown, though the crystal structure of bovine BLG is consistent with a role in vitamin A transport (1). We previously reported that mice carrying a sheep BLG transgene secrete BLG into their milk (2); BLG regulatory regions can direct expression of biomedical proteins into the milk of transgenic mice and sheep (3–5). In this context, it is important that transgene expression is stable. Unstable transgene expression has been described previously; Palmiter *et al.* (6) reported that the level of herpes simplex virus thymidine kinase expression could vary by more than an order of magnitude among progeny of the same founder. Although other transgene insertions express to variable degrees within individual cell lines or transgenic mouse lines (refs. 7–19; M. Mehtali and R.L., unpublished data), there has been no common explanation for the instability of expression. Unstable expression may be due to strong selection against the transgene, for instance by the failure of sperm fertility engendered by testicular thymidine kinase expression (7, 8) or by the toxicity of high-level hepatic expression of plasminogen activator (9). A transgene inserted into the X chromosome (10) or an X-autosome translocation (20) generates mosaic expression due to stochastic X chromosome inactivation. Silencing has also been observed when the transgene integrates into repeat sequence or satellite DNA (11, 12), whereas different levels of transgene expression between animals of the same lineage have been attributed to strain-specific modifier genes (13–15).

Mosaic patterns of expression were also observed in transgenic animals bearing intestinal fatty-acid binding protein fusion transgenes (16). Here, mosaicism was attributed to a deficit of *cis*-acting elements in the transgene. Mosaic expression patterns were also observed in mice carrying different tyrosinase fusion transgenes (17–19); these reports suggested that the striated coat color was an inherent property of the transgene.

To address the stability of transgene expression, we examined three transgenic mouse lines harboring an intact BLG gene. In line 14, BLG expression levels were stable, whereas in lines 7 and 45, BLG levels varied significantly. We examine the nature of this variable expression and show that it reflects variegated patterns of transgene expression within the mammary gland. We demonstrate that variable expression is independent of genetic background and is a property of the transgene locus.

MATERIALS AND METHODS

Mice, Sampling, and Milk Protein Analysis. Transgenic lines 7, 14, and 45 (2) were maintained by systematic crossing to C57BL/6 × CBA F_1 hybrid mice. Litters were standardized to five pups per mother at birth. Milk and tissue samples were collected from 6- to 8-week old transgenic mice (hemizygous for the transgene array) at day 11 of lactation. Backcross experiments used CBA and C57BL/6 mice (Harlan Olac, Bichester, U.K.). Transgenic mice were identified by a PCR assay (21). Milk collection and SDS/20% PAGE analysis of milk proteins were performed as described (22).

DNA Analysis. Genomic mammary DNA was obtained from frozen tissue by standard procedures. *Eco*RI-digested DNA samples (10 μg) were electrophoresed and transferred to Zetaprobe (Bio-Rad) membranes. BLG and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probes were gel purified. A mouse total genomic DNA probe was prepared from nontransgenic mammary tissue. BLG, GAPDH, and mouse genomic DNA probes were labeled with [α-^{32}P]dCTP using a commercial labeling system (Multiprime, Amersham). Following hybridization, results were quantified using a Molecular Dynamics PhosphorImager.

RNA Analysis. Frozen mammary tissue was homogenized in 2 ml RNAzol B (Biogenesis Ltd.) and total mRNA extracted. Samples (10 μg) were electrophoresed on 1.5% denaturing Mops/formaldehyde agarose gels, transfered, and hybridized with BLG and GAPDH probes.

Tissue *in Situ* Hybridization. *Pst*I fragments from the BLG (424 bp; ref. 23) and β-casein (440 bp; ref. 24) cDNA,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: BLG, β-lactoglobulin; FISH, fluorescence *in situ* hybridization; FITC, fluorescein isothiocyanate; c.v., coefficient of variation.

§To whom reprint requests should be addressed.

respectively, were subcloned into pGEM-4Z in both orientations. dUTP[^{35}S]-labeled single-stranded antisense probes and control sense probes were synthesized by *in vitro* transcription using standard procedures (25). Mammary tissue was collected from hemizygous mice at day 11 of lactation and processed for *in situ* hybridization (26).

Chromosome Fluorescence *in Situ* Hybridization (FISH). Lymphocyte cells from mouse spleens were cultured in RPMI 1640 medium with 10% fetal calf serum and 50 μg/ml lipopolysaccharide, at a density of 1×10^7 cells/ml for 45 hr at 37°C. These cells were used for Giemsa banding and karyotyping of chromosomes (27). Genomic ovine BLG clone pSS1tgXS (2) was nick-translated with biotin-16-dUTP (28) and used to assign transgenes to specific metaphase chromosomes from hemizygous line 7, 14, and 45 mice using avidin-fluorescein isothiocyanate (FITC) (29). Digital images were obtained with a Bio-Rad confocal microscope. Double-color FISH used BLG and mouse major satellite DNA (30) probes (see Fig. 4 legend for details) and results were visualized as described (31).

