Variegated Transgene Expression in Mice

Margaret L. Opsahl

Ph. D

University of Edinburgh

2001

EX BIBL. UNIV. EDIN.

DECLARATION

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

Margaret L. Opsahl

ABSTRACT

Transgenic animals harbour exogenous DNA, which is transmitted to their progeny. Many transgenes exhibit appropriate spatio-temporal patterns of expression. However, expression levels often differ between individuals of the same transgenic line. This thesis investigates expression instability, focusing on mice carrying a transgene encoding sheep β-lactoglobulin (BLG), the major whey protein. BLG transgenes inserted as multicopy arrays within the mouse genome are robustly expressed in mammary epithelium and the product is exported into milk (Simons *et al* 1987). In two of the three lines studied, milk BLG levels differed markedly between individuals. This was attributed to mosaic expression due to stochastic inactivation of the transgene (Dobie *et al* 1996), and suggestive of position effect variegation as seen in *Drosophila*. The highly variegating line, BLG/7, was chosen for further studies.

The original line was created on a mixed genetic background. First, to address a possible link between genetic background and variegation, BLG/7 animals were backcrossed to inbred mouse strains CBA and C57BL/6. Transgene expression continued to variegate after 13 backcross generations. However, inbreeding reduced the absolute expression levels: the mean and variance differed significantly from the parental population. Levels were restored by intercrossing inbred strains. Second, homozygous BLG/7 mice were studied. Variegated expression was maintained, but maximum expression levels were indistinguishable from heterozygote levels. Fluorescent *in situ* hybridisation for nuclear BLG transcription revealed that only one transgene locus is active per cell. This is suggestive of a *trans*-homology effect for the BLG/7 transgene. Previously, similar events have only been reported in plants and insects.

Finally, mosaic expression could reflect clonal expansion of committed progenitor cells. To address this issue, we sought to exploit X-inactivation. A lacZ reporter gene under BLG promoter control was targeted to the X-linked hypoxanthine- phosophoribosyl transferase (HPRT) gene of ES cells. To date, chimeric animals have been obtained; progeny reporter mice should address the issue of clonality of mosaic expression in BLG/7 mice.

ACKNOWLEDGEMENTS

I am very grateful to Dr Margaret McClenaghan (Roslin Institute) and Professor Richard Lathe (Centre for Genome Research, Edinburgh University) for initiating this project, and to Dr C. Bruce A. Whitelaw for taking on the main supervisory role in conjunction with Professor Lathe. I am grateful to all three for their practical advice and theoretical discussions throughout the duration of my studies.

I am also very indebted to all the staff in the Small Animal Unit at Roslin, both past and present, without whose excellent animal husbandry, this project would have been impossible. I am also obliged to Anthea Springbett, for invaluable statistical advice both prior to and on completion of the experimental work.

Whilst working in Professor Lathe's laboratory at the Centre for Genome Research, I have greatly benefited from the advice of Ken Rose and Melville Richardson, as well as informative discussions with Dr Andrew Smith on targeting vector design. At the Roslin Institute I would particularly like to thank John Webster for continued practical advice and support in all aspects of PCR and blotting techniques, Ray Ansell and Dr Jim McWhir for invaluable assistance in the ES cell work, Elena Farini for help with the HC-11 cells and Dr George Russell for protein gel imaging and advice on antibodies. For the FISH I received helpful advice from many different sources, including Drs Wendy Bickmore and Paul Perry (MRC Human Genetics Unit, Western General Hospital), Dr Nestor Lopez (Newcastle University), Dr Jacqueline Smith (Roslin Institute). Laura Dick and Louise Dunn (Roslin Institute) gave valuable help and advice on wax embedding and prehybridisation. Finally, I owe a great debt of gratitude to computing support, in particular Matt Blair, without whom the figures and this thesis would have been a lot harder to write.

I am very grateful for the environment provided by the members of the various divisions of Roslin Institute, where an open and welcoming attitude enabled me to seek advice from members of staff across the departmental divisions. This project was financed through a BBSRC CASE award, in conjunction with PPL therapeutics and Professor Alan Colman.

TO A MOUSE

Wee, slekit, cowrin, tim'rous beastie,
O, what a panic's in thy breastie!
Thou need na start awa sae hasty,
Wi' bickering brattle!
I wad be laith to rin an' chase thee,
Wi' murd'ring pattle!

I'm truly sorry man's dominion,
Has broken nature's social union,
An' justifies that ill opinion,
Which makes thee startle
At me, thy poor, earth-born companion,
An' fellow-mortal!

I doubt na, whiles, but thou may thieve;
What then? poor beastie, thou maun live!
A daimen icker in a thrave,
'S a sma' request;
I'll get blessin wi' the lave,
An' never miss't!

Thy wee bit housie, too, in ruin!
It's silly wa's the win's are strewin!
An' naething, now, to big a new ane,
O' foggage green!
An' bleak December's winds ensuing,
Baith snell an' keen!

Thou saw the fields laid bare an' waste,
An' weary winter comin fast,
An' cozie here, beneath the blast,
Thou thought to dwell,
Till crash! the cruel coulter past
Out thro' thy cell.

That wee bit heap o'leaves an' stibble,
Has cost thee mony a weary nibble!
Now thou's turn'd out, for a' thy trouble,
But house or hald,
To thole the winter's sleety dribble,
An' cranreuch cauld!

But Mousie, thou art no thy lane,
In proving foresight may be vain;
The best-laid schemes o' mice an' men,
Gang aft agly,
An' lea'e us nought but grief an' pain,
For promis'd joy!

Still thou art blest, compar'd wi' me;
The present only toucheth thee:
But och! I backward cast my e'e,
On prospect drear!
An' forward, tho' I canna see,
I guess an' fear!

Robert Burns

To the mice (mercifully few), to whom I owe so much: without whose contribution and sacrifice this thesis could not exist

(Paraphrased from Winston Churchill)

Also, to my family for continued support and encouragement throughout my studies: in return, you get a unique coffee table book with which to bore the neighbours and concuss burglars.

Finally, to Graham, who has been there for me throughout, in the good times and the bad, with excellent advice from someone who has travelled this path before me (some of which I followed), and to whom I owe so much.

CONTENTS

	PAGE
Declaration	ii
Abstract	iii
Acknowledgements	iv
Dedication	v - vi
Table of contents	vii -xiv
List of figures	xv - xvi
List of tables	xvii
Abbreviations	xviii
CHAPTER 1: GENERAL INTRODUCTION	1 - 46
1.1 The influence of vector design on transgene expression	**2 - 5**
1.1.1 Intronic sequences required for correct nuclear processing of transcripts	2
1.1.2 *Cis* regulatory sequences required 5' & 3' of coding sequence	3
1.1.3 Promoter choice is important	4
1.1.4 Sequence composition can also influence expression	4
1.2 How homologies within and between exogenous and endogenous gene sequences can affect expression	**6 - 14**
1.2.1 Homology dependent gene silencing in plants	7
1.2.2 Diploid genomes have 2 copies of most genetic loci: are both expressed and is this the norm for both endogenous and transgenic loci?	10
1.2.3 Paramutation and *transvection*: two unique cases of homology-driven regulation of gene expression	11
1.3 The influence of copy number on transgene expression	**14 - 15**
1.3.1 Transgene repeat arrays are susceptible to silencing	14

1.4 The influence of methylation status on transgene expression **15 - 18**

1.4.1 Methylation & Transgene Expression 16

1.4.2 Methylation & Imprinting 17

1.4.3 Methylation & Background 18

1.4.4 Other candidate epigenetic marks involved in gene silencing 18

1.5 The influence of age, germline transmission and genetic background on both endogenous and transgene expression **19 - 24**

1.5.1 Age–Related Silencing 19

1.5.2 Meiotic Inheritance of Epigenetic Silencing 20

1.5.3 The influence of genotype on phenotype 23

1.6 The influence of integration site on transgene expression **24 - 25**

1.7 Position Effect Variegation **25 - 40**

1.7.1 Heterochromatin function and composition 27

1.7.2 PEV and heterochromatin 29

1.7.3 Links between PEV and control of endogenous genes through modification of chromatin conformation 29

1.7.4 A diverse range of modifiers of PEV have been identified 31

1.7.5 Theoretical explanations proposed to explain PEV 32

1.7.6 Regulation of PEV 35

1.7.7 Mechanistic Models proposed to explain heterochromatic spread in PEV 35

1.7.8 PEV and X-inactivation: a case of shared properties and pathways? 38

1.7.9 PEV and Mice 39

1.8 Transgenic Mouse line BLG/7 **41 - 46**

1.8.1 Mammary Gland Function and Composition 42

1.8.2 BLG/7 lines of investigation 43

CHAPTER 2: MATERIALS & METHODS **47 - 84**

2.1 DNA manipulation **47 - 56**

2.1.1 Miniprep Protocol 47

2.1.2 50 ml Scale-up of Miniprep Protocol 47

2.1.3 Modified Qiagen Maxiprep Protocol 48

2.1.4 Transformations 49

2.1.5 DNA digests 50

2.1.6 Ligation Reactions 51

2.1.7 DNA purification methods 52

2.1.8 Polymerase Chain Reaction for genotyping mouse tail digests 53

2.1.9 X-Gal staining of ears for identification of PGK IRES βGeo mice 55

2.1.10 Mass screening of bacteria for successful ligations (colony lifts) 55

2.2 Cell Culture work **57 - 71**

2.2.1 HC-11 Cells, routine maintenance 57

2.2.2 Transfecting HC-11 cells using Calcium Phosphate method 59

2.2.3 X-Gal Staining of cells for LacZ expression 61

2.2.4 ES cells, Routine Maintenance 61

2.2.5 ES cell targeting work 64

2.2.6 Screening of cells for making chimaeric mice 67

2.2.7 Karyotyping Cells 69

2.2.8 Blastocyst injections 70

2.3 Protein Protocols **72 - 75**

2.3.1 Preparation of milk samples for protein analysis 72

2.3.2 Pouring, running and staining protein gels 72

2.3.3 Image Capture and Quantification 75

2.4 FISH **76 - 84**

2.4.1 Fixation and wax embedding of liver and mammary tissue 76

2.4.2 Sectioning of Wax Embedded Tissues 77

2.4.3 Prehybridisation 78

2.4.4 Hybridisations for mRNA expression 79

2.4.5 Nuclear FISH for intronic sequence detection. 81

2.4.6 Nuclear FISH for genomic sequence detection 82

2.4.7 Antibodies & sera used 83

2.4.8 Probes used 83

CHAPTER 3: ANALYSIS OF BACKCROSSED BLG/7 VARIEGATING MICE **85 - 118**

3.1 Introduction **85**

3.2 AIM **86**

3.3 Experimental Approach **86 - 87**

3.4 Results **87 - 106**

3.4.1 Strategy and critique of backcrossed BLG/7 milk sample analysis 87

3.4.2 Inbred milk samples display down-regulation of all milk proteins 89

3.4.3 Variegation occurs irrespective of genetic background, with mean expression levels remaining stable in an outbred background 92

3.4.4 Endogenous proteins show evidence of inbreeding-depression, which selectively represses the BLG transgene 97

3.4.5 Strain-specific modification of expression is superimposed on inbreeding depression effects, with 13th generation backcross populations differing from each other and outbred BLG/7 populations 99

3.4.6 Inbreeding- and strain-dependent down-regulation of BLG/7 expression is not heritable 101

3.5 Discussion **107 - 118**

3.5.1 Variegation is an inherent property of BLG/7 109

3.5.2 BLG/7 expression is open to modification through inbreeding depression 110

3.5.3 Evidence for genotype-specific modification of BLG/7 113

3.5.4 Epigenetic modifications not inherited 115

3.5.5 Variegation is maintained through the generations 118

CHAPTER 4: THE EFFECTS OF HOMOZYGOSITY ON THE EXPRESSION PATTERNS OF THE HIGHLY VARIEGATING TRANSGENE IN BLG/7 MICE **119 - 186**

4.1 Introduction **119 - 120**

4.2 AIM **121**

4.3 Experimental Approach **121 - 122**

4.4 Results **122 - 152**

4.4.1 Strategy and critique of hemi- and homozygous BLG/7 analysis 122

4.4.2 BLG protein levels suggest that the two homologous loci do not variegate independently 128

4.4.3 The spatial pattern of BLG expression revealed by mRNA FISH cannot discriminate between hemizygous and homozygous samples 134

4.4.4 The homologous loci of homozygous BLG/7 mice are not independent 140

4.4.5 Nuclear positioning of BLG/7 alleles in liver suggestive of early separation of loci 150

4.5 Discussion **152 - 186**

4.5.1 The hypothesis 153

4.5.2 Transgenic BLG protein expression does not double in homozygous population compared to hemizygous results 155
4.5.3 Hemi- and homozygote BLG/7 mice exhibit similar cellular expression patterns 157
4.5.4 Probes for BLG/7 nuclear transcription reveals a single focus of transcription in homozygous samples, whereas both homozygous BLG/45 mice and samples probed for WAP reveal two foci 158
4.5.5 BLG/7 expression is monoallelic 158
4.5.6 Nuclear positioning of BLG/7 alleles 162
4.5.7 Theoretical Model 162

CHAPTER 5: IS BLG/7 MOSAIC EXPRESSION CLONAL OR INSTRUCTIONAL? 187 - 203

5.1 Introduction 187 - 189
5.1.1 Is BLG/7 mosaicism clonally or instructionally derived? 187
5.1.2 How to differentiate between instructional and clonal states 188

5.2 Experimental Design 189 - 190

5.3 Design and assembly of the clonal reporter construct: HBZ 190 - 192
5.3.1 Reporter Choice 190
5.3.2 Promoter Choice 190
5.3.3 Choice of targeting vector 192
5.3.4 Choice of target gene 192

5.4 Testing HBZ Reporter Competency 193

5.5 Targeting into ES cells and generation of chimeric mice 193 - 197
5.5.1 Electroporation and selection for positively targeted cells 193
5.5.2 Southern Blotting confirms targeting 195
5.5.3 Blastocyst injections and generation of chimeras 197

5.6 An alternative strategy: Exploiting an alternative X-linked reporter line containing an IRES-βGeo reporter construct targeted to and under PGK gene control control **197 - 203**

5.6.1 MOP mice 197

5.6.2 MOP milk results 199

5.6.3 MOP mRNA FISH 202

CHAPTER 6: GENERAL DISCUSSION **204 - 224**

6.1 BLG/7 mice are an ideal model system for studying PEV **204 - 205**

6.2 Future work based on the backcross experimental results of Chapter 3 **205 - 208**

6.2.1 Continuation of outcrossing experiment 206

6.2.2 Dissection of multiple modifier loci responsible for inbreeding depression and strain-specific modification 206

6.3 Future work based on the results of the effects of two homologous BLG/7 loci on transgene expression in chapter 4 **209 - 212**

6.3.1 Is the epigenetic component leading to monoallelic expression inherited? 209

6.3.2 Timing of allelic exclusion in BLG/7 homozygous mice 210

6.3.3 Spatial distribution of variegated expression pattern 210

6.3.4 Does BLG/7 differ from BLG/45 210

6.3.5 Investigation of BLG/7 nuclear localisation 211

6.3.6 Expression pattern of endogenous milk proteins in nucleus 211

6.3.7 Endogenous neighbouring gene expression patterns 212

6.4 Future work based on the results of the investigation into the clonal origins of BLG/7 variegated expression patterns in chapter 5 **213**

6.5 The position of Position Effect Variegation in human disease **213 - 215**

6.5.1 The different position effects that can cause human disease phenotypes 214

6.5.2 Position effects in transgenic animals, serendipitous model systems 214

6.6 Using knowledge of modifiers of PEV to improve both transgenic livestock and treatment of human diseases **215 - 218**

6.6.1 Using knowledge of modifier loci to improve the performance of transgenic animals 215

6.6.2 Exploiting a knowledge of modifier loci in the treatment of human diseases 216

6.6.3 A role for somatic mutation in human PEV? 217

6.6.4 Using the knowledge of modifier loci for future treatment strategies 218

6.7 *Trans*-interactions and PEV loci **219 - 220**

6.7.1 Why we need to understand the interactions leading to monoallelic expression 219

6.8 Using BLG/7 as a model system for complex diseases involving both PEV and suspected trans-interactions **221 - 223**

6.8.1 BLG/7 as a model system for autosomal dominant diseases 221

6.8.2 Candidate autosomal dominant conditions that may be the result of *trans*-homology interactions 222

6.9 Concluding Remarks **223 - 224**

References **225 - 250**

Appendix **251 - 257**

A1 Raw data for inbred and outcrossed milk samples 251 - 252

A2 Raw data for hemi- and homozygous milk samples 253 - 254

A3 Raw data for MOP milk samples 255 - 256

B Statistical terms 256

C Mating regimes for BLG/7 line maintenance 256 - 257

C1 Mating regimes for hemi- and homozygous BLG/7 populations 257

C2 Mating regimes for backcrossed BLG/7 populations 257

LIST OF FIGURES

		PAGE
Figure 1.1	Position Effect Variegation (PEV) in *Drosophila*	26
Figure 1.2	Mammary gland development during pregnancy and lactation	44
Figure 3.1	SDS PAGE of milk protein samples for CBA & C57BL/6 13th generation backcross mice	90
Figure 3.2	SDS PAGE of milk protein samples for all 3 genotypes	91
Figure 3.3	Bar charts representing individual BLG milk protein yields in mg/ml	96
Figure 3.4	Scatterplot showing individual BLG milk protein content and mean BLG expression for all 3 genotypes .	100
Figure 3.5	Fluorescent *in situ* hybridisations of backcrossed mammary tissue	102
Figure 3.6	SDS PAGE of milk protein samples for outcrossed mice	104
Figure 3.7	Bar chart representation of individual outcrossed BLG milk protein yields in mg/ml	104
Figure 3.8	Scatterplot showing individual BLG milk protein content and mean BLG expression for all 4 genotypes	117
Figure 4.1	SDS PAGE of milk protein samples for Hemi- and Homozygous BLG/7	123
Figure 4.2	mRNA FISH in lactating mammary tissue	125
Figure 4.3	Hemi- & Homozygous milk expression	132
Figure 4.4	Scatterplot showing individual BLG milk protein content for both genotypes .	133
Figure 4.5	Double mRNA FISH from high BLG/7 hemizygote expressor (H1 = 14 mg/ml)	135
Figure 4.6	Double mRNA FISH from high BLG/7 homozygote expressor (HH17 = 8.8 mg/ml)	136
Figure 4.7	Double mRNA FISH from lower BLG/7 hemizygote expressor (H13 = 6.3 mg/ml)	137

Figure 4.8 Double mRNA FISH from low BLG/7 homozygote expressor (HH16 = 7.7 mg/ml) 138

Figure 4.9 Double mRNA FISH of similar hemi- (A-C) and homozygous (D-F) expressors (H2 = 8.8 mg/ml, HH9 = 9.3 mg/ml) 139

Figure 4.10 A - C Graphical representation of the data from table 4.4 143 - 145

Figure 4.11 Intronic RNA FISH for nuclear BLG/7 transcription in homozygous sample (HH17) 147

Figure 4.12 Intronic RNA FISH for nuclear WAP transcription in homozygote BLG/7 mammary tissue 148

Figure 4.13 Intronic RNA FISH 149

Figure 4.14 Intronic RNA & genomic DNA BLG/7 homozygote FISH 151

Figure 4.15 Theoretical Model explaining BLG/7 homozygote monoallelic expression. 164

Figure 5.1 Cloning of targeting vector 191

Figure 5.2 Induction of HC11 cells tranformed with HBZ construct 194

Figure 5.3 Southern Blot of *Bam*HI cut HBZ clones 196

Figure 5.4 HBZ 1 Chromosome spread and resulting chimeric mice 198

Figure 5.5 A) SDS PAGE of milk protein samples for MOP mice 201

B) Bar chart representing individual BLG milk protein yields in mg/ml 201

Figure 5.6 MOP mammary tissue probed for BLG and LacZ expression 203

LIST OF TABLES

Table 3.1	**A)** BLG measurements for 13th generation BLG/7 CBA backcrosses	93
	B) BLG measurements for 13th generation BLG/7 C57BL/6 backcrosse	94
Table 3.2	Average BLG expression with mean, SD & CV for 13th generation BLG/7 backcrosses on CBA and C57BL/6 backgrounds	95
Table 3.3	Backross, outbred and 3rd generation milk protein results	97
Table 3.4	Micro-Kjeldahl analysis results	98
Table 3.5	BLG measurements for outcrossed BLG/7 mice (=MP)	105
Table 3.6	Average BLG expression with mean, SD & CV for outcrossed BLG/7 mice	106
Table 3.7	Backcross, outbred & outcross milk protein results	103
Table 4.1	**A)** BLG measurements for hemizygous BLG/7 mice	129
	B) BLG measurements for homozygous BLG/7 mice	130
Table 4.2	Average BLG expression with mean, SD & CV for hemizygous (H) and homozygous (HH) BLG/7 mice.	131
Table 4.3	Backross, outbred and 3rd generation milk protein results	134
Table 4.4	Summary of mRNA FISH counts for both hemi and homozygote BLG/7 mice	142
Table 4.5	BLG mRNA FISH counts for hemizygote BLG/7 population	166 - 174
Table 4.6	BLG mRNA FISH counts for homozygote BLG/7 population	175 - 186
Table 5.1	BLG measurements for MOP mice	200
Table 5.2	MOP and Outbred (hemizygous) milk protein results	199

ABBREVIATIONS

BLG	β - lactoglobulin
bp	base pair(s)
cDNA	complementary DNA
DAPI	4, 6–diamidino–2–phenylindolene
DEPC	diethyl pyrocarbonate
DTT	dithiothreitol
EtOH	ethanol
FISH	fluorescence *in situ* hybridisation
g	gram(s)
gDNA	genomic DNA
HPRT	Hypoxanthine Phosphoribosyl Transferase
kb	kilobase
M	Molar (moles/litre)
μ	micro
m	milli
mRNA	messenger RNA
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEV	position effect variegation
RNase	ribonuclease
RT-PCR	reverse transcription PCR
SDS	sodium dodecyl sulphate
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol ($C_4H_{11}NO_3$)

Chapter 1

General Introduction

Transgenic technology has revolutionised biological research, its applications spanning a wide field of interest covering both industry and academia. During the past two decades, a large number of different transgenic plant and animal lines have been created. Although they differ widely with regards to transgene content or endogenous gene modification, successful transgenic lines require that any modification display certain key properties. Crucial to all transgenic applications is stability of integration, expression and inheritance of the transgene or endogenous gene modification (for an overview, see Jablonka 1992). With the increasing numbers of transgenic lines created, it has become evident that not all the transgenes are behaving as expected. One of these phenomena is transgene silencing. Silencing can be progressive, with reports of both endogenous and transgene silencing as a function of time, age or generation (see 1.5). All systems examined to date exhibit silencing phenomena, as the examples in this chapter clearly demonstrate. This suggests that the ability to silence gene expression is an ancient and possibly conserved mechanism.

The loss of expression could be caused by transgene deletion, but recent investigations indicate that most transgenes remain stably integrated and inherited once a line has been established (Aigner *et al* 1999). Silencing must therefore be by some mechanism whose action can be maintained without changing the underlying genetic blueprint. Changes in gene expression without sequence modification is termed epigenetics, a term first introduced almost 60 years ago (Waddington 1942). Although the exact definition of the term is debated (Wu & Morris 2001), I believe that the most generally used and accepted definition derives from Holliday's definitions (1987, 1996), where epigenetics is explained as the study of changes in gene expression/function that are mitotically and can also be meiotically inherited without changing the DNA sequence (Wu & Morris 2001). Methylation (Assaad *et al* 1993, Koetsier *et al* 1996, Schweizer *et al* 1998) and histone acetylation (Cavalli & Paro 1999, Kim *et al* 1999) are two examples of epigenetic marks.

Vector design, copy number, methylation status and site of integration are all known to influence transgene expression. Transgene silencing is now frequently reported, as is variable levels of expression from transgenic lines. Variations in transgene product yields are being observed both within and between some transgenic lines containing the same transgene inserted at different genomic locations (Palmiter *et al* 1982 & 1984, Pittius *et al* 1988, Assaad *et al* 1993, Kunz *et al* 1996, Dobie *et al* 1996, Festenstein *et al* 1996, Gutierrez *et al* 1996, Cerutti *et al* 1997, Guy *et al* 1997, Kieffer *et al* 1997, Ogilvy *et al* 1998, Matzke *et al* 2001). Although several of the published papers on variable transgene expression report differences in expression levels between lines (Pittius *et al* 1988, Robertson *et al* 1996, Kieffer *et al* 1997), in many cases there is no mention of any investigation into the relative expression levels between different transgenic lines (Morley *et al* 1996). There are increasing reports of variable expression levels within lines as well (Palmiter *et al* 1982, Bleck & Bremel 1994, Dobie *et al* 1996 & 1997, Matzke *et al* 2001). In addition to this being an interesting question regarding gene silencing, variable transgene expression has serious implications for the viability of commercial transgenics using plants and animals as bioreactors, as well as the future of effective gene therapy and academic investigations into gene function.

1.1 The influence of vector design on transgene expression

Vector design and delivery can have a profound effect on the expression potential of a transgene. The choice of delivery system and carrier construct determines the amount of DNA that can be transferred into the host system. The size limitations will in turn influence transgene design. Three elements are each credited with increasing the frequency of generation of lines harbouring expression competent transgenes.

1.1.1 Intronic sequences required for correct nuclear processing of transcripts

The first transgene constructs contained cDNA driven by minimal promoters. Despite good *in vitro* results, *in vivo* performances were disappointing (Whitelaw *et al* 1991). Inclusion of intronic sequences was found to enhance in *vivo* expression by

increasing the probability of expression (Whitelaw *et al* 1991, Shani *et al* 1992, Clark *et al* 1993, Clark *et al* 1997).

Intronic sequences may serve a function crucial to transgene expression potential: breaking up the coding sequences, contributing to correct chromatin conformation of the transgene locus (1.2) and avoiding homologous interactions with pseudogenes (Kricker *et al* 1992). A second function has been uncovered by studies on the β-globin locus transcription control, where transcription from the intergenic regions is required to keep the locus open through and accessible to the transcriptional machinery (Gribnau *et al* 2000). This may not be a general function of intronic sequences at other loci.

Additionally, 5' and 3' regulatory sequences are also important for transcript processing (Custodio *et al* 1999). This, and other aspects of RNA processing were not investigated as part of this thesis, and will therefore not be discussed further.

1.1.2 *Cis* regulatory sequences required 5' & 3' of coding sequence

Flanking sequences may play a large role in determining expression patterns of many transgenes, in particular the transgene locus investigated as part of this thesis.

Inclusion of locus control regions (LCRs) is reputed to confer position independent transcription and increased transcription rates (Festenstein *et al* 1996, Kioussis & Festenstein 1997), although later reports suggest that not all negative effects imposed by the genome itself or associated transgenic sequences can be overcome by an LCR (Guy *et al* 1996 & 1997, Alami *et al* 2000). LCRs increase the proportion of expressing recovered transgenic lines, enabling expression by establishing an open chromatin configuration (Grosveld *et al* 1997, Kioussis & Festenstein 1997), which requires the inclusion of the entire LCR (Kioussis & Festenstein 1997). In addition, an LCR can convey tissue specificity (Santoso *et al* 2000).

LCRs should not be confused with enhancers, another regulatory control element often included in transgene constructs. Enhancer elements can increase the number

of expressing cells (Walters *et al* 1995, Sutherland *et al* 1997, Blackwood & Kadonga 1998, Francastel *et al* 1999), acting cumulatively in *cis* to increase the probability of correct spatial and temporal expression (Blackwood & Kadonga 1998) by establishing an open chromatin configuration (Kioussis & Festenstein 1997). Inclusion of functional enhancers have been linked with non-centromeric location of transgene integration in cell culture. Mutated enhancers increase transgene silencing and localisation to centromeric heterochromatin (Francastel *et al* 1999).

Enhancers in turn should not be confused with insulators, also known as boundary elements. Insulators protect the transgene from position effects in single copy inserts (Dorer 1997). A putative boundary element has been identified in mammals: it is speculated that the chromatin structure of the maternally derived 5' H19 locus may contain an epigenetic mark that prevents access by Igf2 to the 3' enhancer elements present on the maternal chromosome. These sequences function as silencers in *Drosophila,* supporting their role as a boundary element (Hark & Tilghman 1998).

Inclusion of matrix attachment regions (also known as scaffold attachment regions), thought to influence the nuclear localisation of the gene, is also reported to enhance expression from transgene constructs (Porter & Meyer 1994, Dorer 1997, Ulker *et al* 1999, Vain *et al* 1999). Another reported effect is increased stability of expression over generations (Ulker *et al* 1999, Vain *et al* 1999). A plant-derived SAR works perfectly in mammalian cells (Dietz *et al* 1994), suggesting functional conservation.

The inclusion of all *cis*-regulatory sequences connected to a transgene in its natural environment appears to be able to confer (in most instances), position independent expression. This is reported for the chicken lysozyme locus (Bonifer *et al* 1996, Huber *et al* 1996), human gp91-phox (Lien *et al* 1997), β-casein (Persuy *et al* 1992) and some metallothionein fusion constructs (Palmiter *et al* 1993). The success of large vectors, such as YACs and BACs, may in part be due to inclusion of all regulatory sequences associated with a gene (Giraldo & Montoliu 2001).

1.1.3 Promoter choice is important

The promoter used in a transgene construct is chosen based on several criteria. Promoter strength, specificity and availability are three major factors that govern promoter choice. The choice of an inappropriate promoter can have disastrous consequences, with plants providing a notorious example. Viral promoters are widely used in transgenic plants, unfortunately, they can trigger the genome defence system, which identifies the promoter and all associated sequences as viral and proceeds to inactivate expression (Vaucheret 1993, Assaad *et al* 1993, Ratcliff *et al* 1997, McCabe *et al* 1999). The serendipitous discovery of the plant genome defence mechanism has been exploited to engineer crops resistant to viral attack, with as little as 90 bp of homology (Vaucheret 1993) sufficient to induce specific resistance (Guo *et al* 1998, Waterhouse *et al* 1999, Baulcombe 1999, Kooter *et al* 1999). In animals, a recent paper has demonstrated that the use of endogenous promoters decreases the variability of expression and increases stability and reproducibility when the construct is re-used, as opposed to retroviral promoters (DeAngioletti *et al* 2001).

1.1.4 Sequence composition can also influence expression

Checking for cryptic splice sites or start codons can prevent unwanted expression or processing of the transgene (Al-Shawi *et al* 1991, Yull *et al* 1995, Haseloff *et al* 1997), whilst codon use can be as influential as promoter choice in determining transgene expression potential since prokaryotic and eukaryotic amino acid triplet codon preferences differ. By changing the codon usage in a sequence to match that of the intended host species, the efficiency of protein production from the mRNA is significantly enhanced, as demonstrated for the green fluorescent protein reporter derived from jellyfish (Zolotukhin *et al* 1996). Ensuring that the nucleotide composition of the transgene construct resembles the profile of the host genome is also important for a successful transgene. In plants, transgene sequences with a base content whose pattern diverges from the host are vulnerable to targeted silencing by the host genome defence system (Meyer *et al* 1993, tenLohuis *et al* 1995). The sequence divergence may cause an identifiable structural abnormality.

Evidence for operational genome defence systems in animals

Plant transgenics has suffered from the prevalence of viral promoters in their constructs (1.1.3). Viral and bacterial derived material is also used in animal transgenics, although the sequences used are more commonly reporter genes or 3' processing signals. Although there is little evidence for viral sequences triggering silencing events in mammals, there are growing numbers of articles reporting silencing by prokaryotic reporter sequences such as chloramphenicol acetyltransferase (Clark *et al* 1997) and β-galactosidase (Cohen-Tannoudji *et al* 2000 a & b, Montoliu *et al* 2000, Ramirez *et al* 2001). There is also evidence for an alternative "policing" system in mammals that can identify and eliminate foreign sequences (Scrable & Sambrook 1999). *LacI* sequences of bacterial, but not mammalian, origin were targeted for somatic elimination (Scrable & Sambrook 1999). This highlights the desirability of using mammalian promoters and gene sequences. The presence in both plants and animals of a genome surveillance system suggests an ancient shared mechanism responsible for protecting genome integrity (Matzke *et al* 2000).

1.2 How homologies within and between exogenous and endogenous gene sequences can affect expression

Transgenes are frequently present as multicopy arrays inserted at one or more loci in the genome and often exhibit silencing in one form or another. This silencing phenomenon is termed homology-dependent gene silencing (HDGS). Under the umbrella heading of HDGS, a number of widely different functional silencing routes take shelter (Wu & Morris 1999). HDGS encompasses all epigenetic inactivation based on sequence recognition at either the DNA or RNA level, with silencing taking place at the transcriptional (TGS = transcriptional gene silencing) or post-transcriptional (PTGS = post-transcriptional gene silencing) stages (see Hsieh & Fire 2000 for review). Homologies leading to gene silencing can be found:

A) Within a multi-copy array at a single site (Assaad *et al* 1993, Garrick *et al* 1998, Birchler *et al* 2000) (see 1.3 for further detail),

B) Between a transgene and endogenous genes (Davies *et al* 1997, Jorgensen 1995, Seo *et al* 1995, Smith *et al* 1990, tenLohuis *et al* 1995, van Blokland *et al* 1994) and

C) Between transgenes at linked or unlinked loci (Jorgensen 1990, English *et al* 1997, Neuhuber *et al* 1994, Park *et al* 1995, Vaucheret 1993).

The amount of homology and level of silencing varies. Plants dominate reported HDGS phenomena, both by the sheer bulk of reports, and also because only plants have so far given multiple confirmed results of HDGS covering every sub-category of the umbrella term.

1.2.1 Homology dependent gene silencing in plants

Gene silencing in plants can be triggered by sequence homologies between the transgene and endogenous genes, or between homologous transgenes at separate loci (Flavell 1994, Matzke & Matzke 1995, Itoh *et al* 1997, Matzke & Matzke 1998). This phenomenon is clearly connected, at least mechanistically, to the virally induced silencing phenomenon mentioned in 1.1.3 (Waterhouse *et al* 1999, Baulcombe 1999). Promoter homology is sufficient to induce silencing (Park *et al* 1996), which can occur at the transcriptional or post-transcriptional level (Kooter *et al* 1999), the latter believed restricted to plants and fungi. Transcriptional gene silencing is shared with animals and insects, and will be discussed together. Post-transcriptional gene silencing is unique, and will therefore be discussed in detail below.

Post-transcriptional gene silencing (PTGS)

Natural occurrence of PTGS to control gene expression has only been demonstrated in plants and fungi (Maessen 1997,Cogoni *et al* 1996). In plants, PTGS is believed to be part of an endogenous "immune-system" that fights viral infections and other assaults on cellular and genome integrity (Matzke *et al* 2000) and is also used to control endogenous gene expression (Maessen 1997). A threshold level of mRNA production probably has to be reached before the silencing machinery becomes

activated, and may explain why homologies to certain endogenous genes only trigger silencing of gene expression in certain cell types or developmental stages (Smith *et al* 1990, Boerjan *et al* 1994, Meyer & Saedler 1996). Aberrant or overly abundant mRNA sequences are capable of triggering silencing (Smith *et al* 1990, Vaucheret 1993, Boerjan *et al* 1994, Mueller *et al* 1995, Meyer & Saedler 1996), a combination of a strong transgene promoter and abundant endogenous transcripts triggering the sensing system that targets excessive transcripts of any mRNA species for degredation. This system may be used to both combat viral infection and control endogenous gene expression through a negative feedback loop

The first report of transgene PTGS was for a chalcone synthase (CHS) gene in Petunia (de Carvalho *et al* 1992, Jorgensen 1995). PTGS has helped explain the inherent instability of many plant transgenes, and provided scientists with an excellent tool to generate crops resistant to viral infection (Mueller *et al* 1995). This method can also target endogenous genes, successfully demonstrated by the double silencing of tomato polygalacturonase and pectinerase (Simons & Tucker 1999).

PTGS can be inherited

Despite using a post-transcriptional silencing route, there is evidence that this silencing phenomenon can be inherited. Resistance to potato virus Y infection is maintained through 4 generations of inbreeding (Han *et al* 1999). The same stability is seen in rice with engineered resistance to rice yellow mottle virus (Pinto *et al* 1999). Others have been less successful. Homozygocity for the transgene locus is need to provide complete resistance to pepper mild mottle tobamovirus, with unstable inheritance of the phenotype typical (Tenllado & Diaz-Ruiz 1999) whilst *Brome mosaic virus* resistance phenotypes are variable in progeny from R1 resistant parents (Iyer & Hall 2000). Evaluating the use transgenes to generate resistance to plant viruses is difficult as all the examples test different viruses, with target choice, vector design and delivery influencing effectiveness. Importantly, virus ability to circumvent the induced resistance mechanism will also differ.

Maintenance is either through continued degradation of transcripts or transcriptional silencing following the initial PTGS event. The former precludes meiotic inheritance of the resistance phenotype. The latter can initiate inherited resistance (Vaucheret 1993; 1994), where initial transcription from the transgenic locus silences the locus and all associated loci in *trans*. The transcriptional silencing following PTGS may be initiated by the RNA transcripts themselves interacting with the DNA locus, possibly initiating methylation at the homologous sites (Van Houdt *et al* 2000). A mammalian equivalent is the involvement of Xist RNA in X-chromosome inactivation (Willard & Salz 1997). *Drosophila* provide an alternative suggestion, where RNA-protein complexes are implicated in dosage compensation (Willard & Salz 1997).

Post-transcriptional gene silencing can be induced in mammalian cells

Although confirmed natural PTGS is currently limited to plants and fungi, there is some suggestion that mammalian systems may be capable of PTGS in a manner not induced by antisense-RNA (Bahramian & Zarbl 1999, Sharp 1999).

All silencing pathways may share a common ancestral link, which has evolved into the multitude of silencing phenomena seen today. Some of the pathways may no longer be used in all organisms, but the components remain present and capable of reactivation. The prevalence of PTGS in plants and fungi may relate to their more sedentary lifestyle, which renders them particularly susceptible to viral attack. Late germline segregation may also render the genetic blueprint vulnerable and in need of extra protection from viruses and transposable elements. Identification and silencing of aberrant sequence transcription may be crucial to maintenance of species identity and survival, and has thus been retained through the forces of natural selection.

Animals segregate their germline early in development, protecting integrity of transmission. Mobility and a complex efficient immune system may explain why PTGS was not maintained in animals. The resurfacing of this phenomenon (albeit in an *in vitro* environment) through transgenic manipulation processes suggests that the mechanism lies dormant but can be activated when normal defence barriers against an assault on the genome are bypassed.

1.2.2 Diploid genomes have 2 copies of most genetic loci: are both expressed and is this the norm for both endogenous and transgenic loci?

Endogenous autosomal genes are normally present in a double dose in diploid genomes. Expression from both copies of the gene is presumed, though I am unaware of any study that has investigated this in great depth. Imprinted genes (Tilghman 1999) and those linked to the sex chromosomes (Panning & Jaenisch 1998) are the general exceptions known to exhibit mono-allelic expression. Recently, proof that autosomal genes can also be monoallelically expressed has emerged (Hollander *et al* 1998, Rhoades *et al* 2000). Both papers report apparent regulation of interleukin 2 (IL-2) expression levels through random inactivation of one of the two alleles.

Few transgenic loci are made homozygous, and most of the examples are transgenic plants were homozygocity is a relatively easy state to achieve through selfing. In plants, silencing phenomena are quite diverse and transgenes can exhibit different phenotypes within a single line (Tenllado & Diaz-Ruiz 1999, Iyer & Hall 2000). Homozygocity can induce increased silencing (Tenllado & Diaz-Ruiz 1999), or epigenetic mosacism (Iglesias *et al* 1997). The silent phenotype can be transmitted through the germline (Iglesias *et al* 1997), and reactivation of gene expression once the gene loci are separated can in certain cases take several generations (Jakowitsch *et al* 1999). One paper even reports probable monoallelic expression in homozygous transgenes, where a silent allele is transmitted to 50% of the progeny (Matzke *et al* 2001). Investigations into the effects of gene knockout and connected applications appear to be the predominant reason for generating homozygous transgenic animals (Kabra *et al* 2001, Adam *et al* 2000, Philbrick *et al* 1998, Urbanek *et al* 1997). When a functional transgene has been generated in a homozygous state, few investigators appear to have investigated comparative expression levels of the transgene (at any rate, there is no mention of it in most papers). Two of the few authors to address this are the Matzkes (Matzke *et al* 2001). By extrapolation from their work on transgenic plants, one could estimate that up to ¾ of stably expressing lines should give rise to a doubling in expression level when compared to hemizygotes. This does not include variegating lines (see 1.7).

1.2.3 Paramutation and *transvection*: two unique cases of homology-driven regulation of gene expression

Paramutation and *transvection* are two similar phenomena reported in plants and *Drosophila* respectively (Henikoff & Comai 1998, Wu & Morris 1999). Each involves the interaction of two (usually) homologous gene loci, leading to a modification in gene expression. Somatic pairing between homologous genes is a requirement of both paramutation and *transvection*. The pairing enables the one site to influence the transcriptional activity of the other site (Scholz *et al* 1993, Dernburg *et al* 1996, Matzke *et al* 2001), possibly through transfer of activating or silencing complexes and/or changing nuclear localisation of the loci (Sass & Henikoff 1999).

Transvection

Transvection is a rare phenomenon that can mediate complementation between paired alleles, with the promoter region apparently important for the regulation of the interaction (Morris *et al* 1999). *Transvection* was first used to describe some unexpected interactions between mutant alleles belonging to the bithorax complex (E.B. Lewis 1954, Judd 1988, Tartof & Henikoff 1991). The term *transvection* was used to define the process of homologue pairing leading to altered gene activity. *Transvection* in *Drosophila* has as its most notable example certain mutations of the *brown* gene (Henikoff & Dreesen 1989, Dreesen *et al* 1991, Martin-Morris & Henikoff 1995, Csink & Henikoff 1996, Dernburg *et al* 1996, Sass & Henikoff 1999). Other genes subject to *transvection* phenomena have also been reported in *Drosophila* (Georgiev & Corces 1995, Gdula *et al* 1996, Donaldson & Karpen 1997, Dorer & Henikoff 1997, Morris *et al* 1999). Studies of the *pointed* gene indicate that both *trans*-activation and *trans*-inactivation can occur (Scholz *et al* 1993). One of the few transgenic examples of transvection has been reported for *white* gene (Martin-Morris *et al* 1997). The *trans*-inactivation of a paired homologue can also inactivate genes in close proximity to the insertion site. Despite the vast numbers of reports on silencing, dominant trans-inactivation remains a rare phenomenon, even in *Drosophila*, the most studied of genetic tools.

Paramutation

First described 70 years ago for the maize r1 gene (Winkler 1930), paramutation involves interactions between two gene alleles. The paramutable allele is the sensitive allele, and upon interaction with the dominant paramutagenic allele adopts the paramutagenic transcriptional status (Matzke & Matzke 1995). Separation of the two loci can lead to reversion of the paramutable allele to "wildtype" expression, or permanent transformation to a functional paramutagenic allele. This indicates that there are at least two pathways in operation, with epigenetic modifications that can be maintained through meiosis. Transgenic paramutation was first described in Petunia (Meyer *et al* 1993), and can involve homologous sequences at non-homologous sites, with pairing essential as trisomies led to inactivation of only one of the two paramutable alleles. These results suggest that pairing and "information exchange" occurs during a limited window of the cell cycle. The cell cycle, especially length of S- and G1/S-phase transition, may influence silencing of genes subjected to phenomena such as position effect variegation (Fox & Rhine 1996, Seum *et al* 1996), indicating that "decisions" on gene activation status occur only at certain time points.

The influence of pairing

Somatic pairing is proposed to be a prerequisite for *trans*-interactions (Henikoff & Dreesen 1989, Meyer *et al* 1993, Golic & Golic 1996, Henikoff 1997) and may play a role in the control of some human genes (Tartof & Henikoff 1991). In fungi, pairing of alleles of an endogenous gene has been demonstrated to be essential for the correct regulation of expression and thus sexual sporulation (Aramayo & Metzenberg 1996). The first article suggesting the possibility of a *trans*-sensing mechanism in fungi was published a year earlier (Woloshuk *et al* 1995), describing the dominant inactivation of an aflatoxin biosynthetic pathway gene.

Many examples of pairing dependent changes in transcription have been documented in *Drosophila*, with the subsequent emergence of a large number of similar terms: such as *transvection*, *trans*-interaction, pairing–sensitive repression and *trans*-repression (Henikoff 1997). Many describe silencing of gene expression, with

interactions between allelic and non-allelic loci. *Transvection* is sometimes described as a directional interaction between two unlike mutant alleles, *trans*-inactivation is a directional interaction between two normal identical alleles, whereas pairing-dependent repression is described as a reciprocal interaction between two similar alleles (Henikoff & Comai 1998). Henikoff and Comai (1998) suggest the use of the general term *trans*-sensing effects. I would prefer to use the term of *trans*-homology interactions, which indicates some sequence similarity but does not presume inactivation or changes in transcription levels of both alleles.

The effects of pairing interactions may, in some examples be perpetuated through the germline (Hollick *et al* 1997, Matzke *et al* 2001). Circumstantial evidence suggests that this may be the cause of certain unexplained phenotypes in human diseases (Sabl & Laird 1992, Bennett *et al* 1997). It is not known how the activational history of a gene may be transmitted, but with a confirmed role in imprinting (1.4.2), methylation has been implicated as a possible candidate. As methylation is not implicated in all paramutation events and *Drosophila* lack methyltransferases, other candidates include transfer of heterochromatic proteins and/or histone deacetylases between paired loci (Henikoff & Dreesen 1989, Ronsseray *et al* 1998). Alternatively, pairing may localise the genes to a different part of the nucleus (Dernburg *et al* 1996, Sass & Henikoff 1999). Nuclear localisation can play a crucial role in determination of gene activity (Henikoff 1997, Brown 1999, Francastel *et al* 2000, Wallrath 2000), even prokaryotes appear to organise their cells into distinct domains (Kim & Wang 1999). As current experimental data indicate genomic location and availability of heterochromatic proteins are major factors determining PEV (Wakimoto & Hearn 1990, Talbert *et al* 1994, Gerasimova & Corces 1998, Gerasimova *et al* 2000, 1.7), a combination the two is therefore a strong candidate for the basic machinery governing *trans*-homology interaction events. Additionally, sequence elements within the gene may be able to convey epigenetic signals. One such candidate is the Fab-7 element of *Drosophila* (Cavalli & Paro 1998).

The evidence in 1.1 and 1.2 suggests that inclusion of *cis*-regulatory- and intronic sequences may increase the probability of transgene expression, but cannot always overcome the vulnerability of multi-copy integration (Simon & Knowles 1993).

1.3 The influence of copy number on transgene expression

Repeated sequences are vulnerable to heterochromatin formation (Dorer 1997, Dorer & Henikoff 1997) and silencing (Wolffe 1998). Heterochromatin is most commonly associated with the centromeric and telomeric chromosomal regions, rich in repeated sequences and poor in functional genes. A recent paper (Ahmad & Henikoff 2001) using a conditional system to probe for heterochromatin at a reporter gene site suggests that variegated expression is caused by heterochromatin inhibiting exposure of transcription factor binding sites. The exact nature of repeat susceptibility to heterochromatin formation, and how clusters of multiple rRNA genes avoid silencing is unknown. rRNA gene clusters may contain sequences that attract protective protein complexes, keeping the sites open (Munakata *et al* 2000) and preventing pairing and loop formation whose structural anomalies may attract the protein complexes that form heterochromatin. The heat shock protein loci in *Drosophila* are an analogous system where binding of GAGA and RNA Pol II keep the site open at all times (Cryderman *et al* 1999, Thomas & Elgin 1988). Alternatively, controlled nuclear localisation may hold the key (Akhmanova *et al* 2000).

1.3.1 Transgene repeat arrays are susceptible to silencing

Pronuclear injection of DNA is a widely used method for making transgenic animals. The resulting animals usually contain the transgene integrated as a multicopy array. Copy number dependent expression was hoped for, the data suggested otherwise (Wolffe 1997, Birchler *et al* 2000). High copy number integration is more often associated with transgene silencing or variable expression (1.6 & 1.7), thought to be the result of heterochromatin formation at the transgene array (Dorer & Henikoff 1994). Multicopy plant transgenes are also silenced (Assaad *et al* 1993, tenLohuis *et al* 1995), with timing and extent of silencing linked to copy number (Bucherna *et al* 1999) or expression levels (Chareonpornwattana *et al* 1999). Documentation of rescued expression through copy number reduction confirmed that the array itself can

be responsible for silencing or variable expression of the transgene in mammals (Garrick *et al* 1998), although subsequent papers suggest that this may be dependent on additional factors such as site of integration (Ramirez *et al* 2001) (1.6).

Fungi possess a unique sensing system capable of identifying and inactivating duplicated sequences. Two related mechanisms both involve methylation of homologous sequences. One is methylation induced premeiotically (MIP) in *Ascobulus immerses* (Meyer & Saedler 1996), where duplicated sequences are silenced through methylation. This system can also identify and silence artificially duplicated sequences (Flavell 1994). The other is repeat induced point mutation (RIP) in *Neurospora crassa,* where repeats trigger methylation and subsequent CG – AT mutation, reducing homologies (Cambareri *et al* 1990, Foss *et al* 1991, Singer *et al* 1995). Silencing of duplicated sequences takes place in haploid nuclei during the sexual cycle (Selker 1997). MIP and RIP may have evolved to combat assault on genome integrity from transposable elements and other exogenous mobile elements. Although efficient in fungi, such a mechanism does not appear to be operational in higher eukaryotes.

1.4 The influence of methylation status on transgene expression

Methylation is undeniably linked to repression of gene transcription (Boyes & Bird 1991, Bird & Wolffe 1999, Ng & Bird 1999, Fuks *et al* 2000), though the debate still rages as to its exact role. Does the epigenetic modification initiate changes in chromatin conformation, or is methylation a secondary event (Ng & Bird 1999, Fuks *et al* 2000) in a manner analogous to that of the polycomb group of regulatory proteins? (see 1.6). The evidence is conflicting and reminiscent of the "chicken & egg" scenario. CpG methylation may also be important in the evolution and stabilisation of chromosomal structure, marking duplicated sequences for diversification and protecting against homologous recombination that could disrupt the functional structure of the genome (Kricker *et al* 1992). The RIP and MIP of fungi are examples of this (1.3.1).

Methylation is implicated as a part of the genome defense system, and also controls the temporal & spatial expression of certain genes (Meyer *et al* 1993, Flavell 1994, tenLohuis *et al* 1995, Selker 1997, Bird & Wolffe 1999). Organisms that have adopted methylation as part of their gene control system are apparently incapable of managing without. Mice lacking methyltransferases never survive (Jaenisch 1997, Okano *et al* 1999). Conversely, introduction of a mammalian methyltransferase with subsequent genomic methylation is lethal in *Drosophila* (Lyko *et al* 1999). The problems of Large Animal Syndrome and other defects associated with nuclear transfer are attributed, in part, to incorrect patterns of gene methylation that may be caused by the experimental procedure (Sinclair *et al* 2000, Young *et al* 2001).

1.4.1 Methylation & Transgene Expression

Methylation is also implicated in the unstable expression of some transgenes. Methyltransferases can recognize foldback DNA, the tandem repeats of multicopy transgene arrays susceptible to methylation and thus silencing due to their propensity to form such secondary structures through pairing (Meyer & Saedler 1996, Stam *et al* 1998). This is seen even for transgenes containing a housekeeping gene promoter (Mehtali *et al* 1990), although the CAT reporter may have contributed to the reduced expression seen in high copy number lines.

Transgene expression and methylation levels are inversely correlated in both plants and animals

Despite the MOM paper (Amedeo *et al* 2000), methylation is implicated as a major cause of epigenetic silencing in plants. Disruption of the DNA methylation locus (DDM1) leads to partial reversal of silencing of a previously methylated tryptophan biosynthetic transgene in F1 homozygote *Arabidopsis* (Jeddeloh *et al* 1998). Full reversion to transcriptional activity in the 3rd generation of inbreeding is correlated with a loss of transgene methylation. Loss of methylation leads to transcriptional reactivation of a previously silenced hygromycin phosphotransferase transgene (Scheid *et al* 1998), whereas increased methylation results in progressive silencing of a kanamycin resistance transgene over 4 generations of inbreeding (Kilby *et al* 1992). Methylation silences expression in three rice lines containing transgenic

chalcone synthase genes (Davies *et al* 1997). Outcrossing results in variegated expression, even in the absence of the transgene locus, with methylation maintained at the endogenous gene locus.

Correlations between mouse transgene expression and methylation status have been reported (Palmiter *et al* 1982, Mehtali *et al* 1990, Hayashizaki *et al* 1994), with some evidence that transgene hypermethylation only occurs when the transgene is completely silenced in the line (Sutherland *et al* 1997). This supports the hypothesis that methylation is a secondary event that "locks" the site into a more permanent non-permissive state. Transgenic zebrafish are also vulnerable to methylation-induced silencing (Collas 1998), confirming that this is not a phenomenon confined to mammals in the animal kingdom. Histone acetylation may be an accessory to methylation-led silencing, with deacetylation associated with inactive loci (De Rubertis *et al* 1996, Ekwall *et al* 1997, Svensson *et al* 1998, Cavalli & Paro 1999, Kim *et al* 1999).

Hypomethylation and hypermethylation of transgene loci have been linked to active and inactive transgenes respectively (van Blokland *et al* 1997, Jeddeloh *et al* 1998, Scheid *et al* 1998). Heterogeneous methylation levels of a transgene in periferal blood T-cells have been linked to heterogeneity in expression (Zhu *et al* 1999) with oscillations in methylation associated with oscillating transcription in cell culture (Feng *et al* 2001), strongly implicating methylation as a player in control of gene expression. Even post-transcriptional silencing events have been associated with subsequent nuclear methylation (Ingelbrecht *et al* 1994, Wassenegger *et al* 1994), linking post-transcriptional and transcriptional gene silencing.

1.4.2 Methylation & Imprinting

Imprinting is a well-documented case of using methylation to mark epigenetically and inactivate loci dependent on parent of origin. The best-studied imprinted locus is the Igf2r and H19 locus (Reeve 1996). Parent of origin has been shown to influence transgene methylation state (Allen *et al* 1990), possibly in conjunction with the transgene sequence and its integration site (Hayashizaki *et al* 1994). Genomic

imprinting in *Drosophila* (Bishop & Jackson 1996, Lloyd *et al* 1999) indicates that methylation is not the only epigenetic mark capable of meiotic propagation.

1.4.3 Methylation & Background

Different genetic backgrounds can display significant qualitative and quantitative differences in proteins affecting gene expression. Combined with mode of germline transmission, background can influence transgene expression potential (Schweizer *et al* 1998, Morgan *et al* 1999). Single modifier loci can convey strain-specific methylation of transgenes and surrounding sequences in *cis* (Engler *et al* 1998), different backgrounds increasing or decreasing methylation (Engler *et al* 1991). This is not exclusive to mammals, with reports of zebrafish displaying genotype specific modification to methylation patterns of a transgene, with sex and environment contributing to the phenotypic outcome (Martin & McGowan 1995).

1.4.4 Other candidate epigenetic marks involved in gene silencing

The *Drosophila* genome is not methylated, but is capable of being imprinted. Other epigenetic modifications therefore exist that are capable of both mitotic and meiotic propagation. Histone acetylation can maintain gene silencing through a repressed chromatin state (Ekwall *et al* 1997). Both methylation and acetylation may ultimately work through recruitment of heterochromatic protein components to the site. A recent paper reports discovery of a histone methyltransferase and the stable methylation of Lys 9 at the N termini of histones H3 & H4, a modification that creates a binding site for heterochromatin-associated proteins (Jenuwein 2001). This epigenetic mark could be used to transmit chromatin states.

Another candidate is the MOM gene product. Silenced transgenes have been reactivated through disrupting MOM (mutation in a "morpheus" molecule) expression in *Arabidopsis* (Amedeo *et al* 2000). MOM is required for transcriptional gene silencing, has no effect on methylation, but shares similarities with the ATPase region of the SWI2/SNF2 complex, suggesting that MOM can influence chromatin conformation. Conservation of these proteins argues for a functional link between mammalian, *Drosophila*, yeast and plant silencing phenomena.

1.5 The influence of age, germline transmission and genetic background on both endogenous and transgene expression

Silencing phenomena described in this section are found in both plants and animals, despite the many differences between them. Could basic gene regulation mechanisms remain conserved or are the observed similarities the result of convergent evolution?

1.5.1 Age–Related Silencing

Ageing may trigger methylation changes that can be both gene and tissue-specific. Increased age appears to be a major contributing factor in hypermethylation and silencing of certain genes in cancers (Ahuja *et al* 1998, Toyota & Issa 2000). Conversely, upregulation of gene expression is also linked to ageing (Villeponteau 1997). This article suggests that heterochromatin loss occurs with increased age and with it the loss of heterochromatin–induced silencing. Yeast systems corroborate Villeponteau's theory, documenting age-dependent loss of transcriptional silencing at heterochromatic sites (Jazwinski 1999), and suggest a possible role for histone deacetylases in ageing control (Kim *et al* 1999). This silencing is not confined to endogenous genes. Cohn *et al* (1991) report progressive silencing of a mouse colon-specific transgene, each crypt functioning as a single expression unit. Silencing has also been reported for globin transgenes in mice, the proportion of expressing cells decreasing with age (Robertson *et al* 1996, Sutherland *et al* 1997). In plants, silencing of an S-Adenosyl-L-methionine synthetase transgene in tobacco occurs first in the veins of leaves and later in the whole leaf in new leaves on plants already exhibiting silencing (Boerjan *et al* 1994). In an *in vitro* system, retroviral vectors used to target haematopoietic stem cells and improve oxygen production are silenced over time. Evidence suggests that this silencing is also progressive, with each silencing event apparently irreversible (Zentillin *et al* 2000). Methylation is a suggested silencing event, either as the primary silencer or a secondary epigenetic "lock" (Antequera *et al* 1990, Dieguez *et al* 1998, Sutherland *et al* 1997).

The common theme in the articles above is deregulation of gene expression with increasing age. Perhaps this is a reflection of an organism's finite lifespan. One of the signs of ageing could be loss of regulatory control, leading to a breakdown of cell

function followed by death. Human cancers are more prevalent in older individuals, and could be interpreted as the manifestation of a general degeneration of cellular function. In my study, the BLG/7 mice (1.8) are sampled in their prime. Serial lactations indicate that BLG protein expression does not differ significantly between individual lactations (Dobie *et al* 1996), suggesting that the observed variegation is not age-related. BLG/7 transgene expression is lactation specific, whereas other studies use transgenes containing constitutively active promoters, at least for their cell type (Cohn *et al* 1991, Robertson *et al* 1996, Sutherland *et al* 1997). The limited window of expression differentiates our mice from the others, and may render our transgene immune to time or age-related repression effects.

1.5.2 Meiotic Inheritance of Epigenetic Silencing

Maintenance of epigenetic marks through the germline defies the principle that DNA alone conveys the blueprint of life to the next generation. Research outlined below suggests inheritance can be infinitely more complex. For ease of presentation, the next section has been divided into units that broadly cover the main species used in transgenic research. Such a discussion will eventually focus on plants due to the large amount of data for this phylum.

Drosophila

The Fab–7 trithorax *cis*-regulatory element is a recent *Drosophila* example where the epigenetic expression state can be inherited. Meiotic inheritance of the active state may reflect inheritance of transcriptionally competent chromatin at the gene locus, with H4 hyperacetylation a putative epigenetic tag (Cavalli & Paro 1999). A different example involves the I-element transposon, where complete suppression takes up to 10 generations (Jensen *et al* 1999). This phenomenon is reminiscent of PTGS in plants: transcription is needed to induce silencing; success is proportionally linked to I-element sequence transcript length and dosage. Silencing is cumulative over the generations and transferred through the maternal germline, even non–transgenic offspring display silencing if the maternal parent is a silenced transgene carrier. This suggests the maintenance/transfer of an epigenetic mark that is established very early in embryonic development, before the degradation of maternal mRNA. Paternal

transmission erases genomic silencing, strengthening speculation that maternal mRNA could be the catalyst for silencing of I-element associated sequences.

Fish

Transgene expression can remain stable in zebrafish (Caldovic *et al* 1999). Others have been less successful, silencing of transgenic green fluorescent protein seen by F2 in 20% of zebrafish lines (Amsterdam *et al* 1995).

Plants

Even non-transgenic plants exhibit silencing phenomena; some may be caused by mixed genotypes, with a recent description of the domination of one parental set of rRNA over the other in a hybrid plant after two generations (Chen *et al* 1998). The direction of dominance can be switched in this example, suggesting that the balance of parental factors probably determines the direction of dominance rather than the rRNA. The progressive silencing of a kanamycin resistance gene over 4 generations of inbreeding may be linked to a parental genotype (Kilby *et al* 1992).

A single gene is able to generate several phenotypes. An investigation into the segregation of drug resistance phenotypes through 7 generations in *Arabidopsis* reports resistant parents consistently segregating 3 offspring phenotypes: sensitive, partially resistant or resistant (Assaad *et al* 1993). Silencing could be progressive over the generations, coinciding with increased methylation, but was also reversible. The most varied example is reported for a transgenic construct containing a kanamycin resistance gene (Schmulling & Rohrig 1995). The locus generated antisense transcripts that silenced expression with progressive severity over the generations, but never exhibits complete penetrance. Two different types of silencing were observed. The first was early silencing with methylation and meiotic resetting. The second was a slow stochastic onset of silencing without methylation, exhibiting stable meiotic transmission and with reactivation capable of being sustained over several generations. Some plants exhibited mosaic expression patterns with different intensities reflecting different levels of transgene expression. Introduction of a second transgene, sharing homologies with the kanamycin gene, induced increased

silencing over the generations. Re-establishment of kanamycin expression required separation of the two transgenes for 3 generations. Genetic background could influence silencing, which varied from 24 – 46% in the F2 backcrosses.

That a single transgene locus can cause multiple phenotypes argues for the presence of complex silencing mechanisms in plants. Some of these mechanisms may be unique to plants; perhaps a consequence of the plasticity that can support polyploidy, aneuploidy and hybrid plants in a manner not seen in animals.

Mammals

Plants segregate their germline later in development than animals. This increases the probability that epigenetic mutations are passed on to future generations. This may provide plants with an evolutionary advantage, but no such advantages are apparent for animals (John & Surani 1999).

Most papers on epigenetic inheritance in mammals describe imprinted loci. Parent of origin determines expression, with methylation the apparent epigenetic mark involved. Work has suggested that there are allelic variations in methylation between homologous human chromosomes, which are both tissue specific and capable of germline transmission (Silva & White 1988). Could this indicate that epigenetic inheritance and differentiation between homologous sites is widespread? In humans, the inheritance of epigenetic modification is well known for imprinted loci, as are the conditions that occur when the imprinting process malfunctions. There is now intense speculation that inheritance of epigenetically modified, but otherwise fully functional gene loci, may explain some inexplicable occurrences of disease in genotypically normal individuals. This is the hypothesis put forward to account for cases of Huntingdon disease (Sabl & Laird 1992).

Coat colour is an easy phenotype to study. An early report on putative epigenetic inheritance involves the star coat-marking gene in foxes (Belyaev *et al* 1981). Transmission of the active gene was linked to selection for domestic behaviour in females. In mice, the inheritance of epigenetic changes linked to the agouti coat

colour has been extensively studied. Offspring phenotype is linked not only to the maternal phenotype, but also to that of the maternal grandmother. It is suspected that epigenetic marks are incompletely erased during germline formation, giving rise to a variegated expression pattern resulting in coat colours and patterns covering the entire colour spectrum (Morgan *et al* 1999).

Nuclear transfer has enjoyed success in large animals, but side effects are associated with the technique (Young *et al* 2001). The same phenotypes are now being seen in mice generated by nuclear transfer, with changed expression profiles for a range of genes that are transferred through the germline to the next generation. One line exhibits progressive methylation with increasingly severe phenotypes with each generation. The epigenetic inheritance lasts at least two generations, with studies continuing (Roemer *et al* 1997). These reports raise serious questions over the safety of future cloning ventures, and provide the definite argument against human cloning.

1.5.3 The influence of genotype on phenotype

In plants, genotype influencing gene expression is suggested for hermaphrodite mutants containing mutations in the gynoecium-suppressing function region in white campion (Lardon *et al* 1999). Additionally, studies into hybrids and their inbred parental strains have revealed both qualitative and quantitative differences in gene expression (Sun *et al* 1999).

More information is available for mammals, where genetic background has long been suspected to play a part in influencing expression levels. Dominant strain specific modifiers are documented (McGowan *et al* 1989, Elliot *et al* 1995), with over-expression of a single heterochromatic protein component capable of modifying the extent of transgene variegation in mice (Festenstein *et al* 1999). Genetic background can influence methylation and expression of a variegating transgene (Koetsier *et al* 1996), with increased levels of methylation and silencing associated with a C57BL/6 background. This background was also associated with *de novo* methylation of a hepatitis B virus surface antigen (HBsAg) transgene (Schweizer *et al* 1998). HBsAg expression levels differed depending on the strain, backcrossing

enhancing the effects. Segregation of strain-specific factors meant that crosses produced offspring exhibiting both up- and downregulation of expression, due to the cooperative effects of different combinations of modifier proteins. Intriguingly, background effects were only observed when the transgene was transmitted through the male germline. Silent transgenes can also be reactivated by crossing into a background conducive to expression (Sutherland *et al* 2000), with a C57BL/6 background increasing reactivation of the transgene by decreasing methylation. The authors could select for high expressors, but the overall trend was for progessive silencing with continued variegation. Even high expressors generated low- and non-expressing progeny.

The transgene behaviour reported above suggest that strain specific factors do not act in isolation. Integration site with associated site-specific factors, as well as residual effects derived from the background on which the line originated, probably combine with strain-specific factors to determine the transgene expression profile.

1.6 The influence of integration site on transgene expression

Injected transgene constructs are supposedly integrated at random into the genome, although there is suggestive evidence that integration may be less random than previously thought. Some genomic sites appear more "accessible" than others (Rijkers *et al* 1994) with 5' and 3' sequences capable of targeted insertions (McFarlane & Wilson 1996). Homologies to endogenous sequences can drive targeted insertions; a property exploited in embryonic stem (ES) cell work to generate animals with single copy transgene integrations at defined genomic loci (Bronson *et al* 1996). The limitation is a lack of available ES cells for species other than mouse; therefore, pronuclear injection is still extensively used despite the associated problems of multilocus multicopy random integration.

Random integration can render a transgene vulnerable to what is termed position effects. These effects are either stable, where expression levels are reduced but is seen in all relevant cells; or variegated, where expression levels are maintained but only a proportion of cells are active. "Defensive" structures, such as LCRs, can

increase the probability of transgene expression (1.1.2), but cannot protect against strong position effects (Guy *et al* 1996). Sex chromosomes integration will lead to variable expression due to dosage compensation mechanisms such as X-inactivation (Haller & Woodruff 2000), as will integration into an imprinted region (Duvillie *et al* 1998). Strong endogenous enhancers or silencers close to the integration site may also induce aberrant transcription (Milot *et al* 1996). The effects of neighbouring sequences are believed to cause many human diseases, such as cancers, where chromosomal translocations have been linked to the disease phenotype (Brock 1993).

A widely studied phenomenon of site-dependent control of gene expression is Position Effect Variegation (PEV), where translocation or insertion of a transgene array inside or close to endogenous heterochromatin leads to mosaic transgene expression at the cellular level due to variable spread of heterochromatin.

1.7 Position Effect Variegation

The first description of Position Effect Variegation (PEV) was published over 70 years ago (Muller 1930). This paper and the follow up report (Muller 1932) describe a mutation-induced translocation of the *white* eye colour gene of *Drosophila* from its normal euchromatic site to a position close to centromeric heterochromatin. This juxtaposition led to patchy *white* gene expression where the physical manifestation is a mottled eye colour, combining patches of the usual red eye colour with patches of white (Figure 1.1). Since then, variegating gene expression has been found in many other organisms. Variegation has been reported in plants, insects, yeast, fungi, fish and mammals (Muller 1930, McGowan *et al* 1989, Stuart *et al* 1990, Allshire *et al* 1994, Tartof 1994, Cogoni *et al* 1996, Dobie *et al* 1996, Festenstein *et al* 1996, Meyer & Saedler 1996, Robertson *et al* 1996, Collas *et al* 1999, Giraldo *et al* 1999, Benedict *et al* 2000, Sutherland *et al* 2000, Matzke *et al* 2001).

PEV has two distinct physical manifestations: the mosaic pattern of gene expression visible at the cellular level and the cytological observation of heterochromatin spreading into the variegating gene locus. This increased heterochromatinisation is widely accepted as the cause of silencing (Epstein 1992, Orlando & Paro 1995,

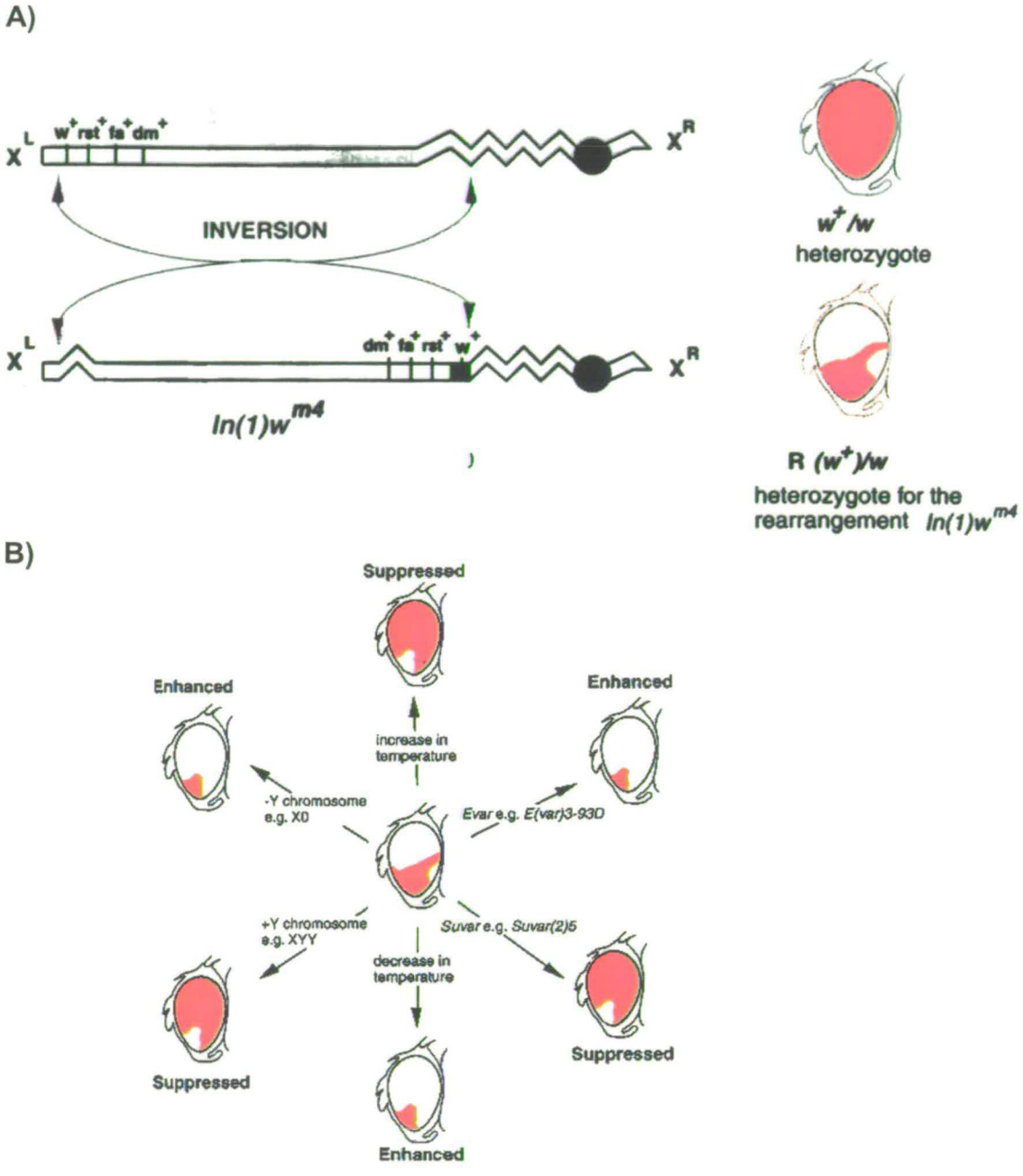


Figure 1.1 Position Effect Variegation (PEV) in *Drosophila*
Reproduced with the kind permission of P.B. Singh (from Singh 1994)

A) Schematic illustration of chromosome inversion leading to variegation of the white (w+) gene due to new position close to constitutive heterochromatin.

B) PEV can be affected by a number of factors that are able to modify determination and degree of variegation within a given tissue. These can be environmental conditions, protein modifiers or other factors that determine availability of heterochromatin.

McCall & Bender 1996). Mosaic expression is explained as the manifestation of variable heterochromatin spreading, with discrete patches of cells sharing the same expression profile due to clonal expansion after "fixation" of the expression state (Paro & Hogness 1991). A large number of variegating genes have been studied and factors that can modify their expression have been identified. Many modifiers are chromatin components (Singh 1994), strengthening the theory that variable spread of heterochromatin causes variegation. Homologues of these modifier genes have been identified in several systems, and experimental evidence suggests that many are probably capable of functional substitution between systems (Stankunas *et al* 1998, Aagaard *et al* 1999, Festenstein *et al* 1999, McMorrow *et al* 2000). In order to understand PEV, it is necessary to have an understanding of heterochromatin.

1.7.1 Heterochromatin function and composition

Heterochromatin is a term used to describe a densely packed higher order chromatin structure. Morphologically, heterochromatin describes deeply staining compact bodies of nuclear material that remain condensed and visible throughout the cell cycle (Heitz 1928, from Weiler & Wakimoto 1995) and are the last to undergo replication (Holmquist 1987). Heterochromatin is a heterogeneous component of higher order chromatin structure (Lloyd *et al* 1997, Wallrath 2000), with some shared features. The tails of histones H3 & H4 are hypoacetylated compared to their euchromatic counterparts (Braunstein *et al* 1996), and nucleosome spacing patterns are more regular (Wallrath & Elgin 1995). Hypermethylation is associated with heterochromatin (Bongiorni *et al* 1999), and there is speculation that hypermethylation may act to stabilise the chromatin state, either by itself or through recruitment of chromatin remodelling complexes (Henikoff 2000).

Although structurally similar, centromeric and telomeric heterochromatin protein composition is different (Wallrath 2000) and thus can react differently to external influences. Common to both is the presence of repeated sequences, although the repeats themselves differ (Cryderman *et al* 1999_B). Introduced repeats can induce heterochromatin formation, and may in part explain the propensity of transgenes to become silenced (1.3.1). That constitutive heterochromatin remains confined to a

defined area of the genome in the absence of structural alteration argues for the presence of structural boundary elements. Such putative boundary elements have been identified and their role experimentally verified (1.7.6). Facultative heterochromatin is formed at discrete genomic sites and may not remain cytologically visible throughout the cell cycle, and include sites occupied by the Polycomb family of regulatory proteins (1.7.3). Heterochromatin structural components remain conserved, with heterochromatic protein 1 (Singh *et al* 1991, Cowell & Austin 1997, Wang *et al* 2000) currently the best characterised.

How repeat structures initiate heterochromatin formation is unknown, but a mechanism capable of identifying repeated structures could exist. Transgenic experiments may have uncovered a hitherto unknown branch of the genome defence system against exogenous attack. Whilst the immune system protects the host body from micro-organisms, a genomic defence system defends the genome at both the somatic and germline level against retroviral, transposon and other assaults on host genome integrity. Heterochromatinisation would prevent both sequence activation and pairing leading to *trans*-activation. Methylation epigenetically marks the sequences, maintaining the heterochromatic state of the locus and aid in c → t mutations (Henikoff & Matzke 1997, Yoder *et al* 1997, Dimitri & Junakovic 1999). That a large proportion of the human genome is comprised of retrotransposons and repeat elements provides supportive evidence for such a functional role for heterochromatin in genome defence (Smit 1996, Henikoff *et al* 1997).

The use of heterochromatin in silencing endogenous genes reveals a role in controlling gene expression as well as protecting genome integrity (Brown *et al* 1997, Brown *et al* 1999). Terminal differentiation of many cell types is associated with progressive genomic heterochromatinisation (Lundgren *et al* 2000), suggesting a role for heterochromatin in determining and maintaining cellular fate and identity. The discovery of cross-species conservation of some heterochromatin proteins, or their functional domains, has encouraged speculation that these proteins may have an even wider function. Yeast silencing proteins appear to have conserved homologues from bacteria to humans (Sherman & Pillus 1997). Swi6 is a Yeast protein

homologous to HP1 (Wang *et al* 2000) involved in imprinting of the mating-type locus and contributing to the cellular memory responsible for maintenance of the silenced state (Grewall 2000). Could mammalian HP1 have a similar functional role?

1.7.2 PEV and heterochromatin

The discovery of PEV and subsequent investigations has ensured that heterochromatin is no longer considered merely an "inert" packaging material. Heterochromatin is now known to play a role in chromosome pairing and centromere function during cell division (Dernburg *et al* 1996, Renauld & Gasser 1997). This ties with earlier observations that heterochromatic regions are late replicating (Holmquist 1987). The pairing capabilities of heterochromatin may also explain the rare cases of dominant PEV, with pairing driving heterochromatinisation of both loci and/or physical removal to a heterochromatic area of the nucleus (Henikoff & Dreesen 1989, Ronserray *et al* 1998, Sass & Henikoff 1999). Functional genes do reside within constitutive heterochromatin (Sinclair *et al* 2000), with this chromatin environment needed for their correct expression (Lu *et al* 2000, Wakimoto & Hearn 1990). Translocations into euchromatic parts of the genome leads to disruption of gene expression in a manner not dissimilar to the position effect variegation displayed by euchromatic genes in a heterochromatic environment (Weiler & Wakimoto 1995). Heterochromatic genes subjected to PEV may also be sensitive to pairing; in this case pairing can rescue expression (Weiler & Wakimoto 1995).

1.7.3 Links between PEV and control of endogenous genes through modification of chromatin conformation

Development from fertilisation through to adulthood is a strictly controlled process of gene activation and silencing. The pattern of gene expression defines a cell and determines its fate. What controls the patterns of expression, changing cells from totipotent to terminally differentiated? The field of developmental biology has some pieces of the puzzle, but the picture is by no means complete.

The pattern of gene expression is determined early in development, but the factors that lay down the initial patterning are not responsible for the long-term maintenance of these patterns; that is the job of a secondary set of controlling factors (Mahmoudi & Verrijzer 2001). The best-studied group of secondary developmental regulators are the Polycomb group (PcG) and trithorax group (trxG) families of regulatory proteins. They control homeotic gene expression in *Drosophila*. Polycomb proteins act as repressors whilst trithorax proteins are activators (Epstein 1992, Kennison 1995). The trxG and PcG complexes serve as a cellular memory of earlier regulatory factor binding (Orlando & Paro 1995). It is believed that many of the binding sequences involved in homeotic gene control may be bound by either PcG or trxG complexes, with the Fab-7 element a prime example (Cavalli & Paro 1998). The site is capable of binding both Pc and GAGA factor (a trxG protein), enabling switching of expression states as well as conveying a memory of past activation that can survive female germline meiosis. A functional link has been established between PEV and homeotic gene expression control through the identification of some of the enhancers and suppressors of variegation as members of either the trx or PcG families. Other modifiers of PEV contain structural similarities to trx & PcG functional domains (Hazelrigg *et al* 1984, Paro & Hogness 1991, Bhadra & Birchler 1996, Carrington & Jones 1996, Larsson *et al* 1996, Platero *et al* 1996, Bhadra *et al* $1997_{A \& B}$, Laible *et al* 1997, Meyer 1999).

Although many PcG proteins do not modulate PEV, their binding creates repressive chromatin complexes that aggregate in subnuclear domains coincident with neither heterochromatin nor transcriptional activety (Buchenau *et al* 1998). Pc-mediated repression blocks pol II transcription, but allows T7 RNA polymerase access (McCall & Bender 1996). This links with heterochromatin-mediated silencing where *trans*-acting factors are blocked from *cis*-acting regulatory sequences (Cryderman *et al* $1999_{A \& B}$), suggesting that PEV and PcG mediated repression involves similar higher order chromosomal structuring (Boivin & Dura 1998).

1.7.4 A diverse range of modifiers of PEV have been identified

PEV was first described in *Drosophila,* with much information arising from experiments in this system, especially concerning factors that can modify PEV. With homologues identified in other animal and plant systems, PEV could be a powerful tool for unravelling the complex process of gene expression control in higher eukaryotes. Modifiers are now counted in their hundreds (Bhadra *et al* 1997_A), and whilst many are components of heterochromatin, increasing numbers are not. Modifiers can be grouped into different classes, which are outlined below:

- Components of heterochromatin; e.g.
    - HP1 (Clark & Elgin 1992, Eissenberg *et al* 1992)
    - Su(var)3-7 (Reuter et al 1990, Cleard *et al* 1997)
    - Enhancer of *zeste* (Carrington & Jones 1996)
- Modifiers of transcription of heterochromatin components; e.g.
    - Lom (Bhadra *et al* 1997_A)
    - Dmrnahel (Eberl *et al* 1997)
- Transcriptional regulators; e.g.
    - Regena (Frolov *et al* 1998)
    - Zeste (Seum *et al* 1996, Gubb *et al* 1997)
    - E(var)3-trl (=GAGA factor) (Seum *et al* 1996, Wallrath 2000)
    - Histone deacetylases, HDAC 1 (Mottus *et al* 2000)
    - EBF (Lundgren *et al* 2000)
- Enzymatic protein modifiers; e.g.
    - Ubiquitin (Henchoz *et al* 1996)
    - PP1 phosphatase (Seum *et al* 1996)
- Cell cycle proteins or developmentally regulated proteins (Lu *et al* 1996); e.g.
    - E2F
    - PCNA
    - Su(var)3-9

The majority of modifiers act in a dosage dependent fashion, with a few exhibiting opposing effects when present in limiting or excessive dosages (Reuter *et al* 1990,

Eissenberg *et al* 1992, Henikoff 1996, Cleard *et al* 1997, Gubb *et al* 1997, Sass & Henikoff 1998, Festenstein *et al* 1999, McMorrow *et al* 2000, Lundgren *et al* 2000). External factors may also affect variegation, examples being temperature (during critical windows of development) and proximity to MARs/SARs (Girard *et al* 1998).

1.7.5 Theoretical explanations proposed to explain PEV

Three main theories have been advanced to account for PEV of transgenes:

1. Somatic mutation or elimination of the transgene causes the variegated expression profile (Karpen & Spradling 1990)

2. Altered chromatin packaging causes PEV through variable spread of the silencing heterochromatic chromatin (Ananiev & Gvozdev 1974, Tartof *et al* 1984, Wallrath & Elgin 1995)

3. Altered nuclear compartmentalisation of the transgene due to its chromosomal position causes PEV, due to limited supplies of factors needed for expression (Wakimoto & Hearn 1990, Talbert *et al* 1994, Gerasimova & Corces 1998)

Theory 1 is disputed by several groups (Ahmad & Golic 1996, Wallrath *et al* 1996). Experiments that refute this theory include examples where a variegating or silent transgene construct has been excised as expression competent ectopic circular DNA (Ahmad & Golic 1996), or re-integrated into the genome and successfully expressed (Al-Shawi *et al* 1990). Although it cannot be wholly ruled out that a rare example of variegation is caused by somatic mutation/elimination (Scrable & Stambrook 1999), I believe the majority of examples have a different underlying cause.

Cytogenetic observations of heterochromatin spread, different nuclease accessibility of active and silent genes, as well as characterisation of modifiers of PEV encompassing heterochromatic proteins and modifiers supports the validity of theory 2 (Clark & Elgin 1992, Belyaeva *et al* 1993, Wallrath & Elgin 1995, Garrick *et al*

1996, Gdula *et al* 1996, Henchoz *et al* 1996, Seum *et al* 1996, Bhaddra *et al* 1997_A, Cleard *et al* 1997, Frolov *et al* 1998, Festenstein *et al* 1999, McMorrow *et al* 2000).

Evidence for nuclear compartmentalisation, with delineation of transcriptionally active and repressed areas, support the third theory of PEV (Brown *et al* 1999, Wallrath 2000). Even different chromosomes appear to occupy distinct nuclear positions in mammalian interphase cells (Haaf & Schmid 1991). In addition, numerous papers report that nuclear localisation can play a crucial role in determination of gene activity (Henikoff 1997, Brown 1999, Francastel *et al* 2000, Wallrath 2000). Even prokaryotes appear to organise their cells into distinct domains (Kim & Wang 1999). There is thus plenty of evidence to support theory 3.

Theories 2 & 3 are not mutually exclusive as there is experimental evidence supporting both (Cryderman *et al* 1999_B). Two connected phenomena link both chromatin structure and nuclear localisation with control of gene expression: *transvection* and paramutation (1.2.3). Transfer of heterochromatic components and/or histone deacetylases during pairing is a prime candidate for the mechanism operating in these phenomena (Henikoff & Dreesen 1989, Ronserray *et al* 1998). An alternative mechanism is localisation of the genes to a different nuclear compartment as a result of the pairing interaction (Dernburg *et al* 1996, Sass & Henikoff 1999).

Changes in chromatin structure can influence nuclear localisation of a gene in interphase cells, with circumstantial evidence linking paramutation and repressive chromatin structures. Efficiency of paramutable allele silencing by the paramutagenic allele is apparently influenced by chromosomal location (Meyer 1999). Heterochromatic gene variegation can also be eliminated/reduced through pairing with the untranslocated homolog, enabling correct nuclear localisation to a heterochromatic compartment (Weiler & Wakimoto 1995). Similarly, insertion of heterochromatin into a eukaryotic gene induces PEV and association of both the mutated and normal gene with heterochromatin through *transvection* (Csink & Henikoff 1996). The best-known example of *transvection* is the $brown^D$ mutation in *Drosophila* (Dreesen *et al* 1991). Studies suggest that associations with

heterochromatin and nuclear localisation change during development, both *brown*^{D} and wildtype associating with centromeric heterochromatin due to homology and chance pairing during interphase (Csink & Henikoff 1996, Dernberg *et al* 1996). Heterochromatin involvement in chromosomal pairing and segregation during mitosis confirms the pairing properties of heterochromatin (Kellum & Alberts 1995, Elgin 1996, Csink & Henikoff 1996) and may explain how cell cycle can influence PEV.

Current experimental data indicate that two of the major factors determining PEV are genomic location of the transgene and heterochromatic protein availability (Wakimoto & Hearn 1990, Talbert *et al* 1994, Gerasimova & Corces 1998, Gerasimova *et al* 2000). Dissociation of silencing and nuclear localisation (Sass & Henikoff 1999, Francastel *et al* 1999, Lundgren *et al* 2000) suggests that control of gene expression is much more sophisticated than the current theories propose, with nuclear localisation only one event in a multi-step process involving many factors. Several layers of complexity are involved in gene silencing, with some genes controlled by "simple" mechanisms such as formation of heterochromatin at the site and/or nuclear localisation to a silencing compartment. Other genes may require more complex regulation through a pathway containing several checkpoints. It is speculated that nuclear repositioning of a gene to centromeric heterochromatin may be a feature of heritable gene silencing (Brown *et al* 1999), possibly controlling terminal differentiation?

An alternative hypothesis suggests that a silent state is the default state, with genes destined for activation epigenetically marked during metaphase (Michelotti *et al* 1997), with DNA transcriptional start sites conformationally distorted in genes due to be expressed. Maintenance of this distortion, or use of it in epigenetic tagging may be through protein binding to ssDNA. Transcription factor dosage can influence expression and nuclear localisation of a variegating gene (Lundgren *et al* 2000), suggesting that promoter potentiation could influence localisation of the gene to a nuclear position where activation is possible. Alternatively, genes destined for silencing may be marked by a conformational distortion during metaphase. The

identification of an ssDNA binding protein associated with pericentromeric heterochromatin (Cortes *et al* 1999) adds weight to this hypothesis.

1.7.6 Regulation of PEV

I cannot envision any reason why several mechanisms cannot operate in the control of gene expression. Shared components and overlapping pathways could account for the bewildering and occasionally conflicting reports on the causes and influences of PEV. Two papers have investigated the effects of introducing two different variegating loci into the same genome (Belyaeva *et al* 1993, Lloyd *et al* 1997). Both used *Drosophila* as the experimental system. If the two loci are subjected to the same constraints, competition for limiting factors should disrupt the usual expression profile of one or both genes. Independent regulation of the two loci would lead to the continuation of the normal expression profiles. Both outcomes were observed, depending on the combinations of loci used, confirming that PEV can be regulated by several different pathways that may or may not overlap. HP1 is a prime candidate protein that may be involved in all PEV pathways as it is a component of both telomeric and centromeric heterochromatin and able to modify PEV at both locations (Festenstein *et al* 1999, McMorrow *et al* 2000). Telomeres and centromeres also support non-overlapping pathways operating in PEV (Talbert *et al* 1994, Wallrath & Elgin 1995, Cryderman *et al* 1999_B). It is therefore difficult to decide whether to treat centromeric and telomeric variegation as separate phenomena, or focus on shared traits in the hope that these are the keys to understanding PEV.

1.7.7 Mechanistic Models proposed to explain heterochromatic spread in PEV

There is currently much debate concerning the precise mechanism(s) involved in the manifestation of PEV, and if it is possible to predict the outcome of harbouring two variegating loci in the same system. Whilst PEV is frequently described as the variable spread of heterochromatin into a normally euchromatic translocated close to or inside a heterochromatic site (Epstein 1992, Orlando & Paro 1995, McCall & Bender 1996), it is not strictly accurate, as PEV is also observed when heterochromatic genes are translocated to a euchromatic site (Schultz & Dobzansky 1934, Wakimoto & Hearn 1990). Loss of heterochromatin is probably responsible for

the instability in this case. As this latter example of PEV is relatively rare, all the mechanisms below discuss the former example of PEV.

Although great progress is being made unravelling the protein composition of heterochromatin, the manner in which spreading of heterochromatin from a euchromatic/heterochromatic breakpoint junction occurs remains open to speculation. Why are some sequences more susceptible than others to position effects? What determines how far heterochromatin can spread into adjacent euchromatin? How are the euchromatic and heterochromatic boundaries defined and maintained in normal cells? Two similar models have been proposed to explain and predict the spread of PEV into transgenes inserted into heterochromatic areas (Balasov & Makunin 1996, Clarke *et al* 1996). Both models assume that there are "binding/initiation" sites for heterochromatic complexes distributed across the whole genome, with proximity of binding sites and availability of heterochromatin components determining the spread of heterochromatin and thus silencing.

Balasov & Makunin's Model

Balasov & Makunin (1996) have proposed a theory that can account for all the questions regarding susceptibility to position effects, heterochromatin spread and delineation. The theory, as I interpret it, is an expansion upon the "Spreading Effect Rule" proposed by Spofford (1976). Spofford suggests that heterochromatin spreads progressively into the euchromatic site at the break junction (a presumed boundary element delineating heterochromatic and euchromatic areas having been removed by the translocation), with the likelihood of variegating expression decreasing with increased distance from the breakpoint junction. It has been noted that heterochromatin can spread up to 1000 Kb from a breakpoint, in either a continuous or discontinuous pattern (Belyaeva & Zhimulev 1991). Local concentrations of heterochromatin components would determine the length of spread, and this would vary between cells, giving rise to the characteristic "speckled" expression pattern of PEV. This theory does not explain discontinuous spread of heterochromatin, which is where Balasov & Makunin (1996) expand. Their PEV model is based on the statistical distribution of compactisator protein (CP) molecules around presumed compactisation initiation centres (CICs). The CP molecules and CICs probably vary

in composition, giving rise to distinct combinatorial subsets conferring both tissue- and stage-specificity to heterochromatin formation. There are four basic assumptions made in this model:

1. CICs are found in both euchromatin and heterochromatin

2. CP molecules are heterogeneous and bind both to CICs and to each other, forming multimeric complexes.

3. There are many different CICs, with differint affinities for the different CP molecules. CIC distributions vary across the genome and account for the differences in heterochromatin protein composition between different heterochromatic domains.

4. Levels of CP molecule binding and distances between the CICs determine the level of cooperation between sites and also determine if the threshold level for heterochromatin formation is reached.

At a euchromatic-heterochromatic breakpoint, CPs spread into the euchromatic area, as there is no longer a boundary element present to prevent it (see Sun & Elgin 1999 for a review on boundary elements, and van der Vlag *et al* 2000 for experimental testing of boundary element strength). The relative concentrations of CICs and CPs, in conjunction with the distances between the CICs, will act to determine where and when heterochromatinisation takes place. Some parts of the euchromatic locus may not have CICs close enough together for formation of heterochromatin, but if there are CIC sites located distally which are closer together and can be reached by the CPs, then they could spread there and give rise to discontinuous compaction.

Experimental results, mainly in *Drosophila*, bolster the validity of this theory. Most convincing are the Polycomb response elements (PREs). Initially identified as part of the Bithorax–C regulatory domains (Simon *et al* 1990), PREs are targets for Polycomb group (PcG) proteins (Simon *et al* 1993). The elements appear diverse in

composition, exhibiting different preferences for PcG complexes, and can easily be equated with CICs alongside PcG proteins as the CPs. PREs act to maintain correct expression patterns via the action of bound PcG and possibly also trithorax group (trx) complexes (Busturia & Bienz 1993, Chan *et al* 1994, Chiang *et al* 1995, Gindhart & Kaufman 1995, Muller & Bienz 1991, Simon *et al* 1993), with several of the sequences identified as putative PRE's (or associated with them) apparently combining both a silencer and a boundary function. The Fab-7 element (Hagstrom *et al* 1996 + 1997) and Mcp element (Muller *et al* 1999) are two examples. In addition, they also contain properties that can convey an active chromatin configuration upon associated sequences (Cavalli & Paro 1999, Muller *et al* 1999). This indicates strongly that trx complexes can bind the recognition site as well. Yeast experiments reinforce the universality of the theory, Sherman & Pillus (1997) suggesting that competition between loci and availability of protein complexes may explain variable silencing at different genomic loci.

The Markov Model

Clarke *et al* (1996) proposed a mathematical model based on conditional probabilities with which to calculate the assembly of heterochromatin into a euchromatic region at a given heterochromatic-euchromatic breakpoint. The model is known as: "A Markov Model for the Assembly of Heterochromatic Regions in Position Effect Variegation" and appears to share the same basic assumptions as Balasov & Makunin. An attempt is made to create a mathematical model upon which to predict the behaviour of one or more variegating loci in the same system. The design is for only one allele of a pair to be variegating, two variegating loci are presumed to be two separate genes at different loci. The model's central tenet is also that the relative concentrations of CICs and CPs together determine the outcome for a given locus' expression state.

1.7.8 PEV and X-inactivation: a case of shared properties and pathways?

A PEV model involving sites responsible for binding/initiation of heterochromatin formation is reminiscent of the spread of heterochromatin during X-inactivation. Both PEV and X-inactivation share many similarities: both determine the activation

state of a defined site. This determination is reputed to take place early in development and maintained in a clonal fashion (Lu *et al* 1998), leading to the speculation that both processes may share some structural if not mechanistic components. Although the mosaic pattern of expression seen for genes subject to PEV is highly suggestive of a clonal route of signal propagation, some reports suggest that this may not be the case in all examples (Kunz *et al* 1996, Voinnet & Baulcombe 1997). A partial relaxation in the silencing structure of heterochromatin (Lu *et al* 1996) that is cell independent may be one explanation. This is seen for the Smcx gene on the X-chromosome, where variable levels of expression and non-clonal release from silencing can be observed (Sheardown *et al* 1996).

1.7.9 PEV and Mice

Although *Drosophila* are excellent and extensively used tools for genetic investigations, they have their limitations, especially when mammalian disease models or proteins are needed. The mouse is a natural choice to complement and extends the research done in *Drosophila* and single celled organisms such as Yeast.

Transgenic technology has been used to create mouse models of human disease and also model methods of producing therapeutic proteins for harvesting (Andres *et al* 1987, Archibald *et al* 1990, Matsuoka *et al* 2000). Other mice have been created in order to investigate the function of specific proteins or gene regulatory sequences (Caterina *et al* 1994, Bonifer *et al* 1996, Festenstein *et al* 1996, Corcoran *et al* 1998). In the course of the generation of all these mice, many of the transgenes have displayed unexpected behaviour linked to any one or a combination of the abovementioned sources of gene expression disruption (1.1-1.7). A chromosomal position effect was demonstrated as the cause of aberrant expression of a HSV-TK transgene, though without mapping of the insertion site classical PEV could not be confirmed (Al-Shawi *et al* 1990). The same transgene contained a cryptic promoter responsible for aberrant tissue expression (Al-Shawi *et al* 1991). Position effects were also deemed to cause the variable expression of lacZ reporter under β-globin promoter and LCR control, with the unusual observation of qualitative and

quantitative expression differences within both lines and individual mice (Guy *et al* 1997).

Mammalian transgenic papers are still rare compared with *Drosophila* and Plants, with PEV phenomena in mammals only reported and acknowledged five years ago (Dobie *et al* 1996, Festenstein *et al* 1996). 18 000 hits are generated by a literature search using transgenic animals as the key, shrinking to 55 when the additional terms PEV or copy number are included. The majority of mammalian PEV papers have until recently not addressed the mechanism(s) surrounding PEV. The identification of mammalian homologues of *Drosophila* proteins known to modulate PEV (Singh *et al* 1991, Cowell & Austin 1997) have generated two papers which investigate the effects that changing modifier concentration has on variegating transgenes in mice (Festenstein *et al* 1999, McMorrow *et al* 2000). PEV studies in mammals have also investigated copy number effects by generating multicopy transgene array inserts that induce variegation, and then reducing the copy number in order to study the effects of reduction of the array size. The first paper reported that reduction in copy number abolished PEV and led to even expression in all cells (Garrick *et al* 1998), resulting in the current mantra: "More is Less". The second paper reports the exact opposite (Ramirez *et al* 2001). Variegation still prevailed and the level of variegation generally remained unchanged, with only an occasional incidence of quantitative reduction in expression from the cells that were expressing the transgene.

With few reports of confirmed classical position effect variegation in mice (variegation within a line, with integration close to constitutive heterochromatin), the first such reported line (Dobie *et al* 1996, containing a milk protein transgene) remains an important source of information, combining the means to study a phenomenon impacting on both commercial production of therapeutic proteins and human diseases where position effects are implicated in the phenotype.

1.8 Transgenic Mouse line BLG/7

BLG/7 is a unique transgenic resource

In variably expressing transgenic mouse lines, silencing is frequently seen at the cellular level, with the extent of silencing variable between related individuals for lines exhibiting within line variations in expression levels (Dobie *et al* 1996, Sutherland *et al* 2000), and between individuals of different lines when the variation in expression is between two lines (Robertson *et al* 1996). In many of the reported cases, the relative transgene expression levels from individual cells will not vary (Graubert *et al* 1998), differences in the number of cells expressing the transgene are responsible for the variable amount of protein production. These qualitative, rather then quantitative, differences in transgene expression generate a mosaic pattern of expression when *in situ* hybridisations for mRNA are performed (Dobie *et al* 1996). This phenomenon therefore bears at least a superficial resemblance a position and/or copy number effect known as Position Effect Variegation (PEV) in *Drosophila*. BLG/7 is one such line exhibiting this phenomenon where centromeric integration has been linked to the variegated phenotype, this being the first report linking genomic location and variegated expression in a mirror image of that reported for PEV in *Drosophila*.

Other mice expressing transgenes under milk promoter control exhibit within line variegation (Bleck & Bremel 1994, Gutierrez *et al* 1996, Barash *et al* 1999), without linking it to insertions close to constitutive heterochromatin, although there was the suggestion that high copy number decreased expression (Barash *et al* 1999). Even in transgenic terms, BLG/7 are a unique resource as the variegation observed may be the result of both copy number and position effects, justifying a more in-depth study. These mice have formed the central focus of the work described in this thesis.

BLG/7 creation and initial characterisation

BLG/7 is a line of transgenic mice created over a decade ago (Simons *et al* 1987). The line contains approximately 25 copies of an ovine milk protein gene called β-lactoglobulin (BLG) (Dobie *et al* 1996). BLG is the major whey protein in ruminants

and has no rodent counterpart. The transgene array has integrated as a single block close to centromeric heterochromatin of mouse chromosome 15 (Dobie *et al* 1996).

BLG/7 was created as part of a study into the feasibility of creating transgenic animals able to secrete pharmaceutically important proteins in their milk, a process known as "pharming". Three transgenic mouse lines containing BLG displayed correct temporal and spatial expression of the transgene. Unexpectedly, two of the lines displayed considerable within line variation in expression levels. This was demonstrated to be due to mosaic expression, with variable numbers of expressing and silent cells correlating with the variable amount of transgenic protein produced in the milk. Correlation between integration near centromeric heterochromatin and high transgene copy number with variegated expression was established (Dobie *et al* 1996). Images presented in the following chapters may be easier to interpret if a brief description of the mammary gland is given at this point.

1.8.1 Mammary Gland Function and Composition

The mammary gland is a complex organ whose basic structure is laid down in the foetus. Hormonal cues delivered during puberty and pregnancy ensures development of a functional organ. Both gland development and protein expression is under the control of placental lactogens, growth factors and signals emanating from the extracellular matrix (ECM). There are differences in expression between the various proteins that comprise the milk, and these differences are attributed to differences in *cis*-regulatory elements and interactions with the hormones and ECM signals.

The gland contains a large number of different cell types, only a subset of which are responsible for the production and secretion of milk proteins. These specialised cells are known as the secretory epithelial cells, and their numbers increase enormously in rodents between parturition and mid-lactation. This increase in cell number is needed to enhance the milk yield and feed the growing litter. Believed to originate from a modified sweat gland, the parts of the mammary gland that comprise the secretory pathway include the aforementioned secretory epithelial cells that line individual alveoli, into which the milk proteins are secreted. The alveoli drain into secretory

tubules that in turn drain into lobular ducts that eventually connect to the teat. Structural cells and other vital components needed to contain the gland and ensure an adequate supply of nutrients and hormonal signal cues to all cells surround the alveoli. A simple example is to liken the structural composition of the mammary gland to a bunch of grapes. Each alveolus can be equated to a single grape in the bunch; with the secretory epithelial cells the skin of the grape. The stems represent the secretory network, congregating in the teat represented by the main stem that connects the entire bunch (see Figure 1.2).

Investigations into expression at the cellular level involve sectioning the mammary tissue and probing for milk gene expression. As only secretory epithelial cells express milk protein genes, the probe used identify no other cell types. In *situ* hybridisations for milk protein gene expression therefore generate images of circularish structures with only the outside edges visible (see Figure 1.2). The mammary gland is infinitely more complex than described here, but this simplified explanation is sufficient for my purposes. I am only concerned with expression from secretory epithelial cells, and provided the endogenous milk genes are expressed normally, I do not consider further examination of the intricacies of mammary gland biology part of the scope of this thesis. Suffice to say that previous investigations into transgene expression in BLG/7 mice revealed a link between the variable amount of protein secreted into the milk and mosaic mRNA expression at the cellular level. This phenotype was further correlated with the transgene multicopy array and its integration close to centromeric heterochromatin (Dobie *et al* 1996). This is analogous to the effects of Position Effect Variegation seen in *Drosophila.*

1.8.2 BLG/7 lines of investigation

Three lines of mice carrying a β-lactoglobulin transgene integrated at different genomic locations with varying numbers of gene copies at those sites have been investigated (Dobie *et al* 1996). Despite evidence that the integrations remained stable without any obvious signs of rearrangements, two of the three lines exhibited variegated levels of expression within the line. Both lines had the transgene integrated close to centromeric heterochromatin of chromosomes 5 (BLG/ 45) & 15

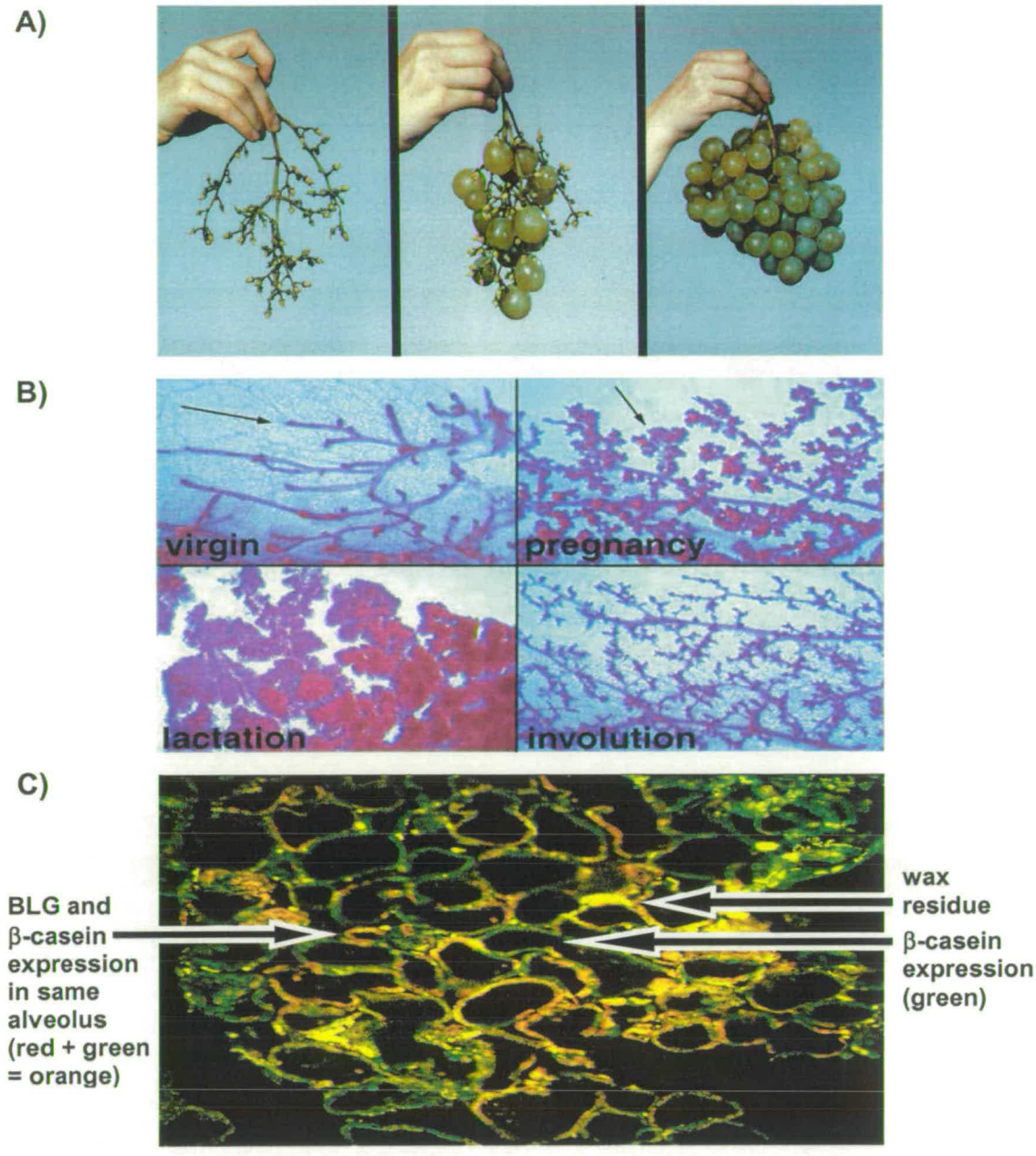


Figure 1.2 Mammary gland development during pregnancy and lactation

A) Representation of mammary gland expansion during pregnancy. The grapes represent the individual alveoli (courtesy of Lothar Henninghausen, NIH).