RESULTS

Lines 7 and 45 Exhibit Variable Expression of BLG. Lines 7 and 14 harbor an identical 16 kb BLG transgene, whereas line 45 carries an 11 kb 3′ truncated transgene; the 3′ truncation does not affect the transcription unit nor transgene function (2). Hemizygous transgenic female mice were mated, milk collected at day 11 of lactation, and BLG protein levels determined.

The ranges of BLG expression levels were different among the three lines (Fig. 1). Line 14 expressed BLG in a stable fashion with only 3.5 mg/ml separating the lowest and the highest expressor. However, BLG levels in milk of line 45 animals ranged from 16 to 30 mg/ml, whereas in line 7 there was a 8-fold difference between the lowest and highest expressors (3 to 23.9 mg/ml). Table 1 shows the standard deviations (SDs) and coefficients of variation (c.v.) for BLG protein expression in the three lines. As expected, lines 7 and 45 had larger SD values than line 14. The c.v. for line 45 was similar to that of line 14, but this reflects its higher average expression level and also a possible upper limit to protein production. We conclude that the particular transgene insertion locus in line 7 determines variation in expression levels of BLG protein, for the identical transgene resulted in stable expression in line 14 animals.

To establish whether variation occurred at the steady-state mRNA level, quantitative Northern blot analyses using BLG and control (GAPDH) hybridization probes were performed. BLG mRNA levels were more variable in lactating gland from animals of lines 7 and 45 than in line 14 animals (Table 1). The correlation coefficients for protein and RNA levels were 0.44, 0.54, and 0.54 for lines 14 ($n = 9$), 45 ($n = 9$) and 7 ($n = 8$, after the exclusion of one outlier), respectively. The results show variable levels of RNA expression in lines 7 and 45; the relationship between protein and RNA was similar in all three lines.

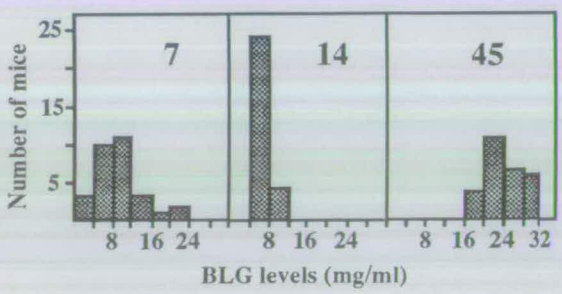


FIG. 1. BLG expression level profiles in milk collected from line 7, 14, and 45 hemizygous mice.

Variable Expression Is Not Due to Rearrangement of the Transgene Locus. We were concerned that variable expression might be due to modifications to the structure of the transgene locus. Therefore, we performed Southern blot analysis on DNA prepared from lactating mammary tissue. Lines 7 and 45 carried 25 and 17 transgene copies, respectively, whereas line 14 harbored 2 copies (32). In all samples, the BLG probe identified a single major 4.3-kb band with no evidence of variations in banding patterns to suggest rearrangement of the transgene arrays. Quantitative hybridization using probes corresponding to ovine BLG and to total genomic DNA revealed that the transgene copy number was indistinguishable within lines, even among mice from line 7 exhibiting the full spread of BLG levels. All lines had similar c.v. values (Table 1) and this contrasts with the protein and RNA data in which the SD and c.v. values differed considerably for the three lines. Furthermore, the c.v. values for protein and RNA in line 7 exceeded that for DNA. While these analyses would not reveal subtle changes in the structure of the transgene array, we surmise that variable expression is not due to stochastic loss or rearrangement of the transgene array in different individual line 7 and 45 mice.

BLG Expression Does Not Vary Between Lactations. To determine whether expression levels exhibit temporal variation within individual line 7 animals, milk BLG protein levels were measured in 28 individuals at day 11 of two successive lactations (Fig. 2*A*). The close correlation between BLG levels measured in two lactations argues that individual levels are fixed by day 11 of the first lactation.

Variable Expression Is Not Due to Genetic Background. The transgenic mice studied were of a C57BL/6 × CBA hybrid background. Variable transgene expression among line 7 mice might be due to segregation of alleles of a modifier gene(s). We therefore backcrossed transgenic animals from line 7 to C57BL/6 and CBA animals for three generations; milk samples were collected at day 11 of lactation and the levels of BLG determined (Fig. 2*B*).