B) Morphology of mammary gland changes between virgin, pregnant, lactating and post-lactational stages (courtesy of Lothar Henninghausen, NIH).

C) Fluorescent *in situ* hybridisation of a wax embedded mammary gland from BLG/7 hemizygous female taken at mid lactation and probed for BLG (Red/Orange) and β-casein (green) mRNA expression. Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

(BLG/7) as multicopy arrays of approximately 17 (BLG/45) and 25 (BLG/7) copies, with the variegation in milk expression evident as mosaic expression at the cellular level. The third line exhibited even expression, both in milk and mammary tissue, with two copies of the gene integrated mid-arm on chromosome 7 (BLG/14).

As BLG/7 was the line with the highest level of variegation, it was investigated further. Backcrosses were performed to two different genetic backgrounds for three generations to check if the variegated expression was caused by segregation of genotype-specific modifier alleles (the mice were created on a mixed F1 background). No significant evidence was found, although there was a weak trend evident in one of the backgrounds (CBA) for reduced expression and variance compared to the original samples. Serial lactations showed that the levels of BLG expression do not vary between lactation, and must therefore be fixed and maintained at some point in advance of the first sampling (11^{th} day of lactation). Location of the transgene close to the centromere suggested that the mosaic expression pattern seen at the cellular level was due to an analogous mechanism that is responsible for position effect variegation in *Drosophila,* with the multicopy nature of the integration site playing a probable additional role (Dobie *et al* 1996).

Taking the project further

The investigation into the variegated expression profile of BLG/7 mice uncovered more questions than it answered. Three natural lines of investigation were incorporated into this project, attempting to answer three specific questions:

- **Are multiple recessive modifier loci responsible for BLG/7 variegation?** The observation that the CBA background exhibited a weak trend towards a lowered level and spread of expression initiated maintenance of the backcrossing regime for a further 10 generations to investigate the possible presence of minor modifier loci affecting BLG/7 expression (Chapter 3)

- **Would an increase in copy number through breeding to homozygocity uncover limiting factors involved in variegation control?**

  The phenotypic links to PEV and copy number dependent silencing phenomena suggested that this would be a good source of information. The behaviour of two identical variegating loci could indicate whether the variegation is subject to any constraints, such as availability of modifier proteins, and if the two loci can interact and affect each other. Homozygous BLG/7 mice were sampled and transgenic milk protein content and cellular expression compared to hemizygous BLG/7 mice (Chapter 4)

- **Are the mosaic patterns of expression clonally or instructionally derived?**

  The discrete patches of cells expressing BLG/7, and similar expression patterns seen for PEV affected genes, are assumed to be clonally derived. An alternative explanation is that the pattern is the result of an instructional state derived from inter-cellular communication (a distinct possibility in mammary tissue, where expression of milk proteins is hormonally controlled). The origin of the expression pattern of BLG/7 was therefore chosen as a third area for investigation. Mice with a lactation specific marker of clonality do not exist, and would have to be made before further investigations (Chapter 5.)

Chapter 2

Materials & Methods

Chemicals are from BDH (Sigma if not stocked by BDH) unless otherwise indicated.

2.1 DNA manipulation

Bacterial Based Manipulations

2.1.1 Miniprep Protocol

1. Grow up 3ml culture over night at 37 °C, or temperature suited to bacteria (DH5α were generally used).
2. Spin1.4 ml in eppendorf at 8000 rpm for 5 min.
3. Remove supernatant and resuspend pellet in 100 µl TE for bacteria (25 mM Tris.HCl, 10 mM EDTA pH 8).
4. Add 200 µl alkaline SDS (0.25 M NaOH, 1% SDS), mix gently, solution should become clear and viscous.
5. Add 150 µl 3M Na or K acetate pH 5, mix by inversion. A white precipitate should appear.
6. Centrifuge 10 minutes at 13000 rpm.
7. Remove supernatant (approx. 400 µl) and extract once with an equal volume of 1:1 phenol:chloroform (centrifuge 3 min at 13000 rpm).
8. Add 1 ml EtOH, mix, precipitate in –20 °C freezer for 30 - 60 minutes.
9. Centrifuge 10 minutes at 13000 rpm, remove supernatant.
10. Wash pellet with 70% EtOH and leave to airdry for 10 minutes.
11. Rehydrate in 50 µl TE with RNase (20 µl 10 mg/ml Ribonuclease A[a] in 0.5 ml).

(a = Sigma)

2.1.2 50 ml Scale-up of Miniprep Protocol

1. Centrifuge culture in falcons for 10 minutes at 4000 rpm
2. Resuspend pellet in 4 ml TE (25 mM Tris, 10 mM EDTA pH 8)
3. Add 8 ml Alkaline SDS (0.25 M NaOH, 1% SDS)
4. Add 6 ml 3M Na or K acetate
5. Centrifuge 10 min at 4000 rpm

6. Filter through sterile fishwool at bottom of 20ml syringe with plunger removed
7. Extract with an equal volume 1:1 phenol:chloroform (here 15 ml)
8. Precipitate 30 min at -70 °C with 2.5 volumes EtOH.
9. Centrifuge down 10 minutes at 4000 rpm
10. Rinse pellet in 70% EtOH and leave to dry, resuspend in 0.5-1 ml TE with RNase (20 µl 10 mg/ml Ribonuclease A[a] in 0.5 ml TE).

(a = Sigma)

2.1.3 Modified Qiagen Maxiprep Protocol

(To maximise DNA yield. Kit catalogue number 121 63)

1. Grow up large (250 ml x 3) cultures at appropriate temperatures overnight
2. Centrifuge 10 minutes 5000 rpm at room temperature
3. Resuspend pellet in 10 ml buffer P1 which contains RNase A
4. Add 10 ml buffer P2 at 37 °C for 5 minutes (invert tube to mix, do not vortex).
5. Add 10 ml chilled buffer P3, mix gently and leave on ice for 20 minutes.
6. Centrifuge 30 min 13 000 rpm at 4 °C
7. Filter supernatant through miracloth
8. Equilibrate Qiagen tip 500 adding 10 ml buffer QBT, allow to empty by gravity (use 4 tips to maximise yield and minimise loss through column blockage).
9. Add supernatant to equilibrated tips and wait until it has flowed through
10. Fill to top with Buffer QC (approx. 15 ml) allow to flow through, then repeat.
11. Elute DNA with 15 ml of Buffer QF.
12. Collect the eluate in 10 or 30 ml small baked glass centrifuge tubes (corex).
13. Add 0.7 volume isopropanol to the eluate (10.5 ml), mix and centrifuge for 30 minutes using a fixed rotor set for 7 – 8000 rpm at 4 ^{0}C.
14. Wash pellet with 5 ml 70% EtOH and recentrifuge as above.
15. Dry pellet briefly, then resuspend using 0.5 ml TE buffer.

Buffer P1: 50 mM Tris.Cl (pH 8), 10 mM EDTA, 100 µg/ml Rnase A.

Buffer P2: 200 mM NaOH, 1% SDS

Buffer P3: 3 M potassium acetate, pH 5.5

Buffer QBT: 750 mM NaCl, 50 mM MOPS pH 7, 15% isopropanol

Buffer QC: 750 mM NaCl, 50 mM MOPS pH 7, 15% isopropanol, 0.15% TritonX-100

Buffer QF: 1.25 M NaCl, 50 mM Tris.Cl pH 8.5, 15% isopropanol

2.1.4 Transformations

(Insertion of plasmid DNA into new host bacteria)

1. Thaw competent cells slowly on ice
2. Add 100 µl of the cells into a clean eppendorf, add 5 µl of a ligation mix (or appropriate amount of plasmid DNA) and mix gently (include no DNA controls)
3. Leave on ice for 20 minutes
4. Heat shock for 90 seconds at 37 or 42 °C
5. Ice for 2 minutes
6. Add 400 µl Luria-Bertani (LB) medium (no drugs), incubate on shaker at 37 °C for 20 - 30 minutes
7. Centrifuge slowly for 1 minute (8000 rpm)
8. Remove nearly all supernatant, leaving approx. 50 µl for resuspension
9. Plate on agar plates with appropriate drug resistance
10. Incubate overnight at 37 °C

Luria-Bertani (LB) medium

Tryptone	10 g/L
Yeast Extract	5 g/L
NaCl	10 g/L

Agar

Bacterial grade agarose in LB medium 15 g/L

Ampicillin plates made to concentration of 1 mg/L (100 µl of 100mg/ml stock)

Kanamycin plates made to a concentration of 0.3 mg/L

Same concentrations of antibiotics used in LB medium.

Glycerol stocks for transformed bacteria

Mix glycerol and bacteria in a ratio of 300: 700 in 1 ml.

2.1.5 DNA digests

General Digests

Enzymes: 1 μg of (clean) DNA should be digested completely by 1 unit of enzyme in an hour. Enzyme volume should not exceed 10% of total digest volume. Enzymes were usually sourced from Boehringer, with New England Biolabs the next choice if the enzyme needed was not available from Boehringer.

Buffers: Use buffer appropriate for enzymes in reaction. Buffers usually supplied with enzyme as 10x strength, so constitute $1/10^{th}$ of total reaction volume

DNA: 1-2 μl for diagnostic test of miniprep DNA, 0.5 – 1 μg for plasmid DNA.

Volume: Diagnostic reactions can be as little as 10 μl, depending on quality of DNA and complexity of reaction (larger volumes are recommended if several enzymes are used, especially if a buffer change is needed).

Blunt ending: Use Klenow enzyme to create blunt ends from previously cut DNA.

Southern Blot Digests

Larger reactions and high concentration enzymes needed when using genomic DNA.

25 μl reactions:

10 μg DNA

2.5 μl High Conc. Enzyme (40 U/ml)

2.5 μl 10 X Buffer

Incubate at relevant temperature (usually 37 ^{0}C) overnight

Buffer Changes

1. Equilibrate a 0.025 μm VS Millipore paper by floating on distilled water for 10 minutes (shiny surface facing up)
2. Place digest on surface for 20-30 minutes, salts should be drawn into the water
3. Remove droplet containing digestion divested of salt buffer and proceed with second digestion, adding required buffer and enzyme.

Can also use to remove small DNA fragments from a digest by dialysing overnight.

2.1.6 Ligation Reactions

Heat inactivate any digestion enzymes for 10 minutes at 65 °C to avoid interference with ligation reaction. Large fragments (15 Kb and above) are difficult to clone.

10 μl reactions:

Vector: Insert	(vary ratio from 1:0 to 1:4)
10 X Buffer	1 μl
10 mM ATP	1 μl 100 mM stock, final concentration 1 mM
T4 DNA Ligase	1 μl
$H_2O_{(distilled,autoclaved)}$	make up to total volume of 10 μl

Ligate overnight at approximately 14 °C.

Insert concentrations needed for cloning dependent on insert and vector sizes:

Insert size (Kb)	Vector concentration (μg/ml)		Insert concentration (μg/ml)
	4 Kb vector	8 Kb vector	
1	125	250	30
2	30	60	25
6	12	25	20
9	6	12	15
15	2	4	10
20	1	2	5

When cloning linker additions, use a 10:1 molar ratio of linker sequence to vector, heat denature at 70 °C for 10 minutes, then ice 2 minutes before adding ligase.

EX BIBL. UNIV. EDINB

2.1.7 DNA purification methods

Dialysis Method

(suited to high molecular weight DNA)

1. Cut out desired DNA band from gel
2. Wash dialysis tube (approx. 10cm) in distilled water and seal one end with a clip.
3. Add 400 µl of buffer to the tube (vary it depending on size of gel band)
4. Add gel to tubing, seal other end with clip and ensure no air bubbles or leakage.
5. Return to gel tank, keep tube immersed using glass plate, dialyse 2 hrs at 100V.
6. Reverse current for 1 minute to remove any DNA clinging to sides of tubing.
7. Remove liquid, avoid including agarose remnants, transfer to fresh eppendorf.
8. Extract with an equal volume phenol:chloroform, then ethanol precipitate DNA
9. Rehydrate in a suitable volume TE.

Dialysis tubing: 10 cm cut tubes in 5 mM EDTA (5 ml 0.5 M EDTA pH 8 in 500 ml distilled water), autoclave before use and store at 4 °C.

Freeze Method

(suitable for small to medium molecular weight DNA)

1. Cut out desired band from the gel.
2. Place in an eppendorf and put in - 70°C freezer overnight
3. Centrifuge for 10 minutes at 14000 rpm
4. Remove liquid and extract with an equal volume phenol:chloroform
5. Ethanol precipitate DNA and resuspend in a suitable volume TE

Fishwool Gel Extraction Method

(suited to smaller molecular weight DNA)

1. Cut out desired DNA band from agarose gel and place in a petridish
2. Remove cap from PCR tube
3. Puncture bottom with fine needle, place fishwool at the bottom using fine forceps
4. Cut gel on petridish into very small pieces and stuff into PCR tube
5. Place in 1.5ml eppendorf with lid removed, centrifuge 13000 rpm for 10 minutes.
6. Transfer liquid to new eppendorf and repeat step 5 centrifuge

7. Phenol:Chloroform and EtOH precipitate the DNA from the liquid
8. Re-hydrate the pellet in 10µl TE

Qiaex Gel Extraction Method

(Qiaex Gel Extraction kit 150, cat. 20021. Suited to smaller molecular weight DNA)

1. Cut out desired DNA band from agarose gel.
2. Add 3 x gel volume of buffer QX1
3. Leave at 50 °C for 10 min. until gel is dissolved. Mix well.
4. Add 10 µl Qiax II (vortex first), leave 10 min at 50 °C, shaking every 2 minutes
5. Centrifuge for 30 seconds and remove supernatant
6. Add 0.5 ml QXI and resuspend pellet
7. Centrifuge for 30 seconds and remove supernatant
8. Repeat steps 6-7
9. Add 0.5ml wash buffer (with EtOH added) and resuspend pellet
10. Centrifuge 30 seconds and remove supernatant
11. Repeat steps 9-10
12. Remove all supernatant, dry pellet at room temperature for 10 - 15 minutes
13. Add 10 µl TE, vortex & mix, Leave at room temperature for 5 minutes
14. Centrifuge 30 seconds, transfer supernatant to fresh eppendorf as contains DNA
15. Repeat steps 13-14

2.1.8 Polymerase Chain Reaction for genotyping mouse tail digests

Tail Digestion

Tails Proteinase K digested overnight (200mg/ml) in 0.3 M NaAc, 10 mM Tris.HCl pH 7.9, 1 mM EDTA and 1% w/v SDS at 37 °C, then briefly vortexed and frozen.

Tail PCR for genotyping

1. Thaw tail digests
2. Vortex briefly to mix sample.
3. Centrifuge at 1300 rpm for 10 minutes at 4°C, 15 minutes for frozen digests.
4. Transfer 20 µl of supernatant into labelled eppendorfs (optional step).
5. Make up PCR mix, don't add taq polymerase until ready to add to samples.

6. Add 1 drop mineral oil/PCR tube, pipette 1µl tail DNA/tube underneath oil.
7. Denature samples on PCR machine (15 min at 95 °C, then cooled to 65 °C).
8. Add 50/25 µl PCR reaction mix with taq polymerase to DNA under oil.
9. Run BLG PCR programme (30 sec at 92 °C, 5 min at 65 °C, cycled 30 x).
10. 15 µl samples of the PCR reaction are run on a 2% agarose gel at 100 V.

Standard mix for BLG/HPRT primed tails for 50 µl reactions:

mix/tube : 5 µl 10 x Mg^+ reaction buffer

8 µl 1.25mM dNTPs (dATP, dGTP, dTTP and dCTP in equal ratio)

5 µl DMSO[a] (dimethyl sulphoxide C_2H_6OS)

0.5 µl 10 mM BLG 5' + 3' primers

0.5 µl 10 mM HPRT 5' + 3' primers

0.15 µl Taq DNA polymerase enzyme

31 µl H_2O (distilled and autoclaved)

* 25 µl reactions work equally well

Alternative Method for PCR

- Make up ready reaction mix, aliquot 25 µl per tube, add oil and 1 µl tail digest
- Run straight on BLG programme without denaturing step

200 X reaction mix:

10xMg^+ reaction buffer	500 µl
1.25mM dNTPs	800 µl
DMSO[a]	500 µl
10mM BLG 5' + 3' primers	50 µl
10mM HPRT 5' + 3' primers	50 µl
H_2O (distilled & autoclaved)	3 100 µl

(a = Sigma)

Ready Reaction mix:

Taq DNA polymerase enzyme	3 µl/1000 µl reaction mix

2.1.9 X-Gal staining of ears for identification of PGK IRES βGeo mice

1. Rinse ears in 1XPBS, each ear in a well on a 48 or 24 well plate on ice.
2. Fix 5 minutes in 2% Formaldehyde: 0.2% Glutaraldehyde in PBS in fridge
3. Rinse three times in PBS with 2 mM MgCl
4. Add stain solution, cover plate and leave overnight in a box at 37 ^{0}C.
5. Examine under microscope, blue ears are easily identifiable. Can change staining solution for PBS to make detection easier. Most results can be seen by eye.

Fix Solution: 2.5 ml 37 – 40% Formaldehyde
0.4 ml 25 % Glutaraldehyde
PBS to 50 ml total volume

Stain Solution: 1 mg/ml X-Gal dissolved in DMSO
5 mM Fe3+/Fe2+ ($K_3Fe(CN)_6$ + $K_4Fe(CN)_6 3H_2O$)
2 mM $MgCl_2 6H_2O$
Make up in 1 X PBS
Mix and filter before use

X-Gal = 5-bromo-4-chloro-3-indolyl-b-D-galactosidase, from Melford Laboratories.

2.1.10 Mass screening of bacteria for successful ligations (colony lifts)

1. Prepare duplicate sets of plated bacterial colonies, use sterile toothpicks to plate out colonies in a pattern that can be used to orientate the plate
2. Cool plates in coldroom for a minimum of 30 minutes before transfer
3. Label each filter in pencil (0.45 μm Nytran filter D-37582 Dassel from Schleicher & Schuell, 82 mm)
4. Place filters pencil side up on plates and leave 1 minute, use a pin to stab orientation holes through the filter into the agar plate
5. Remove filter and place on 3MM double layer of whatman paper soaked in denaturing solution (0.5 M NaOH, 1.5 M NaCl) for 5 minutes, plaque side up
6. Transfer filter to 2 X 3MM whatman paper soaked in neutralisation solution (0.5 M TrisHCl pH 7-8, 1.5 M NaCl) for 5 minutes

7. Transfer to a 2 X SSC (tri sodium citrate) bath, submerge and agitate until bacterial debris is washed off filter
8. Dry filter plaque side up on 3MM paper, then crosslink at 1200 J
9. Wet filter in 2 X SSC and place plaque side up in hybridisation tube, add 20 ml of hybridisation buffer and prehybridise at 65 °C for 20 – 30 minutes before adding probe (see 2.2.5). Hybridise overnight.
10. Rinse with wash buffer, add 20 ml of wash buffer and return to oven for 20 minutes. Repeat twice
11. Wrap filters in cling film, put in cassette, add film and leave overnight in –70 °C freezer.

Hybridisation buffer (Church-Gilbert buffer):

250 ml Phosphate buffer (1 M = 70.9 g Na_2HPO_4, 4 ml 85% orthophosphoric acid in 1 L)
35 g SDS
5 g BSA (fraction V)
1 ml 0.5 M EDTA

make up to 500 ml with distilled water, heat until ingredients are in solution

Wash buffer (1L):

40 ml Phosphate buffer
2 ml 0.5 M EDTA
10 g SDS
958 ml distilled water

2.2 Cell Culture work

Sterile techniques observed at all times. Fume hoods are category 1 containment level. Cell culture designated labcoat and non-powdered gloves worn. 70% Ethanol used as general disinfectant, Microzid used to clean fume hoods.

2.2.1 HC-11 Cells, routine maintenance

HC-11 cells are immortalised mammary epithelial cells, described in detail by Burdon and colleagues (1994).

RPMI complete medium (for HC – 11 cells)

RPMI 1640 medium[a]	500 ml
FCS[b] (foetal calf serum)	50 ml
L-Glutamine[c] (50 μg/ml)	5.6 ml
Penicillin/Streptomycin[a] (10 000 U/ml)	5.6 ml
EGF[a] (epidermal growth factor, 10 ng/ml)	500 μl
Insulin[d] (5 μg/ml)	500 μl

Filter sterilise additions to medium before addition. Store in fridge.

(**a** = Sigma, **b** = Globepharm, **c** = Life Tech., **d** = Gibco BRL)

Recovery of Frozen Cells

1. Remove cell vial from liquid nitrogen, place in hot water to thaw quickly.
2. Transfer cells to 15ml falcon with 5ml medium, centrifuge 5 min at 1000 rpm
3. Remove supernatant, resuspend cell pellet in 5ml RPMI complete medium
4. Transfer to a T25 flask and incubate at 37 °C, 5% CO_2

Routine Maintenance

Medium volumes:

T25	5 ml
T 75	10 – 15 ml

Medium changes:

- Rinse with PBS if there is much cell death before adding fresh medium.
- Check cells daily and change medium every two days or more often, a lighter medium colour indicated need for change.
- Seed at lower densities with larger medium volumes if keeping cells unattended over weekend.

Trypsinisation of cells:

1. Wash cells with PBS
2. Add 1ml TEG (trypsin-EGTA solution) prewarmed to 37 °C to a T25 (2ml for a T75) and return to incubator for 10 minutes.
3. Slap bottle to dislodge cells, then add 3-5 ml medium. This stops reaction.
4. Transfer liquid to a falcon and centrifuge for 5 minutes at 1000 rpm
5. Remove supernatant, resuspend cells in 0.5 ml, make up volume to 10 ml and transfer to T75 (or split into smaller flasks as ¼ or desired cell seeding density)
6. Gently agitate new flask to ensure even distribution and return to incubator

TEG: 3.15 g/L NaCl, 0.06 g/L Na_2HPO_4, 0.108 g/L KH_2PO_4, 0.166 g/L KCL, 0.45 g/L D-glucose, 1.35 g/L Tris, 0.45 ml/L 1% Phenol Red, 50 ml/L 10 x 2.5% trypsin (ICN FLOW), 0.2 g/L EGTA, 0.05 g/L PVA. Filter before use, store at – 20 °C.

Freezing Cells

1. Trypsinise cells, centrifuge and resuspend in 3 ml RPMI complete medium
2. Add 3 ml of freezing mix gently to the cells, and dispense as 1ml aliquots into pre-chilled labelled cryogenic screwtop tubes
3. Place in styrofoam box that has been pre-chilled in normal freezer and put in –70 °C freezer for a few days before transferral to liquid nitrogen

Freeze Mix	10 ml
RPMI complete	6 ml
DMSO	2 ml
FCS (filter first)	2 ml

2.2.2 Transfecting HC-11 cells using Calcium Phosphate method

Need:

RPMI complete medium

DMEM complete medium: **D**ulbecco's **M**odified **E**agle's **M**edium[a]
10% FCS[b]
L-Glutamine[c] (5.6 ml in 500 mls)

2M $CaCl_2$ solution

2 X Hepes Buffered Saline (HBS): 50 mM Hepes[e], pH 7.1
280 mM NaCl
1.5 mM Na_2HPO_4

HEPES = N-2-Hydroxyethylpiperazine-N'-2-ethane sulphonic acid ($C_8H_{18}N_2O_4S$)
(**a** = Sigma, **b** = Globepharm, **c** = Life Tech., **e** = Merc)

Day –1

- Seed 5 X 10^5 cells in T75 flask, usual medium and culturing conditions apply.

Day 0 (transfection day)

1. Change medium at 9 am, use 15 ml DMEM complete medium, incubate for 3 hrs to ensure peak growth rate.
2. Defrost $CaCl_2$, 2 X HBS, DNA solutions at room temperature
3. Start preparing transfection mixes from 11 am, 0.5 ml 2 X HBS per transfection
4. Mix plasmid DNAs with $CaCl_2$, make up to 0.5 ml with $H_20_{dda.}$ 0.5 ml contains 3-5 μg pSVNeo, 20 μg plasmid DNA, 62 μl $CaCl_2$ (final concentration 250 mM)
5. Add the DNA-$CaCl_2$ solution drop-wise to the 2 X HBS.
6. Leave to rest at room temperature for 30 min.
7. Add the 1 ml transfection solution dropwise to the cells in the T75 flasks and gently shake the medium to mix, return to incubator and leave 5 hours.
8. Wash cells with DMEM complete medium, add 10 ml RPMI complete medium, leave to incubate overnight.

Day +1

- Leave cells in RPMI complete medium without selection

Day +2

- Change medium to 10 ml RPMI complete with 200 μg/ml of G418 (Geneticin[c])

Day + 3 - +14

- Continue selection, change medium every two days. Check for resistant colonies

Day + 15 - + 17

- Collect and pool the colonies from each flask, divide between 2 T25 flasks.
- Culture in RPMI complete medium without G418 for 2 days after they have reached confluence.

Day +18 - + 19 : Induction of expression

- Culture the cells with RPMI low serum for 2 days.

Day +20 - +24

- Apply induction medium and culture cells for 4 days. Culture non-induced control cells in parallel, usual culture conditions for both.

Day +25

- Induction of expression complete
- Harvest cells and stain for expression

Induction Medium

RPMI 1640 medium[a]	100 ml
FCS[b]	3 ml (final conc. 3%)
L-Glutamine[c]	1.1 ml (final conc. 50 μg/ml)
Pen/Strep[a]	1.1 ml (final conc. 10 000 U/ml of each)
Dexamethasone[a]	100 μl (final conc. 4 μM)
Insulin[d]	100 μl (final conc. 5 μg/ml)
Prolactin[a]	200 μl (final conc. 5 μg/ml)

Low Serum Medium

RPMI 1640 medium[a]	100 ml
FCS[b]	1 ml (1% final conc.)
Pen/Strep[a]	1.1 ml (final conc. 10 0000U/ml of each)
Insulin[d]	100 μl (final conc. 5 μg/ml)

(**a** = Sigma, **b** = Globepharm, **c** = Life Tech., **d** = Gibco BRL)

2.2.3 X-Gal Staining of cells for LacZ expression

1. Wash cells 1 x with 1 X PBS
2. Overlay with 10 ml fixative (volumes for T75), leave at 4°C for 10 minutes.
3. Wash cells 3 x with 1 X PBS
4. Overlay with staining solution (10 ml for T75), incubate overnight at 37°C with flask tightly closed.
5. Remove staining solution and overlay with 1 X PBS.
6. Visualise cells and staining under microscope
7. Photograph cells

Fixative:

2% Formaldehyde(40 % stock = 5ml in 100 ml)

0.2% Gluteraldehyde (25% stock = 0.8 ml in 100 ml)

Make up to desired volume with 1 X PBS

Staining solution for β-Gal:

1mg/ml X-Galactose	(20 mg/ml stock = 5 ml in 100 ml)
5mM $K_3Fe(CN_6)$	(0.5M stock = 1 ml in 100 ml)
5mM $K_4Fe(CN_6)3H_2O$	(0.5M stock = 1 ml in 100 ml)
2mM $MgCl_26H_2O$	(1M stock = 200 µl in 100 ml)

2.2.4 ES cells, Routine Maintenance

- ES cells divide every 18 – 24 hours and should have their medium changed every day. Normal feeding is 5 – 7 ml per T25, 16 – 20 ml per T75, 60 ml per T180.
- Cells passaged every 2-3 days, ¼ flask used for seeding. 1/6 –$1/8^{th}$ for weekend survival, add extra medium (T25 = 10 ml, T75 = 25 ml, T180 = 100 ml).
- Cells are grown at 37 °C in 5% CO_2

Medium

BHK – 21 Medium[a] (Glasgow MEM)	500 ml bottle
Foetal Calf Serum[b]	28 ml
New Born Calf Serum[b]	28 ml
L- Glutamine[c]	5.6 ml
Non Essential Amino Acids[c]	5.6 ml
Sodium Pyruvate[c]	5.6 ml
β- mercaptoethanol[a]	1.1 ml
ESGRO[f] (LIF = leukaemia inhibitory factor)	0.3 ml

Filter sterilise through 0.22 μm filter + 0.8 μm prefilter before adding to medium.

(**a** = Sigma, **b** = Globepharm, **c** = Life Tech., **f** = Chemicon)

Passaging cells

Cells should be passaged at 85 – 90 % confluency and survive better if medium is changed 2 – 3 hours prior to trypsinisation.

1. Change medium on cells, gelatinise new flask (5 ml 0.1% gelatin for T25, needs 20 minutes to set)
2. After 3 hours wash cells with 1 X PBS and then add trypsin solution (TEG)
    - Need 1 ml for a T25
    - 1.5 ml for a T75
    - 2 ml for a T180
3. Leave trypsin on for one minute (can put in incubator), smack flask and check under the microscope that the cells are dislodged without clumping.
4. Add medium to inactivate trypsin: 5 ml for a T25
   8 ml for a T75
   16 ml for T180
5. Pipette the cell suspension using glass narrow holed pipettes, this ensures single cell suspension. Centrifuge in falcon tubes at 900 – 1000 rpm for 5 minutes.
6. Remove almost all supernatant, flick tube to resuspend cells, add fresh medium and use glass pipette to create single cell suspension. Transfer to fresh flask.

Resuscitating cells

1. Remove vial of cells from LN_2 tank and thaw rapidly in warm water.
2. Resuspend in 5 ml of complete ES medium; centrifuge 5 minutes at 1000 rpm.
3. Aspirate most of medium and resuspend pellet in remnant. Add 5 ml fresh medium and create single cell suspension using glass pipette.
4. Transfer to gelatinised T25 and incubate as normal

Freezing cells

1. Change medium 3 hours in advance, freeze only 80% + confluent cells.
2. Make up or thaw out 2 X Freezing mix: 10ml: 6 ml Complete Medium
   2 ml FCS
   2 ml DMSO
   (Filter before use)
3. Label cryotubes and place in styrofoam box in freezer to cool down
4. Trypsinise and resuspend in 1 or 2 ml (T25/T75) complete medium.
5. Slowly add 1ml (T25) or 2 ml (T75) 2 X FM (kept on ice) dropwise, mix once before quickly dispensing into cryotubes (1ml per tube) using glass pipettes.
6. Place box of tubes in – 70 °C freezer, leave 24 hours before transferring to LN_2

Freezing down cells in 24 well plates

1. Make up 1 X Freezing mix: 10 ml: 8 ml Complete Medium
   1 ml FCS
   1 ml DMSO
   (Filter before use and chill)
2. Change medium 3 hours prior to freezing
3. Remove 24 well plate from incubator and put on ice
4. Aspirate medium
5. Add 300 µl 1 X FM to each well gently down the side
6. Put on lid and tape round it with parafilm, securing with autoclave tape.
7. Place in – 20 °C freezer until all wells are frozen (approx 30 minutes)
8. Plate is placed in precooled padded bag, sealed, labelled and placed in – 80 °C freezer. Should keep for 4 weeks.

Resuscitating cells in 24 well plates

1. Warm medium to 37 °C
2. Remove plate from – 70 °C, add 1 ml warm medium down side of wells, leave a few minutes at 37 ^{0}C for cells to defrost, aspirate carefully from side of well.
3. Add fresh 2 ml of medium and incubate for rest of day
4. Change medium again late in afternoon
5. Change medium again the following day

An alternative method is to suck up all medium, centrifuge and replate cells.

2.2.5 ES cell targeting work

ES cells: RI p14 (129 Ola)

Targeting Construct: HBZ (approx. 21 Kb)

- 4.3 Kb BLG promoter from pBJ39 (Sal I/RV)
- 4.6 Kb LacZ w. Poly A signal (Clontech pβGal-Basic)
- M44.8 (A. Smith HPRT targeting vector):
  9Kb gHPRT from exon 6 (with some of intron 5)
  3 Kb pBS backbone

Reporter inserted into unique exon 6 site (as Sal I fragment) with targeting construct linearised using Xho I.

Reaction Conditions:

- 2 lots of 200 μg linearised vector in 200 μl HBS (from 4mg/ml stock that has been Qiagen purified)
- 2 lots of $2x10^7$ cells in 600 μl HBS
- BioRad Gene Pulser set at: 800 V, 3 μF, T_c=0.1

Culture of Cells Prior to Electroporations:

- RI (p13) vial of ES cells defrosted and cultured for 14 days prior to electroporations. Cells are now p14.
- Split 7 times prior to electroporations, cells cultured in T180 flasks from day 3 post-resuscitation.

Electroporation & Selection of Targeted Cells:

Day 0

1. DNA spun, EtOH washed and resuspended in 2 aliquots of 200 µl 1xHBS
2. RI cells trypsinised and spun down as normal, the pellet resuspended in 15 ml 1xHBS using a glass pipette to ensure a single cell suspension
3. The cells were counted using a graticule. The squares of the grid were used to count the cells. The number of cells/ml is the number of cells counted in a grid $x10^4$.
4. 2 batches of $2x10^7$ cells were prepared in 600 µl 1xHBS each, and mixed with the 200µl aliquot of DNA.
5. The DNA/ES Cell suspension was pipetted into a 0.4cm electrode cuvette (BioRad cat. no: 115 – 2088) ensuring that no bubbles formed.
6. The BioRad Gene pulser is set at 800V and left to warm up.

   3 µF

   T_c=0.1

7. The cuvette is placed into the chamber and pushed into place until the electrodes surround it.
   - Both charge buttons are then pushed simultaneously until a beep is heard and the "charging" icon has vanished from the screen
   - A check is made to ensure that the electroporation parameters have not been altered before the cuvette is removed and left to stand for 10 minutes
8. Each of the two electroporations were resuspended in 100 mls medium.
9. The resuspended cells were plated out on circular dishes (10 mls/dish = 20)
10. Cells need to be cultured for 6 days before selection can proceed (endogenous HPRT to be degraded completely prior to selection).

Day 6 (1 of selection)

- Selection with 6TG (2-amino-mercaptopurine, from Sigma) started (1.67 µg/ml) filtered in normal medium, change every two days.

Day 16 (10 of selection)

- Colony Picking: 48 colonies picked, half of each to 2 x 24 well plates
- The appearance of colonies and the time to pick them will vary between experiments, depending on cell growht, selection strength and means.

1. 48 eppendorfs and 48 wells were labelled, each well filled with 1-2ml selection medium.
2. A P200 set at 120 was used. Holding the plunger down and using the tip to dislodge the pellet, the colony was lifted off the plate and placed in the labelled eppendorf.
3. The pellet was carefully pipetted and approximately half the suspension was transferred to the relevant well.
4. The 24 well plates were placed in the incubator to expand the colonies, the cells in the eppendorfs were spun down and resuspended in 1xPCR buffer (Mg+) with Proteinase K added and incubated for 2 hours at 65 °C:

2 ml:	10x PCR buffer (Mg+)	200 µl
	Proteinase K (20 mg/ml)	2 µl
	H2Odda	1.798 ml

5. The Proteinase K was inactivated by incubation at 90 ^{0}C for 15 minutes.
6. 15 – 20 µl of the lysate was used for a routine PCR assay screening for the BLG promoter. The active components of the reaction mix were increased by 2/5 (2.1.7) as total volume was increased to 50 µl with 20 µl lysate.

- Fixed and stained the leftover plates of unpicked colonies (methanol fixed, stained with 10 % Giemsa):
    1. Wash with 1xPBS
    2. Add fresh PBS and left at room temperature for 10 –15 minutes
    3. Remove PBS, overlay with Methanol, leave 10-15 min. at room temperature.
    4. Remove Methanol, add 10% Giemsa, leave 10 min. at room temperature
    5. Pour off Giemsa stain and rinse very gently with tapwater.
    6. Leave to air dry, colonies visible as purple spots.

2.2.6 Screening of cells for making chimaeric mice

DNA Extraction for Southern Blotting

1. Genomic DNA prepared by digesting the cells overnight with proteinase K cell lysis solution on a shaker, then isopropanol precipitating the DNA (Laird *et al* '91). The precipitated DNA was washed in 70% EtOH and left to air-dry for a few minutes before rehydration in 0.5 ml TE.
    - The resulting DNA was of such high molecular weight that it had to be physically sheared using a needle (26G ½, 0.45 x 13) and further diluted before digestion for Southern Analysis could take place. .
2. Digests performed overnight using high concentration enzyme with top up first thing and the digest proceeding for further two or three hours before being run on a 0.8 % agarose gel.

Blotting

1. Digest gDNA in as small a volume as possible and ensure thoroughly digested.
2. Ran finished digests on 0.8 % agarose gel, either during day at 80 – 100V or O/N at 20 V. Gel should be no more than 7mm thick.
3. Stain gel with ethidium bromide after run finished and photograph.
4. Depurinate gel before setting up transfer: 0.25M HCl on shaker until blue loading dye has turned yellow
5. Rinse with distilled water
6. Set up blot:
    - 0.4 M NaOH is used as transfer buffer and Zetaprobe or Hybond N+ is membrane used for transfer
    - Buffer is poured into metal tray and a cleaned glass sequencing plate is placed over it.
    - 4 sheets of 3MM whatman paper is cut to the approximate width of the gel (a little larger) and the full length of the paper. These are soaked in the buffer and left with the ends folded under the glass plate to form a wick.
    - Use a glass pipette to roll the paper flat

- Place gel face down on the wetted paper (leaves a smooth surface for the transfer). Cover gel completely with cling film, including sides. Ensure that there is plenty of buffer on top of gel.
- Place pre-wetted (water or transfer buffer) membrane cut to fit on gel starting from the middle. Do not remove once in place. Avoid air bubbles, roll with glass pipette for smooth finish.
- Soak 2 3MM whatman paper cut to fit the gel exactly and place on top of transfer membrane. Place 2 dry whatman on top
- Make stack of approx 5 cm using green towels, ensuring that they are evenly stacked and that bottom rows cover the blot evenly.
- Place tray on top of stack and weigh with 250 ml bottle on top.
- Blot O/N or minimum 5 hours (if target size is over 20 % or gel concentration exceeds 1%: always blot O/N)

7. Dismantle blot:
    - Dismantle blotting apparatus, carefully remove membrane and rinse in 2 X SSC
    - Dry membrane between two sheets of Whatman at 65 ^{0}C for an hour (do not UV crosslink using these membranes) At this point the membrane may be stored dry between whatman paper in a plastic bag at room temperature. Drying is not required for Hybond N+, it can be transferred straight to prehybridiation.
    - Prehybridise in 20 ml of hybridisation solution for 5 - 30 minutes before adding probe so that membrane is evenly coated. Ensure no bubbles form between the membrane and the glass walls of the hybridisation tube, and that there are no creases in the membrane.
8. Probe blot:
    - Place membrane in 20 ml of prewarmed hybridisation solution (65 ^{0}C) in hybridisation tubes in oven.

      Hyb Solution:

      0.25 M Na_2HPO_4, pH7.2 (25 ml of 1 M stock in 100 ml)

      7% SDS (35 ml 20 % stock in 100 ml)

      1mM EDTA (0.1 ml of 1M stock in 100 ml)

- Inject probe into tube and leave to hybridise O/N with rotation
- Wash twice for 20 minutes (65 ^{0}C) with wash solution A: 40 mM Na_2HPO_4, 5% SDS (40 ml 1M in 1L, 25 ml 20% in 1L).
- Wash twice for 20 minutes (65 ^{0}C) with wash solution B: 40 mM Na_2HPO_4, 1% SDS (40 ml 1M in1L, 5 ml 20 % in 1L).
- Optional high stringency wash for 20 minutes (65 ^{0}C): 20 mM Na_2HPO_4 0.5% SDS (20 ml 1M in 1L, 2.5 ml 20% in 1L)
- Cover in double layer of cling film and place down to image. Do not allow to dry, especially if want to strip and reprobe blot.

Making Probe

1. Boil 25 ng of probe DNA in total volume of 11 µl for 2 minutes. Probe length is suggested to be between 200 – 1000 bp
2. Centrifuge down briefly, ice for 1 minute
3. Add: 4 µl High Prime
   5 µl (50 µCi) α ^{32}P dCTP
4. Leave at room temperature for 10 minutes
5. Stop reaction with 5 µl 2M NaOH and leave for a few minutes. This has the additional effect of denaturing the probe for you.
6. Increase volume to 100 µl or more with hybridisation buffer or NaOH.
7. Inject into hybridisation tubes via cap hole and run down glass wall not occupied by blot membrane.

2.2.7 Karyotyping Cells

Preparation of mitotic spreads

1. Use all of T25 50% confluent flask
2. Feed cells 1 – 1.5 hrs before trypsinisation, adding 10µl/5µl medium of Colcemid (KaryoMax Colcemid stock at 10 µg/ml from GIBCO BRL)
3. Trypsinise as per usual and centrifuge as normal etc
4. Resuspend in 1ml hypotonic solution, dropwise mixing by shaking at room temperature.
5. Add excess of hypotonic solution to 5 ml and invert once to mix
6. Leave at room temperature for 10 minutes, centrifuge for 5 min at 1000rpm

7. Aspirate.
8. Flick tube to disrupt cells and add freshly prepared ice-cold fixative 1 ml dropwise as for step 4. If a pellet settles, give it a gently flick to resuspend.
9. Add another 5 mls iced fixative, invert, leave at room temperature for 5 minutes
10. Centrifuge 500 rpm for 5 minutes
11. Repeat steps 7 – 10 twice
12. Resuspend final pellet in 1 ml of fixative (can now be kept at – 20 °C if not used on day, but fixative will have to be changed immediately before use).
13. Slides are kept under MeOH in fridge. Wipe dry and label with pencil.
14. Drop droplets of the solution on to the slides from as high as possible.
15. Dry and check for decent spreads under microscope.
16. Stain slides in 10% Giemsa stain for 10 min at room temperature
17. Count chromosomes using microscope.

Hypotonic Solution:

0.56% w/v NaCitrate (KCl)

Fixative:

18 ml MeOH: 6ml Glacial Acetic Acid

2.2.8 Blastocyst Injections

Day 1

Resuscitate 1 vial of ES cells (into T25 flask)

Day 4

AM

1. Feed ES cells, flask should be 75-80% confluent
2. Make up ES medium with HEPES (200 µl 1M HEPES per 10 ml medium) Need at least 1-2 ml per mouse, plus 5 – 10 ml for ES cells and injection procedure.
3. Make up droplets ready for collected embryos on a small round plate and cover with oil (embryo oil saturated with CO_2) Place in incubator.
4. Make spare collection needles for blastocyst collection

5. Start flushing embryos from 10.30
    - Kill by neck dislocation, cut abdominal skin and expose abdominal cavity. Open cavity and gently push aside guts etc. Ovary is found just by the kidney, uterus is usually a yellowish-white tubing thing. Trace downwards until reach T-junction where the other uterine tract is found.
    - Cut across ovary and then downwards towards junction and cut just above the join ---- place on a paper towel with the fat downwards and the ovary side facing the wall. Repeat for second ovary.
    - Lift ovary side of tract with forceps, trim fat sticking to tissue as well as blood vessels. Place on paper towel with ovary facing towards you.
    - Fill 5 ml syringe with HEPES buffered ES medium and ensure that there are no airbubbles. Attach brown needle (26 G ½, 0.45 x 13).
    - Pick up oviduct and stretch uterus, insert needle carefully at top by uterus and hold in place with forceps. Flush out embryos using 0.2-0.5 ml medium into sterile bottle. Repeat for second oviduct.
    - Once bottle filled, cap and keep warm until can transfer to incubator.
    - Blastocysts sorted under microscope in hood and transferred by mouth pipetting to the collection droplets under oil

PM

1. Blastocysts injected with ES cells from 13.00, so T25 is trypsinised at 12.30 and spun down as normal. ½ of cells used for injections, resuspended in 4ml total HEPES buffered ES medium.15-20 ES cells per blastocyst injected.
2. Injected blastocysts transferred to recipient female, 12 + transfers per female preferable, expect about 50% not to implant.

2.3 PROTEIN PROTOCOLS

2.3.1 Preparation of milk samples for protein analysis

1. Thaw milk samples on ice.
2. Vortex samples briefly to mix milk.
3. Rinse Hamilton syringe with sterile water before taking a 30/50 µl milk sample.
4. Rinse the syringe with water 4-5 times between each sample.
5. Dilute the samples 1/5 in sterile water (30/150 = 120 µl)(50/250 = 200 µl).
6. Vortex and centrifuge briefly (let machine gain full speed, then switch it off).
7. Remove defatted milk (140/200 µl), disrupt mycelles without disturbing the fat.
8. Pipette the defatted milk into a new eppendorf.
9. Dilute 1/5 with loading buffer (50 µl DF milk and 200 µl LB)
10. Boil samples for 4 minutes (pin a hole in the tubes, or they will explode).
11. Measure sample volume and adjust for any evaporation with loading buffer.
12. Dilute a sample of 1/25 stock with loading buffer to make a 1/250 dilution.
13. Load 40/50 µl sample using Hamilton syringe (rinse between samples). Load empty wells with loading buffer to ensure a smooth and uniform run.

Milk Reducing/Sample Buffer	**16 ml**
0.5M Tris.HCl pH 6.8	2 ml
10% v/v Glycerol[g]	1.6 ml/2.09 g
10% w/v SDS	3.2 ml
5% v/v β-mercaptoethanol[a]	0.8 ml
0.1% w/v Bromophenol Blue[a]	2 ml
H_2O(distilled, autoclaved)	6.4 ml

(a = Sigma, g = BDH)

2.3.2 Pouring, running and staining protein gels

1. Pour resolving gel the afternoon before use, cover with water and cling film to avoid evaporation. The gel should come up to a level of 1-2 cm below the comb.
2. Remove water, place combs at an angle before pouring stacking gel (prevents air bubbles forming underneath the teeth). Leave to polymerise for 30-45 minutes.

3. Make running buffer
4. Remove combs, rinse wells with distilled water with final rinse of running buffer.
5. Set up protein rig, fill upper chamber with buffer to check for leakage (wipe surface with buffer to ensure good seal between chamber and gel plate).
6. Load wells using Hamilton Syringe before placing into buffer-filled lower chamber. Top up upper chamber buffer if necessary.
7. Run stacking gel at 30 mA (V_{max} = 500, W_{max} = 30), approx. 1 hour.
8. Run resolving gel at 60 - 70 mA. This holds for 2 gel. Adjust downwards one gel, upwards for 4. Run will take approx. 2.5 hours.
9. Stop run 15 minutes after dye has run off bottom. Remove from tank, cool.
10. Open plate and remove stacking gel using spacers. Wet gloves in running buffer and lift gel from plate and place in plastic box.
11. Add fix/stain, place in 50 °C oven for rest of day (enhances stain penetration), place on shaker at room temperature overnight (ensures even stain uptake).
12. Remove stain/fix from gels and add destain until just covering gel. Return to shaker and change destain solution every hour until suitably destained.
13. Use destain without acetic acid for final destain (make up volume with water).
14. Place finished gel on a cleaned glass plate and cover with cling film. Remove cling film for image capture. Store at 4 °C.

Alternatively, make gel during day and run overnight at 25mA. The resolving gel should set in 4 hours; the separating gel in 30-45 minutes. Leave gel in oven all day, then shake overnight. This method was preferred as it gave tighter, cleaner gels.

Resolving gels (makes two 20 cm gels)	**17.5% gel**	**20% gel**
Acryl.Bisacryl.[h] (37.5:1)	34.965ml	39.960 ml
H_2O (distilled, autoclaved)	9.135 ml	4.140 ml
1.5M TrisHCl pH 8.8	15 ml	15 ml
w/v SDS	600 µl	600 µl
APS (ammonim persulphate)	300 µl	300 µl
TEMED[a]	30 µl	30 µl

Stacking gels	**10 ml for 2 gels**
30% Acryl.Bisacryl.[h] (37.5:1)	1.3 ml
H_2O (distilled, autoclaved)	6.1 ml
0.5M TrisHCl pH 6.8	2.5 ml
10% w/v SDS	100 µl
10% APS	50 µl
TEMED[a]	10 µl

TEMED = N,N,N',N'-tetramethyl ethylene diamine. $C_6H_{16}N_2$

Protein Running Buffer	**5 L 1x buffer**	**1 L 10x buffer**
25 mM Tris	15 g	30 g
0.2 M Glycine	72 g	144 g
0.4 mM SDS	5 g	10 g
H_2O distilled	5 L	1 L

Make 5L 1x buffer on day of use, 10x stock can be kept for at least 2 weeks.

Protein fix/stain:	**1L**
50% MeOH	500 ml
7.5% Acetic Acid (glacial)	75 ml
2% w/v TCA (trichloro acetic acid)	20 g
0.04% Coomassie brilliant blue[g] (G-250)	0.4 g
H_2O distilled	425 ml

Destain	**1L**
23% MeOH	230 ml
7% Glacial Acetic Acid	70 ml
H2O distilled	700 ml

(**a** = Sigma, **g** = BDH, **h** = Anachem)

3.3.3 Image Capture and Quantification

1. Capture image of gel using lighbox (not UV) and camera attached to computer running Molecular Analyst version 1.1 (BioRad). Indicate that image is captured using white light setting.
2. Process using Molecular Analyst programme with settings indicating that gel is coomassie stained.
3. All BLG and β-casein bands are boxed in manually, using local background type as the setting.
4. The three BLG standards are identified and used to generate a standard curve upon which the volume report is based (i.e. protein concentration of all boxed samples on gel).
5. The volume report provides the raw data for the BLG samples.
6. The BLG samples are adjusted for loading differences using the β-casein bands of the transgenic samples standardised to the b-casein of the non-transgenic control sample: Adjusted BLG = raw BLG x (transgenic β-casein/control β-casein)
7. The adjusted BLG is the value used to calculate the actual transgenic protein concentration in the milk samples:

   (Adjusted BLG/volume loaded) x 250 (dilution factor)

   This formula gives the protein concentration in μg/μl = mg/ml.

2.4 FISH

Ensure that all instruments required are baked overnight at 180 °C, and that solutions are made up in advance and DEPC (diethyl pyrocarbonate, $C_6H_{10}O_5$, Sigma) treated.

2.4.1 Fixation and wax embedding of liver and mammary tissue

Day 1

1. Make 4% Paraformaldehyde in DEPC 1 X PBS (25 ml/tissue) (takes 2 hrs)
2. Pour into 25 ml sterile tubes and keep on ice.
3. Label 4 cryotubes/mouse (2 for mammary tissue, 2 for liver).
4. Check mice are ready for milking and sacrifice:
    - Pups removed 3 hours prior to oxytocin injection that induces lactation. 10 minutes after administration of oxytocin a general anaesthetic is given.
    - Trained staff milk the mice, leaving one gland untouched for fixation.
    - Mice are killed humanely by neck dislocation immediately prior to dissection once they have been milked.
    - Unmilked gland is dissected and cut in half before placing into paraformaldehyde solution. One lobe of liver is fixed (cut into strips to aid fixation). Remaining tissue is frozen in liquid nitrogen for RNA/DNA.
5. Tissue in paraformaldehyde fixed overnight at 4 °C on shaker.

Day 2

1. Wash fixed tissues, using shaker for all steps
    - 1 X PBS 30 min at 4 °C
    - 30 % EtOH 15 min at room temperature
    - 70 % EtOH 30 min at room temperature
    - 70 % EtOH 30 min at room temperature

Tissues can be kept in final 70% EtOH wash at 4 °C and wax embedded later.

2. Prepare tissues for histological treatment and wax embedding
    - Label histology cages with tissue data in pencil
    - Remove tissue samples and place in prelabelled cages.

- Place cages in histology basket, run RNA programme (takes 16 hours)

o 70 % EtOH	10 min	room temperature
o 70 % EtOH	1 hr 30 min	room temperature
o 64 OP (alcohol over proof)	1 hr 30 min	35 °C
o 64 OP	1 hr 30 min	35 °C
o 74 OP	1 hr 30 min	35 °C
o 74 OP	1 hr 30 min	35 °C
o 74 OP/xylene	1 hr 30 min	35 °C
o xylene	1 hr 30 min	35 °C
o xylene	1 hr 30 min	40 °C
o xylene	1 hr 30 min	45 °C
o WAX	1 hr 30 min	53 °C
o WAX	1 hr 30 min	60 °C

2.4.2 Sectioning of Wax Embedded Tissues

TESPA treatment of slides

- Oven bake slides at 180 ^{0}C overnight (or minimum of 4 hours).
- Place in baked glass racks (can be baked with slides in them).
- Treatment A:

10% HCl:75% EtOH	10 seconds
DEPC water	10 seconds
100 % acetone	10 seconds

AIRDRY

- Treatment B:

2% TESPA in acetone	10 seconds
100% acetone	10 seconds
100 % acetone	10 seconds

AIRDRY

Can be stored for up to 4 weeks with dessicant. Otherwise, redo step B.

All solutions should be made fresh on day used.

TESPA = 3-aminopropyltriethoxysilane, from Sigma.

Sectioning of tissues

- Keep parafin blocks chilled on ice throughout
- Trim off any excess wax
- Wipe waterbath and all instruments with 100 % EtOH, fill waterbath with DEPC treated H_2O and heat to 42 –45 ^{0}C.
- Mount block on microtome with label pointing towards you, section 10 micron sections until you hit the tissue (angle set at 10).
- Once tissue is reached, remove block, clean excess wax from microtome blade and wipe again with EtOH.
- Change microtome thickness to 5 microns.
- Cut a smooth ribbon of about 8 sections, catching end with forceps and extending. Use brush to remove strip from blade and place on a clean slide (NOT TESPA treated!).
- Squirt 30% EtOH under ribbon to smooth out creases.
- Dip slide vertically into waterbath and let section float off.
- Allow wax to melt (1 min), use tweezers to separate individual sections.
- Dip TESPA coated slides vertically into water and catch section.
- Leave slides until water evaporates, label and bake overnight at 60 ^{0}C.
- Store at 4 ^{0}C.

2.4.3 Prehybridisation

All dishes baked and washed with 100% EtOH before use. PFA (paraformaldehyde), Proteinase K and TEA/TCA with acetic anhydride prepared fresh each time

NB! Protocol has to be modified if want to use for genomic FISH (2.4.6)

1. 100% Xylene	20 min
2. 100% Xylene	20 min
3. 100% Xylene	5 min on shaker
4. 100 % EtOH	2 min
5. 100 % EtOH	2 min
6. 90 % EtOH	2 min

7. 70 % EtOH	2 min
8. 50 % EtOH	2 min
9. 30 % EtOH	2 min
10. 1 X PBS	2 min
11. 4% PFA	10 min
12. 1 X PBS	2 min
13. 1 X PBS	2 min
14. Prot K solution	7min 30 seconds
15. 1 X PBS	1 min
16. 4% PFA	2 min
17. DEPC $H_2O_{distilled}$	10 seconds
18. 0.1 M TEA or TCA (steps 18-20)	30 seconds
19. 0.1 M TCA + acetic anhydride (625 µl)	5 min
20. 0.1 M TCA + acetic anhydride (625 µl)	5 min
21. 1 X PBS	2 min
22. 0.85 % NaCl	2 min
23. 30 % EtOH	2 min
24. 50 % EtOH	2 min
25. 70 % EtOH	2 min
26. 90 % EtOH	2 min
27. 100 % EtOH	5 min
28. 100 % EtOH	5 min
29. 100 % EtOH	5 min

Air dry slides. Can be stored dessicated for several months before use.

2.4.4 Hybridisations for mRNA expression

- The probes used were cocktails made up of 5 30'mer oligos (sense or antisense) labelled 5' & 3' with Fluorescein, Digoxigenin or Cy3 (5' only). The oligos were synthesised by MWG.

- The oligo mix was made from stocks, taking 2 μg of each and combining in 1 ml of water. This mix was aliquoted into 200 μl batches and stored at – 20 ^{0}C to avoid contamination. 20 μl of the cocktail was used in 1ml buffer solution to give a final concentration of 40 ng/ml.

- Hybridisations were performed at 37 ^{0}C.

Protocol for mRNA FISH

1. Prehyb with buffer alone is not necessary
2. Heat sections with probe overlaid (approx 100 μl per section) for 6 minutes at 90 ^{0}C on a heating block (keep out of light).
3. Cool for one minute on ice
4. Cover with parafilm cut to size instead of usual glass coverslips
5. Place in a damp chamber and hybridise at 37 ^{0}C overnight
6. Wash sections in covered coplin jars:

   2 X SSC at 37 ^{0}C 10 min
   2 X SSC at 37 ^{0}C 30 min

   0.1 X SSC at room temperature 10 min
   0.1 X SSC at room temperature 30 min

7. Dry slides with soft tissue, avoid tissue
8. Fix tissue using Vectashield Mounting medium (Vector Laboratories) with antifading agent (no DAPI), attach methanol-cleaned coverslip, fix to slide using colourless nail varnish.
9. Visualise and capture image using FISH microscope

Hybridisation solution

(modified from protocol in "Boehringer Mannhein non radioactive *in situ* hybridisation application manual, 2nd edition 1996, p. 130)

	10ml
40 % Formamide	4 ml (100% stock)
2 X SSC (tri sodium citrate)	1 ml (20x stock)
1X Denhardt's solution	0.2 ml (50x stock)
10 % Dextran Sulphate	2 ml (50 % stock)
50 mM Phosphate Buffer	0.5 ml (1M stock)
50 mM DTT (dithiothreitol, Melford Labs.)	0.5 ml (1M stock)
0.250 mg/ml tRNA	0.25 (10 mg/ml stock)
0.5 mg/ml denatured salmon sperm DNA	0.5 ml (10 mg/ml stock)
H_2O DEPC	1.05 ml

Denhardt's Solution: 0.5 g Ficoll 400 (Pharmacia)
0.5 g Polyvinylpyrrolidone
0.5 g BSA, fraction V
50 ml DEPC distilled autoclaved H_2O

Do not autoclave or DEPC treat once made as this will destroy the proteins.

2.4.5 Nuclear FISH for intronic sequence detection.

Secondary detection by anti- DIG/Fluorescein with suitable fluorescent label

1. Hybridise overnight as for mRNA FISH
2. Wash slides as for mRNA FISH
3. Overlay with incubation buffer without antibody to block for 10/15 minutes
4. Add incubation buffer with antibody (1/125 or 1/250 dilution anti DiG, 1/100 dilution of anti-fluorescein), incubate at room temperature for 30 – 45 minutes.
5. Rinse sections 3 X 5 minutes in wash buffer.
6. Mount using Vector Mounting medium for fluorescence with DAPI counterstain to identify intact nuclei.

Incubation Buffer

100 mM Tris-HCl pH 7.4

150 mM NaCl

1% Blocking Reagent (serum, same as antibody source)

Wash Buffer

100 mM Tris-HCl pH 7.4

150 mM NaCl

2.4.6 Nuclear FISH for genomic sequence detection

As the genomic target is less abundant than transcribed RNA, the signal has to be considerably amplified (6 – 9 times). The genomic DNA is also much less accessible than the RNA, probably due to the chromosomal proteins. The prehybridisation protocol therefore has to be modified for the probe to gain access to the target.

1. Follow normal prehybridisation protocol up to and including step 21 then:

    a. 1 X PBS at 37 °C for 2 minutes
    b. 0.1% w/v Pepsin in 0.2M HCl 37 °C for 15 minutes (may need to vary depending on tissue, have used between 10 and 20 min.)
    c. 1 X PBS at room temperature for 1 minute
    d. 4% PFA at 4 °C for 2 minutes
    e. H_2O 10 seconds
    f. Repeat steps 18 – 21 and carry on with rest of protocol as normal

2. Hybridise and amplify signal as for intronic FISH with antibodies to Fluorescein (Rat anti-FITC conjugated to FITC, Serotec). The signal is amplified 6-9 times to get a good image.

2.4.7 Antibodies & sera used

Antibodies:

Against	Hapten	Host	Fragment	Source
DiG	Rhodamine	Sheep	Fab	Boehringer 1 207 750
DiG	Fluorescein	Sheep	Fab	Boehringer 1 207 741
Biotin	Fluorescein	Goat	Fab	Vector Labs SP-3040
Biotin	Cy3	Goat	Fab	Rockland 700-104-098
Fluorescein	Fluorescein	Rat	IgM	Serotec MCA 1332 F
Rhodamine	none	Rabbit	IgG	Molecular Probes A-6397
Rabbit IgG	Cy3	Goat	IgG H & L	Rockland 611-104-122

Sera:

Rat	Serotec C 138 B
Goat	Diagnostics Scotland T 028
Sheep	Diagnostics Scotland T 031

2.4.8 Probes used

BLG: Exonic and intronic probes based on sequences in GenBank X12817, genomic/promoter probes based on sequences in X68105

EXON PROBES:

1. Position 1621 – 1250
2. Position 1681 – 1710
3. Position 2591 – 2620
4. Position 3851 – 3880
5. Position 4581 – 4610

INTRON PROBES:

1. Position 1801 – 1830
2. Position 1861 – 1890
3. Position 1981 – 2010
4. Position 2041 – 2070
5. Position 2161 – 2190

GENOMIC/PROMOTER PROBES:

1. Position 3545-2574
2. Position 2656-2685
3. Position 2769-2798
4. Position 2957-2986
5. Position 3079-3098
6. Position 3145-3174
7. Position 3243-3272
8. Position 3323-2252
9. Position 3445-3474
10. Position 3594-3623

β-casein: Exonic and intronic probes based on sequences in GenBank X13484

EXON PROBES:

1. Position 9291 – 9320
2. Position 9411 – 9440
3. Position 9471 – 9500
4. Position 9531 – 9560
5. Position 9621 – 9650

INTRON PROBES:

1. Position 5041 – 5070
2. Position 5141 – 5170
3. Position 5201 – 5230
4. Position 6311 – 6340
5. Position 6381 – 6410

WAP: Intronic probes based on sequences from U38816

INTRON PROBES:

1. Position 2651 – 2680
2. Position 3121 – 3150
3. Position 3591 – 3620
4. Position 4031 – 4060
5. Position 4621 – 4650
6. Position 4931 - 4960

LacZ: Exonic probe based on sequences from GenBank U13184

EXON PROBES:

1. Position 861 – 890
2. Position 921 – 950
3. Position 1381 – 1410
4. Position 1441 – 1470
5. Position 1501 – 1530
6. Position 1621 – 1650
7. Position 1681 – 1710
8. Position 1851 – 1880
9. Position 1951 – 1980
10. Position 2071 – 2100

Chapter 3

Analysis of Backcrossed BLG/7 Variegating Mice

3.1 Introduction

Most examples of variegation are caused by mutations (van Deutekom *et al* 1996, Spiegel 1997, Gottlieb *et al* 2001), translocations (Muller 1930, Milot *et al* 1996) or transgene insertions into unfavourable positions in the genome (Dobie *et al* 1996, Henikoff 1996). The latter two examples are known as position effect variegation (PEV). *Drosophila* has dominated the research into PEV, a small genome and fast generation time facilitating analysis of many well-characterized and distinct strains. The work has uncovered many modifiers of PEV (Henikoff 1990, Singh 1994). Advances have enabled studies in other organisms (most notably plants, yeast, fungi and rodents), and homologues of modifiers of PEV already identified in *Drosophila* have been found (Singh *et al* 1991, Cowell & Austin 1997). Some of these homologues are interchangeable between species and can exert a modifying effect on variegating genes in place of the endogenous modifier gene (Stankunas *et al* 1998, Aagaard *et al* 1999, Festenstein *et al* 1999).

The relative abundance of modifier proteins can influence transgene expression (Festenstein *et al* 1999, McMorrow *et al* 2000). Modifiers can vary between strains, both qualitatively and quantitatively, resulting in strain-specific modification of expression (Schweizer *et al* '98). Transgenic mouse line BLG/7 was created on a mixed F1 CBA:C57BL/6 background (Dobie *et al* 1996). BLG/7 mice express the ovine β-lactoglobulin transgene during lactation. Although correctly expressed both temporally and spatially, transgene expression levels differ between siblings, due to variegated expression at the cellular level in the mammary gland (Dobie *et al* 1996). The integration of the transgene as a concatemer of approximately 25 intact copies close to centromeric heterochromatin suggested that the observed variegation could be related mechanistically to PEV. Alternatively, the observed variegation could be caused by the presence of a mixed set of modifier genes from the two different

genetic sources that make up the F1 background of the line. As a C57BL/6 background has been demonstrated to influence transgene expression levels for more than one transgene (Koetsier *et al* 1996, Schweizer *et al* 1998, Sutherland *et al* 2000), strain-specific modification of expression resulting in variegated BLG/7 expression was a strong possibility.

Ideally, strain specific effects on transgene expression should be investigated in a congenic background. Backcrossing was therefore initiated into the two respective strains (see appendix C3 for mating regimes). At the 3rd generation, there was little evidence for any strain–specific effects affecting gene expression (Dobie *et al* 1996), which suggested that no dominant single modifier locus was responsible for the variegated expression. There was, nonetheless, a small trend towards a lower mean expression level of the transgene for the CBA backcross population, suggesting the possible existence of minor modifier loci. This trend was also weakly apparent for the C57BL/6 backcross population. Although the observed trend was not statistically significant for either backcross population, it did not rule minor modifier loci as a causative agent of variegation, so backcrossing was continued for a further 10 generations. The 13th generation backcross mice are considered functionally congenic (99.4%) (M. A. Suckow, P. Danneman & C. Brayton, "The Laboratory Mouse", CRC Press).

3.2 AIM

The aim of this investigation was to establish conclusively whether previous work on 3rd generation BLG/7 (Dobie *et al* 1996), which concluded that variegation was a product of the transgene itself at its site of integration, held true in the 13th generation backcross mice.

3.3 Experimental Approach

We wished to evaluate the effect of backcrossing into a congenic background on BLG/7 transgenic protein expression levels in the milk of lactating females.

A backcrossing regime to inbred CBA and C57BL/6 colonies was therefore set up for BLG/7. The line was also maintained on the original F1 background in parallel with the backcrosses, enabling sampling of this group to establish if any changes in expression levels occurred after multiple passages through the germline. BLG protein content in the milk samples was determined by quantitative SDS gel electrophoresis as described earlier (2.3).

There are several reasons why we sampled our mice at the 13th generation. Generation time was a major factor. 13 rounds of backcrossing take 6½ years, assuming a generation turnaround of 6 months. Additionally, maintenance becomes increasingly difficult with a limited number of inbred transgenic mice. The equivalent BLG/7 F1 outbred mice were also sampled so that the expression levels of all three groups could be compared to each other and to the original samples from the 3rd generation of transgenic mice, both backcrossed and F1 genotypes. At the 13th generation of inbreeding, the mice are considered congenic (99.4%). Apart from material close to the transgene integration site (up to 0.6% of genome), all loci (99.4% of genome) should contain genetic material solely derived from a single inbred background. Any strain-linked loci that may be responsible for the observed variegated expression should therefore be evident at this stage. The two strains in this case are CBA and C57BL/6.

3.4 Results

3.4.1 Strategy and critique of backcrossed BLG/7 milk sample analysis

Sample generation

Expression levels of BLG peak at mid-lactation, which in mice is the 11th day of lactation (Whitelaw *et al* 1992). The day of birth is counted as day 0. The majority of milk samples were taken at this time point, although a few were taken at either day 10 or day 12.

An intensive mating regime was set up to generate 13th generation backcross females from both backgrounds and an equivalent number of control outbred BLG/7 F1 females. The control females served a dual purpose. The main function of the control females was as an outbred comparison to the two inbred backgrounds. The second function was as a means to test for any changes in BLG expression connected to generation time, when compared to the expression levels of the original outbred mice (Simons *et al* 1987, Whitelaw *et al* 1992, Dobie *et al* 1996). All three lines have been maintained in parallel through male germline transmission. An equal number of generations separate the lines and the founder mice, which were generated by microinjection (Simons *et al* 1987).

Milk samples at the 11th day of lactation were collected from 21 C57BL/6 and 18 CBA BLG/7 females. A control group of 19 F1 outbred BLG/7 females were also sacrificed for milk at the same lactational stage. The females were all between 3 and 4 months old, enabling us to discount any age-related factors (1.5.1) when analysing the protein content in their milk. Although the age and lactational stages were the same for all females, milk collection spanned several months due to the reluctance of the backcross mice to breed. Backcross mice gave rise to much poorer milk yields than their F1 counterparts. About 1/5th of the females did not release any milk from their teats, even after chemical induction of milk release using oxytocin. This necessitated the setting up of more breeding couples in order to generate enough lactating females to ensure that sufficient milk samples were provided for the successful conclusion of the experiment. Since collection was spread over a considerable time, milk samples were frozen fresh. As soon as all the samples had been collected, they were processed together. This treatment of the samples was designed so that any variability in expression could not be attributed to differences in the buffers and solutions, which could arise if they were processed separately.

Sample processing

All sample processing is described in the materials & methods section (2.3). The BLG standards were from the same source, as were all chemicals used in the protein gels, buffers, fixes and stains. The BLG standards were purified from sheep milk and

quantified using the micro–Kjeldahl technique (Aschaffenburg & Drewry 1957) prior to the commencement of the project.

Quantification of protein concentration was done in triplicate sets for each sample, using pixel intensity comparisons to the three BLG standards run on every protein gel. Analysis of the gels used the Multi Analyst Programme version 1.1 (BioRad). Three dilutions of BLG standard provided an accurate standard curve, with all BLG bands boxed in manually. Loading differences were corrected for using β-casein as an internal control standardised to the non-transgenic sample present on every gel. The same non-transgenic control sample was run for each line. The arithmetic mean of the three measurements was used in determination of BLG expression levels for each individual sample.

The backcross milk samples were compared to each other and to the outbred results. The expression profiles of all three groups were compared to earlier published data for this line (Simons *et al* 1987, Whitelaw *et al* 1992, Dobie *et al* 1996).

All gels run contained three BLG standards, one non-transgenic control milk sample and nine transgenic milk samples. A total of 9 C57BL/6 BLG/7 protein gels, 6 CBA BLG/7 gels and 9 outbred protein gels were run. A gel from each backcross group is reproduced in Figure 3.1. Figure 4.1 contains an example of the outbred protein gels (hemizygotes). Although the protein gels were Coomassie stained, greyscale depictions were preferred, as the protein bands are more clearly defined. All analysis of the images was done with the gels defined as coomassie stained by the software package (2.3.3).

3.4.2 Backcross milk samples display down-regulation of all milk proteins

The first observation of differences between the backcross and F1 outbred milk samples occurred before any quantification of BLG expression levels took place. Trial gels were run to determine the ideal sample loading volume, non-transgenic control and BLG concentrations for the standard curve (the BLG standards had to fall within the expression range of the samples and below the saturation point of the

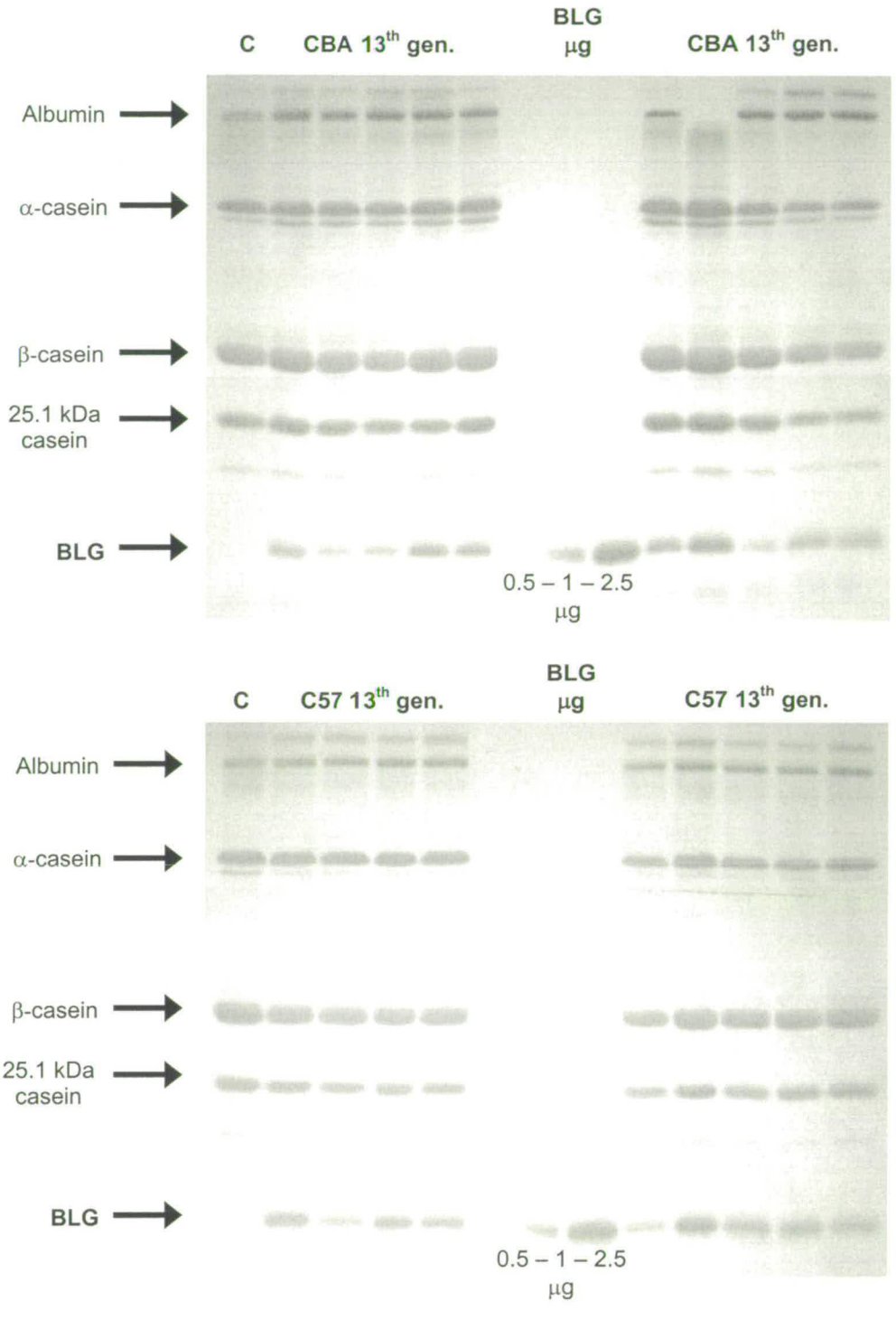


Figure 3.1 **SDS PAGE of milk protein samples for CBA & C57BL/6 13th generation backcross mice** C = non-transgenic control milk, BLG = β-lactoglobulin standard. Transgenic samples were loaded 50 µl of 1/250 dilution.

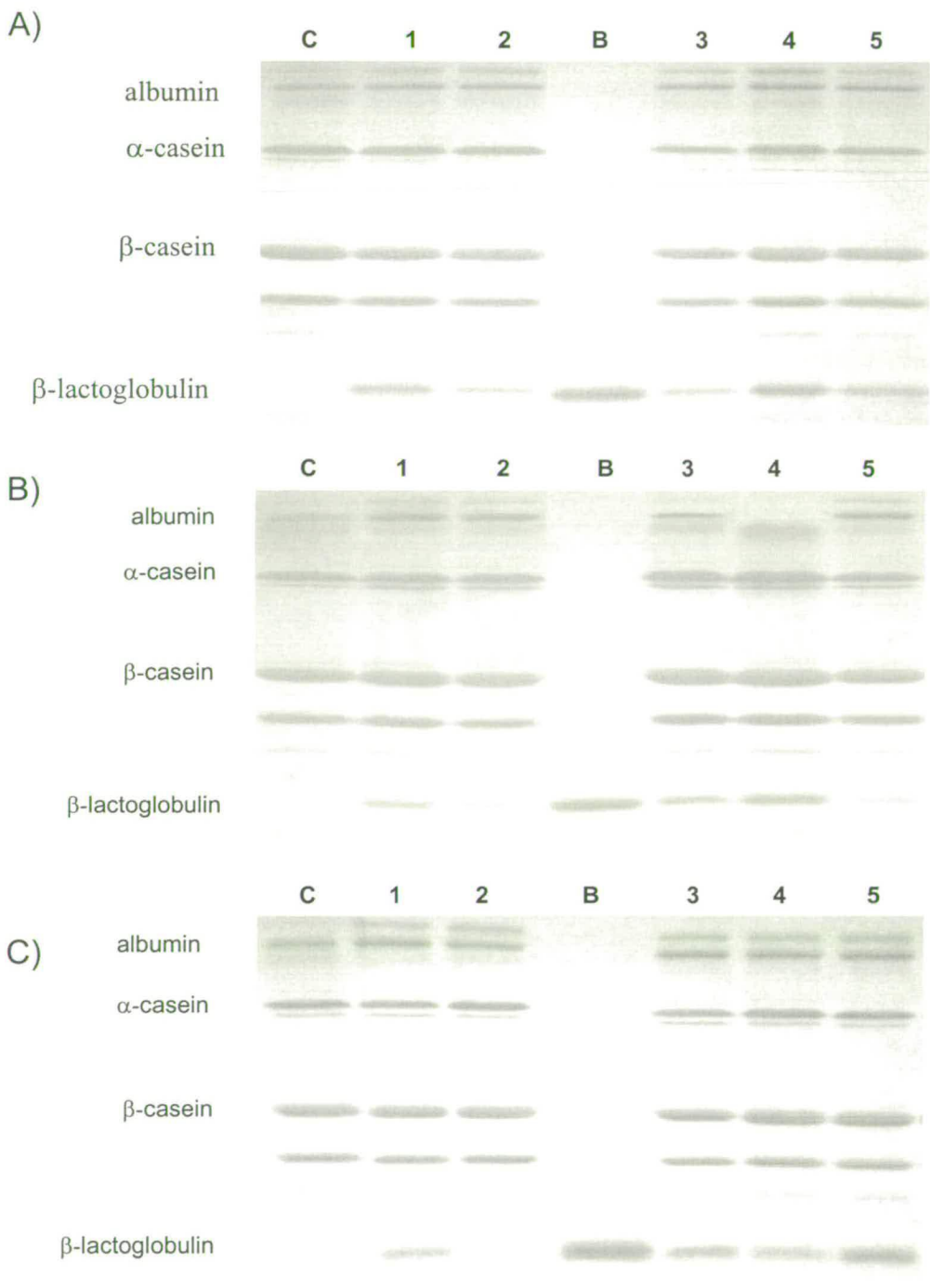


Figure 3.2 **SDS PAGE of milk protein samples for all 3 genotypes.**

A) is 13th generation CBA backcross samples, B) is 13th generation backcross C57BL/6 and C) is outbred samples. C = non-transgenic control milk, B = β-lactoglobulin standard: 2 μg for A) + B), 3 μg for C). A) + B) transgenic samples were 50 μl of 1/250 dilution, C) had 40 μl loading

image capture system). An immediately noticeable characteristic of the protein gels was that the inbred milk sample proteins were consistently less abundant than in equivalent F1 outbred samples (Figure 3.2). This down-regulation was not limited to BLG, but appeared to include all the endogenous milk proteins. Comparative protein band intensities were only achieved when 50 µl of a 1/250 dilution were loaded for the backcross milks, as opposed to 40 µl for the rest. This gives a rough estimate of $1/5^{th}$ reduction in overall protein expression in the backcrossed samples.

3.4.3 Variegation occurs irrespective of genetic background, with mean expression levels remaining stable in an outbred background

The second observation made by visual examination of the protein gels from all three backgrounds was that BLG expression varied between individuals within each group (Figure 3.1, 3.2 & 4.1). Within a genotype, the endogenous milk proteins did not appear to vary. Both backcross genotypes exhibited high levels of variegation for BLG expression. It could be suggested this provides conclusive proof that the observed variegation is an inherent property of the transgene at the integration site, independent of genetic background. The continued variegation of BLG expression within the F1 outbred mice also indicates that passage number through the germline does not affect the variegating properties of the transgene.

Raw data of both BLG and β-casein concentrations were collected in triplicate sets (Appendix A1) and BLG protein levels quantified (2.3.3). This gave three separate BLG concentration measurements for each milk sample in mg/ml (Table 3.1). The arithmetic mean of the three measurements was used for the calculation of individual expression levels. A representation of the BLG protein yields for all three groups of mice is shown in Table 3.2 and Figure 3.3 (a-c).

The data generated for BLG expression levels (Table 3.2) was subjected to statistical analysis. Mean expression, standard deviation and coefficient of variation was compared between the three groups, and comparisons were additionally made to previously published data (Dobie *et al* 1996) for the 3^{rd} generation backcrosses (Table 3.3). No significant differences were found between the 3^{rd} generation inbred

	BLG measured mg/ml	BLG average mg/ml		BLG measured mg/ml	BLG average mg/ml
CBA 7I 1	3.26	3.15	**CBA 7I 11**	3.07	2.88
	2.73			2.84	
	3.34			2.72	
	3.25				
			CBA 7I 12	3.55	4.54
CBA 7I 2	2.71	2.52		5.55	
	2.07			4.52	
	2.77				
	2.52		**CBA 7I 13**	4.57	4.08
				3.98	
CBA 7I 3	3.49	3.50		3.9	
	3.4			3.85	
	3.62				
			CBA 7I 14	3.84	4.84
CBA 7I 4	4.79	4.73		5.04	
	4.82			4.47	
	4.59			5.36	
				5.47	
CBA 7I 5	4.05	4.17			
	4.25		**CBA 7I 15**	4.74	4.70
	4.2			4.98	
				4.38	
CBA 7I 6	3.45	3.76			
	3.9		**CBA 7I 16**	3.72	4.78
	3.93			5.26	
				4.75	
CBA 7I 7	5.13	6.26		5.39	
	6.72				
	6.58		**CBA 7I 17**	4.17	4.18
	6.59			4.62	
				3.04	
CBA 7I 8	2.47	2.68		4.88	
	2.78				
	2.78		**CBA 7I 18**	4.73	3.66
				3.29	
CBA 7I 9	5.31	5.38		3.4	
	5.35			3.21	
	5.49				
CBA 7I 10	3.25	5.59			
	6.17				
	6.46				
	6.49				

Table 3.1 **A)** BLG measurements for 13th generation BLG/7 CBA backcrosses

	BLG measured mg/ml	BLG average mg/ml		BLG measured mg/ml	BLG average mg/ml
C57 7I 1	6.24	6.22	**C57 7I 11**	6.23	6.23
	6.26			6.13	
	6.16			6.33	
C57 7I 2	4.10	4.13	**C57 7I 12**	5.71	5.88
	4.14			5.97	
	4.16			5.98	
C57 7I 3	5.83	5.88	**C57 7I 13**	8.25	8.34
	5.90			8.40	
	4.86			8.37	
	5.90				
			C57 7I 14	0.12	1.06
C57 7I 4	4.12	4.14		1.14	
	4.22			1.93	
	4.09				
			C57 7I 15	6.08	6.27
C57 7I 5	3.76	3.78		6.29	
	4.00			6.44	
	3.59				
			C57 7I 16	5.06	5.18
C57 7I 6	7.47	7.43		5.40	
	7.37			5.09	
	7.45				
			C57 7I 17	5.61	5.83
C57 7I 7	4.72	4.77		5.98	
	4.69			5.91	
	4.89				
			C57 7I 18	6.45	6.57
C57 7I 8	5.23	5.29		6.64	
	5.28			6.61	
	5.36				
			C57 I7 19	9.80	9.45
C57 7I 9	5.88	5.71		9.96	
	5.31			8.59	
	5.74				
	5.89		**C57 I7 20**	5.49	5.51
				5.67	
C57 7I 10	4.48	4.54		5.38	
	4.63				
	4.50		**C57 I7 21**	9.79	9.61
				9.84	
				9.21	

Table 3.1 **B)** BLG measurements for 13th generation BLG/7 C57BL/6 backcrosse

	Average BLG mg/ml	Mean mg/ml	STD	CV
CBA 7I 1	3.28	**4.31**	**1.18**	**27.30**
CBA 7I 2	2.67			
CBA 7I 3	3.50			
CBA 7I 4	4.73			
CBA 7I 5	4.17			
CBA 7I 6	3.76			
CBA 7I 7	6.63			
CBA 7I 8	2.68			
CBA 7I 9	5.38			
CBA 7I 10	6.37			
CBA 7I 11	2.88			
CBA 7I 12	4.55			
CBA 7I 13	3.91			
CBA 7I 14	5.29			
CBA 7I 15	4.70			
CBA 7I 16	5.13			
CBA 7I 17	4.56			
CBA 7I 18	3.30			

	Average BLG mg/ml	Mean mg/ml	STD	CV
C57 7I 1	6.22	**5.80**	**1.92**	**33.07**
C57 7I 2	4.13			
C57 7I 3	5.88			
C57 7I 4	4.14			
C57 7I 5	3.78			
C57 7I 6	7.43			
C57 7I 7	4.77			
C57 7I 8	5.29			
C57 7I 9	5.71			
C57 7I 10	4.54			
C57 7I 11	6.23			
C57 7I 12	5.88			
C57 7I 13	8.34			
C57 7I 14	1.06			
C57 7I 15	6.27			
C57 7I 16	5.18			
C57 7I 17	5.83			
C57 7I 18	6.57			
C57 I7 19	9.45			
C57 I7 20	5.51			
C57 I7 21	9.61			

Table 3.2 Average BLG expression with mean, SD & CV for 13th generation BLG/7 backcrosses on CBA and C57BL/6 backgrounds

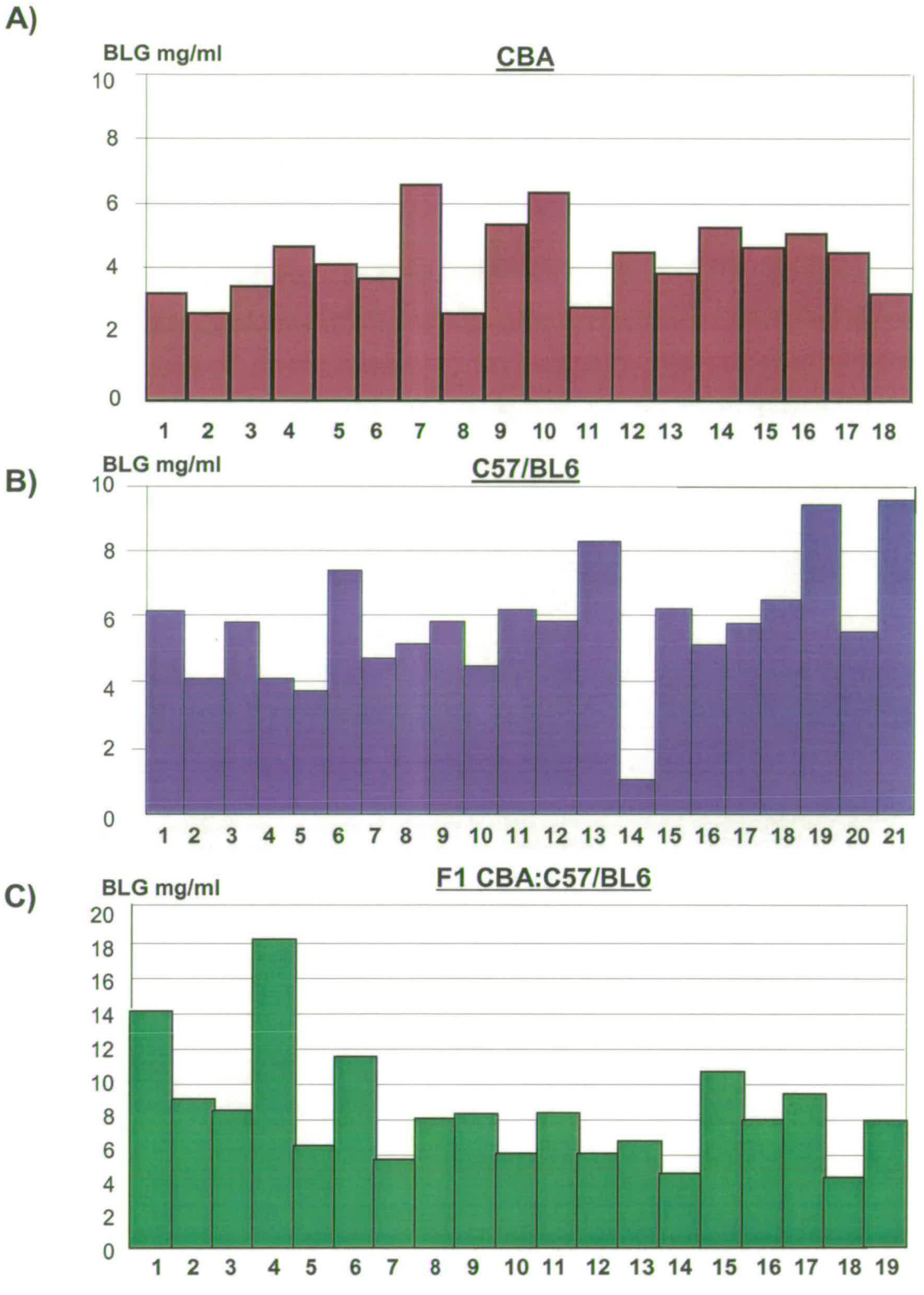


Figure 3.3 **Bar charts representing individual BLG milk protein yields in mg/ml**

A = CBA backcross samples, **B** = C57BL/6 backcross, **C** = outbred samples

Each bar corresponds to the mean BLG expression of the individual numbered below (i.e, in chart A individual 1 = CBA 1 etc)

mice and my F1 CBA:C57BL/6 outbred BLG expression values. The mean expression value of 8.2 mg/ml compares well with the two previously published results for BLG/7 F1 mice of 7.4 mg/ml +/- 1 (McClenaghan *et al* 1995) and 9.5 mg/ml +/- 4.9 (Dobie *et al* 1996). Comparisons of the coefficient of variegation for my F1 BLG/7 outbred CBA:C57BL/6 mice (42%) and both 3rd generation backcross (45% + 54%) (Table 3.3) and F1 (52%) data (Dobie *et at* 1996) also fail to unearth any significant differences.

MICE	MEAN	*SD* (standard deviation)	CV (coefficient of variation)
CBA 13th gen. backcross	4.2 mg/ml	1.03	24.5%
C57BL/6 13th gen. backcross	5.8 mg/ml	1.92	33.2 %
Outbred F1 CBA:C57BL/6	8.2 mg/ml	3.46	42%
CBA 3rd gen. backcross*	6.3 mg/ml	2.8	45%
C57BL/6 3rd gen. backcross*	9.7 mg/ml	5.3	54%

* = from Dobie *et al* 1996 & K.W. Dobie PhD Thesis 1996.

Table 3.3 Backross, outbred and 3rd generation milk protein results

Transgene expression levels are therefore maintained in the parental population. The result confirms that variegation is successfully maintained over many generations without any progressive silencing through repeated rounds of germline transmission. This indicates strongly that both the cause and mechanism responsible for variegation at this locus remain unchanged.

3.4.4 Endogenous proteins show evidence of inbreeding depression, which selectively represses the BLG transgene

Initial observations of protein gel staining intensities suggested that endogenous milk protein production was reduced in the backcross populations (3.4.2). The reduction in BLG expression levels and variance around the mean (Table 3.3) for the backcross samples was therefore expected. As the outbred BLG/7 population did not vary from results gained in earlier experiments (3.4.3), samples from all three populations were

sent out for micro-Kjeldahl analysis in order to determine total protein content and enable comparisons between the populations. The technique works by measuring the total nitrogen content of a sample, from which an estimation of the total protein content can be made.

Micro-Kjeldahl analysis requires a minimum of 100 µl per sample for analysis, with all sample measurements duplicated. Poor milk yields from the backcross mice resulted in milk samples that averaged less than 100 µl from each mouse, with 30 µl of this used in the initial sample preparation. It was therefore necessary to pool the samples for each group. Duplicate samples from each of the two pools (CBA and C57BL/6 13th generation BLG/7) and individual outbred F1 BLG/7 samples were sent to the Hannah Institute, where the analysis of total protein content was performed by Dr Eleanor Noble. The mean BLG expression level is already known for each pool and the individual outbred F1 BLG/7 samples (Tables 3.2 & 3.3), so endogenous protein content estimation was easily calculated by subtracting the known BLG content from individual results (Table 3.4).

	Total Protein	- BLG	Endogenous protein expression	BLG protein expression
CBA 13th gen. backcross	85 mg/ml	81 mg/ml	**85%**	**53%**
C57BL/6 13th gen. backcross	87 mg/ml	81 mg/ml	**85%**	**71%**
Outbred F1 CBA:C57BL/6	103 mg/ml	95 mg/ml	**100%**	**100%**

Table 3.4 Micro-Kjeldahl analysis results

Compared to the F1 outbred results, both backcross populations had their endogenous milk protein levels reduced by 15% (Table 3.4). This ties in neatly with my estimate of 1/5th reduction based on visual examination and loading adjustments (3.4.2). I have attributed the down-regulation of endogenous milk protein production to the effects of inbreeding depression as both genotypes exhibited identical levels of depression. The depressed expression of endogenous milk proteins may extend to all endogenous protein expression, and could be a major contributory factor responsible

for the loss of fecundity in the backcross populations. Others have reported loss of both fecundity and viability with increased levels of inbreeding in other systems (Morjan *et al* 1999).

Next, the reduction in BLG content in the backcross populations was compared with the outbred F1 BLG/7 mice. Unexpectedly, the reduction in BLG expression exceeds 15% in both backcross populations (Table 3.4). This result suggests that the factors contributing to inbreeding depression may selectively suppress the transgene.

3.4.5 Strain-specific modification of expression is superimposed on inbreeding depression effects, with 13th generation backcross populations differing from each other and outbred BLG/7 populations

The reduction in BLG expression in the backcross mice does not only exceed the 15% reduction of endogenous gene expression, but the level of additional reduction varies considerably between the two genetic backgrounds (Table 3.4). The C57BL/6 13th generation backcrossed BLG/7 mice exhibit a 29% reduction in BLG expression levels, whereas the CBA background exhibits a staggering 47% reduction in mean expression. This clearly demonstrates that there are strain-specific differences between the two backgrounds that can exert an effect on transgene expression.

The backcross mice results differ significantly ($p<0.01$), both from outbred BLG expression values and from the outbred variance (see Figure 3.4). Mean expression values and coefficient of variance values are lowered for both backcross groups compared to the F1 outbred mice (Table 3.3). This means that the 13th generation backcross populations also differ significantly from the 3rd generation backcrosses. A mild trend towards a reduction in mean expression and variegation apparent in the 3rd generation of CBA backcrossed BLG/7 mice is substantially enhanced in the 13th generation CBA BLG/7 backcross mice. Although no such trend towards lowered expression and variance was observed in the 3rd generation of C57BL/6 backcrossed BLG/7 mice, 13th generation backcross mice provide ample evidence that down-regulation of expression and variance also takes place in this background. More surprisingly, the backcross mice also differ significantly in mean expression and

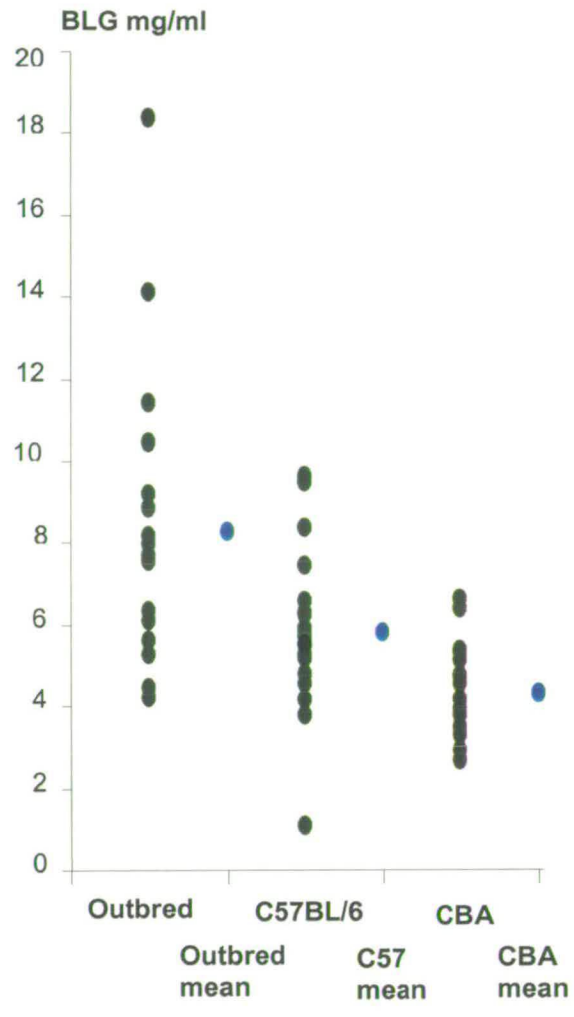


Figure 3.4 Scatterplot showing individual BLG milk protein content and mean BLG expression for all 3 genotypes .

variance from each other. Mean C57BL/6 BLG expression levels are significantly higher than the levels observed in CBA mice ($p<0.01$). The variance of BLG protein levels in the C57BL/6 population is also significantly greater than in the CBA population ($p<0.05$). Therefore, although BLG/7 still variegates after extensive backcrossing into the CBA and C57BL/6 genetic backgrounds, significant differences in expression levels between these two populations have been uncovered.

Fluorescent *in situ* hybridisation for BLG and β-casein expression (2.4.4) was performed on mammary tissue from backcrossed C57BL/6 BLG/7 mice. The results demonstrate clearly that the observed variable expression of BLG within the backcross population is due to variegation, with only a proportion of the mammary epithelial cells expressing BLG whilst all of them express β-casein (Figure 3.5). This had already been demonstrated previously for the parental population (Dobie *et al* 1996), but the result confirmed that the mechanism through which the variegation operates had not changed.

3.4.6 Inbreeding- and strain-dependent down-regulation of BLG/7 expression is not heritable

The probable multilocus effects of inbreeding and background could reasonably be attributed to subtle changes and differences in protein concentrations and composition. How these multilocus effects act on the transgene locus is unknown, though it is highly possible that some of them modify the locus epigenetically. Direct investigation into the nature of epigenetic modification at the locus would be very difficult, as the locus itself is highly methylated even in the active state (Stephen Kwan & C.B.A. Whitelaw, personal communication). The variegated nature of gene expression in the mammary tissue would also confuse any results, as dissection of expressing cells from non-expressing cells is extremely technically demanding, and outside available resources. One practical avenue of investigation was open, the possibility that the epigenetic modifications may be inherited through the germline. Inbreeding depression, progressive silencing or a combination of the two could be causing the reduction in transgene expression in the backcrossed populations. To discriminate between these possibilities, transgenic progeny from backcrossed

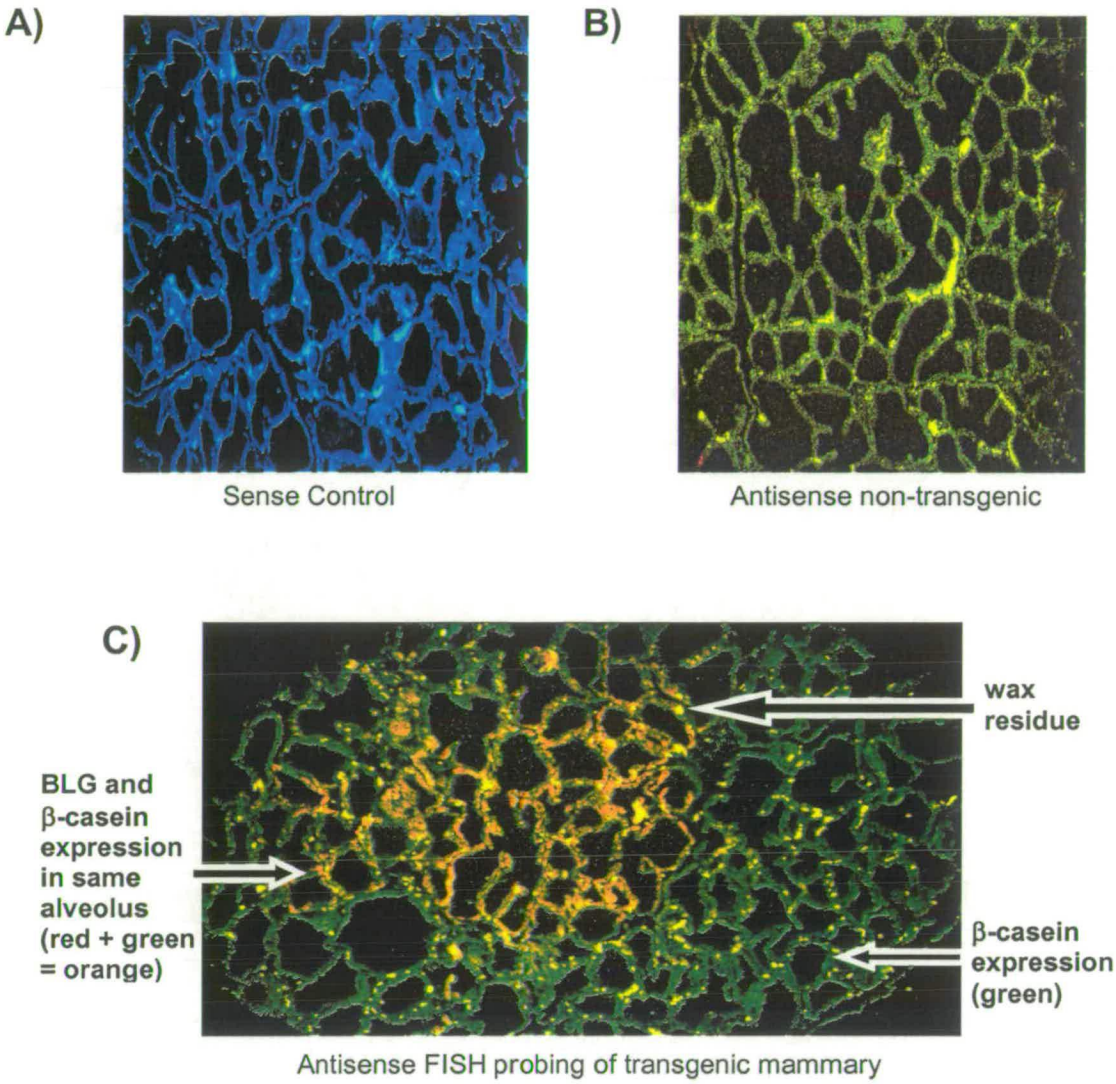


Figure 3.5 Fluorescent *in situ* hybridisations of backcrossed mammary tissue

Probes used were 5 30'mer oligo cocktails of BLG (Red, Cy3 labelled) and β-casein (Green, Fluorescein labelled) sequences. There were control sense probes (A) and antisense probes for transcription (B & C). Tissue from 18th generation backcross.

A) Sense probing of transgenic tissue, both probes. Signal is due to autofluorescence in the blue light range **B)** Antisense probing of non-transgenic tissue, both probes used. Only β-casein is expressed (green signal). **C)** Antisense probing of transgenic backcrossed C57BL/6 tissue, both probes used (β-casein is expressed in all epithelial cells, where BLG is also expressed the signal is orange due to mixing of red and green probe signals. Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals

C57BL/6 BLG/7 males mated to purebred CBA females were examined for BLG/7 expression. Restoration of expression to reference F1 animal levels would confirm inbreeding depression, whilst persistent depression would support progressive silencing with inheritance of epigenetic modification at the transgene locus. The reciprocal cross between CBA BLG/7 and C57BL6 was not performed because the CBA backcrossed line had been discontinued for logistic reasons.

The C57BL/6 BLG/7 males mated to purebred CBA females were 16^{th} and 17^{th} generation backcross. The progeny are of comparable genetic composition (50% CBA, 50% C57BL/6) to reference outbred F1 BLG/7 animals, and to differentiate between the two populations will be referred to as outcrossed F1 BLG/7. Outcrossed F1 BLG/7 females were mated to non-transgenic stock F1 males. Milk samples were collected on day 11 of lactation, and 40 µl of 1/250 diluted samples were run on the protein gels. The endogenous proteins were compared to, and loading adjustments were made using the same non-transgenic control sample that was used for the original F1 outbred samples (Figure 3.6). BLG concentrations were assessed as previously (see Table 3.5 & Figure 3.7). Mean expression, standard deviation and cv were calculated as before (see Table 3.6). The F1 outcrossed BLG/7 results were compared to the F1 outbred BLG/7 data collected previously (Table 3.7). Complete reversion to original BLG mean expression levels was observed for the F1 outcrossed BLG/7 mice. It was evident from the level of staining on the protein gels that expression had reverted to F1 levels for both endogenous proteins and BLG. Variegation of BLG expression was nonetheless very much evident in these mice.

MICE	MEAN	*SD*	CV
CBA 13^{th} gen. backcross	4.2 mg/ml	1.03	24.5%
C57BL/6 13^{th} gen. backcross	5.8 mg/ml	1.92	33.2 %
Outbred F1 CBA:C57BL/6	8.2 mg/ml	3.46	42%
Outcrossed 50:50 CBA:C57BL/6	**8.23 mg/ml**	**1.79**	**21.8%**

Table 3.7 Backcross, outbred & outcross milk protein results

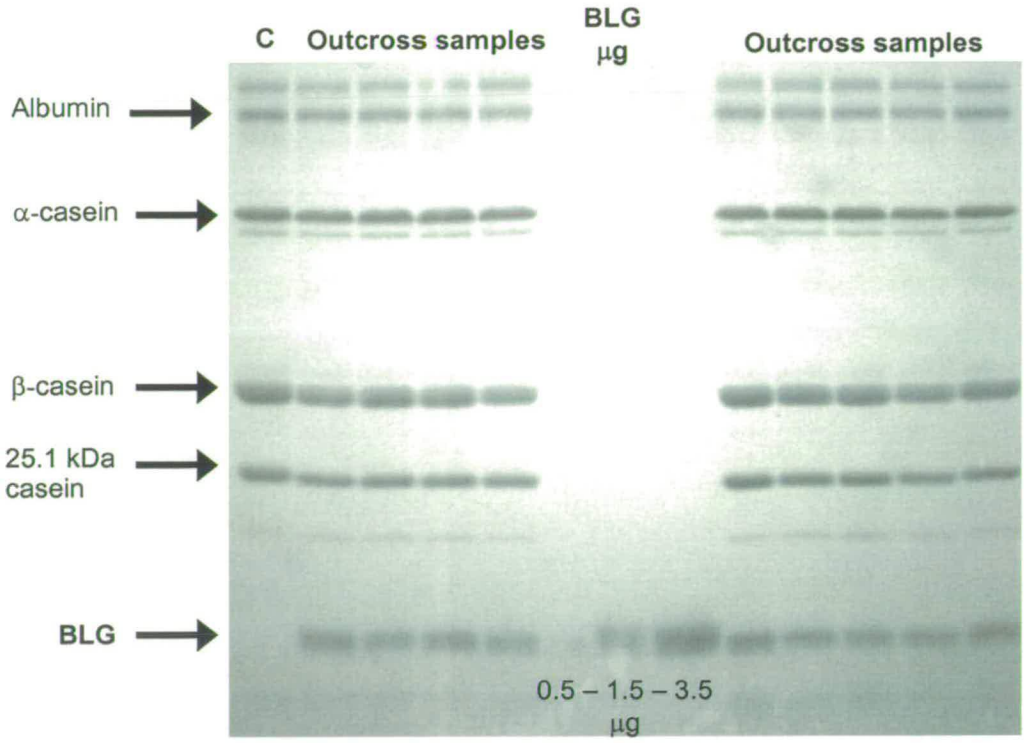


Figure 3.6 **SDS PAGE of milk protein samples for outcrossed mice**

C = non-transgenic control milk, BLG = β-lactoglobulin standard. Transgenic samples were loaded 40 µl of 1/250 dilution.