The variance and mean BLG levels were not significantly different between the third backcross C57 population and the mixed CBA/C57 population ($P > 0.05$). Although variance and mean values were significantly different between the third backcross CBA and CBA/C57 populations ($P < 0.05$), the difference in the range of expression levels was small. We have since generated data from fourth generation backcross animals (96.9% C57, $n = 14$; 96.9% CBA, $n = 13$) and confirm that variance values for both backgrounds are again similar to the original CBA/C57 population ($P > 0.05$). There was also no clear trend in BLG expression levels through sublines within each backcross; for example, a parent expressing 11 mg/ml gave rise to daughters with expression levels of 5 mg/ml and 22 mg/ml.

It is predicted that the presence of single or multiple modifier loci would be revealed by major alterations in both the variance and mean values for expression levels over three generations. Because no such alterations were observed by the fourth backcross, we conclude the variable expression of the transgene within line 7 is not due to heterogenity of genetic background.

Mosaic Expression of BLG in the Mammary Gland. We wished to determine whether levels of BLG expression were reflected in differential expression of the transgene in the mammary epithelium. *In situ* hybridization studies were performed on abutting tissue sections collected at day 11 of lactation from individuals of lines 7, 14, and 45. Endogenous β-casein mRNA was detected throughout the mammary epithelium in all three transgenic lines, demonstrating that all secretory epithelial cells of the gland have the capacity to express a milk protein gene (Fig. 3 *B*, *D*, and *F*). Sections from

Table 1. BLG expression levels and copy numbers within three transgenic lines

Line	Protein $\bar{x} \pm$ SD	Protein c.v.	RNA c.v.	DNA c.v.
7	9.5 ± 4.9	52% (n = 30)	47% (n = 9)	36% (n = 18)
14	6.6 ± 1.1	17% (n = 28)	13% (n = 9)	41% (n = 13)
45	23.7 ± 3.9	16% (n = 28)	26% (n = 9)	42% (n = 13)

$\bar{x}$, mean; c.v. = 100 × SD/$\bar{x}$; n, number of mice. BLG protein $\bar{x}$ values are absolute values (mg/ml); the c.v. measurements for BLG protein were the same if calculated from absolute values or those normalized for the endogenous protein albumin (data not shown). RNA and DNA values are measured in arbitrary units following normalization (within a line) to endogenous GAPDH mRNA or total DNA determined by separate hybridizations to the GAPDH and total genomic probes; for this reason, mean values for RNA and DNA are not comparable between different lines and are not shown.

low-expressing line 7 individuals showed small clusters of cells strongly expressing BLG mRNA, surrounded by nonexpressing cells (Fig. 3*A*). High expressors from line 7 had the opposite pattern of expression with small regions of cells negative for BLG mRNA (Fig. 3*C*). The proportion of positive cells in any one section correlated with the level of BLG in the milk; higher expressors had a greater proportion of cells positive for transgene mRNA. Line 14 was distinctly different because BLG expression was present in all the epithelial cells examined, matching the pattern of β-casein mRNA expression (Fig. 3 *E* and *F*). In lines 7 and 14, the introduced transgene is identical, arguing that mosaic expression is due to the location and/or nature of the transgene array. Line 45 also exhibited mosaic expression of BLG, although less so than in line 7 (Fig. 3*G*). This may reflect the higher average expression level in line 45; individuals from line 7 and 45 with the same levels of BLG protein exhibited a similar degree of mosaicism. Although this analysis does not permit precise quantitation of per-cell expression levels, the results suggest that variations in the level of BLG protein and mRNA expression in line 7 and 45 individuals appear to be due to variations in the proportion of mammary epithelial cells expressing the transgene.

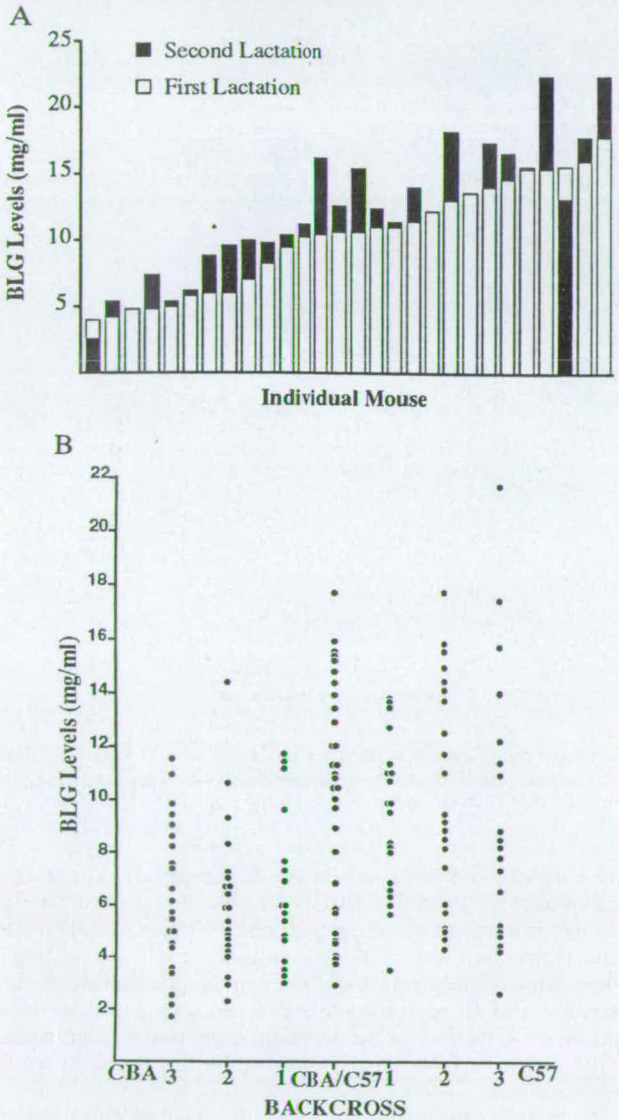


FIG. 2. BLG expression levels in milk of line 7 animals: consecutive lactations and effect of genetic background. (*A*) BLG levels in milk collected at day 11 of lactation in two successive lactations. The average difference between the first and second lactations was 2 (± 0.4) mg/ml. (*B*) BLG levels in line 7 (C57BL/6 × CBA) animals backcrossed to the two parental strains. The CBA/C57 group is mice of C57BL/6 × CBA background (50% genetic contribution from each parental strain); CBA 1–3 and C57 1–3 are mice of three successive crosses to each parental strain (parental genetic background contributions are 75%, 87.5%, and 93.75%, respectively).

The Transgene Locus Lies Near a Centromere in Lines 7 and 45. Variegated expression of endogenous genes (33) and transgenes (34) has been described in *Drosophila* and linked to proximity to the centromere. To determine the subchromosomal localization of the BLG transgenes, we applied *in situ* hybridization to chromosome spreads prepared from animals of different transgenic lines. G-banding and karyotyping in conjunction with FISH using a BLG probe demonstrated that the transgene arrays had integrated into chromosomes 15 (line 7), 7 (line 14), and 5 (line 45). Further, the BLG locus had integrated midarm in line 14 (Fig. 4*a*) but was close to the centromere in lines 7 and 45. To refine the localization of the transgene array relative to the centromere, we used a BLG probe in conjunction with a major satellite (196) probe that predominantly decorates centromeric regions. In all line 7 spreads, the BLG locus was close to but out with the centromere (Fig. 4*b*). A crude estimate from 56 chromatids indicated that the array lies within the 3.5% of the chromosome proximal to centromeric heterochromatin. In line 45 the resolution of the experiment made the array appear to be within major satellite DNA (Fig. 4*c*, chromosome arrow). However, the transgene signal was observed as clearly separate from major satellite DNA in interphase nuclei where the chromosomes are decondensed (Fig. 4*c*, nucleus arrow). By quenching BLG signals, it was clear that integrated arrays were not associated with major translocations of this satellite DNA.

DISCUSSION

The introduction of transgenes to the mouse germline by pronuclear injection of DNA usually results in tandemly repeated and head-to-tail arrays at a single random site within the genome. Though the transgene insertion site can clearly have an influence upon tissue-specificity and level of expression, it has often been assumed that expression of the transgene, once integrated, is stable within a transgenic line. However, processes have been described, such as the "RIP"