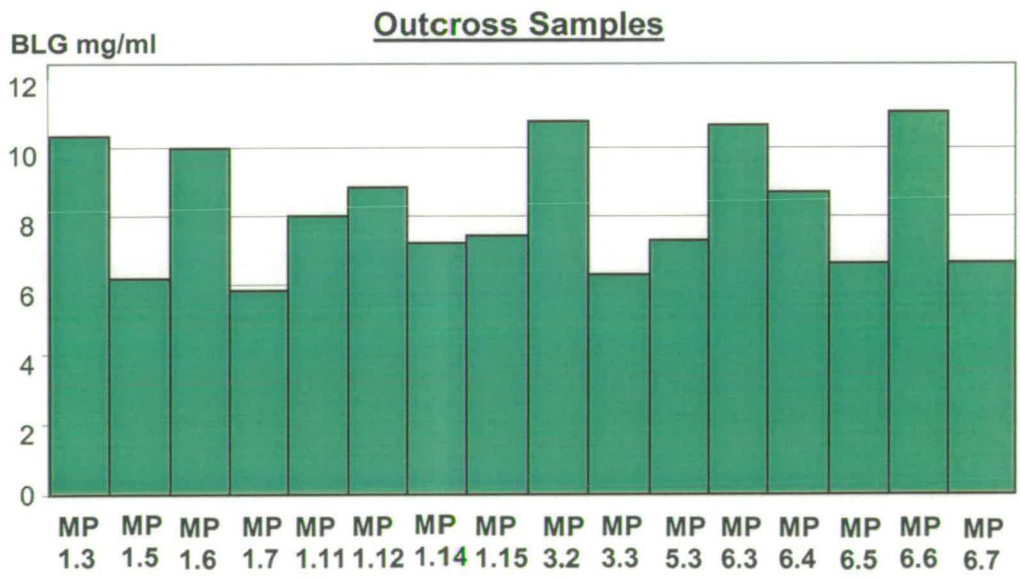


Figure 3.7 **Bar chart representation of individual outcrossed BLG milk protein yields in mg/ml**

	[BLG] mg/ml	Average BLG mg/ml
MP 1.3	10.31	**10.28**
	10.13	
	10.42	
MP 1.5	6.11	**6.20**
	6.23	
	6.26	
MP 1.6	9.92	**9.95**
	9.99	
	9.96	
MP 1.7	5.73	**5.85**
	5.94	
	5.88	
MP 1.11	7.90	**7.99**
	7.88	
	8.18	
MP 1.12	8.74	**8.82**
	9.55	
	8.17	
MP 1.14	7.06	**7.22**
	7.82	
	6.78	
MP 1.15	7.09	**7.44**
	8.10	
	7.14	

	[BLG] mg/ml	Average BLG mg/ml
MP 3.2	10.82	**10.72**
	11.04	
	10.30	
MP 3.3	6.60	**6.31**
	6.07	
	6.28	
MP 5.3	7.43	**7.30**
	7.40	
	7.08	
MP 6.3	10.77	**10.60**
	10.42	
	10.61	
MP 6.4	9.05	**8.68**
	8.66	
	8.32	
MP 6.5	6.70	**6.62**
	6.28	
	6.90	
MP 6.6	11.24	**10.98**
	10.80	
	10.89	
MP 6.7	6.83	**6.66**
	6.38	
	6.77	

Table 3.5 BLG measurements for outcrossed BLG/7 mice (=MP)

	Average BLG mg/ml	Mean mg/ml	STD	CV
MP 1.3	10.28	**8.23**	**1.79**	**21.79**
MP 1.5	6.20			
MP 1.6	9.95			
MP 1.7	5.85			
MP 1.11	7.99			
MP 1.12	8.82			
MP 1.14	7.22			
MP 1.15	7.44			
MP 3.2	10.72			
MP 3.3	6.31			
MP 5.3	7.30			
MP 6.3	10.60			
MP 6.4	8.68			
MP 6.5	6.62			
MP 6.6	10.98			
MP 6.7	6.66			

Table 3.6 Average BLG expression with mean, SD & CV for outcrossed BLG/7 mice

3.5 Discussion

The study of long-term transmission of transgenes in mice supports current beliefs that once stably integrated and proven to transmit intact, a transgene will remain stable (Aigner *et al* 1999). Unfortunately, most papers do not thoroughly examine the stability of their constructs once a line has been confirmed to express the transgene product. One article that does address transgene stability, examines the same transgene construct (α–1 antitrypsin) in both sheep and mice (Carver *et al* 1993). Sheep transmitted and expressed the construct stably, possibly aided by the use of an ovine promoter. Mice displayed variegated expression, unstable transmission or genotypic background effects in ¾ of all lines. A different example is given by a paper reporting stable production of IGF-1 in transgenic rabbits (Zinovieva *et al* 1998). Closer examination of the data presented in the article reveals large differences in individual protein yields, with evidence for decreased expression with increasing generations. The discrepancy between the conclusions of the paper and the data presented is surprising, but may be due to a general reluctance to accept that within-line variegation can exist in mammals. That a variegated phenotype can be maintained through several rounds of cell division is demonstrated in plants, where regeneration from a variegated sector of a tobacco leaf containing a transgenic *Sulfur* gene gave rise to variegated plants (Fitzmaurice *et al* 1999). This suggests that variegated expression can be heritable, with the instability heritably stable.

Transgenic mouse line BLG/7 has previously been shown to exhibit variegated transgene expression analogous to position effect variegation in *Drosophila* (Dobie *et al* 1996). As different mouse genetic backgrounds can exert widely different effects on transgene expression (McGowan *et al* 1989, Allen *et al* 1990, Elliot *et al* 1995, Schweizer *et al* 1998), the generation of BLG/7 on an F1 CBA:C57BL/6 background could be the source of the variegated expression. It can be argued that the mixture of the two different sets of modifier proteins could affect the transgene expression potential, leading to variegated expression. However, analysis of 3rd generation backcrosses provided no evidence for any major strain-specific modifier of variegation affecting the transgene array at this locus in either of the two backgrounds, although there was a small but statistically non-significant trend

towards a decrease in expression (Dobie *et al* 1996). The evidence was therefore suggestive of stable maintenance of variegation, but could not conclusively rule out the influence of strain-specific modifiers.

A C57BL/6 background is documented to affect transgene expression, though the reports are conflicting. Two papers report downregulation of transgene expression with concurrent increased methylation at the transgene locus (Koetsier *et al* 1996, Schweizer *et al* 1998). A third paper reports upregulation of transgene exression with decreased methylation at the locus (Sutherland *et al* 2000). In none were backcrosses continued for more than 4 generations, and strain-specific modifications usually presented themselves within two generations. The different effects on gene expression exerted in the same background are difficult to explain, though I will attempt to forward my own hypothesis. Several different factors may be contributing towards determination of expression. Firstly, local factors that are intrinsic to the insertion site could affect expression potential. Secondly, residual background effects and epigenetic inheritance, varying with the backgrounds the lines were created on may also exert an effect on expression. Thirdly, passage through the germline could also affect transgene loci differently; depending on a combination of insertion site, integration profile and gamete sex. Combine the above with strain-specific differences in modifier protein content and composition, and the outcome will depend on the balance of activating and silencing factors. The trend in the 3rd generation BLG/7 backcross mice suggests that in this instance, C57BL/6 may exert a repressive effect on the transgene locus. The backcrosses into both genetic backgrounds were therefore continued, and sampled at the 13th generation for the presence of minor modifier loci that may be contributing towards the observed variegated expression profile of the F1 and 3rd generation backcross BLG/7 populations.

The aim of this investigation was therefore to establish whether the level of variegation was maintained through the generations in both backcross and F1 outbred backgrounds, and if there was evidence that minor modifier loci could be responsible for the phenotype. Analysis of milk samples demonstrated that Outbred F1 13th

generation BLG/7 mice maintain the same transgenic expression profile as the founder population. 13th generation backcross BLG/7 populations differ both from each other and outbred populations, indicating the presence of background-specific modifier loci affecting BLG expression. In addition, inbreeding depression downregulates all milk protein expression in the backcross populations, with a selective depressive effect on BLG/7. These background- and inbreeding specific effects are not transmitted through the germline when outcrossed to generate an F1 mixed genetic background. The conclusion of all this is that BLG/7 still variegates after successive rounds of germline transmission in both a congenic and mixed genetic background.

3.5.1 Variegation is an inherent property of BLG/7

The BLG/7 locus displays variable expression in both 13th generation CBA and C57BL/6 backcross populations (3.4.2 & 5). This suggests that the mixed background upon which the transgenic line was generated is not responsible for the variegation phenotype, as separation of the different genotypic modifiers by backcrossing has failed to generate a stable expression pattern. Variegation is also observed in F1 outbred BLG/7 mice maintained in parallel with the backcross populations (3.4.2), suggesting that the cause of variegation is maintained through successive germline transmissions. Maintenance of expression levels and cv values in the F1 outbred BLG/7 mice that are comparable to earlier generations of BLG/7 mixed background mice indicate that passage through the germline does not affect the extent of variegation and expression potential of the transgene (3.4.3). The above observations prove that BLG/7 variegation is an inherent property of the transgene at the site of integration. What is observed for BLG/7 is therefore a position effect.

Previous studies of the BLG/7 line has identified that the transgene has integrated as a multicopy array close to centromeric heterochromatin of chromosome 15 (Dobie *et al* 1996). This integration site is highly suggestive of a position effect akin to *Drosophila* PEV acting on the transgene locus of BLG/7. If this is the case, the locus should be open to modification of expression through changes in concentrations of heterochromatic proteins and other modifiers of PEV.

3.5.2 BLG/7 expression is open to modification through inbreeding depression

Although BLG/7 expression continued to variegate on both backcross genetic backgrounds, there were significant differences in expression levels and variance when compared to the F1 outbred BLG/7 data (3.4.5, Table 3.3). Therefore, BLG/7 expression is open to quantitative modification.

Inbreeding depression represses endogenous protein production

Both 13th generation backcross populations exhibited reduced BLG expression levels and a decreased variance around the mean when compared to the outcross F1 BLG/7 mice. Visual observations of the protein gels suggested that endogenous milk protein expression was reduced in the backcross mice (3.4.2). Micro-Kjeldahl analysis confirmed this observation (3.4.4), the results demonstrating that expression was only 85% of the outbred level.

The reduction of endogenous protein expression is almost certainly the manifestation of inbreeding depression effects, as both backcross populations exhibit the same level of depression. Protein expression was only examined in the milk, but there is no reason not to assume that the results observed are representative for all tissues. A frequent manifestation of inbreeding depression is the reduction of viability and breeding fitness, the result of the accumulation of duplicate copies of deleterious gene alleles. Both backcross genotypes exhibit reduced levels of breeding success and poor milk yields.

Selective reduction of BLG/7

Since the endogenous protein production was reduced by 15%, the extent of BLG repression in both backcross populations was compared to the outbred F1 BLG/7 population. The outbred mean of 8.2 mg/ml was set at 100%. Surprisingly, backcross BLG expression levels were depressed by more than 15% (3.4.4). The C57BL/6 BLG/7 mean expression was reduced by 29%, and CBA levels by 47%. This is, to my knowledge, the first reported case of selective "discrimination" of a transgene in an inbred environment.

The results provide clear evidence that whilst inbreeding depression down-regulates all milk protein expression; there is an additional negative bias towards the transgene. There is also a genotype-specific component involved (3.4.5), as demonstrated by the different levels of depression exhibited by the backcross populations, addressed in the next section.

Genotype-specific modifier action may account for the reduction in BLG/7 expression in both backcross populations, with different levels of down-regulation in each background. Alternatively, genotype-specific modifiers may only act on the transgene in the CBA background. The additional 14% reduction in the C57BL/6 backcross population, in excess of the 15% inbreeding-specific reduction, could be the manifestation of an additional selective disadvantage for the transgene in an inbred backcground. This explanation is credible, as the 3rd generation C57BL/6 backcross results do not support decreased transgene expression, whereas a reduction (albeit statistically non-significant) in mean expression was seen in the CBA 3rd generation backcross (Dobie *et al* 1996).

BLG/7 may be at a positional disadvantage

The additional negative bias towards the BLG/7 locus afforded by inbreeding depression effects may be a direct result of its genomic position. The positioning of the transgene array close to centromeric heterochromatin could impose a twofold disadvantage on the construct. Firstly, the location may render it more susceptible to changes in chromatin protein composition (Festenstein *et al* 1999, McMorrow *et al* 2000). The multicopy nature of the transgene integration product could exacerbate the effects afforded by close proximity to heterochromatin. (Dorer & Henikoff 1994, Matzke & Matzke 1995, Hsieh & Fire 2000). Inbreeding may reduce the availability of factors that contribute towards keeping the locus in an open conformation. A reduction in such locus activating complexes would increase the probability of the locus becoming silenced, decreasing the mean expression level. The levels of available transcription complexes may be reduced in an inbred background. This could influence activation levels as transcription factors are capable of initiating changes in higher order chromatin structure during the early stages of gene activation

(Lundgren *et al* 2000, Ahmad & Henikoff 2001). An increase in repressive factors, such as heterochromatic proteins, could also account for the decrease in BLG expression (Festenstein *et al* 1999, McMorrow *et al* 2000).

Secondly, the transgene integration site could render it at a positional disadvantage within the nucleus. The chromosomal position occupied by the transgene may sequester it to a different nuclear compartment to that occupied by the endogenous milk genes. Recent papers highlight the importance of nuclear localisation for gene expression potential (Francastel *et al* 2000, Brown *et al* 2001). Circumstantial evidence suggests silencing factors are non-uniformly distributed throughout the nucleus (Wallrath 2000), and may also hold true for activating factors. There is even evidence that subcellular localisation can influence gene expression in prokaryotic organisms (Kim & Wang 1999). Integration close to centromeric heterochromatin may localise the transgene array away from nuclear compartments where endogenous milk protein gene transcription takes place, making it harder for the BLG/7 transgene to compete for limited resources needed for transcription (should inbreeding lead to a reduction in production of parts of the translational and transcriptional machinery). A combination of nuclear position and regulatory region composition may thus render the transgene selectively disadvantaged.

Whatever the cause, the selective reduction of BLG/7 expression due to inbreeding depression presumably reflects a recruiting disadvantage for the β-lactoglobulin sequences when compared to the endogenous milk genes. An extreme consequence of inbreeding depression could be complete transcriptional silencing of a gene (Matzke *et al* 1993). Although this was not observed in either background at the 13^{th} generation of backcrossing, it is still theoretically possible that this could occur in later generations if the depressive effects are increased. Should the repressive effect on transgene expression of an inbred background be common, then current backcrossing strategies designed to alleviate unstable transgene expression may exacerbate the problem instead of curing it.

3.5.3 Evidence for genotype-specific modification of BLG/7

Genetic background may play a part in influencing expression levels. Dominant strain specific modifiers are documented (McGowan *et al* 1989, Elliot *et al* 1995), with over-expression of a single heterochromatic protein component capable of modifying the extent of transgene variegation in mice (Festenstein *et al* 1999). Genetic background can influence methylation and expression of a variegating transgene (Koetsier *et al* 1996). 4th generation backcross mice continued to variegate, with increased levels of methylation and silencing associated with a C57BL/6 background. This background was also associated with *de novo* methylation of a hepatitis B virus surface antigen (HBsAg) transgene (Schweizer *et al* 1998). HBsAg expression levels differed depending on the strain, backcrossing enhancing the effects. Segregation of strain-specific factors meant that crosses produced offspring exhibiting both up- and downregulation of expression, due to the cooperative effects of different combinations of modifier proteins. Intriguingly, background effects were only observed when the transgene was transmitted through the male germline.

The 13th generation backcross populations differ in both mean expression levels and spread of expression from 3rd generation backcross mice, outbred mixed genotype mice and from each other (Table 3.3). These results suggest that whilst strain-specific modifiers are not primarily responsible for BLG/7 variegation, modifiers capable of down regulating without silencing the transgene exist. The modifiers will vary between strains, with the extent of modification and thus silencing varying between the two genotypes.

Inbreeding depression effects can account for 15% of the reduction in expression in both backgrounds. An additional 14% of the reduction may be assigned to a putative selective depressive effect exerted by inbreeding depression-specific factors on the transgene. All of the reduction in BLG/7 expression in the C57BL/6 background could therefore be solely due to the effects of inbreeding, although it cannot be discounted that a part or all of this reduction may be caused by genotype-specific modifier action. In the CBA background, a reduction in BLG/7 expression of 18% remains unaccounted for from the total reduction level observed (47%). Background-

specific modification of expression is the only remaining explanation for the additional decrease in mean BLG/7 expression level in the CBA 13th generation backcross compared the C57BL/6 background.

Multiple recessive modifier loci must be responsible for BLG/7 background-specific reduction in expression

There are numerous reports of strain-specific modification of transgene expression levels (Allen *et al* 1990, Schwizer *et al* 1998, Mayeaux-Portas *et al* 2000). None examine backcrosses beyond the fourth generation, with the strain-specific modifications generally evident by the 2nd or 3rd generation. Clear segregation patterns emerge that are indicative of a single modifier locus. This is supported by the demonstration that over-expression of a single protein can exert a modifying effect on the expression of a variegating transgene (Festenstein *et al* 1999, McMorrow *et al* 2000).

The BLG/7 transgene locus exhibits no significant changes in expression in the 3rd generation of backcrossing to the CBA and C57BL/6 genetic backgrounds. This led to the conclusion that no single major modifier locus was present that affected BLG/7 expression (Dobie *et al* 1996). The reduction in the 13th generation backcrosses of both mean expression and variability, not wholly explained by the effects of inbreeding depression, suggests that strain-specific modification can take place (3.4.5). As no clear evidence for strain-specific modification was evident in the 3rd generation, this suggests that multiple recessive loci exist. Together, these loci can exert an effect on BLG/7 expression, at least in the CBA background. Each of these loci would have a subtle but cumulative effect on expression as genotypic background purity increases. The different levels of expression and variability between the backcross genotypes indicate that the composition and/or concentration of modifiers differ between the two backgrounds. Whether the cumulative effects would enable silencing of the locus in later generations remains to be seen. If the modifications generated by the modifier loci can be transmitted through the germline, there is increased probability that some individuals will exhibit complete silencing of BLG/7 expression in future generations.

All the results so far provide evidence that both inbreeding- and strain–specific factors can have a quantitative effect on BLG/7 expression. What is equally evident is that none of these factors can affect the underlying mechanism responsible for the variegation phenomenon. In other words, inbreeding– and strain–specific factors appear to exert a quantitative but not a qualitative effect on BLG/7 transgene expression.

3.5.4 Epigenetic modifications not inherited

There are several papers reporting progressive silencing of genes through successive germline transmissions (Kilby *et al* 1992, Lardon *et al* 1999), continued manifestation of a phenotype after removal of the transgene (Davies *et al* 1997, Jensen *et al* 1999) or incomplete erasure of epigenetic modifications in the germline (Morgan *et al* 1999). Although the early separation between somatic and germ cell lineages in animals renders epigenetic inheritance of non-imprinted loci a rarity, there are some examples. Epigenetic inheritance of modifications due to strain-specific and inbreeding factors was therefore investigated by outcrossing backcrossed BLG/7 to the opposite inbred strain in order to generate 1st generation 50:50 CBA:C57BL/6 BLG/7 mice, the genetic counterparts to our outbred F1 BLG/7 mice.

The complete reversion to outbred BLG/7 mean expression levels indicates that inbreeding- and/or genotype-specific modifiers do not convey a heritable epigenetic modification at the BLG/7 locus, at least not in a mixed genetic background through male germline transmission (3.4.6). This is mirrored in a paper examining the effects of inbreeding and outcrossing in ladybeetles (Morjan *et al* 1999). Maintenance of variegated BLG/7 expression provides additional confirmation that this is an inherent property of the transene at the site of integration irrespective of genetic background.

Despite displaying expression levels within the range seen for the outbred F1 BLG/7 mice, the coefficient of variance was considerably reduced in outcrossed BLG/7. Outbred F1 BLG/7 mice have a cv of 42%, whilst the outcrossed F1 BLG/7 mice

only exhibit a cv of 22%. This was unexpected. As the C57BL/6 BLG/7 generation 13 backcross mice have a cv of 33.2%, the outcross results remain especially puzzling. A possible explanation lies in the genotypic composition of the outcrossed versus outbred F1 mice:

The F1 outbred BLG/7 population is maintained on a mixed genetic background generated through F2 x F1 matings, where any given pair of homologus loci can contain three combinations of genetic-background alleles (hemizygous, or homozygous for either CBA or C57 alleles). The greater genetic variability generates greater variance in expression levels between individuals in the group. The outcrossed mice are genotypically 50% CBA and 50% C57BL/6, being first generation outcrosses from a C57BL/6 background to a CBA background. Each duplicated chromosomal locus can only have a single combination of alleles, one each from the CBA and C57BL/6 backgrounds. This reduction of genetic variability would explain the reduction in the coefficient of variation. The higher variability of the backcross parental population may be explained by the effects of inbreeding depression- and background-specific modification. The two sets of modifiers may interact, possibly with random segregation within and between individuals, generating more diversity in a stochastic fashion than the backcrossed genotype might initially suggest possible. The homogeneity of allele combinations in the outcrossed F1 BLG/7 mice greatly reduces their scope for variability, and hence the reduction in the cv value. Mean BLG levels revert to outbred F1 BLG/7 levels due to a combination of heterosis and absence of inherited epigenetic modification, but the range of expression levels is reduced due to the homogeneity of modifier loci in this background. Figure 3.8 depicts all four genotypic backgrounds with expression levels and range. I fully expect that the outcrossed F1 BLG/7 mice will revert to a cv level in line with outbred F1 BLG/7 mice after a few generations of maintenance on the F1 background.

The reversion to parental expression levels and maintenance of variegated expression supports the conclusion that there is no progressive silencing of the transgene through repeated rounds of germline transmission, at least in a mixed genetic

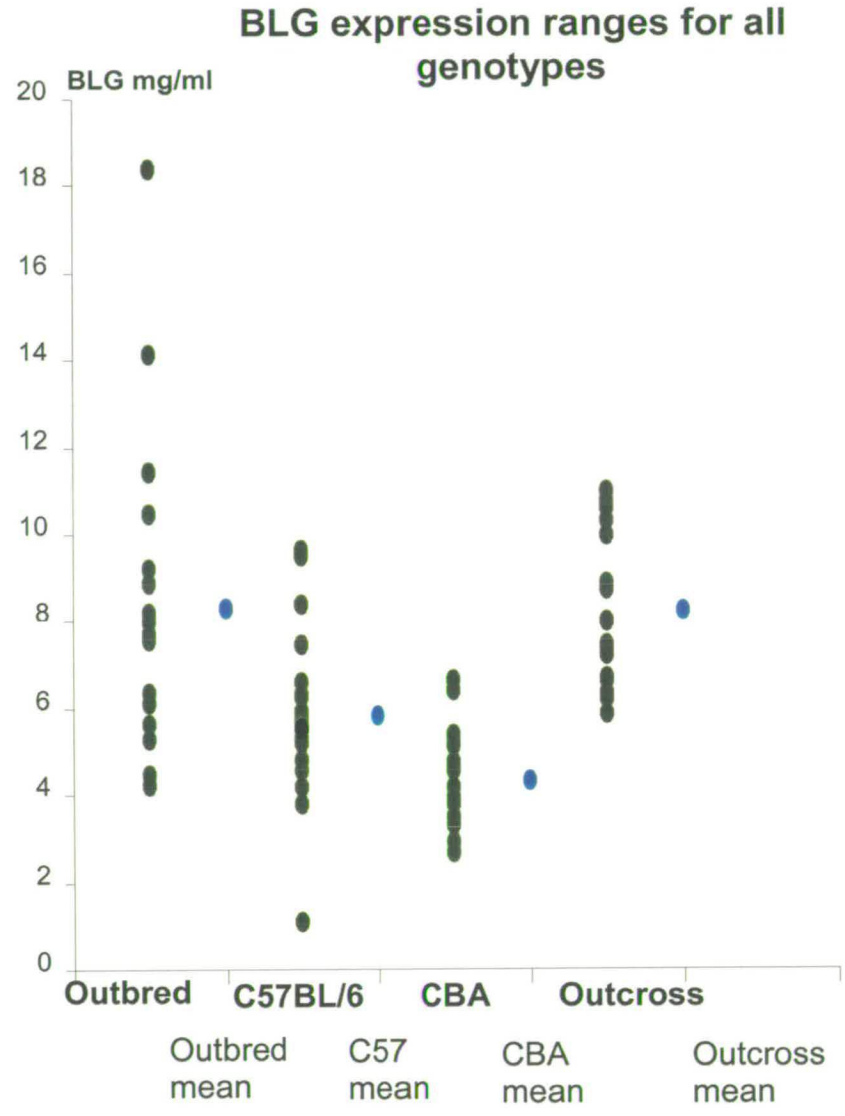


Figure 3.8 **Scatterplot showing individual BLG milk protein content and mean BLG expression for all 4 genotypes**

background. Inheritance of epigenetic modifications or modifier proteins responsible for inbreeding- and background-specific depression of expression can also be discounted by the results of this experiment. This result also excludes germline mutation as a possible cause of reduced expression in the inbred background.

3.5.5 Variegation is maintained through the generations

BLG/7 expression variegates in all genetic backgrounds examined and at all generational stages investigated. The depressed BLG/7 expression in the backcrossed mice is de-repressed upon outcrossing, even after 16/17 generations of backcrossing, and reverts to the expected outbred BLG/7 levels. Expression does not fall with increasing age of an animal, as has been demonstrated for other transgenic mice (Cohn *et al* 1991, Robertson *et al* 1996, Sutherland *et al* 1997). In my study, the BLG/7 mice are sampled in their prime. Serial lactations indicate that BLG protein expression does not differ significantly between individual lactations (Dobie *et al* 1996), suggesting that the observed variegation is not age-related. BLG/7 transgene expression is lactation specific, whereas other studies use transgenes containing constitutively active promoters, at least for their cell type (Cohn *et al* 1992, Robertson *et al* 1996, Sutherland *et al* 1997). The limited window of expression differentiates our mice from the others, and may render our transgene immune to time or age-related repression effects.

These results indicate that there are no generational or age-related effects on BLG/7 expression, and that the transgene locus is not subject to any restrictions imposed by a genome policing system.

Chapter 4

The effects of homozygosity on the expression patterns of the highly variegating transgene in BLG/7 mice

4.1 Introduction

Position effect variegation (PEV) studies in mammals have investigated copy number effects by generating multicopy transgene array inserts that induce variegation, and then reducing the copy number at the site in order to study the effects of a reduction in array size. The first paper outlining such an experiment reported that reducing copy number abolished PEV and led to even expression in all cells (Garrick *et al* 1998). The second paper, published three years later, reports the exact opposite result (Ramirez *et al* 2001). Reduction in copy number did not abolish variegation. The level of variegation remained largely unchanged, with the occasional quantitative reduction in expression from cells that expressed the transgene.

A separate issue was addressed by Belyaeva and co-workers (1993) and Lloyd and co-workers (1997), where the behaviour of two different variegating constructs in the same system was investigated (both groups used *Drosophila*). The two variegating genes may or may not be subject to common silencing mechanisms, so combining both in the same system could give rise to two results. If there is competition for a limiting factor, then the established silencing patterns for one or both genes will be disrupted. The absence of either limiting steps or an overlap of silencing pathways would ensure that both genes variegate as normal independently of each other. Both outcomes were observed (Belyaeva *et al* 1993, Lloyd *et al* 1997), so it is probable that there are limiting factors and mechanistic pathway overlaps operating in the various reported incidences of PEV.

The ability of unlinked variegating loci to affect each other's expression may be connected to certain phenomena so far only reported in non-mammalian species. A series of studies have suggested that the expression state of one allele can

occasionally affect the expression of the other allelic locus. This has been reported in *Drosophila* (Henikoff & Dreesen 1989, Golic & Golic 1996, Henikoff 1997), fungi (Woloshuk *et al* 1995) and plants (Meyer *et al* 1993, Matzke & Matzke 1995) for both endogenous genes and transgenic loci. Pairing of the two loci is a requirement of all the phenomena. In *Drosophila* and fungi, the phenomenon is termed *transvection*, with the interaction able to either activate or inactivate gene expression. Both outcomes can be associated with a single gene, an example being the *pointed* gene of *Drosophila* (Scholz *et al* 1993). A transgenic example of *transvection* in *Drosophila* is provided in the case of the *white* gene (Martin-Morris *et al* 1997). In plants, a similar phenomenon of pairing dependent inactivation of gene expression is termed "paramutation" (Winkler 1930, Matzke & Matzke 1995). This may be the mechanism responsible for the complete silencing of a repeated chalcone synthesis transgene in Arabidopsis in the homozygous state, whilst hemizygous plants display variegated expression (Davies *et al* 1997).

Transvection and paramutation both describe phenomena that influence the transcriptional status of both paired alleles, and may share mechanistic pathways. The mechanisms underlying *trans* effects are not known, but may include the transfer of heterochromatic proteins and/or histone deacetylases between the loci (Henikoff & Dreesen 1989, Ronserray *et al* 1998).

As discussed earlier, the BLG/7 transgene is integrated close to peri-centromeric heterochromatin and is predisposed to silencing. Position effect variegation is suggested as the underlying cause of the variegated expression of the transgene, with variable spread of heterochromatin into the transgene integration locus from the centromere the proposed mechanism (Dobie *et al* 1996). It is entirely possible that the peri-centromeric location causes the physical recruitment of the transgene locus into heterochromatin, with inactivation the result of an environment lacking in activating factors rather than physical heterochromatinisation of the locus itself, as reported in *Drosophila* (Sabl & Henikoff 1996). To my knowledge, no investigations have been made into the effects of a homozygous environment on a variegating transgene in mammals.

4.2 AIM

The aim of this investigation was to test whether the highly variegating transgene locus of BLG/7 would behave independently when in the homozygous state.

4.3 Experimental Approach

We wished to evaluate the possibility that expression of the highly variegating BLG/7 transgene might be influenced by the presence of an identical variegating locus on the homologous chromosome.

As a first approach, we sought to generate BLG/7 homozygous animals and determine the expression levels of BLG protein in the milk. As will be discussed below, the protein expression results initiated an extension of the investigation to include an *in situ* hybridisation analysis of both homozygous and hemizygous transgenic mammary gland tissue.

BLG/7 has been maintained on a mixed F1 CBA:C57BL/6 background (generated through F2 x F1 matings) as both a hemizygous and homozygous colony (see appendix C1 for maintenance mating regime). A mating regime was set up to generate 20 females from each colony for milk and tissue sampling. Samples were collected on the 11^{th} day of lactation (peak of BLG protein production, Whitelaw *et al* 1992), from 19 hemizygous and 19 homozygous BLG/7 females.

Levels of BLG protein were determined by quantitative SDS gel electrophoresis as described earlier (2.3), whilst fluorescent *in situ* hybridisation (2.4) using probes against BLG and β-casein mRNA was used to evaluate the proportion of alveoli producing BLG mRNA. The results of the investigation into protein and mRNA expression led to the examination of nuclear transcription from the BLG/7 loci and localisation of the integrated transgene locus with respect to the primary transcripts. I believe this may be the first successful attempt at performing chromosomal FISH using wax embedded tissue sections. The available protocols describing

chromosomal FISH are all based on chromosomal spreads, or single cell layers using cultured or blood derived cells.

4.4 Results

4.4.1 Strategy and critique of hemi- and homozygous BLG/7 analysis

Estimation of transgenic protein concentration in milk samples

Quantification of protein concentration was done in triplicate sets for each sample, using light pixel intensity comparisons to the standard curve generated by the staining intensities of the three BLG standards run on each protein gel. Analysis of the gels used the Multi Analyst Programme version 1.1 (BioRad). Loading differences were corrected for using β-casein as an internal control standardised to the non-transgenic sample present on every gel. The same non-transgenic control sample was run for both hemi- and homozygous samples. The arithmetic mean of the three measurements was used in determination of BLG expression levels for each individual sample.

18 protein gels were run in total, 9 each for the hemi- and homozygous samples. A gel from each group is reproduced in figure 4.1. Although the protein gels were Coomassie stained, greyscale depictions were preferred, as the protein bands appear cleaner and more clearly defined. All analysis of the images was done with the gels defined as Coomassie stained by the software package (2.3.3).

Fluorescent *in situ* hybridisation (FISH)

Serial sections of 5μm thickness were taken from each individual wax embedded mammary sample. Inexperience in the use of the microtome and water bath meant that initial tissue wastage was high.

Although radioactive *in situ* hybridisation techniques were available for qualitiative investigation into β-lactoglobulin expression in our mammary tissue, a fluorescent alternative was deemed desirable both from the point of view of increased safety and

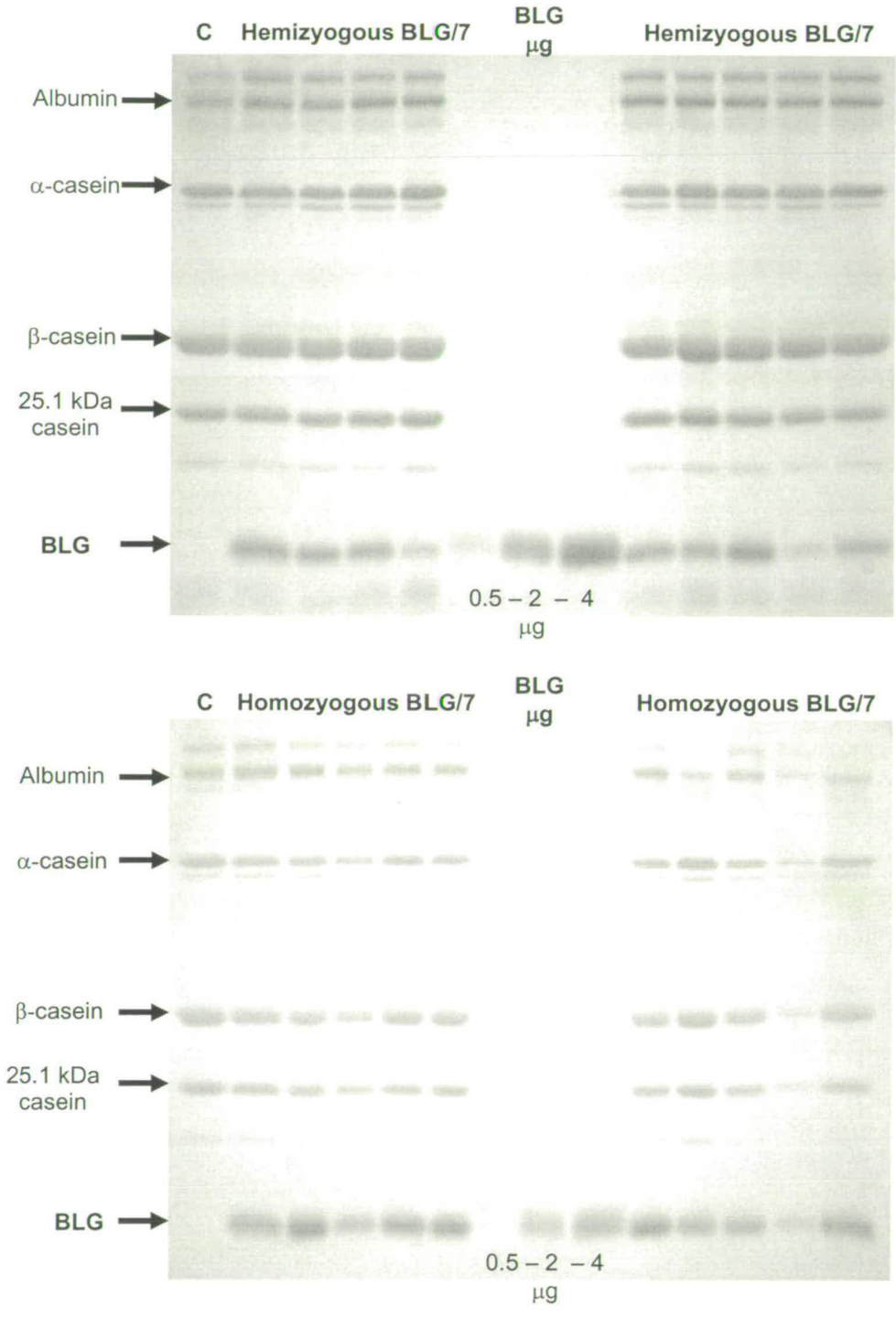


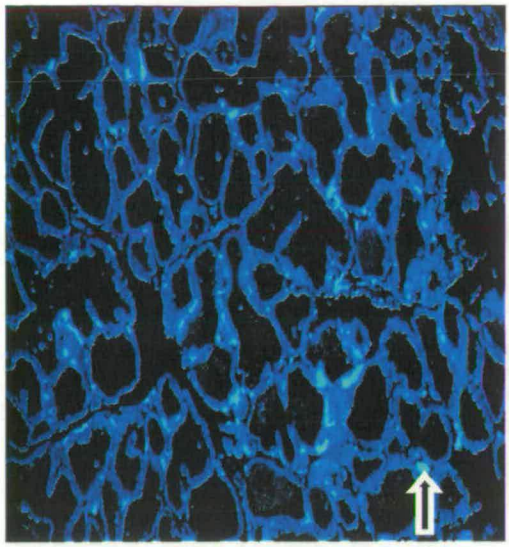
Figure 4.1 **SDS PAGE of milk protein samples for Hemi- and Homozygous BLG/7**
C = non-transgenic control milk, BLG = β-lactoglobulin standard. Transgenic samples were loaded 40 μl of 1/250 dilution.

improved resolution at the cellular level. The development of such a technique would also enable simultaneous probing for expression patterns of several different mRNA transcripts.

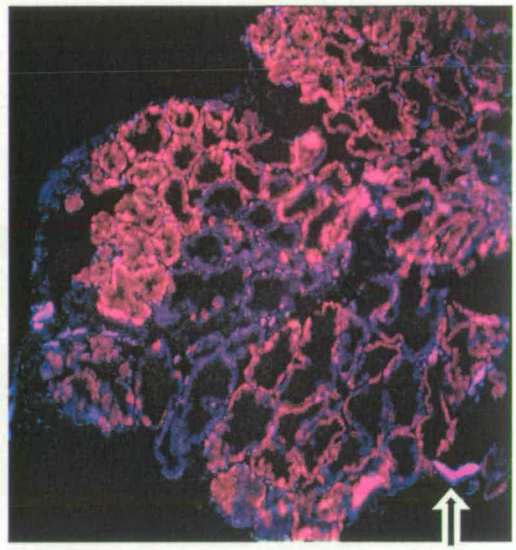
The FISH protocol for wax-embedded tissue sections took six months to develop and is a hybrid of several different published experimental protocols (see 2.4.3 - 4). RNA probes did not provide a sufficiently strong signal, and were exchanged for a cocktail of five 30'mer DNA oligos based on the published BLG and β–casein (control) coding sequences (2.4.8). Oligo sequences were separated by a minimum 30 bases to avoid steric interference between the fluorescent haptens attached at the 5' and 3' ends. SINE and LINE elements were excluded from the probes, as were obvious repetitive sequences. MWG synthesised the oligos and attached the haptens (Cy3 to BLG probes, Fluorescein to β–casein controls). Oligos were ordered in both sense and antisense orientations, enabling probe specificity to be checked.

Prehybridisation follows standard protocols, with an extended xylene dewaxing step (2.4.3). Heating the sections with hybridisation buffer overlaid serves the dual purpose of easing probe access into the tissue and dispersing any residual wax. The freezing step ensures that the wax solidifies into easily identifiable globules (see yellow-white spots in *in situ* figures, e.g. Figure 4.2). Control hybridisations using both sense and anti-sense oligo cocktails for BLG and β-casein probes alone or in combination were performed on transgenic and non-transgenic tissue (Figure 4.2, a-e). The tissues auto-fluoresces weakly in all light ranges due to residual wax, with the strongest fluorescence in the blue light range. This is not a problem as the probes are labelled with either Cy3 (red, BLG probes) or Fluorescein (green, β-casein probes). Sense probes do not generate a signal, and BLG antisense probes only hybridise in transgenic mammary tissue, confirming probe specificity (Figure 4.2). When both sets of probes are used together, the BLG signal appears orange, due to the mixing of two fluorescent signals as all mammary epithelial cells express β-casein mRNA (Figure 4.2). Visualisation was at 10x magnification.

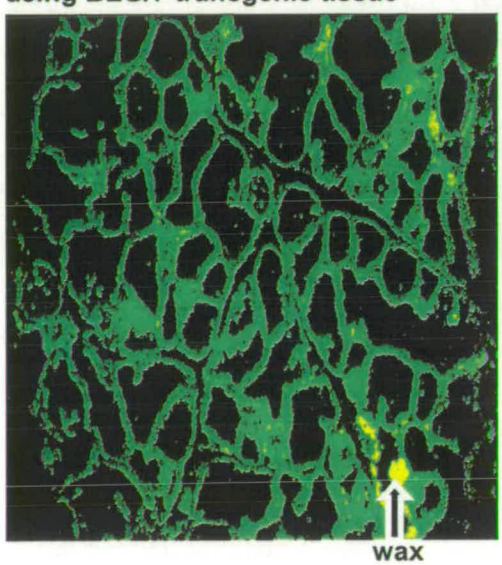
FISH for BLG mRNA in Hemizyogous tissues



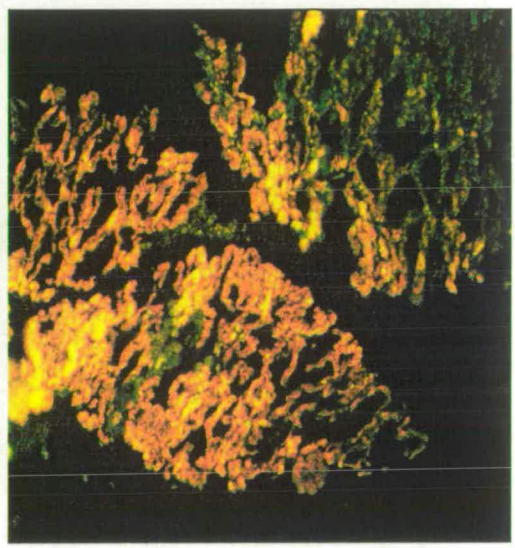
A) Sense control BLG (Cy3, red) & β-casein (Fluorescein, green) probes against DAPI (blue) background using BLG/7 transgenic tissue



C) Antisense Cy3 (red) BLG probe against DAPI (blue) background using BLG/7 transgenic tissue



B) Antisense Fluorescein (green) β-casein & Cy3 (red) BLG probes using non-transgenic control tissue

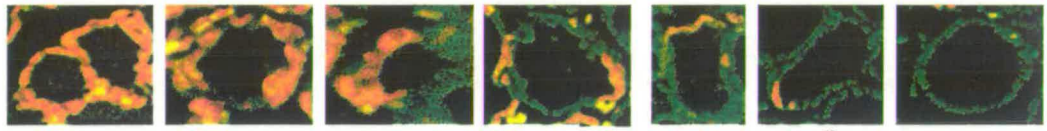


D) Antisense Cy3 (red) BLG & Fluorescein (green) β-casein probes using BLG/7 transgenic tissue (red + green = orange signal)

Figure 4.2 **mRNA FISH (BLG & β-casein) in lactating mammary tissue**
10 X magnification. Scale: 1 cm = 80 μm

Control hybridisations using sense (A) and antisense (B-D) probes for β-casein (A,B & D, green signal = Fluorescein) and BLG (A, C & D, red signal = Cy3) mRNA expression in transgenic (A, C&D) and non-transgenic (B) tissue Bright fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals. Whole tissue autofluoresces in blue light range (DAPI).

Double FISH using cocktails of oligos against both BLG and β-casein exonic sequences were performed in duplicate sets for each individual tissue sample. Ten or more views from each section were digitally captured and the proportion of expressing alveoli in each view was estimated manually. Each alveolus was scored as to whether it was 100% ON, 100% OFF, ¾ ON, ½ ON, ¼ ON, 1/8th ON or 1/16th ON (see below for representative examples). The assignation of individual expression status is to a certain extent subjective, and the error margin must reflect this. By repeating the hybridisations, I have tried to minimise the level of human error. Where the two counts have differed significantly, a third round of hybridisations were performed in order to ascertain a more detailed picture of the relative expression profile.



ON 3/4 1/2 1/4 1/8th 1/16th OFF

Limitations of the FISH analysis

The main two limitations identified for this approach are human error and sampling size. The cause of human error lies in the manual assignation of relative expression levels for each alveolus, a subjective decision, and the repetitiveness of the task. This has been compensated for by the repeated rounds of analysis.

The effects of sampling require further explanation. Ideally, each mammary tissue should be subjected to serial sectioning, with samples covering the entire gland quantified for expressing alveoli. The timescale involved, and the number of samples in need of processing (38), rendered such a detailed investigation impossible. Instead, enough 5μm sections were cut to ensure that a minimum of five slides, each containing 2-3 sections, were available from each tissue sample. As my skill in using a microtome increased, less tissue was needed to generate the same number of slides. Earlier tissue samples cut (hemizygous) thus cover more of the total mammary gland than the later samples (homozygous tissues). Hemizygous counts may therefore be, on an individual basis, slightly more representative of the total tissue than individual homozygous counts. As most expression levels are duplicated, the combined results should be representative of the general trend. The data presented verifies the validity

of this approach (Figure 4.10). I am therefore confident that the data generated by the technique provides an accurate insight into the relationship between transgenic milk protein content and proportion of expressing cells in both groups of mice.

Primary transcript and genomic DNA FISH

The success of the mRNA FISH prompted a similar approach to identify primary transcript generation from the BLG/7 loci. Probes were designed against intronic sequences, again using five 30'mer oligos. Digoxygenin (DiG) was used as a hapten in order to enable signal amplification, as it was considered improbable that sufficient primary transcript would be present to enable visualisation without a secondary amplification step. Hybridisation proceeded as for mRNA detection, with the addition of an antibody detection step. The antibody used was conjugated to either Fluorescein or Rhodamine (Red) (the antibody can be purchased with either hapten). Visualisation was at 100x magnification. Nuclei were DAPI counterstained to enable identification and provide a contrasting background to the transcript signal.

Probes against β-casein intronic sequences labelled first with biotin, then fluorescein, failed to generate a signal. Initially interpreted as a failure of the antibody to biotin, the subsequent failure of the fluorescein labelled probe in conjunction with antibodies to fluorescein prompted the conclusion that the intronic β-casein sequence data may be inaccurate. A second set of intronic probes, against mouse whey acidic protein (WAP), was ordered. Although not as abundant as the caseins, this protein can still be seen as the smallest protein at the bottom of the milk gel, just below the BLG band in the transgenic samples (Figure 4.1). Sufficient intronic sequence data was not available for any of the more abundant milk proteins.

Probe generation for chromosomal FISH was identical to the other oligo designs. Ten 30'mer oligos against the BLG promoter region were made. A fluorescein hapten was attached 5' and 3', with amplification via a fluorescein antibody conjugated to fluorescein (2.4.6). The pre-hybridiation protocol was modified to include a pepsin digestion step (2.4.6), which was essential to ensure probe access to genomic DNA. Pepsin may exert its effect through digesting chromosome-specific proteins or

protein bonds. All chromosomal FISH protocols that include a digestion step use pepsin, which is why it was included in this protocol. The chromosomal FISH protocol took two months to perfect. Concerns that the additional digestion step might eradicate the primary RNA transcript were allayed by excellent test FISH results using intronic BLG probes. The primary transcript signal was, if anything, stronger than before. Double FISH for intronic RNA and genomic DNA also worked well. Primary transcripts were detected using the DiG antibody conjugated to rhodamine, whilst fluorescein was used to amplify the chromosomal signal.

4.4.2 BLG protein levels suggest that the two homologous loci do not variegate independently

Raw data was collected for each protein gel, with both BLG and β-casein levels measured (Appendix A2). Three separate measurements of BLG concentration were generated for each sample (Table 4.1). The arithmetic mean of these measurements was used to calculate individual expression levels, with the results shown in Table 4.2 and Figure 4.3. The BLG expression levels for both populations were subjected to statistical analysis, with mean expression, standard deviation and coefficient of variation values calculated and compared (Table 4.3).

The most noticeable result was the absence of an increased upper limit of BLG expression in the homozygous mice compared to their hemizygous counterparts. This upper limit is most obvious when the data is represented in the form of a scatterplot (Figure 4.4). Neither population exceeded a transgenic protein expression level of 20 mg/ml. The highest hemizygote expression was 18mg/ml, the highest homozygote expression 19 mg/ml. This ceiling of BLG/7 protein expression does not represent a physiological restraint as another line of transgenic mice containing a variegating β-lactoglobulin transgene (BLG/45) essentially doubles transgene output in the homozygous state, achieving an average expression of 32 mg/ml (McClenaghan *et al* 1995). This result shows that the mammary gland is capable of higher levels of BLG production than the maximum achieved by the BLG/7 line. The absence of a doubling of expression in BLG/7 homozygote mice must therefore be caused by some other reason than a pure physiological constraint. The homozygous mice do

	BLG measured mg/ml	BLG average mg/ml		BLG measured mg/ml	BLG average mg/ml
H1	13.76	14.09	**H11**	8.08	7.99
	14.33			7.67	
	14.19			8.23	
H2	8.97	8.82	**H12**	5.82	5.62
	8.50			5.11	
	8.99			5.92	
H3	8.04	8.16	**H13**	6.44	6.35
	8.21			6.26	
	8.22			6.34	
H4	19.13	18.35	**H14**	4.37	4.43
	18.13			4.17	
	17.80			4.74	
H5	5.49	6.08	**H15**	11.14	10.45
	6.43			10.95	
	6.32			9.25	
H6	11.43	11.38	**H16**	7.92	7.61
	11.24			7.36	
	11.45			7.56	
H7	5.36	5.25	**H17**	9.83	9.16
	5.23			8.74	
	5.16			8.91	
H8	7.96	7.70	**H19**	4.61	4.19
	7.44			3.58	
	7.70			4.37	
H9	7.85	7.98	**H20**	7.78	7.54
	8.08			7.47	
	8.02			7.37	
H10	5.46	5.59			
	5.68				
	5.63				

Table 4.1 **A)** BLG measurements for hemizygous BLG/7 mice

	BLG measured mg/ml	BLG average mg/ml		BLG measured mg/ml	BLG average mg/ml
HH1	13.12	13.81	**HH11**	18.10	18.20
	13.63			18.66	
	14.67			17.82	
HH2	12.82	13.97	**HH12**	8.72	8.94
	14.21			8.69	
	14.87			9.41	
HH3	9.25	9.22	**HH13**	11.07	11.12
	9.43			11.00	
	8.98			11.29	
HH4	9.75	9.56	**HH14**	8.94	9.06
	9.56			9.05	
	9.37			9.18	
HH5	9.76	9.74	**HH15**	11.61	11.74
	9.94			11.75	
	9.52			11.86	
HH6	11.56	12.12	**HH16**	7.56	7.68
	12.30			7.44	
	12.50			8.04	
HH7	7.66	7.64	**HH17**	18.47	18.81
	7.78			18.85	
	7.48			19.10	
HH8	11.88	12.45	**HH18**	11.69	11.54
	12.78			11.89	
	12.71			11.04	
HH9	9.29	9.30	**HH19**	12.43	12.42
	9.50			12.32	
	9.12			12.51	
			HH20	12.01	12.03
				11.79	
				12.30	

Table 4.1 **B)** BLG measurements for homozygous BLG/7 mice

	Average BLG mg/ml	Mean mg/ml	STD	CV
H1	14.09	**8.25**	**3.46**	**41.98**
H2	8.82			
H3	8.16			
H4	18.35			
H5	6.08			
H6	11.38			
H7	5.25			
H8	7.70			
H9	7.98			
H10	5.59			
H11	7.99			
H12	5.62			
H13	6.35			
H14	4.43			
H15	10.45			
H16	7.61			
H17	9.16			
H19	4.19			
H20	7.54			

	Average BLG mg/ml	Mean mg/ml	STD	CV
HH1	13.81	**11.54**	**3.08**	**26.71**
HH2	13.97			
HH3	9.22			
HH4	9.56			
HH5	9.74			
HH6	12.12			
HH7	7.64			
HH8	12.45			
HH9	9.30			
HH11	18.20			
HH12	8.94			
HH13	11.12			
HH14	9.06			
HH15	11.74			
HH16	7.68			
HH17	18.81			
HH18	11.54			
HH19	12.42			
HH20	12.03			

Table 4.2 Average BLG expression with mean, SD & CV for hemizygous (H) and homozygous (HH) BLG/7 mice.