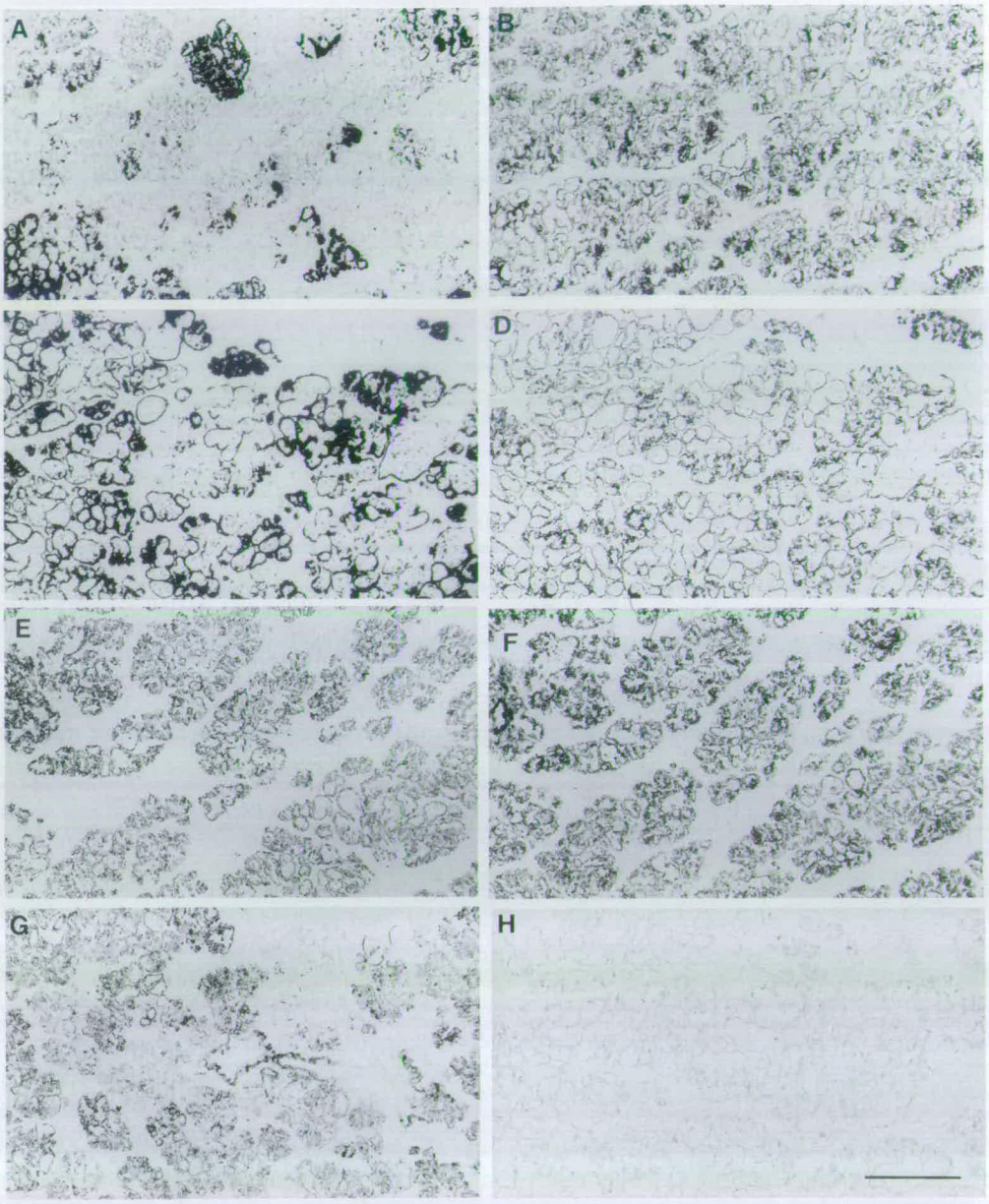


FIG. 3. *In situ* hybridization analysis of BLG and β-casein mRNA expression patterns in mouse mammary tissue. (*A* and *B*) Hybridization of BLG and β-casein probes to tissue from a low-expressing line 7 mouse (4 mg/ml). (*C* and *D*) BLG, β-casein, high-expressing line 7 mouse (17 mg/ml). (*E* and *F*) BLG, β-casein, line 14 mouse (8 mg/ml). (*G*) Line 45 mouse (25 mg/ml). (*H*) Control, hybridization with a β-casein sense probe, line 7 mouse (17 mg/ml). (Bar = 0.5 mm.)

phenomenon in *Neurospora* (35) and cosuppression in plants (36), that result in transgene silencing and variegated expression. We therefore entered into a study of transgene expression stability using three mouse lines that express the ovine BLG gene selectively within the lactating mammary gland. Only one expressed the transgene in a stable manner, whereas wide variations in the individual level of BLG expression were observed among individuals of the other two lines.

In the most variable line (line 7), expression levels are fixed within the individual. Further, the genetic background of this line does not play a significant role in this variation. In this pedigree, as well as in a second line (line 45), mammary tissue *in situ* hybridization experiments revealed mosaic expression of the transgene. In contrast, all the secretory cells within transgenic glands expressed the gene for an endogenous milk protein (β-casein), arguing against mosaicism due to cycling of active and inactive domains under hormonal control (37); recently, uniform expression of four endogenous milk proteins was observed throughout the mouse mammary epithelium by day 1 of lactation (38). While our observations do not rigorously rule out threshold effects due to suboptimal levels of lactogenic hormones leading to stochastic onset of milk protein gene expression, we think this unlikely. First, sections were taken from glands at a time when milk production is near-maximal and all secretory cells are producing β-casein. Second, line 14, harboring the identical construct to that present in the highly variable line 7, shows stable expression throughout the gland.