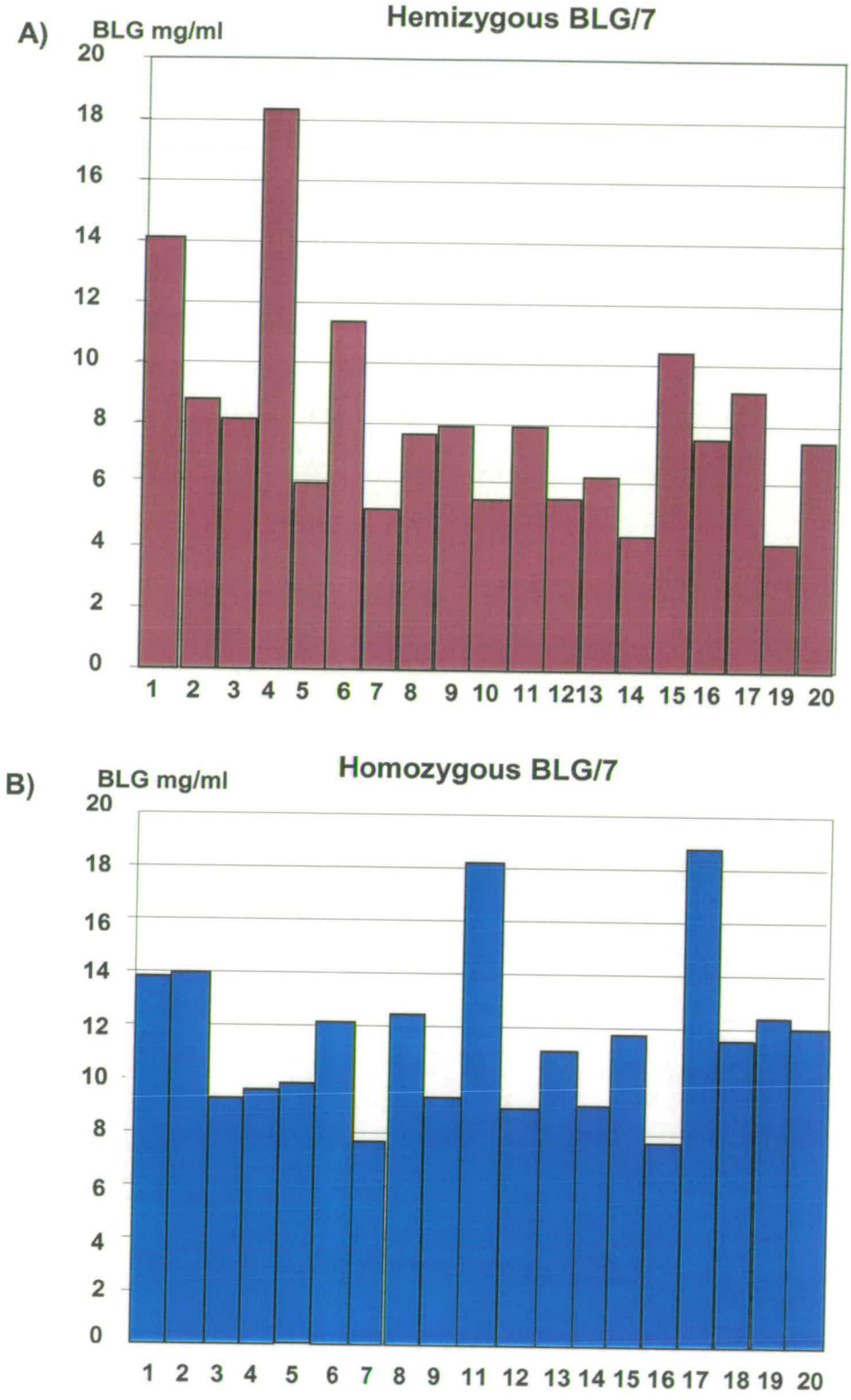


Figure 4.3 Hemi- & Homozygous milk expression
Bar charts representing individual BLG milk protein yields in mg/ml.
A = hemizygote samples, B = homozygote samples

Each bar corresponds to the mean BLG expression of the individual numbered below it (i.e., in chart A individual 1 = H1 etc)

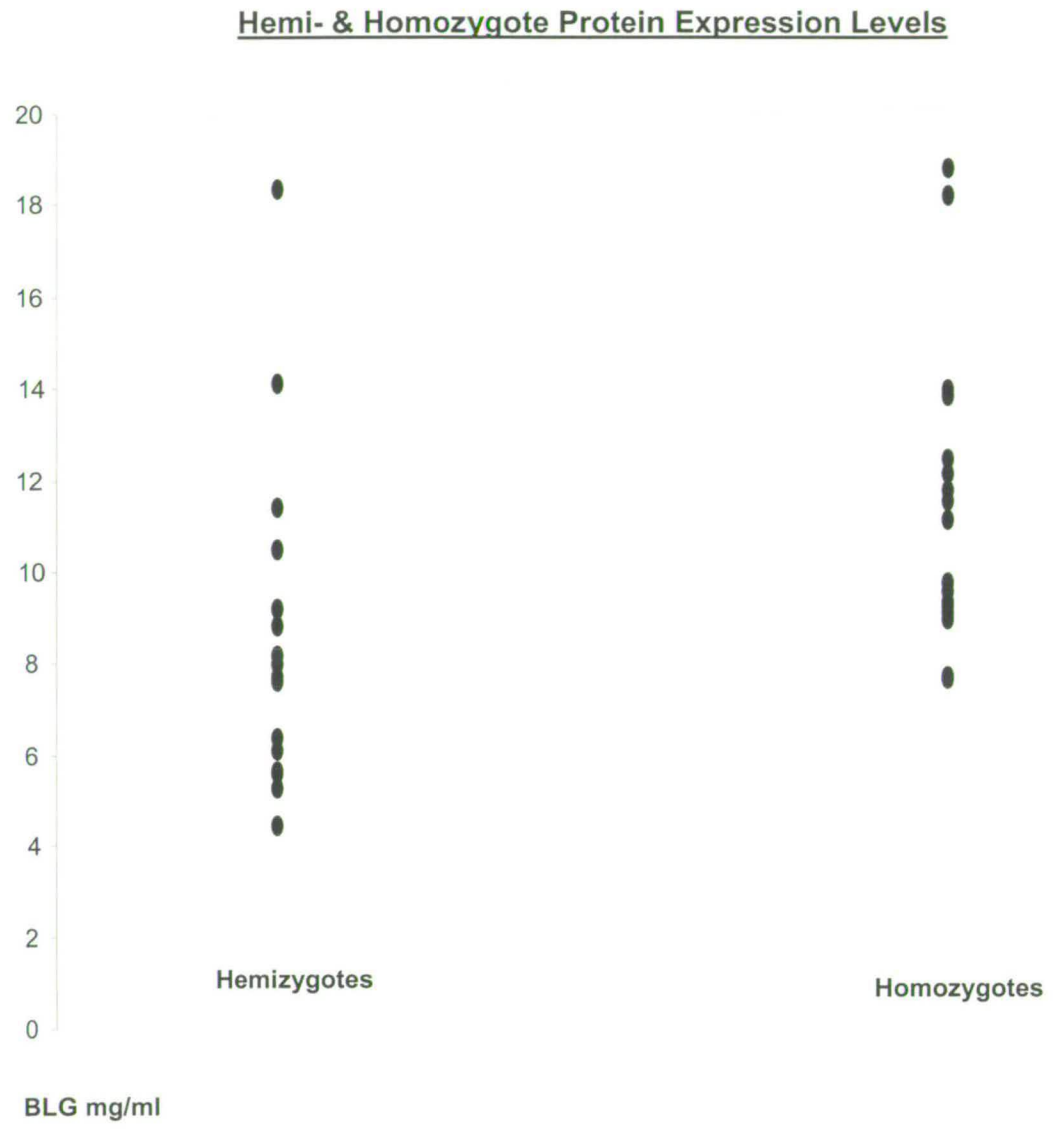


Figure 4.4 **Scatterplot showing individual BLG milk protein content for both genotypes**

Each dot represents one individual sample in the given population. It is clear from this representation that whilst the individuals from both populations with the lowest level of BLG expression have very different amounts of protein in their milk, the highest expressors in both populations display similar BLG milk protein content.

nonetheless exhibit a near doubling of minimum expression levels compared to the hemizygous mice (Figure 4.4).

Hemizygous protein yields exhibited a minimum level of just under 4 mg/ml, whilst the homozygous minimum was 7.5 mg/ml. Mean expression levels were therefore increased in homozygotes (11.5 mg/ml) compared to hemizygotes (8.2 mg/ml), whilst the coefficient of variegation was naturally reduced in the homozygote population (26.7% vs. 42%) (Table 4.3).

MICE	MEAN	SD (standard deviation)	CV (coefficient of variation)
Hemizygous BLG/7 (outbred F1 CBA:C57BL/6)	**8.25**	**3.46**	**41.98**
Homozygous BLG/7 (outbred F1 CBA:C57BL/6)	**11.54**	**3.08**	**26.71**

Table 4.3 Backross, outbred and 3rd generation milk protein results

4.4.3 The spatial pattern of BLG expression revealed by mRNA FISH cannot discriminate between hemizygous and homozygous samples

Initial examination of the mRNA FISH images prior to analysis provided clear evidence that even within a single section there are high levels of variegation, with views exhibiting strikingly different proportions of expressing alveoli. This was evident for both high (Figures 4.5 & 4.6) and low (Figures 4.7 & 4.8) expressing individuals of both hemizygous and homozygous origin. The alveoli rarely appeared to have every cell expressing the transgene, even in high expressors, mirroring previously published images for hemizygous BLG/7 mice using radioactive probes (Dobie *et al* 1996). When examining hemi- and homozygous individuals with equal protein expression levels, it did appear that they also shared roughly the same proportion of expressing alveoli with a similar pattern of expression evident (Figure 4.9).

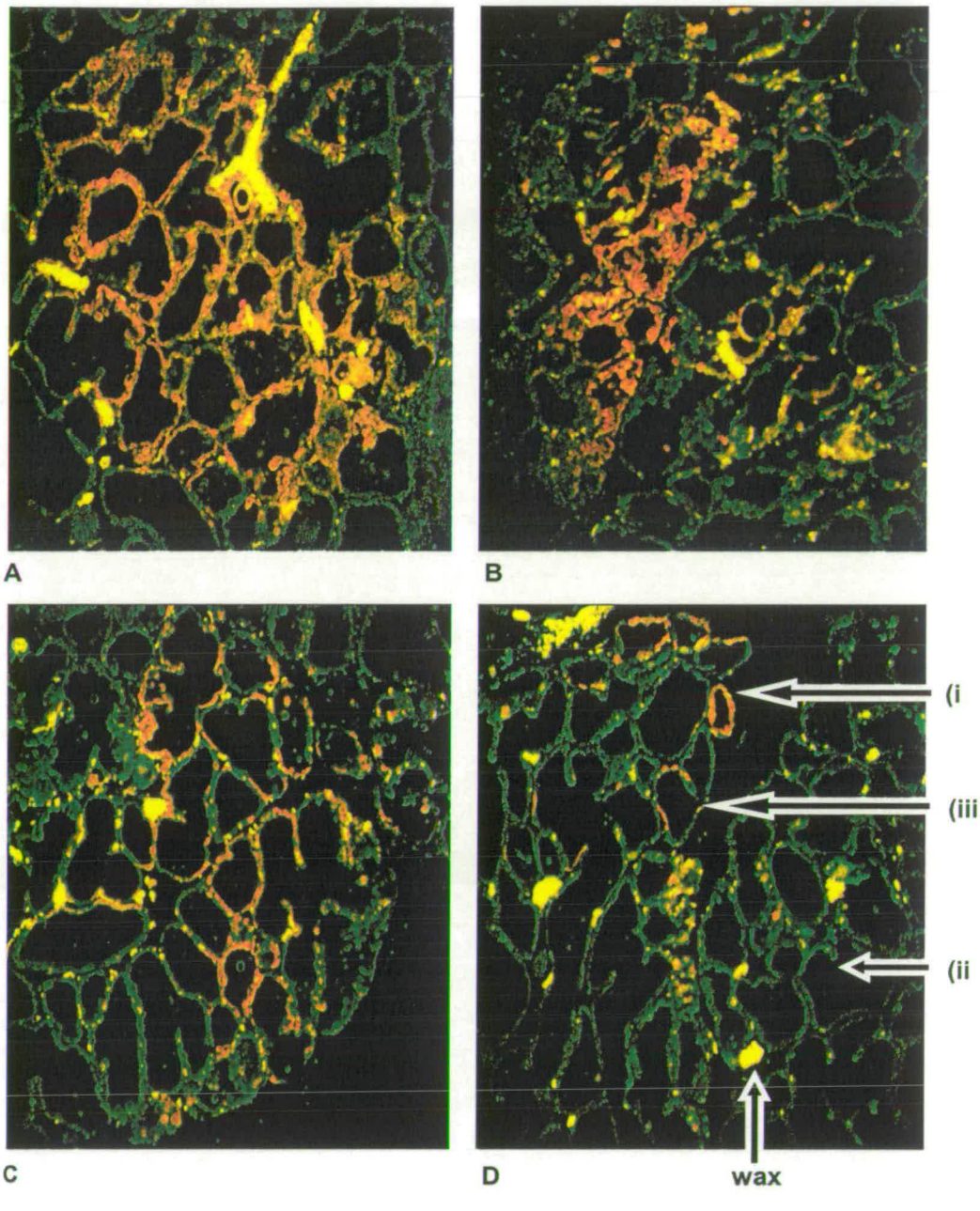


Figure 4.5 **Double mRNA FISH (BLG & β-casein) from high BLG/7 hemizygote expressor (H1 = 14 mg/ml) Scale: 1 cm = 80 μm**

4 different views (A – D) from the same section of lactating mammary tissue shown, magnified 10 X. Probe concentration used was 40 ng/ml. Probes used were BLG (Cy3 = red signal) & β-casein (Fluorescein = green signal). All epithelial cell express β-casein (green), some also express BLG (red + green = orange). Note that alveoli can comprise of cells expressing both genes (**i**), β-casein (**ii**) or a combination of the two (**iii**). Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

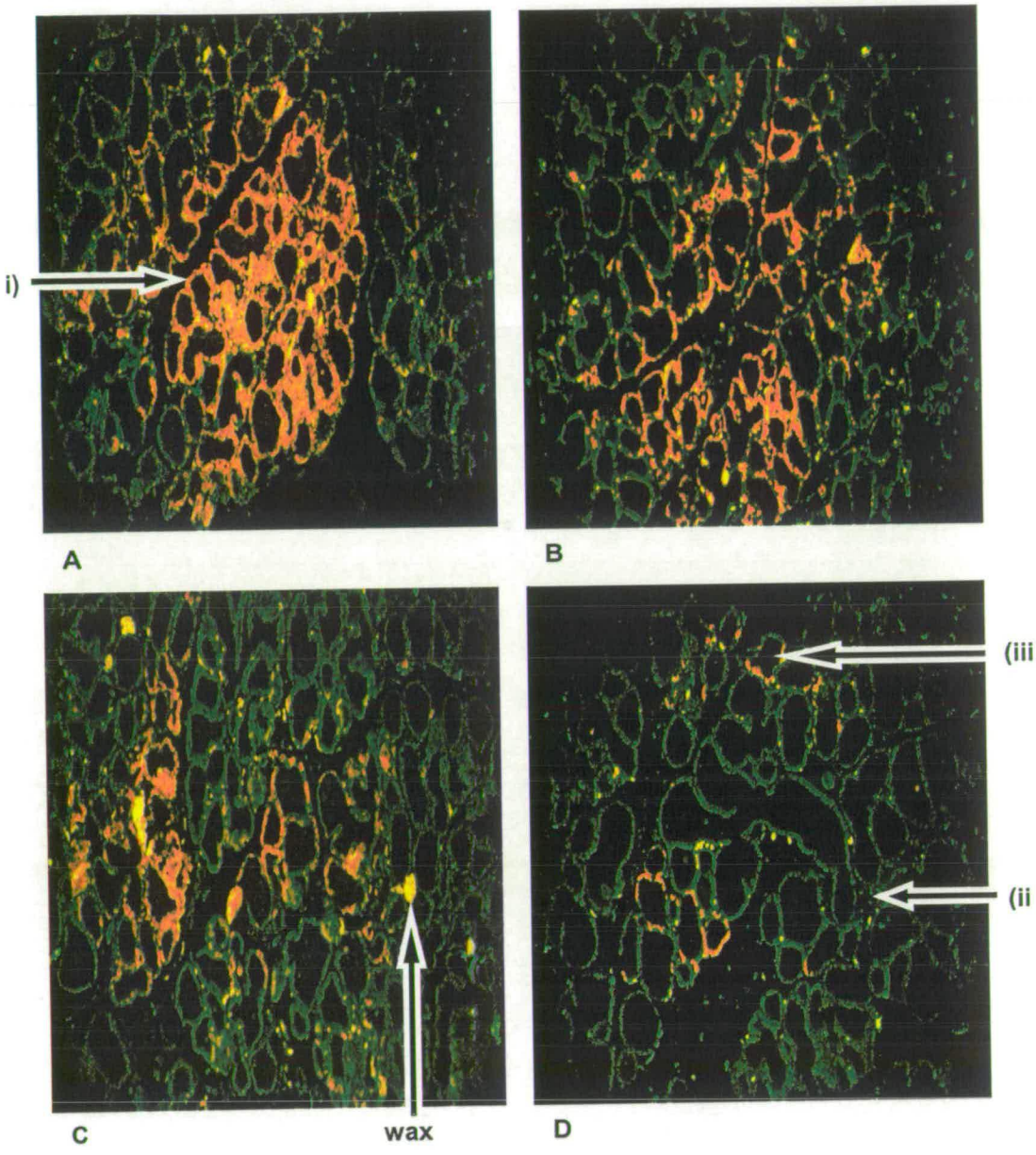


Figure 4.6 **Double mRNA FISH (BLG & β-casein) from high BLG/7 homozygote expressor (HH17 = 18.8 mg/ml) Scale: 1 cm = 80 μm**

4 different views from the same section of lactating mammary tissue shown (A – D), magnified 10 X. Probe concentration was 40 ng/ml. Probes used were BLG (Cy3 = red signal) & β-casein (Fluorescein = green signal) All epithelial cell express β-casein (green), some also express BLG (red + green = orange). Note that alveoli can comprise of cells expressing both genes (i), β-casein (ii) or a combination of the two (iii). Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

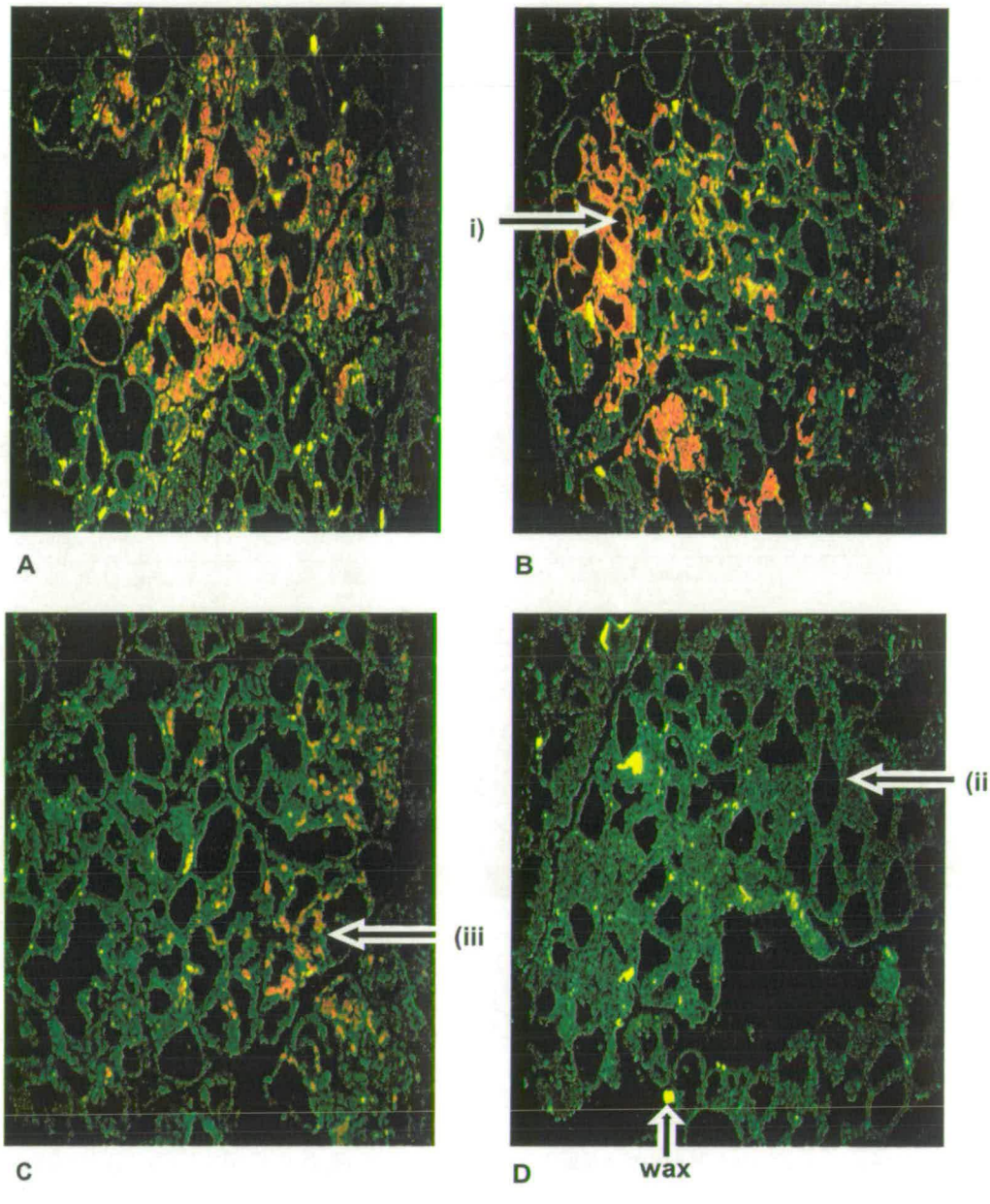


Figure 4.7 **Double mRNA FISH (BLG & β-casein) from lower BLG/7 hemizygote expressor (H13 = 6.3 mg/ml) Scale: 1 cm = 80 μm**

4 different views (A – D) from the same section of lactating mammary tissue shown, magnified 10 X. Probe concentration was 40 ng/ml. Probes used were BLG (Cy3 = red signal) & β-casein (Fluorescein = green signal) All epithelial cell express β-casein (green), some also express BLG (red + green = orange). Note that alveoli can comprise of cells expressing both genes (i), β-casein (ii) or a combination of the two (iii). Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

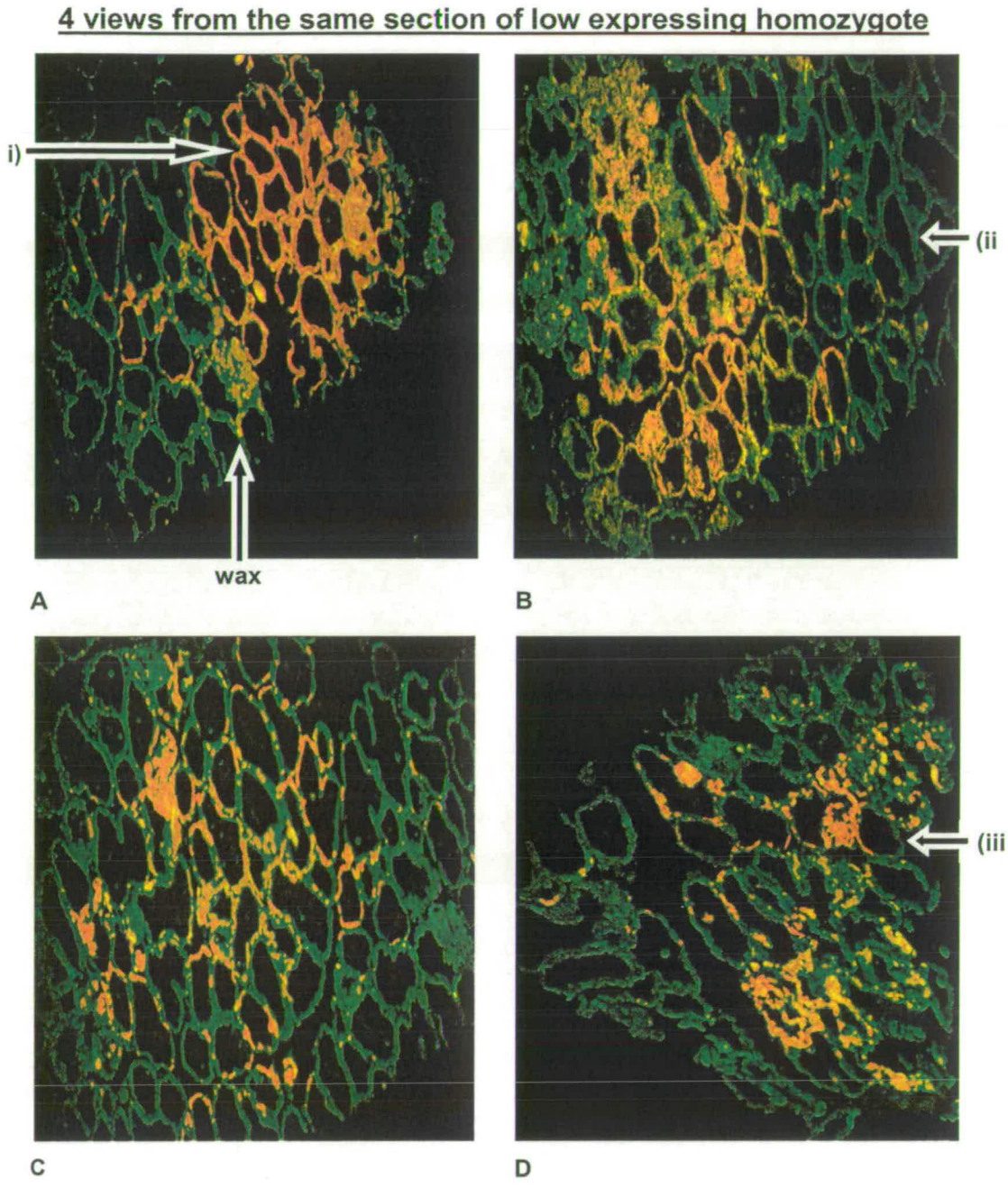


Figure 4.8 **Double mRNA FISH (BLG & β-casein) from low BLG/7 homozygote expressor (HH16 = 7.7 mg/ml) Scale: 1 cm = 80 μm**

4 different views (A – D) from the same section of lactating mammary tissue shown, magnified 10 X. Probe concentration used was 40 ng/ml. Probes used were BLG (Cy3 = red signal) & β-casein (Fluorescein = green signal) All epithelial cell express β-casein (green), some also express BLG (red + green = orange). Note that alveoli can comprise of cells expressing both genes (**i**), β-casein (**ii**) or a combination of the two (**iii**). Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

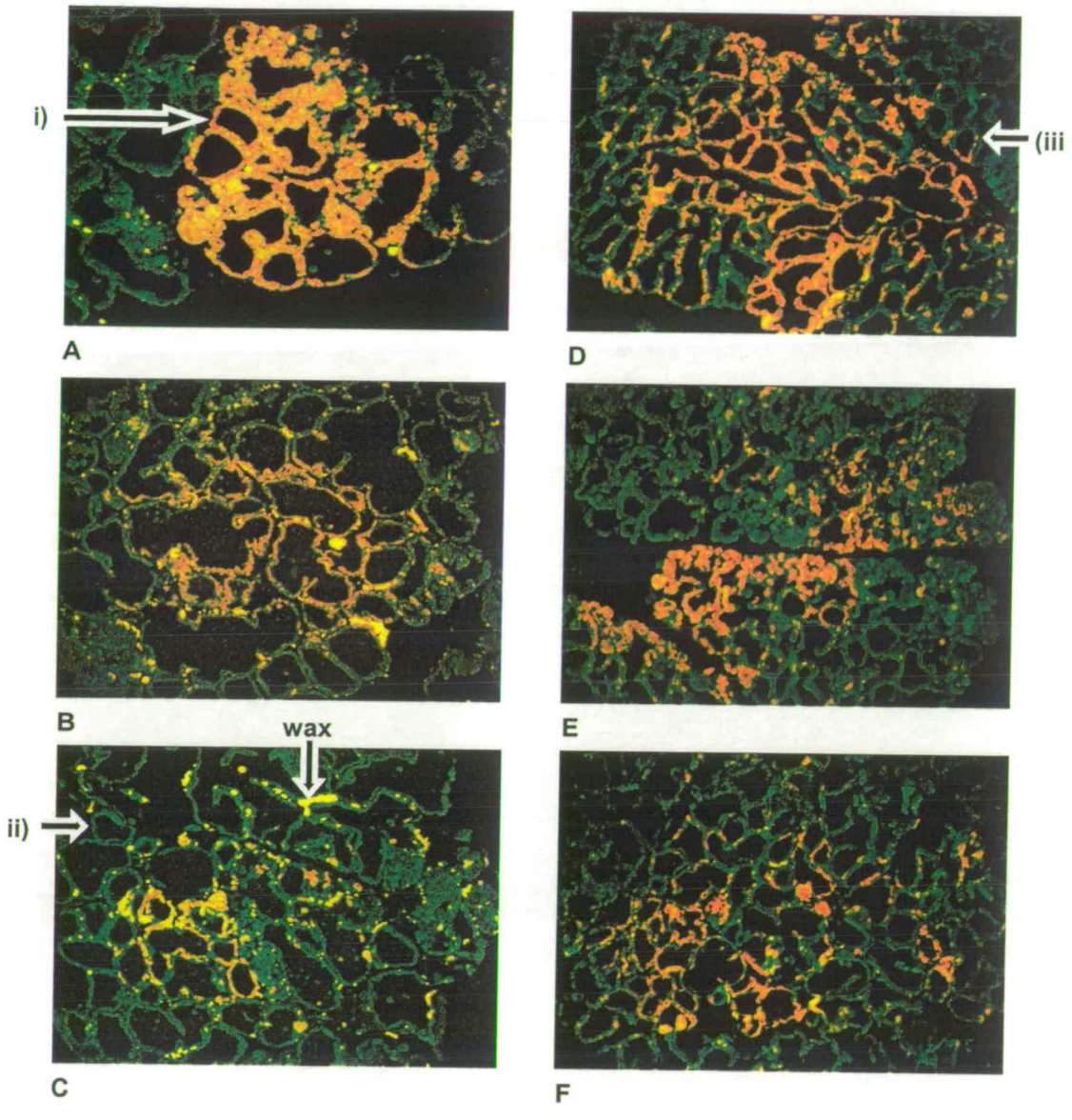


Figure 4.9 **Double mRNA FISH (BLG & β-casein) of similar hemi- (A-C) and homozygous (D-F) expressors (H2 = 8.8 mg/ml, HH9 = 9.3 mg/ml) Scale: 1 cm = 100 μm**

Three different views of the same section of hemizygote H2 (A-C) and homozygote HH9 (D-F) lactating mammary tissues shown, magnified 10 X. Antisense probes for β-casein (Fluorescein = green signal) and BLG (Cy3 = red signal) mRNA expression used at a concentration of 40 ng/ml. All epithelial cell express β-casein (green), some also express BLG (red + green = orange). Note that alveoli can comprise of cells expressing both genes (**i**), β-casein (**ii**) or a combination of the two (**iii**). Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

The compiled results for each round of *in situ* counts are in Table 4.4, with detailed data in Table 4.5 (hemizygous results) and Table 4.6 (homozygous results) at the back of this chapter. These results are also represented graphically, with Figure 4.10 (a - c) showing the results of rounds 1 & 2, 3 & 4, and the combined results with BLG protein expression plotted against the proportion of active alveoli. For both groups, the proportion of expressing cells increases with the concentration of transgenic protein expressed in the milk. Both sets of mice appear to require similar levels of expression, as determined by % of expressing cells, to generate equal levels of protein in the milk (Figure 4.10). In neither population did the proportion of expressing alveoli exceed 50%, the highest were scored at around 40%. There may be a small trend towards a higher proportion of expressing cells in the homozygous samples, but this is not significant when human and sampling error are taken into consideration. Homozygote samples are generally higher expressors than hemizygotes, with subsequent higher probability of overestimation of the number of active cells. The sampling differences due to microtome experience mentioned above (4.4.1) will also contribute to a slight increase in the wobble factor.

4.4.4 The homologous loci of homozygous BLG/7 mice are not independent

The homozyogous BLG/7 population exhibits a doubling in the minimum expression but displays an equal maximum expression level when compared to the hemizgous population. The homozygous population has a larger proportion of higher expressing individuals, which explains the differences in mean expression and variance. The presence of two transgenic loci in the homozygous population doubles the probability that one of the loci will be active, presuming that the two loci act independently of each other, and can account for the near doubling in minimum transgenic protein expression compared to the hemizygous population. The identical expression ceiling in both populations, on the other hand, suggests that the two loci cannot be operating independently of each other. Three testable models were proposed to explain the observations above:

1) Both homologous loci are active, but at half the activity of the individual hemizygous loci. Activation remains a random process.

2) Only one of the two loci is active, with activity on par with hemizygous cells.

3) Both loci are active, but only in half the number of cells. Each locus has the same activity as its hemizygous counterpart.

Models 1 + 2 can be differentiated from model 3 by counting the number of actively expressing cells in the mammary tissue and correlating it to the known protein level in the corresponding milk samples.

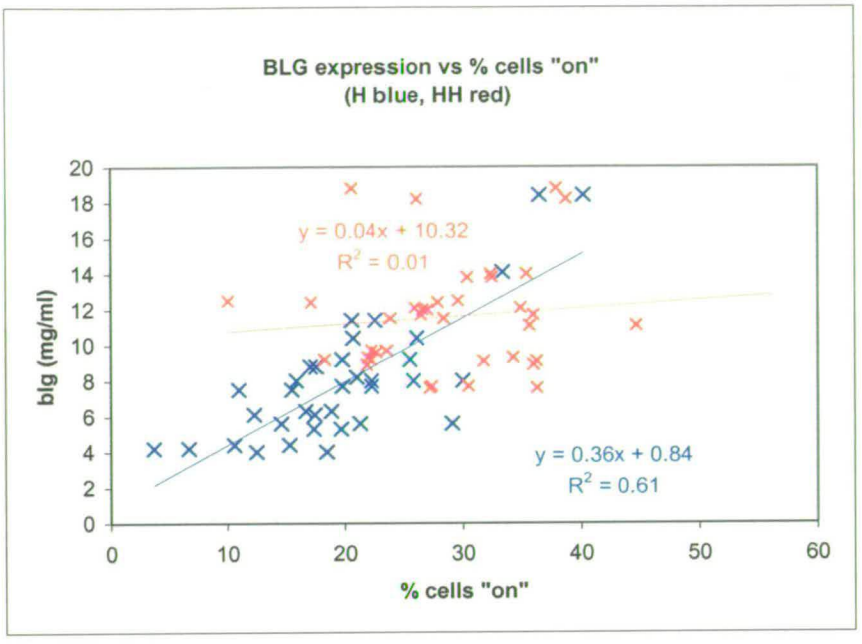
RNA FISH performed on hemi- and homozygous mammary tissue proved that both populations display variegation at the cellular level (Figures 4.5 – 4.9). Correlation of number of expressing cells with protein content in milk established that both populations exhibit similar expression patterns (Figure 4.10). There is no strong trend towards a halving or doubling in the percentage of transgene mRNA expressing cells in the homozygous population compared to the hemizygote results, enabling the elimination of model 3. This established that homozygous expression was either from a single locus in all expressing cells, or that both loci were active but at half the normal activity. This equates to models 1 & 2 and confirms that some form of *trans*-homology interaction is taking place in BLG/7 homozygous mice. To differentiate between the two models, primary transcript FISH was performed on homozygous BLG/7 mammary tissue.

As neither hemizygotes nor homozygotes exhibited more than 40% actively expressing cells, the search for nuclear transcription foci was concentrated in the higher expressing samples. RNA FISH against intronic BLG sequences in both hemi- and homozygous mammary tissue provided nuclei with evidence of transcription. In both cases, an elongated transcription focus was evident (Figure 4.11). This elongated focus suggests that in transcribing cells, transcription may occur from multiple copies of the gene within the single array site. The size of the transcription focus appeared to be the same in both hemi- and homozygous samples, although it is

Hemizygotes	[BLG]	Round 1	Round 2	Round 3	Average Expression
H1	14.1 mg/ml	33.40%	20.90%	28.90%	**27.73%**
H2	8.8 mg/ml	17.60%	17.10%	14.20%	**16.30%**
H3	8.2 mg/ml	21%	11.60%		**16.30%**
H4	18.4 mg/ml	40.20%	36.50%	32.50%	**36.40%**
H5	6.1 mg/ml	17.50%	12.30%		**14.90%**
H6	11.4 mg/ml	22.60%	20.60%	12.90%	**18.70%**
H7	5.3 mg/ml	17.40%	19.70%	10.80%	**15.97%**
H8	7.7 mg/ml	19.80%	22.30%		**21.05%**
H9	8.0 mg/ml	22.30%	15.90%	15.60%	**17.93%**
H10	5.6 mg/ml	13.00%	7%		**10.00%**
H11	8.0 mg/ml	19.50%	29%		**24.25%**
H12	5.6 mg/ml	14.60%	18.80%		**16.70%**
H13	6.3 mg/ml	16.70%	18.90%	8.70%	**14.77%**
H14	4.4 mg/ml	10.60%	15.30%		**12.95%**
H15	10.4 mg/ml	26.10%	20.70%	15.30%	**20.70%**
H16	4.0 mg/ml	12.50%	17.30%	6.90%	**12.23%**
H17	9.2 mg/ml	19.80%	25.50%		**22.65%**
H19	4.2 mg/ml	3.70%	6.70%	2.70%	**4.37%**
H20	7.5 mg/ml	11%	15.50%	5.70%	**10.73%**
Homozygotes					
HH1	13.8 mg/ml	32.49%	30.40%	31.40%	**31.43%**
HH2	14 mg/ml	32.38%	35.40%	28.40%	**32.06%**
HH3	9.2 mg/ml	18.32%	22.30%	17.80%	**19.47%**
HH4	9.6 mg/ml	22.74%	19%		**20.87%**
HH5	9.7 mg/ml	22.30%	23.60%	13.90%	**19.93%**
HH6	12.1 mg/ml	25.94%	34.90%		**30.42%**
HH7	7.6 mg/ml	23.42%	33.40%		**28.41%**
HH8	12.5 mg/ml	11.97%	29.60%	22.90%	**21.49%**
HH9	9.3 mg/ml	34.27%	22%	19.10%	**25.12%**
HH11	18.2 mg/ml	38.74%	26.60%	33.50%	**32.95%**
HH12	8.9 mg/ml	21.91%	29%		**25.66%**
HH13	11.1 mg/ml	31.58%	30.10%		**30.84%**
HH14	9.1 mg/ml	26.83%	29.20%	18.80%	**24.94%**
HH15	11.7 mg/ml	26.46%	33.50%		**29.98%**
HH16	7.7 mg/ml	25.88%	22.80%		**24.34%**
HH17	18.8 mg/ml	37.93%	20.60%	31.00%	**29.84%**
HH18	11.5 mg/ml	23.88%	28.40%	22.60%	**24.96%**
HH19	12.4 mg/ml	17.69%	27.90%		**22.80%**
HH20	12 mg/ml	27.02%	26.60%	23.80%	**25.81%**

Table 4.4 Summary of mRNA FISH counts for both hemi and homozygote BLG/7 mice
Each round represents one section where the numbers of active cells are evaluated as the proportion of alveoli expressing the transgene versus the total number of alveoli in view. Three separate rounds of mRNA FISH were performed, and the average of the three counts have been used.

1st round of in situ expression level results



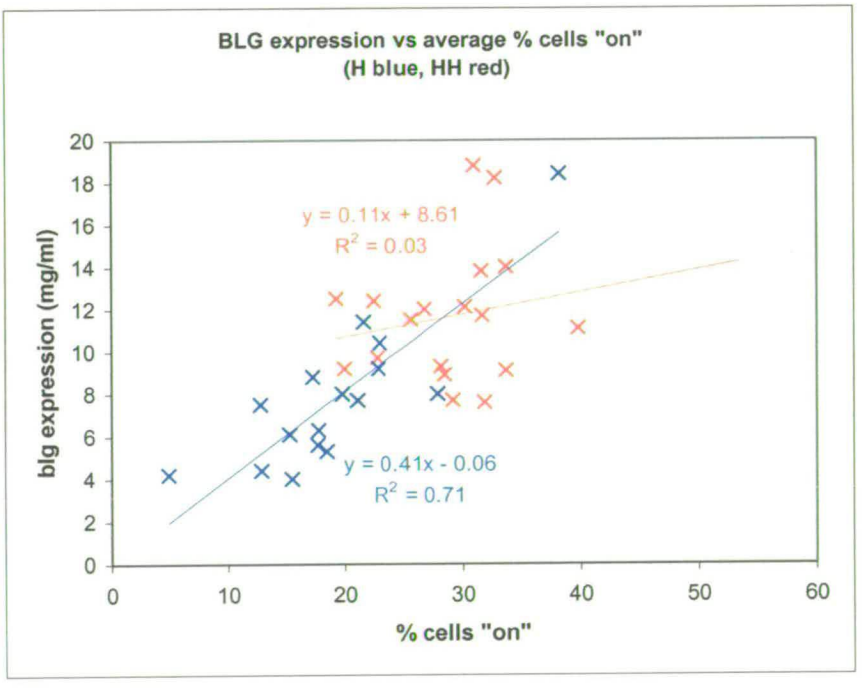
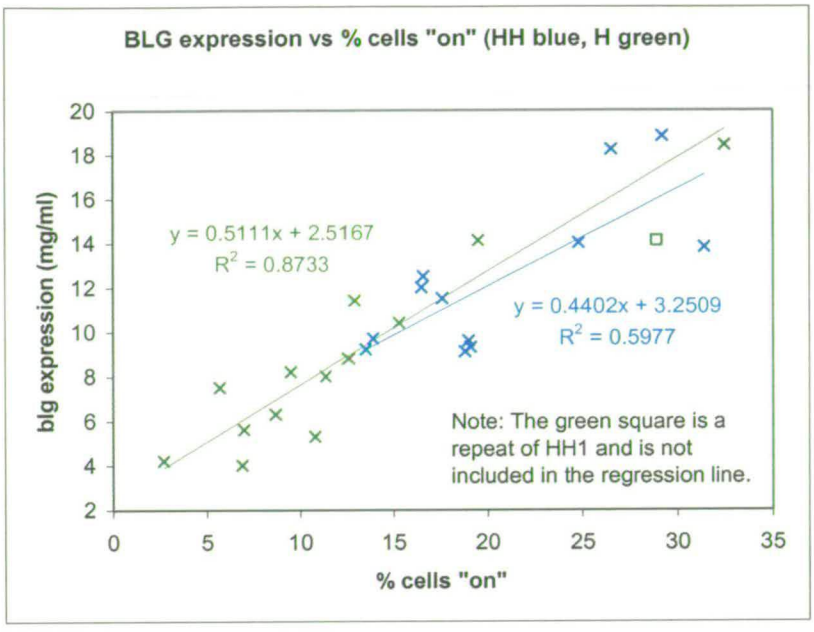


Figure 4.10 A Graphical representation of the data from table 4.4
1st and 2nd rounds of *in situ* count results. The percentage of active cells (as determined by the proportion of alveoli in a given section actively expressing BLG mRNA) is plotted against the level of BLG protein in the milk of an individual. Each cross represents a single individual mouse.

2nd round of *in situ* expression level results



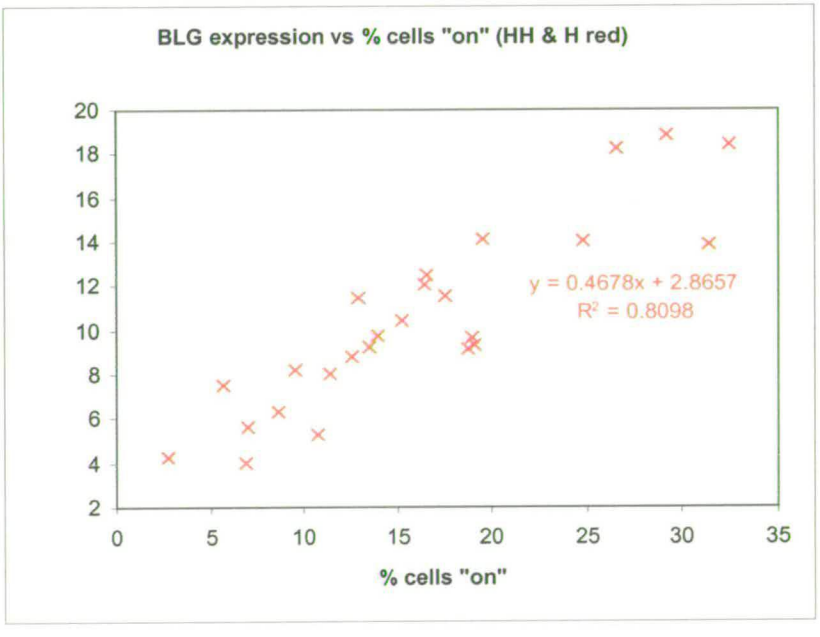
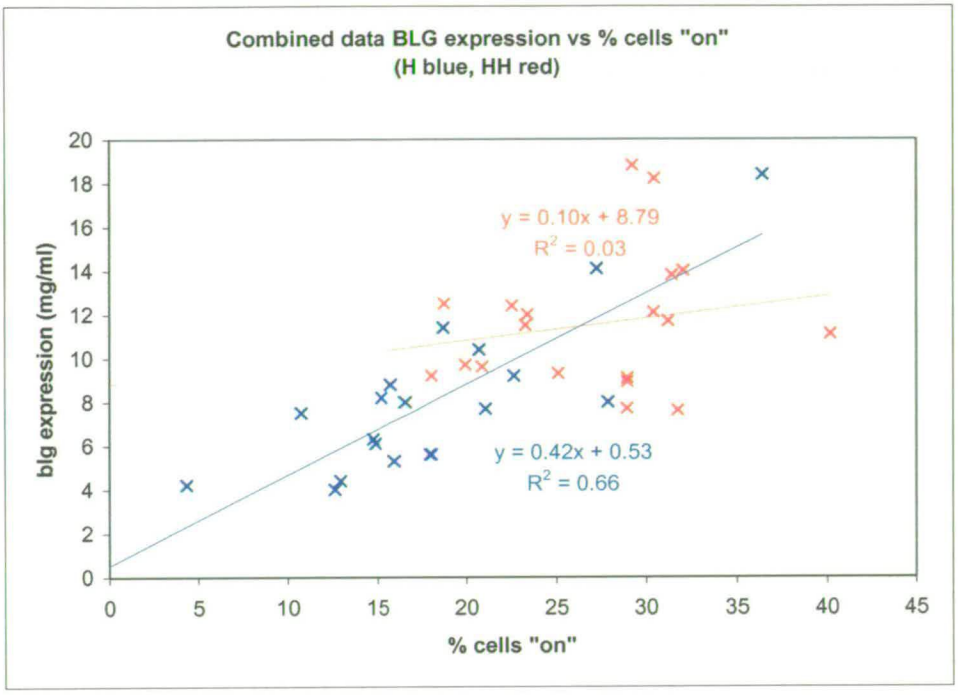


Figure 4.10 B Graphical representation of the data from table 4.4
3rd and 4th rounds of *in situ* count results. The percentage of active cells (as determined by the proportion of alveoli in a given section actively expressing BLG mRNA) is plotted against the level of BLG protein in the milk of an individual. Each cross represents a single individual mouse.

Combined *in situ* expression level results



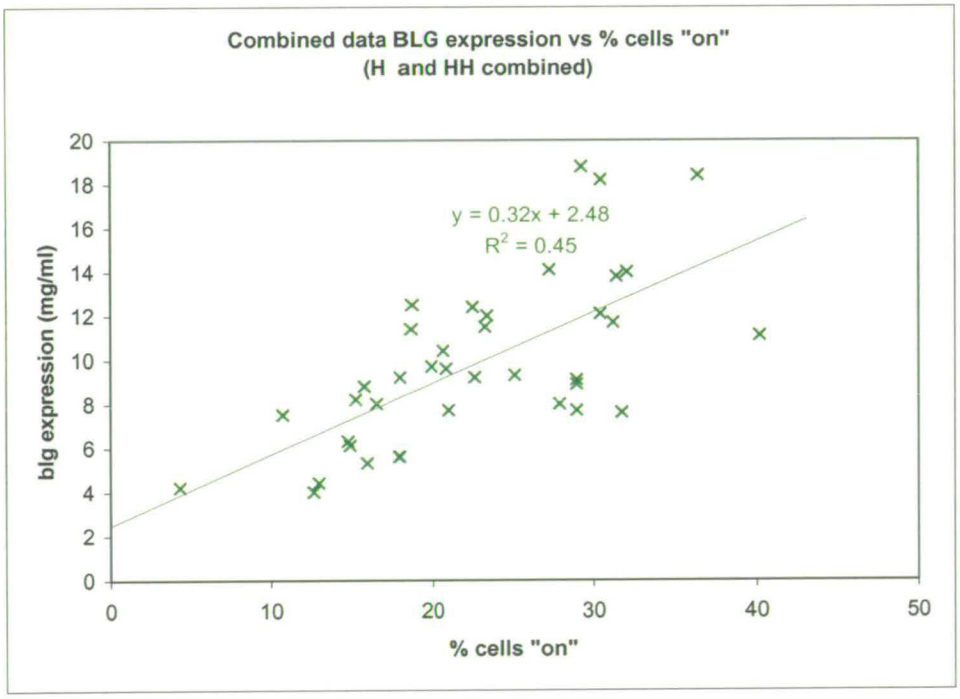


Figure 4.10 C Graphical representation of the combined *in situ* count results from table 4.4.

The percentage of active cells (as determined by the proportion of alveoli in a given section actively expressing BLG mRNA) is plotted against the level of BLG protein in the milk of an individual. Each cross represents a single individual mouse.

difficult to be sure as the size did vary in all samples depending on the cut and orientation of the nucleus. No evidence could be found for a spatially separate second transcription focus in homozygous samples, even when using rhodamine conjugated DiG antibody whose signal is weaker than that provided by the fluorescein hapten (Figure 4.11). It is possible that both alleles are transcribing, with co-localisation rendering it impossible to dissociate the two transcripts, resulting in the observed single elongated transcription focus. One control would be to check the number of expressing foci from an endogenous milk protein gene, although ultimately the intronic transcription would have to be correlated to the nuclear localisation of the two genomic loci in homozygous tissue samples.

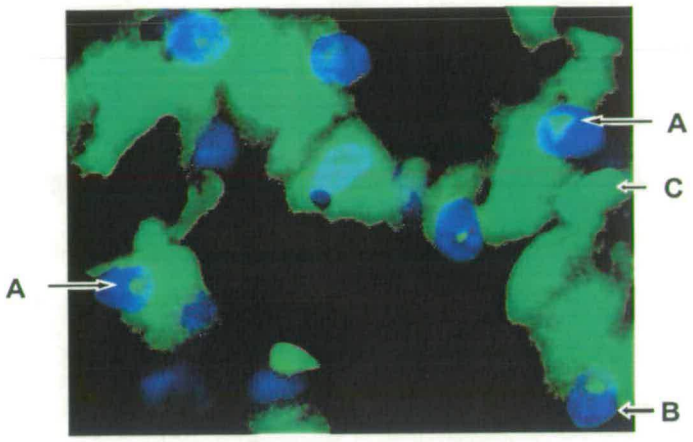
WAP primary RNA FISH suggests that both allelic loci are active for endogenous milk protein genes

WAP was chosen as the control gene for primary transcription in the nucleus. The primary probes used fluorescein haptens. Hybridisations were performed as usual, and a faint signal was discernible after 9 amplification rounds, using the fluorescein-conjugated anti-fluorescein antibody. This signal could be captured by the digital camera after a further 3 rounds of amplification (Figure 4.12). Two small, but distinct foci can be seen for WAP intronic transcription. These foci are close to, but physically separate, from each other. WAP is not highly expressed on day 11 of lactation, so the minute stature of the expression foci was anticipated. An extrapolation of this result would be that all endogenous milk protein genes probably express from both loci.

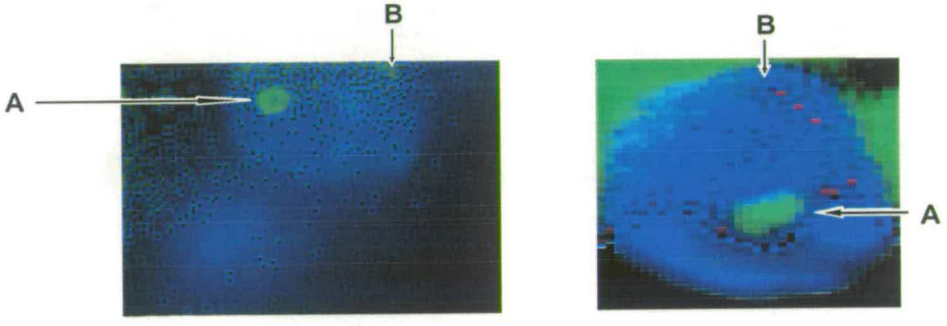
BLG/45 homozyogus individuals express from both genomic loci

Whilst trying to generate a working probe for endogenous gene intronic signal, samples of mid-lactation mammary tissue from BLG/45 homozygous individuals were collected. These mice are from a transgenic line known to exhibit a doubling of BLG expression in the homozygous population (McClenaghan *et al* 1995). When the tissues were probed for primary transcription, a single large transcription focus was initially revealed (Figure 4.13). Using the antibody conjugated to the less brightly fluorescent rhodamine conjugate, the signal resolved into two discrete foci positioned

Intronic RNA FISH for BLG



Homozygous mammary tissue probed for intronic BLG transcription (Green) showing several actively expressing nuclei. Intronic BLG signal (**A**) is from antibody to DiG (Fluorescein conjusgated = green),which was the hapten of the oligo cocktail used. Nuclei (**B**) are DAPI counterstained (= blue) Note that size and shape vary, which could be due to different orientations of the nuclei relative to the cut of the section. The transcription foci were generally quite elongated, but never resolved into two separate foci. Green fluorescence outside of blue nucleii (**C**)due to autofluorescence of tissue itself and luminescence from outside the focal plane.



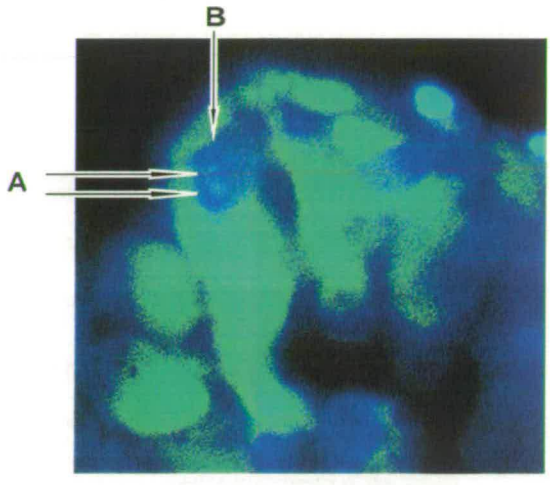
Close-up of transcription focus from a homozygous high expressor. Intronic BLG signal (**A**) is from antibody to DiG (Fluorescein conjugated = green),which was the hapten of the oligo cocktail used. Nucleus (B) is DAPI counterstained (= blue).