We believe that the patching of cells in which the transgene is active or inactive is most consistent with clonal expansion. Our results suggest epigenetic silencing of transcription occurs stochastically in individual progenitor cells, which is then transmitted through cell division to daughter cells, giving rise to mosaic patterns of expression. The closest precedent for our observations is afforded by the striated coat color patterns recorded by Mintz and colleagues (17, 18) in animals harboring tyrosinase fusion transgenes. All lines appeared to reproduce the striped pattern of expression, and animals either fully-

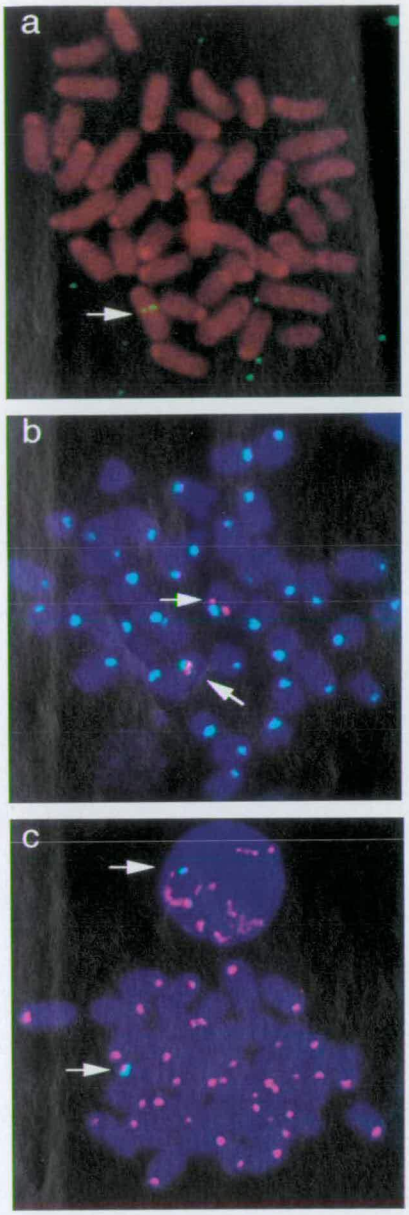


FIG. 4. Localization of BLG and major satellite (196) DNA sequences by chromosomal FISH to metaphase chromosomes. (*a*) Chromosomes from hemizygous line 14 mice were stained (red) using propidium iodide. Biotinylated BLG probe (arrow) was visualized indirectly using avidin-FITC (green). (*b*) Chromosomes from homozygous line 7 mice were stained (blue) with 4′,6-diamidino-2-phenylindole (DAPI). The BLG probe incorporated digoxigenin-11-dUTP and the major satellite DNA probe incorporated biotin-16-dUTP. The BLG probe (arrow) was detected with rhodamine conjugated anti-digoxigenin and Texas red conjugated anti-sheep (red). The major satellite probe was detected using successive layers of avidin-FITC, biotinylated anti-avidin, and avidin-FITC (green). Each chromatid of metaphase chromosome 15 can be seen labeled with the BLG signal in this example. (*c*) Chromosomes and one interphase nucleus from hemizygous line 45 mice; procedures were as for *b* except that the fluorochrome labeling systems were reversed for the BLG probe (arrow) and the major satellite probe.

pigmented or unpigmented were not reported; this contrasts with our observations that a transgene insertion can give rise both to animals in which the transgene is predominantly silent and to animals in which the transgene is expressed in the majority of target cells. Robertson *et al.* (39) have reported differing levels of expression in red blood cells for a globin/*lac* Z construct among transgenic mouse lines; underlying mechanisms may be similar, resulting in phenotypic differences that reflect the clonal origin of mosaic patches (phenoclones; ref. 40).

The phenomenon of position-effect variegation in *Drosophila* affords an informative parallel. Chromosomal rearrangement that relocates an endogenous gene locus close to the centromere can result in silencing of the locus in a variable proportion of cells (33). Silencing is attributed to chromosomal condensation or heterochromatinization brought about by proximity to the heterochromatic centromere. We argue that a similar process may occur in mouse. First, in both mouse lines exhibiting mosaic transgene expression, the array is integrated in the vicinity of the centromere, whereas in the stably-expressing line the transgene is inserted some distance away. Second, gene variegation in *Drosophila* can extend over 50–60 polytene bands (1–2 megabases) from a rearrangement breakpoint (41). While no precise measure can be provided, our estimates from FISH chromosomal spreads suggest that the transgene insertion of line 7 is some 4–5 megabases from centromeric heterochromatin. This is broadly consistent with the situation found in *Drosophila*, particularly if relative genomes sizes are taken into account. Third, Dorer and Henikoff (34) recently reported that heterochromatinization in *Drosophila* is not only dependent upon proximity to the centromere but also on the number of repeats present at the transgene locus. We note that both lines displaying mosaic expression harbor some 20 transgene copies, whereas the stably expressing line 14 contains only two copies. Heterochromatic features of an unusually large (1100 kb) β-globin transgene integrated into a peritelomeric region were reported recently (42). Furthermore, Mehtali *et al.* (43) described increased extinction of transgene activity (as judged by reporter gene expression and DNA methylation measurements) with increased copy number. It is therefore plausible to suggest that both integration site and number of transgene copies may predispose to the mosaic expression phenotype; indeed, proximity to repetitive DNA *per se* (either centromeric satellite DNA or repeat transgene copies) could be responsible for variable expression.