Figure 4.11 Intronic RNA FISH for nuclear BLG/7 transcription in homozygous sample (HH17)

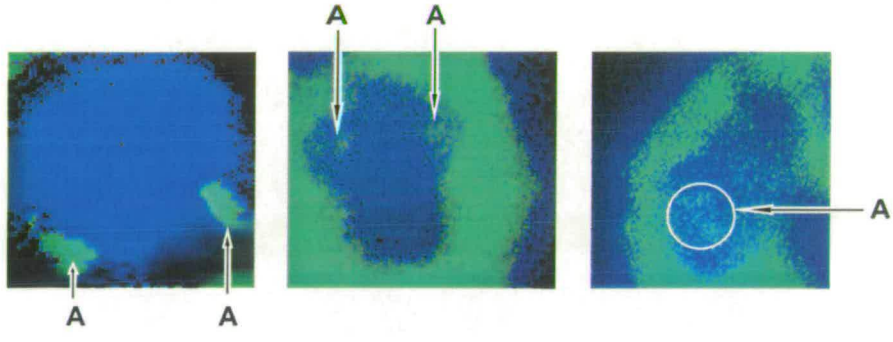
Homozygous high expressors exhibit only a single (green) transcription focus (A) in each nucleus (B) visualised through DAPI staining (blue). Arrows point to representative examples in the images.

Intronic BLG probe is DiG conjugated (40 ng/ml cocktail of 5 30'mer oligos) and visualised using antibody to DiG conjugated to fluorescein (green). Nuclei are counterstained with DAPI (blue) present in mounting medium. Magnification is 100 X. Average nucleus diameter 2 μm.

Intronic RNA FISH for WAP



Intronic WAP transcription: two green (Fluorescein) foci (**A**) are visible in the nucleus (**B**), which is DAPI counterstained (blue). Green fluorescence outside blue nucleus due to autofluorescence of tissue itself and luminescence from outside the focal plane.



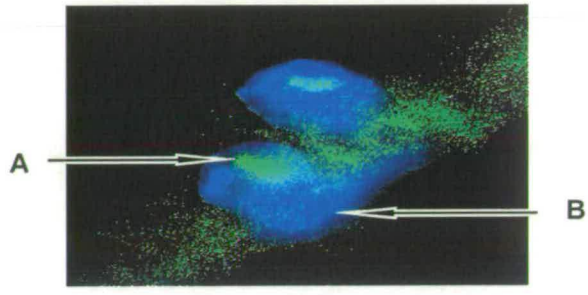
Close-up of transcription focus from a homozygous high expressor. Green intronic WAP signal (**A**) is from antibody to Fluorescein (Fluorescein conjugated), which was the hapten of the oligo cocktail used.The signal shows up as two small but distinct transcription foci. Image capture was difficult due to the small size of transcript foci and relatively faint signal compared to the blue DAPI nuclear stain. Green fluorescence outside blue nucleus due to autofluorescence of tissue itself and luminescence from outside the focal plane.

Figure 4.12 Intronic RNA FISH for nuclear WAP transcription in homozygote BLG/7 mammary tissue

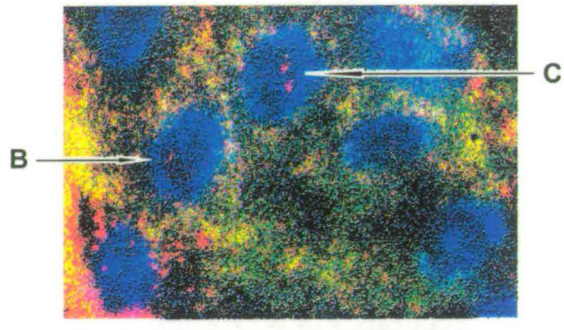
Two small, but distinct transcription foci (**A**) are visible in each nucleus (**B**), evidence that transcription occurs from both genomic WAP loci.

Intronic WAP probe is Fluorescein conjugated (40 ng/ml cocktail of 6 30'mer oligos) and visualised using antibody to Fluorescein conjugated to Fluorescein (green). Nuclei are counterstained with DAPI (blue) present in mounting medium. Magnification is 100 X. Average nucleus diameter 2 μm.

Intronic RNA FISH for BLG in BLG/45 Homozygotes



BLG/45 homozygote mammary probed for intronic BLG expression (= green) using primary probes labelled with DiG, and Fluorescein-conjugated antibody to DiG as secondary probe. The signal (**A**) is very bright, with an elongated transcription focus visible that may or may not come from two genomic loci. Nuclei (**B**) are DAPI counterstained (= blue). Green fluorescence outside of nucleii due to autofluorescence of tissue and luminescence from outside the focal plane.



BLG/45 homozygote mammary probed as above, but this time the secondary probe is conjugated to the less brightly fluorescing Rhodamine (= red) which clearly shows that the elongated transcription focus above is resolved into two distinct foci (**C**) in the DAPI (= blue) stained nucleii (**B**). Transcription is therefore taking place at both genomic loci in BLG/45. Green and red fluorescence outside of nucleii due to autofluorescence and luminescence from outside the focal plane.

Figure 4.13 Intronic RNA FISH for nuclear BLG transcription in homozygous BLG/45 sample using either Fluorescein (green, top image) or Rhodamine (red, bottom image) conjugated antibodies to DiG.

Homozygous BLG/45 exhibit two transcription foci (**C**) in nucleus (**B**) when DiG antibody is conjugated to the less brightly fluorescing Rhodamine hapten (= red). This proves that BLG can be simultaneously transcribed from both homologous genomic loci.

Intronic BLG probe is DiG conjugated (40 ng/ml cocktail of 5 30'mer oligos) and visualised using antibody to DiG conjugated to Fluorescein (green) or Rhodamine (red). Nuclei are counterstained with DAPI (blue) present in mounting medium. Magnification is 100 X. Average nucleus diameter 2 μm.

close together (Figure 4.13). Two foci also became apparent if hybridisations using the fluorescein-conjugated antibody were left a week for the signal intensity to die down. This confirmed that the BLG gene is capable of bi-allelic transcription.

These results suggested that BLG/7 homozygous mice only transcribe from one of the two loci, but this could not be conclusively proven with the data available. RNA FISH protocols would have to be adapted to perform double FISH for both primary RNA transcription and genomic BLG/7 sequence localisation in order to identify the number of active alleles.

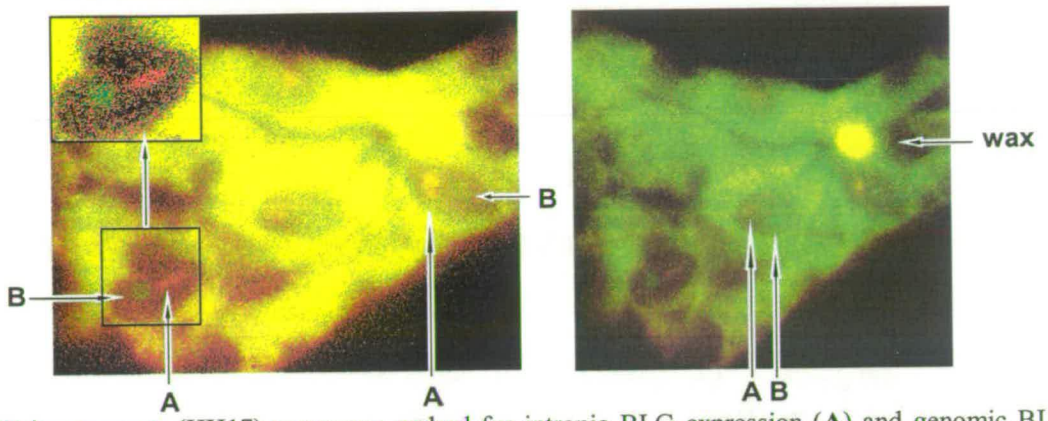
Intronic RNA and chromosomal FISH reveal only one active locus

It can be clearly seen in figure 4.14 that there are two distinct genomic loci, but only one transcription focus for BLG/7 homozygotes. The two genomic loci appear to be in two different focal planes, suggesting that they occupy separate nuclear locations. The single transcription focus co-localises with only one of the two genomic loci, proving that only one of the two loci is transcriptionally active in BLG/7 homozygous mice. This is an example of allelic exclusion, through some form of *trans*-homology interaction, a phenomenon never before reported in transgenic mammals.

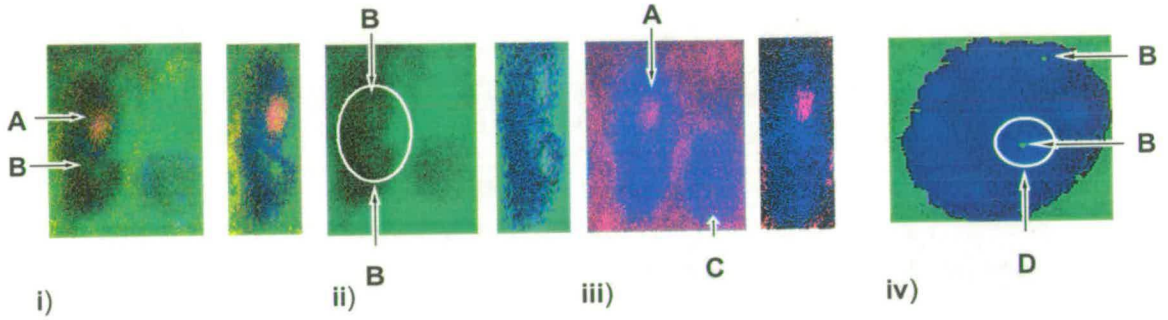
4.4.5 Nuclear positioning of BLG/7 alleles in liver suggestive of early separation of loci

Chromosomal FISH was also performed on liver sections taken from the homozygous BLG/7 mice (Figure 4.14 D). The nuclei appear to be much smaller, as is the chromosomal signal (12x amplification needed), but again there appear to be two distinct signals that may be localised in different nuclear compartments. This is suggestive of early compartmentalisation of the BLG/7 locus, taking place before the liver and mammary lineages part company.

An additional observation derived from the difficulty encountered in capturing images of the genomic BLG/7 location in liver nuclei. The image was captured as a



BLG/7 homozygote (HH17) mammary probed for intronic BLG expression (**A**) and genomic BLG location (**B**). Intronic transcripts (**red**) visualised through secondary detection of DiG labelled primary probes through Rhodamine-conjugated antibody to DiG (**A** = red signal). Genomic location (**green**) visualised through multiple amplification of fluorescein conjugated primary probes and fluorescein conjugated antibodies to fluorescein (**B** = green signal). It is evident that there is separation between the intronic signal (**A**) and a genomic signal (**B**). Green/Yellow fluorescence outside of nuclei (black circles) due to autofluorescence and luminescence from outside the focal plane.



BLG/7 homozygote (HH2) tissue probed as above, with images captured for both probes (**i**), genomic signal (**ii**), and intronic signal (**iii**). This series of images shows conclusively that transcription takes place at only one genomic locus in homozygote mammary tissue, which is covered by the red signal in the double image capture (see **i** vs **ii**). The second genomic locus does not associate with a transcription focus. **iv**) is a liver nucleus from the same animal probed for genomic BLG. The image clearly shows two small green foci that appear to be inhabiting quite separate parts of the nucleus, both co-localising with brigher areas of blue flourescence (**D**) which may be the condensed chromatin that forms the nucleolar bodies.

Figure 4.14 Intronic RNA & genomic DNA BLG FISH in homozygous BLG/7 tissue samples

Homozygous BLG/7 show two clearly separate genomic signals (**green**, **ii**) and only one transcription focus (**red**) which co-localises with one of the two (**green**) genomic signals (**i** & **ii**). Arrows identify nuclear transcript signal (**A** = red) and genomic signal (**B** = green). In i) – iii), nuclei (**C**) are DAPI counterstained (= blue).

Intronic BLG probe is DiG conjugated (40 ng/ml cocktail of 5 30'mer oligos) and visualised using antibody to DiG conjugated to Rhodamine (Red). Genomic BLG probe is fluorescein conjugated (40 ng/ml cocktail of 10 30/mer oligos) and visualised after amplification using antibodies to fluorescein conjugated to fluorescein (Green). Nuclei are counterstained with DAPI (Blue) present in mounting medium. Magnification is 100 X. Average nucleus diameter 2μm.

two-colour image using the DAPI counterstain to identify the nuclei. DAPI binds nucleic acids (predominantly DNA, although some binding to RNA can occur) and fluoreseces in the blue light range. Although the genomic probes (fluorescein tagged) were clearly visible under the microscope, capturing the signal was a problem due to the greater signal intensity of the DAPI signal, especially in the areas of presumed condensed chromatin that comprise the cytological structures termed chromocenters (Singh & Huskisson 1998) or heterochromatin foci (Netzer *et al* 2001). The BLG/7 genomic loci have an apparent tendency to co-localize with these structures, the DAPI signal obliterating the weaker fluorescein signal. This nuclear localisation, together with their apparent physical separation in a tissue where neither is expressed, provides additional circumstantial evidence that BLG/7 expression may be regulated in part by its specific nuclear localisation.

4.5 Discussion

The general assumption has been that if a transgene is expressed in the hemizygous condition, then the homozygous individuals should exhibit a doubling in both transcription and translation; provided the cellular machinery can cope. The validity of this belief has been challenged by reported homology-dependent silencing of both transgenes and endogenous genes (tenLohuis *et al* 1995, Vaucheret 1993, Henikoff 1997, Selker 1999, Birchler *et al* 2000). The majority of papers describe plant systems, with a few depicting similar phenomena in *Drosophila* (Dorer & Henikoff 1994, Pal-Bhadra *et al* 1997, Muller *et al* 1999) and fungi (Selker 1999). For homozygous variegating transgenes, only one paper describes a detailed investigation (Davies *et al* 1997). Arabidopsis containing additional copies of endogenous chalcone synthetase gene variegated when hemizygous, with expression from all loci silenced in the homozygous state. Outcrossing only partially alleviated the silencing, possibly due to incompleter erasure of methylation during meiosis.

To my knowledge, no investigations have been made into the effects of a homozygous environment on a variegating transgene in mammals. Belyaeva and co-workers (1993) and Lloyd and co-workers (1997) looked at the behaviour of two different variegating constructs in the same system (both used *Drosophila*) (4.1). The

results suggested that there probably are limiting factors and mechanistic pathway overlaps operating in the various occurrences of PEV. Homozygous BLG/7 mice contain the transgenic locus on both homologous chromosomes, and the two homologous loci are thus initially subjected to exactly the same external and internal influences.

The aim of this investigation was to test whether the highly variegating transgene locus of BLG/7 would behave independently when in the homozygous state, helping to elucidate whether mammals possess any of the homology-based *trans*-sensing mechanisms described for other systems (Wu & Morris 1999). BLG protein expression in milk was quantified, and mRNA FISH using wax embedded sections of lactating mammary tissue was used to determine the proportion of alveoli actively transcribing the transgene. The results for hemi- and homozygous BLG/7 populations were compared, and indicated that the two homologous homozygous loci do not variegate independently. Subsequent FISH for nuclear transcription and chromosomal FISH proved that BLG/7 expression is mono-allelic.

4.5.1 The hypothesis

As our BLG/7 mice, along with the separate BLG/45 line of mice carrying a similar construct at a different integration site, were the first reported transgenic mammals with PEV (Dobie *et al* 1996), further characterisation of this line was undertaken by breeding homozygous individuals and investigating the effect of two homologous variegating loci on transgene expression. Four possible outcomes were hypothesised:

1. Average BLG expression doubles in the homozygous population, indicating independent operation of each locus without limiting factors at either the transcriptional or the translational stages.

2. Expression is upregulated but not doubled in the homozygous populations, indicating a limiting step present at the transcriptional or translational stages.

3. The homozygous population exhibits expression levels equal to the hemizygote population, indicating some form of homology–effect and/or a rate-limiting step.

4. Expression is downregulated in the homozygous population, indicating that some form of homology-dependent silencing is activated (a phenomenon so far only reported in plants, *Drosophila*, fungi and yeast)

Option 2 could be eliminated immediately, as a second variegating line of mice containing the β-lactoglobulin transgene (BLG/45) can secrete up to 1.5 times more transgenic protein in their milk in their hemizygous state than the highest BLG/7 hemizygous expressors (Dobie *et al* 1996). Homozygous BLG/45 mice secrete a mean value of 33 mg/ml BLG in their milk, close to twice the highest values measured for hemizygous BLG/7 individuals (McClenaghan *et al* 1995, Dobie *et al* 1996). The BLG/45 expression data thus indicate that the cellular machinery is capable of high protein throughput, and should therefore not pose a rate-limiting step for BLG/7 transgene expression. Nevertheless, the different integration locus of BLG/7 may impose a positional disadvantage in recruitment of transcription factors that may only become evident when two loci compete for a limited resource. Support for a possible positional disadvantage is provided by BLG/7 location to chromosome 15, whilst endogenous casein genes reside on chromosome 5 and WAP on chromosome 11. In addition, Wilkie and colleagues (1999) suggest that chromosomal position may determine nuclear localisation of a gene. If this is true in mice as well as *Drosophila*, then it is likely that BLG/7 is not located near any of the endogenous milk protein genes in the nucleus.

Two of the remaining three hypothetical outcomes are loosely based on an extrapolation of Balasov & Makunin's theory (1.7.7). Assuming that the homozygous loci act independently of each other, with heterochromatic proteins in limiting supply, I would expect twice or more than double the amount of expression when compared to the hemizygous population. This scenario equates to option 1. The logic behind this is that two loci double the likelihood that a given cell will contain a transcriptionally active transgene locus compared to single locus animals. If a

limiting heterochromatic factor is titrated out, there is increased chance that both loci will be operational in the same cell, more than doubling the output. This scenario assumes that the homologous loci operate independently of each other. An alternative scenario presents itself where the homologous loci pair and interact. If this occurs, with heterochromatin components in abundance, the CICs of both sites could cooperatively silence both alleles to a greater extent than expected for independent loci, giving rise to decreased expression levels when compared to the hemizygotes. This scenario equates to option 4.

Having eliminated option 2 of the four hypothetical outcomes, and assigned options 1 and 4 to the two scenarios outlined above, only hypothetical outcome 3 remains to be accounted for. This outcome was not considered to represent a likely event, as no evidence could be found in the literature to support it, until the publication of a recent transgenic plant paper. A transgenic line is described where homozygous plants exhibit expression levels identical to that of their hemizygous counterparts (Matzke *et al* 2001). Earlier investigations had been suggestive of such an event, but had failed to characterise the phenomenon in detail (Neuhuber *et al* 1994).

4.5.2 Transgenic BLG protein expression does not double in homozygous population compared to hemizygous results

Both populations display variegated BLG protein expression in the milk (Figure 4.1 & 4.3, tables 4.2 & 4.3). The near doubling in the minimum expression observed for the homozygous population compared to the hemizygous population is not unexpected. Assuming the two homologous loci operate independently of each other, the homozygote mice have twice the likelihood of at least one of the two sites becoming activated than the hemizygote mice have with only one transgenic locus.

The observation of a maximum expression level of 20 mg/ml for both BLG/7 populations is equivalent to hypothetical outcome 3 (4.5.1), where equal expression levels are exhibited by both hemi- and homozygous populations due to either homology effects or a rate-limiting step. As transgenic line BLG/45 demonstrates that expression from this gene can far exceed 20 mg/ml (4.5.1), it was considered

unlikely that the homozygous expression pattern could be attributed to any physiological constraint. This suggested that the absence of doubling of maximum BLG expression in the homozygous population must be due to a property of the transgene array at the site of integration, possibly due to a homology effect.

Three possible explanations for the milk protein results were suggested:

1. Each active homozygous cell displays only half the activity of an equivalent hemizygous cell.

2. Each active homozygous cell displays twice the activity of the equivalent hemizygous cell, with the homozygous nature of the locus rendering it twice as susceptible to inactivation.

3. Each active cell displays an equal level of activity irrespective of genetic composition.

Scenario 1 is unlikely as it cannot account for the doubling of minimum BLG expression levels in the homozygous population. Scenarios 2 & 3 can both account for the observed protein expression patterns, and equate to hypothetical outcomes 3, and 1 & 2 respectively (4.5.1). The two scenarios can be differentiated by determination of the number of expressing cells required to give rise to a given level of protein expression in both populations. Scenario 2 would require that homozygous individuals display half the number of expressing cells than the equivalent hemizygous sample, whereas scenario 3 proposes that equal number of cells will be needed to generate the same protein level in the milk of both populations. To differentiate between these two scenarios, mRNA FISH was performed on wax embedded mammary tissue and the proportion of alveoli actively expressing the transgene were determined for each individual.

The latter two scenarios both hinge on some form of homology effect taking place, either to silence transcription from one or both loci, or to reduce transcription pace. Although there are plenty of reports in the literature citing homology dependent gene

silencing (1.2 & 4.1), these examples are not representative of what is seen for BLG/7 mice. Transcription is not silenced, just maintained at hemizygous levels in both populations. *Transvection* in Drosophila involves pairing of homologues, with inactivation of both a frequent result (1.2.3). This is also the case for most examples of paramutation in plants, although sometimes expression persists at a lower level (1.2.3). These two phenomena involve the pairing of two dissimilar gene alleles, whereas BLG/7 homozygotes contain two identical loci. The involvement of either transfer of epigenetic marks or sequestration to a heterochromatic nuclear compartment in these documented phenomena (1.2), suggests possible mechanistic links to BLG/7 homozygote behaviour. The location of BLG/7 close to centromeric heterochromatin may facilitate pairing and exchange of heterochromatin between the two loci. Therefore, despite the differences in the outcomes of *transvection* and paramutation to BLG/7 homozygote expression patterns, it is possible that an analogous process is involved in BLG/7 expression control.

4.5.3 Hemi- and homozygote BLG/7 mice exhibit similar cellular expression patterns

Despite the manual evaluation of the number of expressing cells, the process remained accurate enough to distinguish between a halving in the number of expressing cells in homozygous samples compared to equivalent hemizygous expressors. The results show conclusively that there is no doubling or halving of the number of expressing cells when the two populations are compared (Tables 4.4–4.6). Homozygous BLG/7 expression is therefore either from both loci at half the normal activity, or from only one of the two loci. Some form of *trans*-homology interaction is thus likely to be taking place in these mice.

The observation that different views from the same section, as well as views from a different section of the same tissue can vary drastically, underpins the variability of the expression patterns seen in BLG/7 mice. A variegated pattern of expression that varies throughout the mammary tissue adds an additional level of complexity to the assignation of correct estimation of the percentage of expressing cells that are

required to generate a given level of protein. This helps explain the "wobble" in the graphs of figure 4.10 (a-c).

4.5.4 Probes for BLG/7 nuclear transcription reveals a single focus of transcription in homozygous samples, whereas both homozygous BLG/45 mice and samples probed for WAP reveal two foci.

RNA FISH against intronic sequences exhibited a single elongated focus of transcription in both hemi- and homozygous mammary tissue (Figure 4.11). The elongated nature of the transcription focus suggests that in active cells, transcription may be taking place from more than one of the gene copies along the transgene array. Although the single focus is suggestive of transcription from one locus, it could not be discounted that the homozygous loci could co-localise.

Detection of nuclear transcription from endogenous whey acidic protein (WAP) revealed two small foci, close to but physically separate from each other (Figure 4. 12). As WAP is not highly expressed relative to the caseins on day 11 of lactation, the minute stature of the transcription foci compared to BLG/7 was anticipated. An extrapolation of this result would be that endogenous milk protein expression probably takes place from both alleles. Although it is presumed that most genes exhibit bi-allelic expression in all tissues, there have to my knowledge not been any reports into concerted examination of multiple endogenous gene expression patterns at the nuclear level.

Examination of homozygote samples from the line of transgenic BLG mice known to exhibit a doubling in expression levels from homozygote mice (McClenaghan *et al* 1995) enabled resolution of what was initially seen as a single large transcription focus into two distinct foci (Figure 4.13). This confirmed that the BLG gene was capable of bi-allelic transcription.

4.5.5 BLG/7 expression is monoallelic

Double FISH for nuclear transcription and genomic DNA revealed a single focus of transcription that co-localised with only one of the two genomic signals (Figure

4.14). This proved that some form of *trans*-homology interaction occurs in BLG/7 homozygous mice, which leads to allelic exclusion. In other words, in cells that are capable of actively expressing BLG, some mechanism operates to prevent expression from more than one of the two loci. This is the first reported *trans*-homology effect in transgenic mammals.

The observation that the two genomic signals appear to occupy different nuclear locations (focusing so that both signals could be imaged together proved difficult, suggesting that they do not occupy the same focal plane) suggests that both active and silent states are irrevocably committed in the cell. I saw no evidence for switching of expression states between the alleles, as seen for globin genes (Wijgerde *et al* 1995) where signals from two loci were sometimes retained in the same nucleus. It cannot be totally discounted that switching does not take place in my samples, the combination of low levels of cells expressing the transgene (40 % or less), with the low probability of capturing such a switch taking place, makes it possible that switching could be overlooked. The spatial separation of the two homologous loci in the nucleus does, however, argue against a model where expression switches from one locus to the other. It cannot, therefore, be discounted that expression from both loci may occur early in lactation, with one of the two loci inactivated in an analogous process to that determining either X-chromosome inactivation or post-transcriptional gene silencing (1.7.8, 1.2.1).

Monoallelic expression from endogenous genes

Monoallelic expression is already well documented for imprinted and X-linked genes (Tilghman 1999, Panning & Jaenisch 1998). There is also mounting evidence that other genes can display monoallelic expression (Chess 1998). IL-4 alleles in CD4+ T cells are independently regulated, and display both mono- and biallelic expression. This pattern of expression is "fixed" and clonally inherited, with the suggestion that IL-4 gene expression is random and established at a late stage of development (Bix & Locksley 1998). Another interleukin, IL-2 is also reported to exhibit monoallelic expression in some T-cells (Hollander *et al* 1998, Rhoades *et al* 2000), with other T-cells exhibiting bialllelic expression (Rhoades *et al* 2000). Randomised monoallelic

expression may also be the route used to generate functionally distinct olfactory neurons (Chess *et al* 1994). These reports suggest that there are more endogenous genes than previously thought that may exhibit monoallelic expression, at least in certain cell types. The monoallelic expression seen in homozygous BLG/7 mice may therefore be the result of its integration site rather than the transgene array *per se*. This is supported by BLG/45, where expression is from both loci (Figure 4.13). What none of the papers address is the actual mechanism by which one allele is functionally silenced, but not the homologue. Imprinting may hold a clue, at least as a plausible explanation for BLG/7 expression.

Although there are no known imprinted regions on mouse chromosome 15, where BLG/7 is integrated close to centromeric heterochromatin (somewhere between cytological bands B2 – second half of A2), it is possible that some of the genes close to the integration site will be imprinted and thus influence BLG/7 expression. There are precedents for location in imprinted regions influencing gene expression (Tilghman 1999, Jeong *et al* 2000), and genes can be influenced by events that take place at a considerable distance. Deletions up to 1Mb away have been documented to affect expression of imprinted genes in Praeder-Willy syndrome (Tilghman 1999).

There is also the possibility that the genes close to the transgene integration site are monoallelically expressed as part of a dosage compensation mechanism. This scenario appears increasingly plausible when one discovers that the integration site must be close to a cluster of hormone receptor genes that include Ghr, IL-7r, Lifr, Npr 3, Osmr, Prl r, Ptgerep 2 & 4 (Jackson Lab Mouse Genetics Database). There is evidence for the involvement and synthesis of growth hormone receptor in differentiating and lactating mammary (Feldman *et al* 1993, Glimm *et al* 1990). The involvement of the prolactin receptor in mammary function is well documented. It is not impossible that at least these two genes, whose involvement in mammary gland development and function is known, contribute to the regulation of these events through the mechanism of dosage compensation. BLG/7 may exhibit monoallelic expression due to close proximity to Ghr and Prlr, with shared temporal activation an accessory to the effect. As at least two interleukin genes are monoallelically

expressed, it is possible that control of the effects of the hormones is also exerted through receptor availability, therefore the IL-7r may also be monoallelically expressed. In this scenario, integration close to the IL-7r gene could also be responsible for the monoallelic expression pattern of BLG/7.

The monoallelic expression of genes close to the BLG/7 integration site may exert an effect on BLG/7 by ensuring that one of the two loci is automatically silenced along with the endogenous gene locus chosen for inactivation. The homologous BLG/7 locus could then be incorporated into a potentially active domain, but the final transcriptional status of the locus is controlled by factors not related to the control of expression of the monoallelically expressed endogenous gene. The potentially active locus is regulated by the mechanism responsible for the inherent variegating phenotype of BLG/7, which are independent of any other regulatory mechanisms in operation on nearby endogenous genes.

An alternative variation on the theme above would be that there is a selective pressure to keep one of the loci active, due to the presence of a vital gene near the integration site. The cells where both loci are inactivated could be rescued by intercellular interactions with active cells. I find this improbable, as some homozygote tissues exhibit no more than 10-15% active cells. The active cells are usually located in discrete patches, with large areas of the mammary gland lacking any cells actively expressing BLG. I cannot envision that the remaining 85-90% of mammary epithelial cells can be rescued by the expression levels of the few active cells.

Array structure as causative agent in *trans*-homology effect

The multicopy nature of the BLG/7 insertion may initiate heterochromatin formation in *cis*, with the close proximity to centromeric heterochromatin increasing the probability that heterochromatin will form at the site. Each site may be independent of the other for formation of a repressive chromatin structure, but there may also be intermediate formation of heterochromatin at the site that does not constitute a repressive structure (due to the limited availability of repressive factors). Only if both

sites are activated, are the two homologous loci exposed to each other. This contact, in the form of pairing, ensures that the sub-critical amounts of heterochromatic formation at both loci also come into contact, and transfer of heterochromatin from one site to the other enables inactivation of one of them (Balasov & Makunin's model operating in *trans*, 1.7.7).

The pairing of the two sites may be facilitated by the proximity to centromeric heterochromatin. It is believed that the heterochromatin plays a central role in the meiotic chromosomal pairing process (Renauld & Gasser 1997). The presence of repetitive sequences such as LINEs, SINEs, multiple pseudogenes and rRNA gene clusters may also contribute to this process.

4.5.6 Nuclear positioning of BLG/7 alleles

Homozygous BLG/7 tissues probed for genomic BLG/7 localisation exhibit two distinct signals that appear to be physically separate (Figure 4.14). This holds true for mammary tissue, where the active and silent loci occupy different nuclear locations. Prelimenary investigations in liver samples suggest that this physical separation of the loci also occurs in tissues where milk protein expression does not take place (Figure 4.14).

These observations suggest that the decision with regards to the activational status of the transgene locus may occur early in development (before cells differentiate into liver and mammary lineage pathways), with nuclear localisation a possible key player in maintaining the decision. These are prelimenary observations and further investigations are needed using samples from other tissues, as well as trying to establish their relative positions in the nucleus with regards to key factors such as heterochromatic proteins and milk protein genes.

4.5.7 Theoretical Model

Irrespective of the details of how monoallelic expression is determined, BLG/7 expression remains complex even when in the hemizygous state. Any theory as to when and how silencing of the transgene array is achieved has to be able to explain

both hemizygous and homozygous expression profiles. Both variegation and allelic exclusion have to be catered for, as well as the observation that there is expression from less than 50% of all cells, even in the highest expressing individuals. There is no clear way to reconcile these observations with current knowledge of gene expression control, I have therefore attempted to suggest a possible scenario. My model encompasses a three-step process, the first two steps common to both populations, with the third step exclusive to the homozygote condition (Figure 4.15).

A detailed explanation of the theoretical model depicted in figure 4.15

The first step takes place early in development, possibly coincidental with X-inactivation determination. A random process permanently inactivates the BLG/7 locus in approximately 50% of all the cells. There is no information in the literature that can cast a light on what such a process may be, although it could be linked to monoallelic expression of one or more of the endogenous genes close to the integration site. The remaining 50% contain a potentially active locus. In homozygous cells, inactivation of one locus also inactivates the homologous locus. This may be a *trans*-homology interaction step, and I envision a mechanism analogous to *trans*-inactivation of the *brown* locus in *Drosophila*. The multicopy nature of the transgene, coupled with the integration site close to centromeric heterochromatin, may render both sites susceptible if one of the pair is inactivated through ectopic modification. The modification and proximity to the centromere may drag both loci into a heterochromatic compartment in the nucleus, so that although only one of the sites may be physically silenced, the other copy, by its nuclear location, cannot become activated either.

The second step occurs later in development. A random but stochastic process proceeds to determine whether the potentially active (potentiated) loci will remain active or silenced. Once this decision is made, the cell is committed and passes on the decision to all its progeny. This potentiation versus activation step has some experimental evidence to back it up (Bulger & Groudine 1999). In the hemizygous mice, this final event leads to the observed variegated expression pattern.

Theoretical Model Depicting Inactivation events for Hemi- (H) and Homozygous (HH) BLG/7

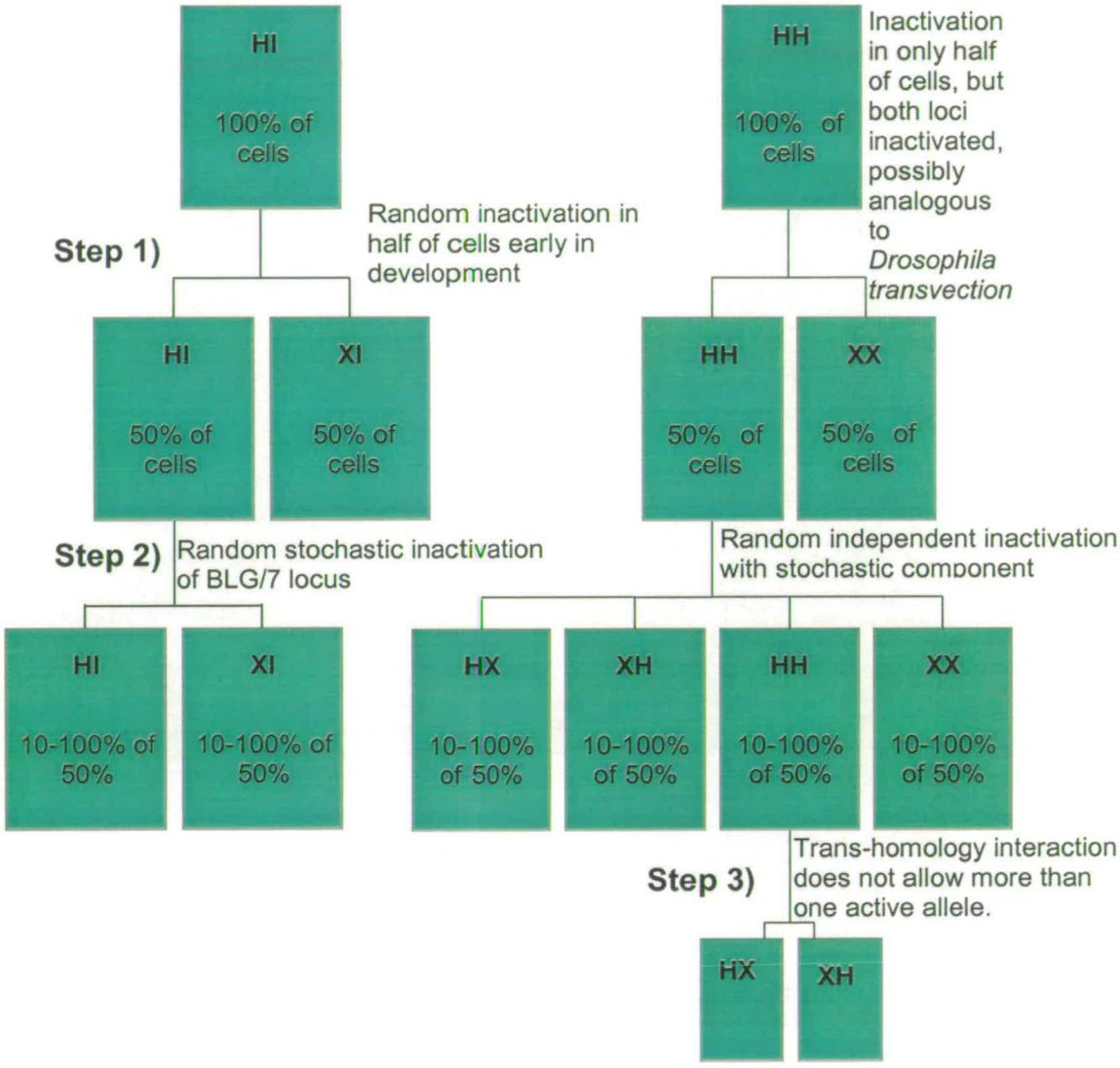


Figure 4.15 Theoretical Model explaining BLG/7 homozygote monoallelic expression.

H = BLG/7 locus, X = inactivated BLG/7 locus, I = locus without BLG/7 insertion

In the homozygous mice, each potentiated cell has two potentially active loci. The random process of step 2 has four possible outcomes, as seen in figure 4.15. One of the possible outcomes is that both loci are activated, but the results from the FISH for both mRNA and intronic transcription indicate that this does not take place. I have therefore included a putative 3rd step which is only operational when both alleles are activated in a single cell. This step, like one of the earlier steps, involves *trans*-homology interactions, but in this case, the outcome is radically different. One allele is inactivated whilst the other remains active. Perhaps the open configuration of both paired loci contains only enough bound heterochromatic protein components between them to inactivate one locus. The locus with the highest concentration of heterochromatic proteins may, through increased density of bound heterochromatin, possess a higher affinity for the available heterochromatic complexes. A directional flow of heterochromatin ensues until the one locus has become entirely heterochromatinised, disrupting the pairing interaction. The disruption of pairing allows the sequestration of the inactivated locus to a heterochromatic region of the nucleus, spatially separate from the active locus.

This model is the only one that I can think of which accounts for that no more than 50% of cells are transcriptionally active in either population, the within line variegation, the identical maximum expression level for both hemi- and homozygotes and the doubling in the minimum expression levels.

There is only one other report of a similar occurrence of transgenic allelic exclusion, and that is in transgenic plants (Matzke *et al* 2001). Whereas I have FISH results to prove that only one of the two alleles in the homozygous mice remains active, the Matzke paper relies on the results of an outcrossing experiment. The outcrosses from the homozygous plants segregate 50:50 active and silenced progeny, which suggests that only one of the two loci is active in the parental population, and that this epigenetic modification is maintained through the germline. As the animal germline is segregated much earlier in development, there may not be any germline transmission of BLG.7 locus activation states, although the relevant outcrossing experiments should still be done to eliminate this (see chapter 6).

1st set		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H1	a	27	5	7	13	2	2	11	233	33.4
	b	0	8	0	0	1	0	35		
	c	7	11	5	9	0	6	17		
	d	0	7	3	1	0	0	21		
	e	3	8	2	6	0	0	16		
	SUM	37	39	17	29	3	8	100		
	%	15.9	16.7	7.3	12.4	1.3	3.4	42.9		
H2	a	0	0	0	0	0	0	17	128	17.6
	b	1	13	5	4	0	0	7		
	c	0	4	5	1	0	5	11		
	d	0	4	5	0	0	0	14		
	e	0	5	6	1	0	0	9		
	f	0	0	0	0	0	0	11		
	SUM	1	26	21	6			69		
	%	0.8	20.3	16.4	4.7	0.0	0.0	53.9		
H3	a	4	6	2	3	0	0	8	176	21.0
	b	0	14	0	2	0	0	20		
	c	0	4	2	2	0	0	45		
	d	2	12	5	13	0	0	6		
	e	0	10	0	0	0	0	16		
	SUM	6	46	9	20	0	0	95		
	%	3.4	26.1	5.1	11.4			54.0		
H4	a	5	9	7	22	0	0	13	196	40.2
	b	0	14	5	5	0	0	16		
	c	3	5	4	8	0	0	1		
	d	0	5	9	5	0	0	2		
	e	1	22	6	14	0	0	15		
	SUM	9	55	31	54	0	0	47		
	%	4.6	28.1	15.8	27.6	0.0	0.0	24.0		
H5	a	1	9	8	3	0	0	29	220	17.5
	b	0	25	5	3	0	0	10		
	c	0	5	3	1	0	0	29		
	d	0	10	6	1	0	0	19		
	e	1	2	3	7	0	0	40		
	SUM	2	51	25	15	0	0	127		
	%	0.9	23.2	11.4	6.8	0.0	0.0	57.7		
H6	a	0	8	5	9	0	0	18	167	22.6
	b	0	11	1	1	0	0	6		
	c	0	12	1	5	0	0	16		
	d	3	4	0	6	0	0	17		
	e	0	6	3	5	0	0	30		
	SUM	3	41	10	26	0	0	87		
	%	1.8	24.6	6.0	15.6	0.0	0.0	52.1		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 1/9

1st set		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H7	a	0	4	4	7	6	19	19	209	17.4
	b	0	2	1	1	3	3	14		
	c	0	13	5	4	4	8	21		
	d	1	17	4	4	0	0	18		
	e	0	4	2	1	4	9	7		
SUM		1	40	16	17	17	39	79		
%		0.5	19.1	7.7	8.1	8.1	18.7	37.8		
H8	a	0	7	0	0			16	224	19.8
	b	0	7	5	1			23		
	c	0	17	6	3			17		
	d	3	31	5	5			19		
	e	0	13	11	3			32		
SUM		3	75	27	12			107		
%		1.3	33.5	12.1	5.4			47.8		
H9	a	0	7	5	6			21	211	22.3
	b	2	16	3	5			26		
	c	1	7	2	3			20		
	d	4	10	10	6			17		
	e	0	17	0	1			22		
SUM		7	57	20	21			106		
%		3.3	27.0	9.5	10.0			50.2		
H10	a	0	11	0	0			10	144	13.0
	b	0	9	5	2			31		
	c	0	4	2	2			21		
	d	0	6	0	0			23		
	e	0	5	4	2			7		
SUM		0	35	11	6			92		
%		0.0	24.3	7.6	4.2	0.0	0.0	63.9		
H11	a	0	14	2	2	0	0	12	222	19.5
	b	1	10	5	10	0	0	16		
	c	1	1	16	4			42		
	d	0	16	7	0	0	0	34		
	e	0	6	5	0	0	0	18		
SUM		2	47	35	16	0	0	122		
%		0.9	21.2	15.8	7.2	0.0	0.0	55.0		
H12	a	0	9	6	2	0	0	44	240	14.6
	b	4	10	7	4	0	0	33		
	c	0	8	6	0	0	0	30		
	d	0	5	3	0	0	0	41		
	e	1	8	6	2	0	0	11		
SUM		5	40	28	8	0	0	159		
%		2.1	16.7	11.7	3.3	0.0	0.0	66.3		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 2/9

1st set		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H13	a	0	4	2	3	0	0	28	222	16.7
	b	0	14	0	0	0	0	23		
	c	9	2	2	5	0	0	40		
	d	2	11	6	2	0	0	33		
	e	0	7	8	0	0	0	21		
SUM		11	38	18	10	0	0	145		
%		5.0	17.1	8.1	4.5	0.0	0.0	65.3		
H14	a	0	10	6	4	0	0	50	243	10.6
	b	0	13	3	0	0	0	43		
	c	0	19	5	0	0	0	14		
	d	0	4	0	0	0	0	40		
	e	0	7	5	0	0	0	20		
SUM		0	53	19	4			167		
%		0.0	21.8	7.8	1.6			68.7		
H15	a	0	4	5	0	0	0	6	166	26.1
	b	0	11	4	2	0	0	27		
	c	0	5	5	5	0	0	12		
	d	2	6	4	3	3	4	24		
	e	9	4	5	7	0	0	9		
SUM		11	30	23	17	3	4	78		
%		6.6	18.1	13.9	10.2	1.8	2.4	47.0		
H16	a									12.5
	b	0	12	11	4	0	0	44	210	
	c	0	10	0	0	0	0	36		
	d	0	12	1	0	0	0	32		
	e	0	22	5	1	0	0	20		
SUM		0	56	17	5	0	0	132		
%		0.0	26.7	8.1	2.4	0.0	0.0	62.9		
H17	a	0	0	0	0			30	188	19.8
	b	0	16	8	0			16		
	c	0	9	5	1			7		
	d	0	8	7	3			24		
	e	3	20	10	4			17		
SUM		3	53	30	8			94		
%		1.6	28.2	16.0	4.3			50		
H18	a	0	15	6	3			15	223	22.4
	b	0	17	0	0			31		
	c	7	8	1	1			26		
	d	1	17	5	8			21		
	e	4	6	4	7			20		
SUM		12	63	16	19			113		
%		5.4	28.3	7.2	8.5			50.7		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 3/9

1st set		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H19	a	0	5	0	0	3	6	26	219	3.7
	b	0	5	0	2	3	3	32		
	c	0	0	0	0	0	0	43		
	d	0	0	0	0	0	26	33		
	e	0	0	0	0	7	6	19		
	SUM	0	10	0	2	13	41	153		
	%	0	4.6	0.0	0.9	5.9	18.7	69.9		
H20	a	0	9	2	1	2	6	22	223	11.0
	b	0	5	2	1	0	0	56		
	c	0	3	1	0	4	2	35		
	d	1	5	4	3	3	0	18		
	e	2	3	3	5	0	4	21		
	SUM	3	25	12	10	9	12	152		
	%	1.3	11.2	5.4	4.5	4.0	5.4	68.2		

2nd set		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
H2	a	0	6	0	0	20	178	
	b	0	7	0	0	19		17.1
	c	0	17	0	0	25		
	d	0	14	1	2	15		
	e	0	0	19	0	4		
	f	0	12	10	0	7		
	SUM	0	56	30	2	90		
	%	0	31.5	16.9	1.1	50.6		
H4	a	10	16	9	12	27	230	36.5
	b	0	12	3	2	13		
	c	12	10	5	6	10		
	d	2	11	2	2	23		
	e	3	14	6	9	11		
	SUM	27	63	25	31	84		
	%	12.6	29.3	11.6	14.4	39.1		
H5	a	0	5	1	0	29	165	12.3
	b	2	11	2	1	16		
	c	0	14	1	0	20		
	d	0	12	0	0	17		
	e	0	20	0	0	14		
	SUM	2	62	4	1	96		
	%	1.2	37.6	2.4	0.6	58.2		
H6	a	0	10	0	0	16	170	20.6
	b	0	27	5	4	0		
	c	0	13	2	3	14		
	d	0	25	0	0	18		
	e	0	21	3	1	8		
	SUM	0	96	10	8	56		
	%	0	56.5	5.9	4.7	32.9		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 4/9

2nd set		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
H7	a	1	17	3	3	22	204	19.7
	b	0	14	1	1	16		
	c	3	5	2	3	28		
	d	1	18	3	5	13		
	e	0	23	2	2	18		
	SUM	5	77	11	14	97		
	%	2.5	37.7	5.4	6.9	47.5		
H8	a	0	17	4	3	24	250	22.3
	b	4	26	4	2	8		
	c	0	22	7	3	19		
	d	2	14	7	3	30		
	e	1	24	3	3	20		
	SUM	7	103	25	14	101		
	%	2.8	41.2	10	5.6	40.4		
H9	b	0	10	2	1	16	136	15.9
	c	0	20	3	2	8		
	d	2	4	2	3	13		
	e	0	7	0	0	15		
	f	0	11	2	0	15		
	SUM	2	52	9	6	67		
	%	1.3	34.4	6.0	4.0	44.4		
H11	a	1	7	3	4	30	245	
	b	0	13	4	2	24		29.0
	c	10	12	4	7	16		
	d	7	5	5	11	4		
	e	4	10	7	5	8		
	f	0	6	2	2	32		
	SUM	22	53	25	31	114		
	%	9.0	21.6	10.2	12.7	46.5		
H12	a	0	12	7	0	25	209	18.8
	b	0	15	1	1	17		
	c	0	13	11	2	12		
	d	0	20	0	1	21		
	e	0	23	6	4	18		
	SUM	0	83	25	8	93		
	%	0.0	39.7	12.0	3.8	44.5		
H13	a	0	17	5	2	33	220	18.9
	b	6	12	6	7	20		
	c	0	12	1	4	36		
	d	0	7	0	0	24		
	e	0	10	3	5	10		
	SUM	6	58	15	18	123		
	%	2.7	26.4	6.8	8.2	55.9		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 5/9

2nd set		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
H14	a	3	15	5	3	14	226	15.3
	b	0	8	0	5	39		
	c	0	7	1	0	23		
	d	0	31	2	2	21		
	e	0	17	1	0	29		
	SUM	3	78	9	10	126		
	%	1.3	34.5	4.0	4.4	55.8		
H15	a	0	12	3	1	20	217	20.7
	b	0	31	5	2	6		
	c	0	12	0	6	28		
	d	5	18	3	2	35		
	e	0	10	5	4	9		
	SUM	5	83	16	15	98		
	%	2.3	38.2	7.4	6.9	45.2		
H16	a	0	13	2	0	24	220	17.3
	b	3	18	6	5	25		
	c	3	16	4	2	28		
	d	0	13	1	0	24		
	e	0	11	2	2	18		
	SUM	6	71	15	9	119		
	%	2.7	32.3	6.8	4.1	54.1		
H17	a	0	31	4	1	13	221	25.5
	b	0	24	5	4	7		
	c	3	15	4	9	12		
	d	3	20	7	7	13		
	e	0	8	0	0	31		
	SUM	6	98	20	21	76		
	%	2.7	44.3	9.0	9.5	34.4		
H18	a	0	12	8	0	16	201	14.6
	b	0	7	0	0	39		
	c	3	12	1	5	28		
	d	0	12	5	1	25		
	e	0	9	2	1	15		
	SUM	3	52	16	7	123		
	%	1.5	25.9	8.0	3.5	61.2		
H19	a	0	7	2	0	20	141	6.7
	b	0	0	0	0	37		
	c	0	12	0	0	15		
	d	0	0	0	0	18		
	e	0	7	1	2	20		
	f	0	0	0	0	15		
	SUM	0	26	3	2	110		
	%	0	18.4	2.1	1.4	78.0		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 6/9

2nd set	ON	1/4 ON	1/2 ON	3/4	OFF	SUM	% ON
H20 a	0	9	0	0	9	147	15.5
b	0	12	1	2	11		
c	0	7	1	1	18		
d	2	10	7	6	12		
e	0	0	0	0	24		
f	0	0	0	0	15		
SUM	2	38	9	9	89		
%	1.4	25.9	6.1	6.1	60.5		

3rd set	ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H1 3 a	0	4	2	1	2	2	25	292	20.9
3 b	0	5	3	2	6	6	6		
3 c	1	5	2	4	7	4	10		
3 d	1	9	5	9	1	3	9		
3 e	1	7	4	5	2	5	16		
3 f	0	5	2	1	3	2	17		
3 g	2	7	2	3	2	3	21		
3 h	2	6	3	6	1	4	8		
3 I	0	4	1	1	1	2	9		
SUM	7	52	24	32	25	31	121		
%	2.4	17.8	8.2	11.0	8.6	10.6	41.4		
H2 3 a	0	2	0	0	0	5	19	366	14.2
3 b	3	4	2	3	5	2	21		
3 c	0	3	2	3	1	1	32		
3 d	0	3	1	0	2	1	27		
3 e	0	1	0	0	2	3	23		
3 f	0	4	2	1	1	1	21		
3 g	0	2	1	1	1	1	12		
3 h	1	12	0	1	2	2	13		
3 I	1	6	4	0	6	3	15		
3 j	8	6	3	10	1	0	18		
3 k	0	3	0	1	0	20	11		
SUM	13	46	15	20	21	39	212		
%	3.6	12.6	4.1	5.5	5.7	10.7	57.9		
H3 3 a	0	2	1	0	1	8	18	291	11.6
3 b	0	0	0	0	0	5	32		
3 c	0	2	1	1	2	1	42		
3 d									
3 e	0	2	0	1	2	7	33		
3 f	0	7	2	1	0	5	11		
3 g	4	6	8	10	4	1	8		
3 h	0	3	2	2	3	6	16		
3 I	0	7	1	0	1	1	21		
SUM	4	29	15	15	13	34	181		
%	1.4	10.0	5.2	5.2	4.5	11.7	62.2		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 7/9

3rd set		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H4	3 a	9	8	3	9	0	0	6	338	32.5
	3 b	1	12	2	2	2	2	8		
	3 c	1	9	1	3	6	4	12		
	3 d	4	9	2	6	2	3	10		
	3 e	4	6	4	5	2	4	16		
	3 f	12	4	5	6	3	5	6		
	3 g	3	5	5	4	2	3	20		
	3 h	1	2	3	3	4	5	18		
	3 I	6	8	6	5	6	3	8		
	SUM	41	63	31	43	27	29	104		
	%	12.1	18.6	9.2	12.7	8.0	8.6	30.8		
H6	3a	0	1	0	0	0	2	51	213	12.9
	3 b	0	3	3	2	6	1	15		
	3 c	0	4	1	2	0	20	10		
	3 d	0	3	2	4	0	1	8		
	3 e	0	6	0	1	0	17	7		
	3 f	3	5	2	6	1	5	21		
	SUM	3	22	8	15	7	46	112		
	%	1.4	10.3	3.8	7.0	3.3	21.6	52.6		
H7	3 a	1	9	3	7	1	2	17	287	10.8
	3 b	2	6	2	5	1	2	38		
	3 c	0	7	3	4	2	3	11		
	3 d	0	1	0	0	0	2	41		
	3 e	0	1	1	1	0	2	25		
	3 f	0	5	2	0	4	0	45		
	3 g	0	0	1	0	3	1	26		
	SUM	3	29	12	17	11	12	203		
	%	1.0	10.1	4.2	5.9	3.8	4.2	70.7		
H9	3 a	4	2	3	7	1	1	39	314	15.6
	3 b	4	4	5	8	2	6	15		
	3 c	0	0	0	0	0	0	71		
	3 d	0	7	0	4	3	2	31		
	3 e	0	1	3	7	3	3	34		
	3 f	3	5	6	4	0	4	22		
	SUM	11	19	17	30	9	16	212		
	%	3.5	6.1	5.4	9.6	2.9	5.1	67.5		
H10	3 a	0	1	3	1	3	2	12	195	7.0
	3 b	0	3	0	0	2	6	23		
	3 c	0	4	2	1	1	4	18		
	3 d	0	3	1	1	2	5	32		
	3 e	0	2	2	1	4	5	25		
	3 f	0	0	0	0	2	4	20		
	SUM	0	13	8	4	14	26	130		
	%	0.0	6.7	4.1	2.1	7.2	13.3	66.7		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 8/9