Our studies provide the first evidence for variegated expression of a transgene between individuals within a single transgenic line. Such mosaic expression has implications for investigations using gene addition technology. Gene addition is now being tested as a means of causing tissue specific cell ablation in the treatment of brain tumors (44), or in the study of the consequences of the loss of thyroid hormone production without surgery in mice (45). Variegated expression in such experiments would prevent 100% ablation and therefore result in variable numbers of cells remaining. Similarly, genetic manipulation of animal organs for transplant purposes (46) might be prejudiced by unpredictable transgene expression. In biotechnological applications, such as the production of proteins of biomedical interest in the milk of transgenic livestock (3, 4), mosaic expression will reduce the yield of product. Since this project was completed, other colleagues have reported variable expression within transgenic lines (B. Binas and T. Burdon, personal communication; S. Morley and J. J. Mullins, personal communication) suggesting that the phenomenon may not be uncommon. Particularly intriguing is the possibility that some endogenous genes within the mammalian genome, especially those close to heterochromatic regions [e.g., immunodeficiency, centromeric instability, facial anomalies (ICF) syndrome (47)] or present as tandem duplications at a single locus, may be susceptible to the type of mosaic silencing we describe here.

We are grateful to Andy Dawson, Eric Thompson, Roberta Wallace, and Frances Thompson for assistance with experiments. We thank Arthur Mitchell for the mouse major satellite DNA probe, and Norrie Russell, Roddy Field, and Elliot Armstrong for photography and

graphics. We would also like to thank Bert Binas, John West, Duncan Davidson, and Nick Hastie for helpful discussions and for reading the manuscript.

1. Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C. T., Findlay, J. B. C., Sivaprasadarao, R., Jones, T. A., Newcomer, M. E. & Kraulis, P. J. (1986) *Nature (London)* **324,** 383–385.
2. Simons, J. P., McClenaghan, M. & Clark, A. J. (1987) *Nature (London)* **328,** 530–532.
3. Clark, A. J., Simons, J. P., Wilmut, I. & Lathe, R. (1987) *Trends Biotechnol.* **5,** 20–24.
4. Archibald, A. L., McClenaghan, M., Hornsey, V., Simons, J. P. & Clark, A. J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5178–5182.
5. Wright, G., Carver, A., Cottom, D., Reeves, D., Scott, A., Simons, J. P., Wilmut, I. & Colman, A. (1991) *Bio/Technology* **9,** 830–834.
6. Palmiter, R. D., Wilkie, T. M., Chen, H. Y. & Brinster, R. L. (1984) *Cell* **36,** 869–877.
7. Al-Shawi, R., Burke, J., Jones, C. T., Simons, J. P. & Bishop, J. O. (1988) *Mol. Cell. Biol.* **8,** 4821–4828.
8. Wilkie, T. M., Braun, R. E., Ehrman, E. J., Palmiter, R. D. & Hammer, R. E. (1991) *Genes Dev.* **5,** 38–48.
9. Sandgren, E. P., Palmiter, R. D., Heckel, J. L., Daugherty, C. C., Brinster, R. L. & Degen, J. L. (1991) *Cell* **66,** 245–256.
10. Tan, S. S., Williams, E. A. & Tam, P. P. L. (1993) *Nat. Genet.* **3,** 170–174.
11. Butner, K. & Lo, C. W. (1986) *Mol. Cell. Biol.* **6,** 4440–4449.
12. Talarico, D., Peverali, A. F., Gimelli, E., Meneveri, R., Mondello, C. & Valle, G. D. (1988) *Mol. Cell. Biol.* **8,** 1336–1344.
13. McGowan, R., Campbell, R., Peterson, A. & Sapienza, C. (1989) *Genes Dev.* **3,** 1669–1676.
14. Allen, N. D., Norris, M. L. & Surani, M. A. (1990) *Cell* **61,** 853–861.
15. Engler, P., Haasch, D., Pinkert, C. A., Doglio, L., Glymour, M., Brinster, R. & Storb, U. (1991) *Cell* **65,** 939–947.
16. Sweetser, D. A., Hauft, S. M., Hoppe, P. C., Birkenmeier, E. H. & Gordon, J. I. (1988) *Proc. Natl Acad. Sci. USA* **85,** 9611–9615.
17. Bradl, M., Larue, L. & Mintz, B. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 6447–6451.
18. Mintz, B. & Bradl, M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 9643–9647.
19. Porter, S. D. & Meyer, C. J. (1994) *Development (Cambridge, U.K.)* **120,** 2103–2111.
20. Cattanach, B. M. (1974) *Genet. Res.* **23,** 291–306.
21. Whitelaw, C. B. A., Archibald, A. L., Harris, S., McClenaghan, M., Simons, J. P. & Clark, A. J. (1991) *Transgenic Res.* **1,** 3–13.
22. McClenaghan, M., Springbett, A., Wallace, R. M., Wilde, C. J. & Clark, A. J. (1995) *Biochem. J.* **310,** 637–641.
23. Gaye, P., Hue-Delahaie, D., Mercier, J. C., Soulier, S., Vilotte, J. L. & Furet, J. P. (1986) *Biochimie (Paris)* **68,** 1097–1107.
24. Gupta, P., Rosen, J. M., D'Eustachio, P. & Ruddle, F. H. (1982) *J. Cell Biol.* **93,** 199–204.
25. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056.
26. Monaghan, A. P., Davidson, D. R., Sime, C., Graham, E., Baldock, R., Bhattacharya, S. S. & Hill, R. E. (1991) *Development (Cambridge, U.K.)* **112,** 1053–1061.
27. Klever, M., Grond-Ginsbach, C., Scherthan, H. & Schroeder-Kurth, T. M. (1991) *Hum. Genet.* **86,** 484–486.
28. Langer-Safer, P. R., Levine, M. & Ward, D. C. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4381–4385.
29. Fantes, J. A., Bickmore, W. A., Fletcher, J. M., Ballesta, F., Hanson, I. M. & Heyningen, V. (1992) *Am. J. Hum. Genet.* **51,** 1286–1294.
30. Pietras, D. F., Bennett, K. L., Siracusa, L. D., Woodworth-Gutai, M., Chapman, V. M., Gross, K. W., Kane-Haas, C. & Hastie, N. D. (1983) *Nucleic Acids Res.* **11,** 6965–6983.
31. Fantes, J. A., Oghene, K., Boyle, S., Danes, S., Fletcher, J.M., *et al.* (1995) *Genomics* **25,** 447–461.
32. Whitelaw, C. B. A., Harris, S., McClenaghan, M., Simons, J. P. & Clark, A. J. (1992) *Biochem. J.* **286,** 31–39.
33. Spofford, J. B. (1976) in *Genetics and Biology of Drosophila*, eds. Ashburner, M. & Novitski, E. (Academic, London), Vol. 1c, pp. 955–1019.
34. Dorer, D. R. & Henikoff, S. (1994) *Cell* **77,** 993–1002.
35. Singer, M. J., Marcotte, B. A. & Selker, E. U. (1995) *Mol. Cell. Biol.* **15,** 5586–5597.
36. Matzke, M. & Matzke, A. J. M. (1993) *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* **44,** 53–76.
37. Molenaar, A. J., Davis, S. R. & Wilkins, R. J. (1992) *J. Histochem. Cytochem.* **40,** 611–618.
38. Robinson, G. W., McKnight, R. A., Smith, G. H. & Hennighausen, L. (1995) *Development (Cambridge, U.K.)* **121,** 2079–2090.
39. Robertson, G., Garrick, D., Wu, W., Kearns, M., Martin, D. & Whitelaw, E. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5371–5375.
40. Mintz, B. (1970) *Symp. Int. Soc. Cell Biol.* **9,** 15–42.
41. Hartmann-Goldstein, I. J. (1967) *Genet. Res.* **10,** 143–159.
42. Manuelidis, L. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1049–1053.
43. Mehtali, M., LeMeur, M. & Lathe, R. (1990) *Gene* **91,** 179–184.
44. Culver, K. W., Ram, Z., Wallbridge, S., Ishii, H., Oldfield, E. H. & Blaese, R. M. (1992) *Science* **256,** 1550–1552.
45. Wallace, H., McLaren, K., Al-Shawi, R. & Bishop, J. O. (1994) *J. Endocrinol.* **143,** 107–120.
46. Cozzi, E. & White, D. J. G. (1995) *Nat. Med.* **1,** 964–966.
47. Gimelli, G., Varone, P., Pezzolo, A., Lerone, M. & Pistoia, V. (1993) *J. Med. Genet.* **30,** 429–432.