3rd & 4th	set	ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
H13	3 a	0	2	2	2	4	8	36	353	8.7
	3 b	1	9	6	3	2	2	49		
	3 c	0	0	0	0	0	0	61		
	3 d	1	4	1	4	1	4	30		
	3 e	0	4	9	6	3	1	49		
	3 f	0	4	0	0	1	5	39		
	SUM	2	23	18	15	11	20	264		
	%	0.6	6.5	5.1	4.2	3.1	5.7	74.8		
H15	3 a	0	8	6	0	1	2	32	307	15.3
	3 b	0	2	7	3	2	2	12		
	3 c	0	6	4	1	1	9	23		
	3 d	0	5	5	2	0	8	25		
	3 e	2	4	6	0	3	1	29		
	3 f	1	1	8	8	4	5	21		
	3 g	0	6	6	1	4	4	27		
	SUM	3	32	42	15	15	31	169		
	%	1.0	10.4	13.7	4.9	4.9	10.1	55.0		
H16	3 a	0	0	0	0	0	2	36	255	6.9
	3 b	0	3	2	1	2	3	42		
	3 c	0	4	2	0	4	7	23		
	3 d	0	8	8	2	3	4	42		
	3 e	0	4	2	1	3	4	43		
	SUM	0	19	14	4	12	20	186		
	%	0.0	7.5	5.5	1.6	4.7	7.8	72.9		
H19	3 a	0	6	2	0	1	1	17	151	2.7
	3 b	0	1	0	0	1	1	21		
	3 c	0	1	1	0	1	2	31		
	3 d	0	0	0	0	0	0	30		
	3 e	0	0	0	0	0	0	34		
	SUM	0	8	3	0	3	4	133		
	%	0	5.3	2.0	0.0	2.0	2.6	88.1		
H20	3 a	0	0	0	0	0	0	35	327	5.7
	3 b	0	3	3	1	2	7	24		
	3 c	0	4	4	2	1	3	37		
	3 d	0	7	0	0	1	1	34		
	3 e	0	6	0	0	1	2	32		
	3 f	0	4	3	1	6	2	20		
	3 g	0	5	0	0	2	5	32		
	3 h	0	1	0	0	2	1	33		
	SUM	0	30	10	4	15	21	247		
	%	0.0	9.2	3.1	1.2	4.6	6.4	75.5		
H1	4 a	0	9	6	5	6	3	13	334	28.9
	4 b	1	5	1	14	2	5	15		
	4 c	1	2	3	2	1	1	33		
	4 d	4	3	3	14	2	4	16		
	4 e	3	2	7	14	3	1	7		
	4 f	1	11	2	4	0	1	12		
	4 g	0	4	1	4	1	2	20		
	4 h	2	4	9	10	2	3	15		
	4 I	0	4	1	4	2	0	4		
	SUM	12	44	33	71	19	20	135		
	%	3.6	13.2	9.9	21.3	5.7	6.0	40.4		

Table 4.5 BLG mRNA FISH counts for hemizygote BLG/7 population page 9/9

1st round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 1	a	8	6	9	6	14		32.49
	b	5	33	15	8	18		
	c	6	26	16	10	10		
	d	0	10	4	3	24		
	e	0	15	8	3	20		
	SUM	19	90	52	30	86	277	
	%	6.9	32.5	18.8	10.8	31.0		
HH 2	a	4	20	9	10	8		32.38
	b	0	14	0	0	14		
	c	0	9	4	2	8		
	d	0	13	3	6	12		
	e	5	18	10	7	7		
	SUM	9	74	26	25	49	183	
	%	4.9	40.4	14.2	13.7	26.8		
HH 3	a	0	23	4	7	30		18.32
	b	0	14	5	2	49		
	c	1	13	2	5	46		
	d	0	12	7	5	33		
	e	6	15	4	9	26		
	SUM	7	77	22	28	184	318	
	%	2.2	24.2	6.9	8.8	57.9		
HH 4	a	0	3	4	3	12		22.74
	b	0	17	5	12	36		
	c	3	25	5	8	18		
	d	3	7	2	1	41		
	e	3	7	4	8	27		
	SUM	9	59	20	32	134	254	
	%	3.5	23.2	7.9	12.6	52.8		
HH 5	a	0	21	3	5	53		22.30
	b	0	24	5	4	50		
	c	1	21	3	3	30		
	d	16	18	6	21	31		
	e	2	20	7	4	32		
	SUM	19	104	24	37	196	380	
	%	5.0	27.4	6.3	9.7	51.6		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 1/12

1st round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 6	a	0	15	5	0	28		25.94
	b	1	10	8	4	17		
	c	7	9	5	8	31		
	d	4	13	7	7	19		
	e	1	22	10	9	25		
	SUM	13	69	35	28	120	265	
	%	4.9	26.0	13.2	10.6	45.3		
HH 7	a	1	11	8	5	26		23.42
	b	3	13	6	7	14		
	c	1	22	7	1	16		
	d	0	6	2	0	23		
	e	0	10	4	6	14		
	SUM	5	62	27	19	93	206	
	%	2.4	30.1	13.1	9.2	45.1		
HH 8	a	0	2	2	2	7		11.97
	b	0	12	3	1	36		
	c	0	17	6	5	33		
	d	1	10	2	3	42		
	e	0	10	0	0	44		
	SUM	1	51	13	11	162	238	
	%	0.4	21.4	5.5	4.6	68.1		
HH 9	a	9	14	5	9	28		34.27
	b							
	c	0	19	6	7	18		
	d	12	19	10	10	28		
	e	4	7	2	3	3		
	SUM	25	59	23	29	77	213	
	%	11.7	27.7	10.8	13.6	36.2		
HH 10	a	13	14	5	17	13		56.45
	b	21	3	3	12	8		
	c	19	9	4	7	3		
	d	26	7	6	5	12		
	e	0	25	14	20	9		
	SUM	79	58	32	61	45	275	
	%	28.7	21.1	11.6	22.2	16.4		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 2/12

1st round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 11	a	0	20	11	13	15		38.74
	b	15	33	10	11	3		
	c	3	27	8	12	15		
	d	1	25	3	10	16		
	e	14	32	2	8	15		
	SUM	33	137	34	54	64	322	
	%	10.2	42.5	10.6	16.8	19.9		
HH 12	a	0	0	0	0	31		21.91
	b	0	10	0	0	20		
	c	1	29	7	4	23		
	d	13	25	5	2	19		
	e	0	18	4	6	18		
	SUM	14	82	16	12	111	235	
	%	6.0	34.9	6.8	5.1	47.2		
HH 13	a	1	20	7	3	38		31.58
	b	9	9	8	10	2		
	c	6	20	9	10	26		
	d	0	13	8	9	35		
	e	5	17	14	11	14		
	SUM	21	79	46	43	115	304	
	%	6.9	26.0	15.1	14.1	37.8		
HH 14	a	0	5	0	0	30		26.83
	b	0	13	4	9	9		
	c	0	14	4	3	21		
	d	6	12	9	9	15		
	e	6	13	7	6	23		
	SUM	12	57	24	27	98	218	
	%	5.5	26.1	11.0	12.4	45.0		
HH 15	a	0	12	3	0	4		26.46
	b	3	7	1	10	2		
	c	0	30	4	1	25		
	d	1	22	5	5	15		
	e	2	10	4	4	18		
	SUM	6	81	17	20	64	188	
	%	3.2	43.1	9.0	10.6	34.0		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 3/12

1st round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 16	a	9	21	12	12	18		25.88
	b	0	10	1	1	53		
	c	8	14	6	8	19		
	d	0	21	3	4	18		
	e	1	8	4	7	26		
	SUM	18	74	26	32	134	284	
	%	6.3	26.1	9.2	11.3	47.2		
HH 17	a	0	20	5	5	15		37.93
	b	0	10	7	3	15		
	c	0	15	8	3	28		
	d	14	34	5	22	1		
	e	9	21	15	22	11		
	SUM	23	100	40	55	70	288	
	%	8.0	34.7	13.9	19.1	24.3		
HH 18	a	0	15	1	2	17		23.88
	b	0	18	2	0	25		
	c	3	14	13	10	15		
	d	0	12	4	3	12		
	e	0	10	11	1	12		
	SUM	3	69	31	16	81	200	
	%	1.5	34.5	15.5	8.0	40.5		
HH 19	a	0	9	2	3	34		17.69
	b	2	17	2	11	36		
	c	9	13	5	3	28		
	d	6	10	3	9	14		
	e	0	21	3	4	26		
	SUM	17	70	15	30	138	270	
	%	1.1	25.6	11.5	5.9	30.0		
HH 20	a	0	10	2	2	2		27.02
	b	0	12	4	3	7		
	c	0	20	0	0	8		
	d	4	12	2	6	13		
	e	0	8	2	1	6		
	SUM	4	62	10	12	36	124	
	%	3.2	50.0	8.1	9.7	29.0		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 4/12

2nd round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 1	a	1	30	6	4	20		30.4
	b	4	14	13	14	13		
	c	0	20	4	3	7		
	d	0	11	5	1	8		
	e	3	8	7	1	12	209	
	SUM	8	83	35	23	60		
	%	3.8	39.7	16.7	11.0	28.7		
HH 2	e	3	10	7	7	4		35.4
	f	0	28	9	3	12		
	g	1	10	7	2	8		
	d	3	7	9	8	8	146	
	SUM	7	55	32	20	32		
	%	4.8	37.7	21.9	13.7	21.9		
HH 3	a	1	19	8	2	8		22.3
	b	1	20	2	2	15		
	c	0	15	3	2	34		
	d	1	16	6	3	12		
	f	3	24	5	4	24	230	
	SUM	6	94	24	13	93		
	%	2.6	40.9	10.4	5.7	40.4		
HH 4	a no OK pictures, redo							
HH 5	a	9	29	2	6	11		
	b	0	6	0	0	32		
	c	1	13	4	2	19		23.6
	d	0	20	4	6	13		
	e	0	12	5	2	14	210	
	SUM	10	80	15	16	89		
	%	4.8	38.1	7.1	7.6	42.4		
HH 6	a	3	15	5	13	11		34.9
	b	3	13	6	7	14		
	c	0	25	6	9	12	242	
	SUM	6	53	17	29	37		
	%	4.2	37.3	12.0	20.4	26.1		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 5/12

2nd round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 7	a	3	12	2	7	6		33.4
	b	3	11	1	2	9		
	c	0	9	3	11	3		
	d	1	8	1	5	13		
	e	3	13	3	3	3		
	f	0	8	1	7	6		
	g	0	19	4	4	16	200	
SUM		10	80	15	39	56		
%		5.0	40.0	7.5	19.5	28.0		
HH 8	a	3	15	5	2	26		29.6
	b	0	34	0	0	20		
	c	0	23	4	0	6		
	d	3	20	5	12	6		
	e	9	9	4	6	5	217	
SUM		15	101	18	20	63		
%		6.9	46.5	8.3	9.2	29.0		
HH 9	a	0	17	10	8	10		22.0
	b	0	22	1	2	26		
	c	0	26	6	4	22		
	d	0	18	4	0	3		
	e	0	19	0	0	13	211	
SUM		0	102	21	14	74		
%		0.0	48.3	10.0	6.6	35.1		
HH 10	a	3	20	9	4	0		49.0
	b	8	27	7	12	1		
	c	11	21	8	3	0		
	d	10	14	6	10	8	182	
SUM		32	82	30	29	9		
%		17.6	45.1	16.5	15.9	4.9		
HH 11	a	5	32	2	5	25		26.6
	b	4	32	5	8	10		
	c	3	23	11	6	20		
	d	1	29	1	1	18		
	e	2	14	2	5	24	288	
SUM		15	130	21	25	97		
%		5.2	45.1	7.3	8.7	33.7		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 6/12

2nd round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 12	a	0	22	7	5	26		29.4
	b	6	10	6	17	28		
	c	2	25	10	11	10		
	d	0	24	11	3	23	246	
	SUM	8	81	34	36	87		
	%	3.3	32.9	13.8	14.6	35.4		
HH 13	a	0	7	7	7	25		30.1
	b	0	20	10	3	17		
	c	4	19	14	7	28		
	d	13	25	3	9	10		
	e	2	12	11	6	28	287	
	SUM	19	83	45	32	108		
	%	6.6	28.9	15.7	11.1	37.6		
HH 14	a	2	11	5	7	13		29.2
	b	1	26	4	8	20		
	c	0	18	7	6	7		
	d	0	36	9	1	10		
	e	0	14	7	4	6	222	
	SUM	3	105	32	26	56		
	%	1.4	47.3	14.4	11.7	25.2		
HH 15	a	6	7	2	14	8		33.5
	b	0	21	1	0	14		
	c	4	26	11	12	15		
	d	5	8	9	9	20		
	e	0	8	15	0	17	232	
	SUM	15	70	38	35	74		
	%	6.5	30.2	16.4	15.1	31.9		
HH 16	a	0	19	6	1	16		22.8
	b	1	16	2	4	47		
	c	0	20	8	0	25		
	d	5	12	6	10	6		
	e	0	15	2	4	6	231	
	SUM	6	82	24	19	100		
	%	2.6	35.5	10.4	8.2	43.3		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 7/12

2nd round		ON	1/4 ON	1/2 ON	3/4 ON	OFF	SUM	% ON
HH 17	a	0	31	3	0	33		20.6
	b	0	20	6	2	14		
	c	0	19	0	0	7		
	d	4	11	3	1	15		
	e	2	12	4	0	5	192	
	SUM	6	93	16	3	74		
	%	3.1	48.4	8.3	1.6	38.5		
HH 18	a	6	0	18	0	0		28.4
	b	3	2	2	3	6		
	c	1	12	5	2	6		
	d	0	0	0	0	30		
	e	0	16	0	0	7	119	
	SUM	10	30	25	5	49		
	%	8.4	25.2	21.0	4.2	41.2		
HH 19	a	3	25	11	5	17		27.9
	b	4	20	5	11	20		
	c	2	11	8	3	12		
	d	0	17	9	4	16		
	e	2	29	7	4	30	275	
	SUM	11	102	40	27	95		
	%	4	37.1	14.5	9.8	34.5		
HH 20	a	0	7	2	6	7		26.6
	b	1	14	3	7	6		
	c	0	11	3	8	17		
	d	1	11	4	2	15		
	e	0	21	0	0	9	155	
	SUM	2	64	12	23	54		
	%	1.3	41.3	7.7	14.8	34.8		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 8/12

3rd round		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
HH1	3a	0	4	7	2	4	0	9	143	31.4
	3b	2	8	7	5	2	0	11		
	3c	4	9	6	6	3	0	10		
	3d	2	4	4	3	4	3	10		
	3e	1	1	2	3	2	1	4		
	SUM	9	26	26	19	15	4	44		
	%	6.3	18.2	18.2	13.3	10.5	2.8	30.8		
HH2	3a	3	8	7	3	4	0	24	359	24.8
	3b	4	5	7	5	2	2	19		
	3c	5	4	5	7	4	2	18		
	3d	4	3	3	7	2	4	35		
	3e	6	2	6	15	1	6	41		
	3f	0	5	6	3	4	2	21		
	3g	1	8	3	7	2	1	23		
	SUM	23	35	37	47	19	17	181		
	%	6.4	9.7	10.3	13.1	5.3	4.7	50.4		
HH3	3a	2	8	2	6	5	1	39	613	17.8
	3b	0	6	4	1	2	4	41		
	3c	5	8	5	8	3	10	30		
	3d	4	2	4	6	5	3	35		
	3 e	2	3	7	3	2	5	29		
	3 f	5	8	2	7	4	3	31		
	3 g	4	7	4	6	1	3	33		
	3 h	0	7	8	6	4	4	47		
	3 I	0	4	8	10	2	4	31		
	3 j	1	8	2	2	3	7	37		
	SUM	23	61	46	55	31	44	353		
	%	3.8	10.0	7.5	9.0	5.1	7.2	57.6		
HH4	3a	0	4	5	10	2	5	13	255	19.0
	3b	1	3	3	10	1	5	8		
	3c	0	2	0	7	1	0	16		
	3d	0	8	4	3	2	1	15		
	3e	1	0	2	3	0	0	17		
	3f	2	2	3	0	2	3	23		
	3g	0	2	1	3	1	3	18		
	3h	0	1	1	0	1	2	35		
	SUM	4	22	19	36	10	19	145		
	%	1.6	8.6	7.5	14.1	3.9	7.5	56.9		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 9/12

3rd round		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
HH5	3a	0	6	3	0	1	5	41	244	13.9
	3b	0	12	6	7	0	5	17		
	3c	1	7	2	5	3	4	29		
	3d	0	5	8	0	2	2	24		
	3e	0	6	2	3	2	3	33		
	SUM	1	36	21	15	8	19	144		
	%	0.4	14.8	8.6	6.1	3.3	7.8	59.0		
HH8	3a	1.0	4.0	7.0	7.0	5.0	10.0	38.0	239.0	22.9
	3b	0	3	1	0	2	4	44		
	3c	1	3	2	3	5	6	23		
	3d	12	4	3	10	2	4	3		
	3e	2	2	4	10	3	1	10		
	SUM	16	16	17	30	17	25	118		
	%	6.7	6.7	7.1	12.6	7.1	10.5	49.4		
HH9	3a	1	5	5	8	4	10	29	414	19.1
	3b	5	11	9	12	1	1	32		
	3c	1	7	4	4	3	4	20		
	3d	0	10	7	2	4	3	37		
	3e	0	6	7	8	3	11	35		
	3f	2	4	9	7	6	3	39		
	3g	0	5	3	0	4	4	19		
	SUM	9	48	44	41	25	36	211		
	%	2.2	11.6	10.6	9.9	6.0	8.7	51.0		
HH11	3a	1	12	7	2	1	6	25	311	33.5
	3b	10	1	5	11	3	1	16		
	3c	10	12	9	5	2	4	6		
	3d	5	3	8	14	2	2	8		
	3e	1	4	2	10	0	1	6		
	3f	4	6	2	8	3	5	28		
	3g	0	4	5	4	2	5	20		
	SUM	31	42	38	54	13	24	109		
	%	10.0	13.5	12.2	17.4	4.2	7.7	35.0		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 10/12

3rd round		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
HH11	3a	1	12	7	2	1	6	25	311	33.5
	3b	10	1	5	11	3	1	16		
	3c	10	12	9	5	2	4	6		
	3d	5	3	8	14	2	2	8		
	3e	1	4	2	10	0	1	6		
	3f	4	6	2	8	3	5	28		
	3g	0	4	5	4	2	5	20		
	SUM	31	42	38	54	13	24	109		
	%	**10.0**	**13.5**	**12.2**	**17.4**	**4.2**	**7.7**	**35.0**		
HH14	3a	0	13	5	7	8	4	29	800	18.8
	3b	7	4	9	6	5	6	26		
	3c	0	5	3	2	3	9	50		
	3d	0	11	4	5	2	6	28		
	3e	8	4	3	7	2	3	24		
	3f	5	8	10	7	0	0	36		
	3g	3	3	3	4	3	5	58		
	3h	7	5	8	7	3	1	31		
	3i	0	4	1	0	1	2	54		
	3j	0	3	2	0	2	2	44		
	3k	4	16	5	10	0	5	36		
	3l	0	4	1	0	0	0	42		
	3m	5	9	6	15	3	4	5		
	SUM	39	89	60	70	32	43	463		
	%	**4.9**	**11.1**	**7.5**	**8.8**	**4.0**	**5.4**	**57.9**		
HH17	3a	5	14	9	7	0	4	17	766	31.0
	3b	1	7	13	6	3	2	11		
	3c	10	9	11	18	3	0	11		
	3d	20	8	8	21	1	4	21		
	3e	3	7	9	8	0	0	20		
	3f	7	6	8	17	2	5	20		
	3g	4	4	1	13	7	2	15		
	3h	4	13	7	9	3	4	30		
	3i	1	10	5	7	4	8	25		
	3j	1	5	4	5	4	10	28		
	3k	5	7	10	17	1	11	12		
	3l	0	2	1	1	1	5	47		
	3m	0	3	3	4	6	6	35		
	SUM	61	95	89	133	35	61	292		
	%	**8.0**	**12.4**	**11.6**	**17.4**	**4.6**	**8.0**	**38.2**		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 11/12

3rd round		ON	1/4 ON	1/2 ON	3/4 ON	1/8 ON	1/16 ON	OFF	SUM	% ON
HH18	3a	0	5	3	1	3	3	10	216	22.6
	3b	0	0	0	0	0	8	12		
	3c	5	2	3	2	2	8	12		
	3d	1	9	3	4	1	3	11		
	3e	2	8	11	6	1	3	11		
	3f	3	3	3	2	2	9	13		
	3g	0	4	1	4	2	5	12		
	SUM	11	31	24	19	11	39	81		
	%	**5.1**	**14.4**	**11.1**	**8.8**	**5.1**	**18.1**	**37.5**		
HH20	3a	0	11	2	1	2	3	16	425	23.8
	3b	0	3	2	1	1	9	19		
	3c	0	10	6	16	4	6	26		
	3d	1	4	3	7	3	2	21		
	3e	2	2	1	4	2	3	22		
	3f	10	5	8	13	4	3	25		
	3g	1	4	6	2	3	5	18		
	3h	0	9	7	1	1	2	19		
	3i	2	2	3	9	2	5	14		
	3j	1	4	7	3	1	4	7		
	SUM	17	54	45	57	23	42	187		
	%	**4.0**	**12.7**	**10.6**	**13.4**	**5.4**	**9.9**	**44.0**		

Table 4.6 BLG mRNA FISH counts for homozygote BLG/7 population page 12/12

Chapter 5

Is BLG/7 mosaic expression clonal or instructional?

5.1 Introduction

In variably expressing transgenic lines, silencing is frequently seen at the cellular level. In some cases, transgene expression levels from individual cells do not vary (Graubert *et al* 1998). Instead, differences in the number of expressing cells explain variable protein production. Qualitative differences in transgene expression generate a mosaic pattern of expression revealed by *in situ* hybridisation (Dobie *et al* 1996). This phenomenon bears at least a superficial resemblance to Position Effect Variegation (PEV) seen in *Drosophila* (1.7). BLG/7 is one such line exhibiting this phenomenon (Dobie *et al* 1996), and these mice have formed the central focus of the work described in this thesis.

5.1.1 Is BLG/7 mosaicism clonally or instructionally derived?

The mosaic pattern of expressing and silent cells evident in the mammary tissue of BLG/7 mice probed for BLG mRNA expression (Dobie *et al* 1996) is suggestive of a clonal expansion of cells from a progenitor cell after "fixation" of the decision of the expression status of the BLG transgene. An alternative explanation is that the patterns is the result of an instructional state achieved through inter-cellular communication

Support for clonal origins of mosaic expression patterns

Drosophila based research has demonstrated that position effect variegation is manifested by variable spread of heterochromatin. Once established, the pattern can become stable and clonally inherited (Spofford 1976, Paro & Hogness 1991), giving rise to the variegated pattern of expression characteristic of the phenomenon. Similarities between modifiers of variegation and regulators of homeotic gene expression are also suggestive of a clonal maintenance of activational status. It has been suggested that the mosaic phenotype is determined during differentiation, at least for PEV in *Drosophila* eye development (Lu *et al* 1996).

Support for instructional silencing origins of mosaic expression patterns

The evidence for instructional silencing comes from plant experiments on viral defence. Observations of zones of silencing of viral derived sequences preceding the zone of infection, alongside absence of any viral RNA upon a second infection challenge suggests that a diffusible systemic signal is produced that can move between cells and initiate a response (Voinnet & Baulcombe 1997, Jorgensen *et al* 1998). Post-transcriptional silencing and reactivation of a transgenic chitinase gene in *Nicotiana sylvestris* is reported to involve non-clonal events (Kunz *et al* 1996). Although the promoter used was viral (CaMV 35 S), the variegation appeared to be influenced by developmental stage, suggesting that physiological processes can affect gene silencing (Kunz *et al* 1996). Support for developmental control of heterochromatin-mediated silencing is also found in *Drosophila*, although there is no mention of a non-clonal inheritance of expression patterns in this case (Lu *et al* 1998). A non cell autonomous process is suggested responsible for patchwork expression in hair follicles of pwk/pwk mice (Aubin-Holstein *et al* 1998). As there are a number of hormones involved in the regulation of mammary gland expansion and milk expression (1.8.1), it is possible that some of these exogenous compounds could be involved in the generation of BLG/7 variegation.

5.1.2 How to differentiate between instructional and clonal states

To differentiate between clonal and instructional origins of the BLG/7 variegated cellular expression pattern, cells that share the same clonal history need to be marked. X-inactivation is probably the best example of an activation state decision initiated early in development and clonally maintained throughout all subsequent cellular expansions. A comparison of the expression patterns of cells probed for an X-linked marker gene and BLG/7 mRNA should enable dissection of whether the two patterns overlap or not in mice doubly hemizygous for both transgenic loci. Pattern overlap may also indicate the time-point at which determination of BLG/7 transcriptional state occurs. For example, if a temporal link between X-inactivation and BLG/7 silencing can be established, it would enhance speculation that the two phenomena also share mechanistic links.

There are several caveats to this experimental approach. A central point is the observation that silencing of the X-chromosome occurs at different times in a tissue-specific manner (Tam *et al* 1994, Lu *et al* 1996), though this does not rule out uniform timing of the initial determination event. Any interpretation of results will nevertheless have to bear this observation in mind.

A second potential problem is the experimental evidence that the entire mammary gland can be generated from a single cell. It is thus a theoretical possibility that each gland is derived from a single mammary progenitor cell, though I do not think that this is likely as fully mature mammary glands contain multipotent stem cells throughout (Kordon & Smith 1998).

A third problem is the complexity of the mammary gland, where only a subset of the cells express and secrete milk proteins (1.8.1). BLG/7 mRNA will only be visible in these cells. An x-linked reporter under the control of a ubiquitous promoter will be expressed in all the different cells comprising the mammary gland, rendering reliable interpretation of the expression patterns of the reporter and BLG/7 a difficult task. It was therefore considered essential to use a reporter under the control of a milk protein specific promoter. Mice fitting this description do not exist, so the first objective was the making such reporter mice.

5.2 Experimental Design

The process of X-inactivation marks the clonal origins of cells, and was chosen for exploitation in this project. Mice with an X-linked reporter under milk-protein promoter control had to be generated before any other part of the investigation could take place (5.1.2). Homozygous reporter mice would be used as controls to check that the reporter does not itself variegate, before double transgenic could be made containing both the BLG/7 locus and the reporter gene in a hemizygous state. Mammary tissue from lactating animals would be gathered and processed (2.3), with double mRNA FISH performed (2.4) and the resulting expression patterns analysed for coincidence of the reporter and BLG/7 expression profiles. Matching patterns, or

patterns that exist within each others boundaries, would confirm that both are the result of clonal inheritance of an expression state.

5.3 Design and assembly of the clonal reporter construct: HBZ

The targeting construct, HBZ (= **H**PRT, **B**LG, **L**acZ), was made from 3 separate components. The targeting vector (M44.8) was a kind gift from Andrew Smith (CGR, Univ. Edinburgh). The promoter (from pBJ39), was from the Whitelaw lab (Roslin Institute) and the LacZ reporter sequence from a commercial vector (pβGal Basic, Clontech). Figure 5.1 shows the cloning steps and final reporter-targeting construct.

In order to ensure that the reporter construct is integrated as a single copy insert into a genomic location that is known to be permissive for transcriptional activity in all cells (i.e. the site of a housekeeping gene), embryonic stem cell technology was used (Bronson *et al* 1996).

Construct Components

- Reporter: LacZ (4.6 Kb)
- Promoter: BLG (4.3 Kb)
- Targeting Vector: genomic HPRT (8.9 Kb, intron 5⟶) in pBS (3 Kb)
- Target: HPRT (Hypoxanthine Phosphoribosyl Transferase)

5.3.1 Reporter Choice

LacZ was chosen as the reporter for marking clonally related cells because it can be visualised by staining for protein production and contained a coding sequence of a size that allowed probes to be designed for mRNA FISH. The reporter sequence was also commercially available with a Poly A site at the 3' end (Clontech, pβGal Basic).

5.3.2 Promoter Choice

The milk protein-specific promoter chosen was BLG. BLG was a natural choice since the entire functional sequence was available within the group, and has been used previously to drive transgenic protein expression in the mammary of both mice

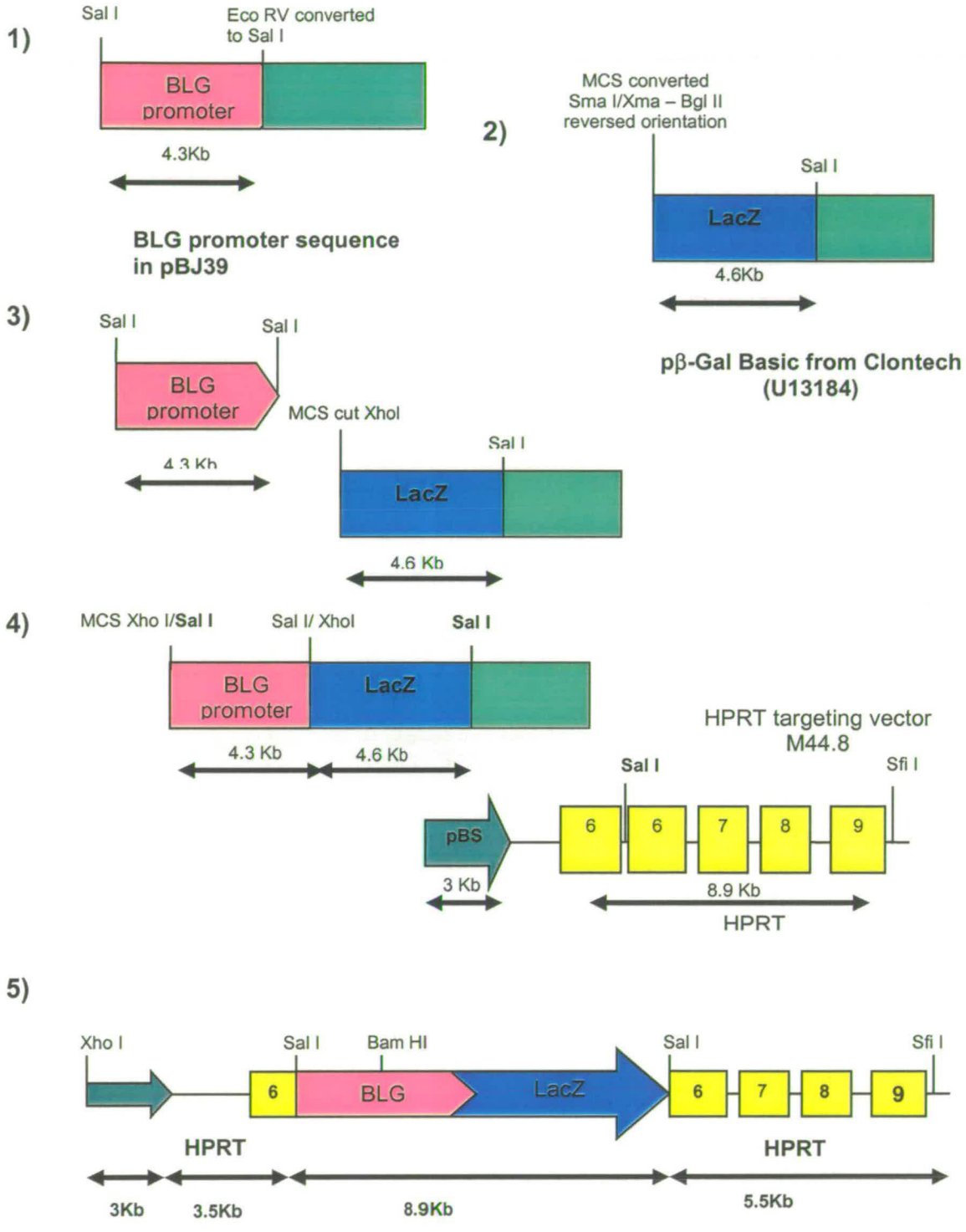


Figure 5.1 **Cloning of targeting vector HBZ (HPRT BLG LacZ)**

and sheep (Archibald *et al* 1990, Wright *et al* 1991, Clark *et al* 1993). The promoter was a 4.3 Kb *Sal* I/*Eco* RV fragment from pBJ39 (courtesy of CBA Whitelaw, Roslin Institute), which was initially inserted into the modified multiple cloning site of the reporter plasmid pβGal Basic (Clontech). See Figure 5.1.

5.3.3 Choice of targeting vector

The targeting vector (M44.8) was a kind gift from Andrew Smith (CGR, Univ. Edinburgh) and comprises of approx. 9Kb of genomic HPRT sequence starting from the latter half of intron 5 until the end of the gene, inserted into a 3Kb pBS backbone. Reporter construct insertion was possible using a unique *Sal*I site in exon 6.

The targeting vector was chosen for two reasons. One was a good record of efficient integration at the target site (A. Smith, personal communication), the other was the target gene: HPRT.

5.3.4 Choice of target gene

The Hypoxanthine Phosphoribosyl Transferase gene (HPRT) catalyses one of the first steps in the purine salvage pathway, is X-linked, expressed at low levels in most cell types and is in this respect very much a functional housekeeping gene.

The position on the X-chromosome and its status as a housekeeping gene are two of the four reasons why the gene was chosen as the target for the reporter gene integration. The third reason was the ease of selection for a functional knockout of the gene using 6-thioguanine, abrogating any need for a positive selectable marker in the targeting construct itself. This is a huge advantage as it reduces the size of the construct and minimises the length of sequence that bears no homology to the target gene, as well as increasing the chance that the reporter will function normally in *vivo* (1.1.4). The fourth and final reason for selecting this gene as a target was the availability of a targeting construct with a good record of integration into the HPRT site (5.3.3)

5.4 Testing HBZ Reporter Competency

Before undertaking the lengthy process of targeting the reporter construct to HPRT in ES cells, it was first tested for expression and induction competency in *vitro*, using HC11 cells.

Transient transfections, where the targeting construct HBZ was co-transfected with a vector containing neomycing for selection purposes were performed. The positive selected cells were pooled, expanded and hormonally induced using lactogenic hormones (2.2.2, 2.2.3). Once induction of milk protein genes was completed, the cells were stained for presence of LacZ, (2.2.3) with comparisons made between cells that had been induced and uninduced. LacZ staining confirmed the transcriptional competency and inducibility of the construct (Figure 5.2). Both induced and uninduced cells exhibited β-gal staining, as the HC11 cells maintain basal expression of milk genes even in the uninduced state (Burdon *et al* 1994). Induction increases the expression of milk-specific genes to between 10-20%, which is why this system is used for testing constructs before their introduction into animals. The construct was now considered ready for the next stage, introduction and targeting into ES cells.

5.5 Targeting into ES cells and generation of chimeric mice

The ES cells used were male (RI from 129/Ola, passage 14), ensuring that there was only one site to be targeted.

5.5.1 Electroporation and selection for positively targeted cells

The Qiagen purified targeting vector was linearised using *Xho*I, and 200 μg used in an electroporation reaction with $2x10^7$ ES cells (2.2.5). The cells were then grown without selection for 7 days (counting the electroporation day) before 6-thioguanine selection was initiated (1.67 μg/ml). The wait was necessary to ensure that all endogenous HPRT had disappeared in the cells where the gene has been correctly targeted and disabled. 10 days of selection were sufficient for single colonies to appear that could be picked and split into two halves, one for genotyping and the other for future use in generating transgenic mice.

A)

B)

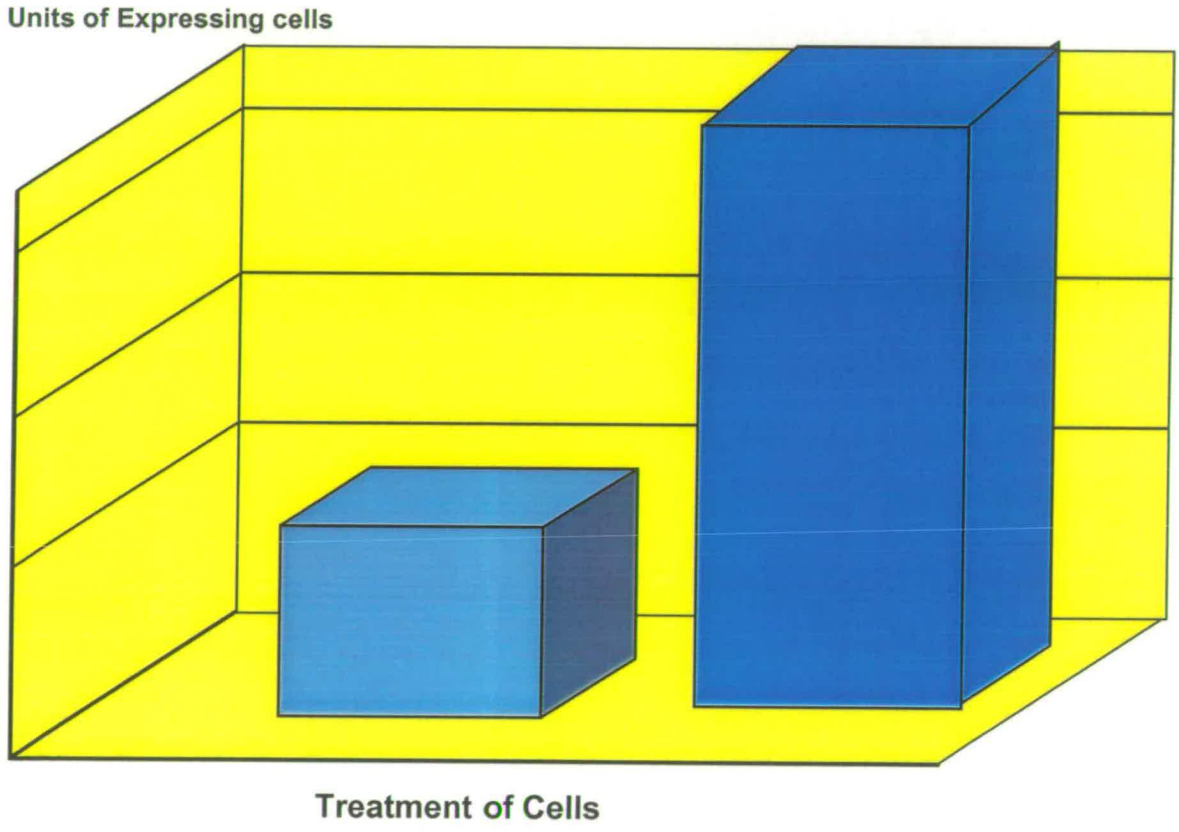


Figure 5.2 Induction of HC11 cells tranformed with HBZ construct

A) Induced transfected HC11 cells stained for LacZ expression

B) Representation of differences in number of staining units in induced and uninduced transfeceted HC 11 cells

The cells destined for genotyping were expanded, spun down and treated with proteinase K before being used in a standard PCR reaction to identify cells positive for the BLG promoter. 48 colonies were picked in all, 44 were confirmed as positive for the transgene promoter by PCR reaction (4 failed to give a PCR result).

There were of course many more colonies on the plates than the 48 that were picked (no more than 6 from each plate), so these plates were methanol fixed and Giemsa stained in order that I could estimate the targeting efficiency (2.2.5). 84 colonies from 15 plates give a targeting efficiency of $5.6x10^{-5}$, which is close to the efficiency seen when a positive selection marker is included (Thomas & Capecchi 1987).

5.5.2 Southern Blotting confirms targeting

32 clones were expanded, with some cells from each kept and used to make genomic DNA (2.2.6). The genomic DNA was digested with *Bam*HI or *Hind*III (two sets of digests, only clones that provided positive results in both sets of reactions would be considered correctly targeted) and RI genomic DNA included as a control. The digests were run on a 0.8% agarose gel and Southern Blotted (2.2.6). The blots were probed using a 600 bp *Sph*I/*Bgl*II fragment from intron 5 of HPRT (Thomas *et al* 1992). Both *Bam*HI and *Hind*III cut inside the reporter construct, so any positively targeted clones should exhibit a lower molecular weight band than the negative control, with non-targeted clones exhibiting two bands.

28 clones were confirmed as correctly targeted by Southern Blotting (Figure 5.3) (the remaining 4 clones proved resistant to analysis due to digestion difficulties). 6 of these clones were chosen as potential candidates for blastocyst injection, and cells were used to make chromosome spreads to check karyotype. The clones picked came from different plates to ensure that they were not derived from the same integration events, clones HBZ 1, 7, 17, 27, 38 & 47 were used, with HBZ 1, 17 & 47 chosen for blastocyst injection.

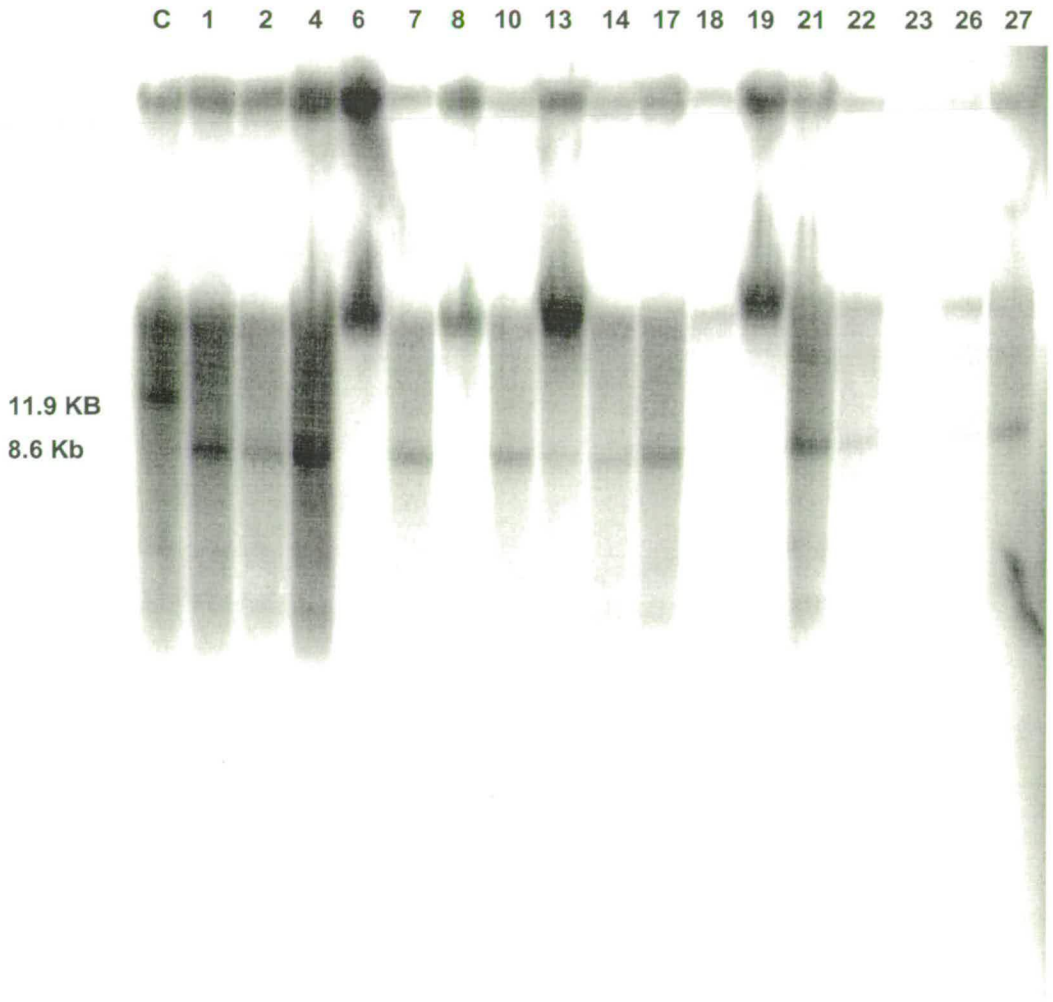


Figure 5.3 Sothern Blot of *Bam*HI cut HBZ clones

C = RI control DNA. Rest of lanes are HBZ clones 1 – 27.

Probe used is 600 bp *Sph*I/*Bgl*II fragment from intron 5 of HPRT (Thomas *et al* 1992)

5.5.3 Blastocyst injections and generation of chimeras

HBZ 1 was the first clone used for blastocyst injections, with 4 chimeric mice resulting (agouti coat colour evidence of chimeric origins, Figure 5.4). Unfortunately, only 2 of the mice bred, neither providing germline transmission. One mouse would not breed, and the other had to be put down due to ill health before reaching sexual maturity.

Circumstances beyond my control meant that attempts to make more chimeras had to be discontinued, and no more were generated. This was unfortunate, but the project design was such that this risk had been foreseen and additional experimental approaches incorporated into the project (chapters 3 & 4). In addition, initial analysis was performed on another line that had become available in the lab containing IRES-βGeo targeted to the X-linked PGK gene.

5.6 An alternative strategy: Exploiting an alternative X-linked reporter line containing an IRES-βGeo reporter construct targeted to and under PGK gene control control

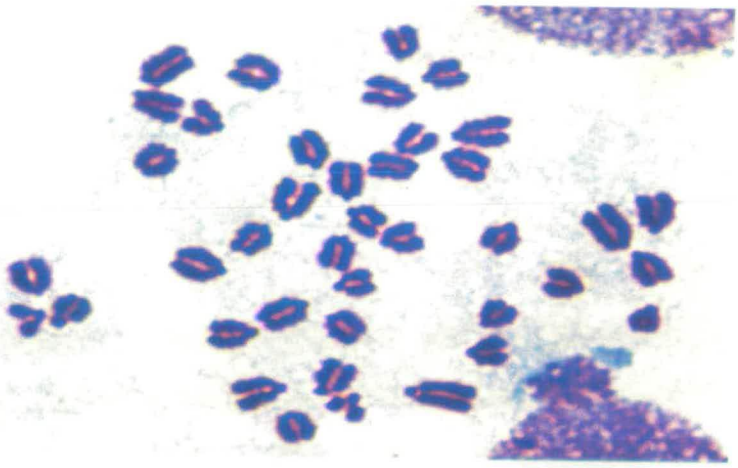
We gained access to X-linked reporter mice under a constitutively active promoter control, and decided to use them to generate double hemizygous transgenic mice containing both the BLG/7 locus and the PGK IRES-βGeo construct. These were designated as MOP mice (**M**argaret **O**psahl **P**GK). The generation of these mice would accomplish two goals:

- Check that LacZ expression does not interfere with BLG/7 expression
- Establish the double mRNA FISH protocols

5.6.1 MOP mice

PGK IRES-βGeo females (courtesy of D. Strathdee and A.J. Clark) were bred to BLG/7 outbred males in order to generate doubly hemizygous females (ear and tail analysis, 2.1.8 & 2.1). 35 MOP females were generated.

A)



B)

1

2

Figure 5.4 HBZ 1 Chromosome spread and resulting chimeric mice

A) Chromosome spread from clone HBZ 1, which was used to make the chimeric mice pictured below.

B) Chimeric mice resulting from blastocysts injections using clone HBZ 1 as a single mouse (B1) or together with non-transgenic litter mates (B2)

The failure of chimeric HBZ mice to transmit the targeted insertion through the germline has meant that this part of the investigation could not be taken further during the course of this project. The generation of the MOP mice ensured that some progress could be made, establishing whether LacZ expression can interfere with BLG/7 expression and if double RNA FISH can generate a clear image of clonality in the mammary gland. It was hoped that even with the obvious background signal from non-epithelial cells, that the question of the clonal origins of BLG/7 variegation may be addressed in these mice.

5.6.2 MOP milk results

Milk and tissue samples were collected at day 11 of lactation from all 35 females. Collection and processing of milk and tissue samples took place over a period of several weeks. An examination of the levels of transgenic protein expression in the milk was made and compared to the results gathered for BLG/7 hemizygous mice (Table 4.3). The same non-transgenic control sample was run for both the MOP and hemizygous BLG/7 samples. The arithmetic mean of three measurements was used in determination of BLG expression levels for each individual sample. The average expression from each individual is shown in Table 5.1. The mean BLG expression level and spread of expression correlate with that seen for hemizygous BLG/7 (Table 5.2)

MICE	MEAN	SD (standard deviation)	CV (coefficient of variation)
MOP	**7.88**	**3.59**	**45.53**
Hemizygous BLG/7 (outbred F1 CBA:C57BL/6)	**8.25**	**3.46**	**41.98**

Table 5.2 MOP and Outbred (hemizygous) milk protein results

	BLG mg/ml	Average		BLG mg/ml	Average		BLG mg/ml	Average
MOP 1.1	15.80	16.13	**MOP 6.3**	5.53	6.05	**MOP 11.1**	6.54	6.90
	16.68			6.79			7.30	
	15.90			5.83			6.86	
MOP 1.2	18.94	19.67	**MOP 6.4**	8.11	8.18	**MOP 11.5**	3.89	4.29
	21.22			8.51			4.90	
	18.87			7.93			4.06	
MOP 1.3	13.95	14.46	**MOP 6.5**	5.88	6.62	**MOP 11.7**	4.30	4.80
	14.56			6.68			4.82	
	14.88			7.29			5.27	
MOP 2.1	5.43	5.34	**MOP 7.2**	6.81	7.56	**MOP 11.8**	4.89	5.26
	5.26			7.53			5.54	
	5.33			8.33			5.36	
MOP 2.2	5.39	5.41	**MOP 7.4**	5.31	5.86	**MOP 12.4**	10.01	10.02
	5.48			5.73			9.99	
	5.36			6.54			10.06	
MOP 2.3	13.19	12.78	**MOP 9.1**	6.25	7.01	**MOP 12.5**	8.15	8.10
	12.91			6.98			8.27	
	12.23			7.79			7.89	
MOP 2.4	10.25	10.26	**MOP 10.1**	5.28	5.66	**MOP 13.2**	7.73	7.81
	10.09			5.01			8.04	
	10.45			6.68			7.65	
MOP 2.5	10.33	9.98	**MOP 10.2**	4.67	5.42	**MOP 13.7**	10.52	10.52
	10.27			5.78			10.44	
	9.34			5.81			10.61	
MOP 3.1	5.43	6.12	**MOP 10.3**	8.55	8.64	**MOP 13.8**	4.33	4.47
	5.45			8.89			4.57	
	7.47			8.48			4.51	
MOP 4.3	8.79	8.68	**MOP 10.8**	4.49	4.83	**MOP 15.2**	4.99	4.99
	8.56			5.03			4.56	
	8.71			4.97			5.42	
MOP 4.4	7.66	7.81	**MOP 10.9**	4.09	4.83	**MOP 16.1**	5.14	5.29
	7.86			4.98			5.24	
	7.92			5.43			5.50	
MOP 5.2	11.59	11.52	**MOP10.12**	2.90	4.51			
	11.61			5.06				
	11.37			5.58				
	11.56							

Table 5.1 BLG measurements for MOP mice

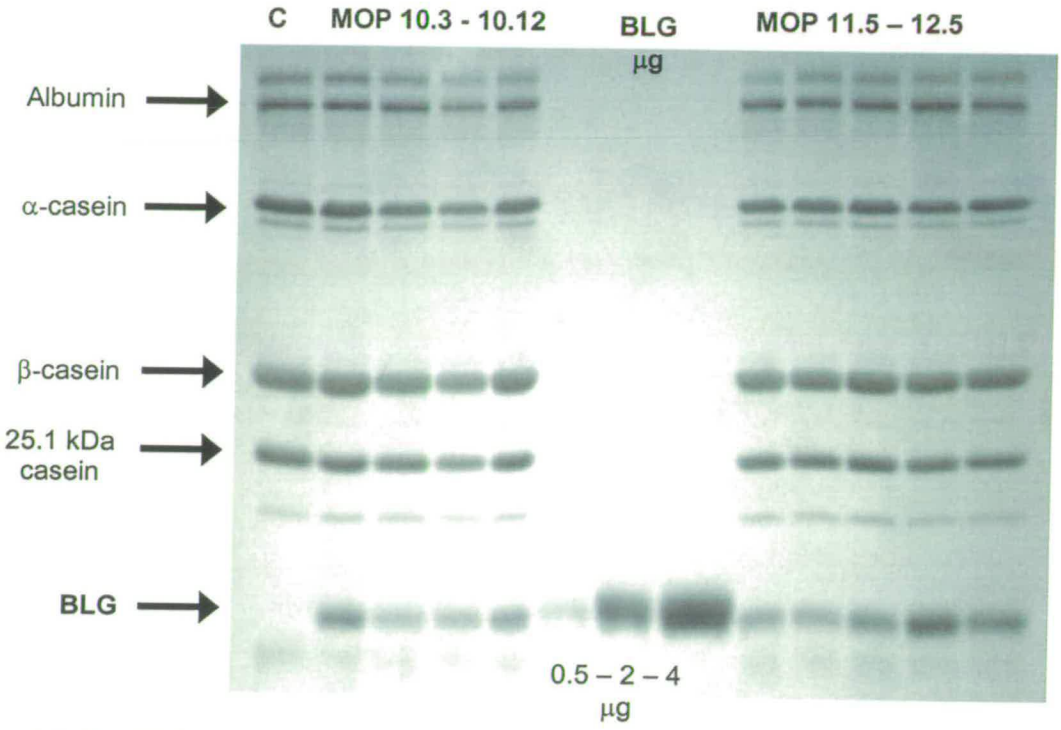


Figure 5.5 A) SDS PAGE of milk protein samples for MOP mice

C = non-transgenic control milk, BLG = β-lactoglobulin standard. Transgenic samples were loaded 40 µl of 1/250 dilution.

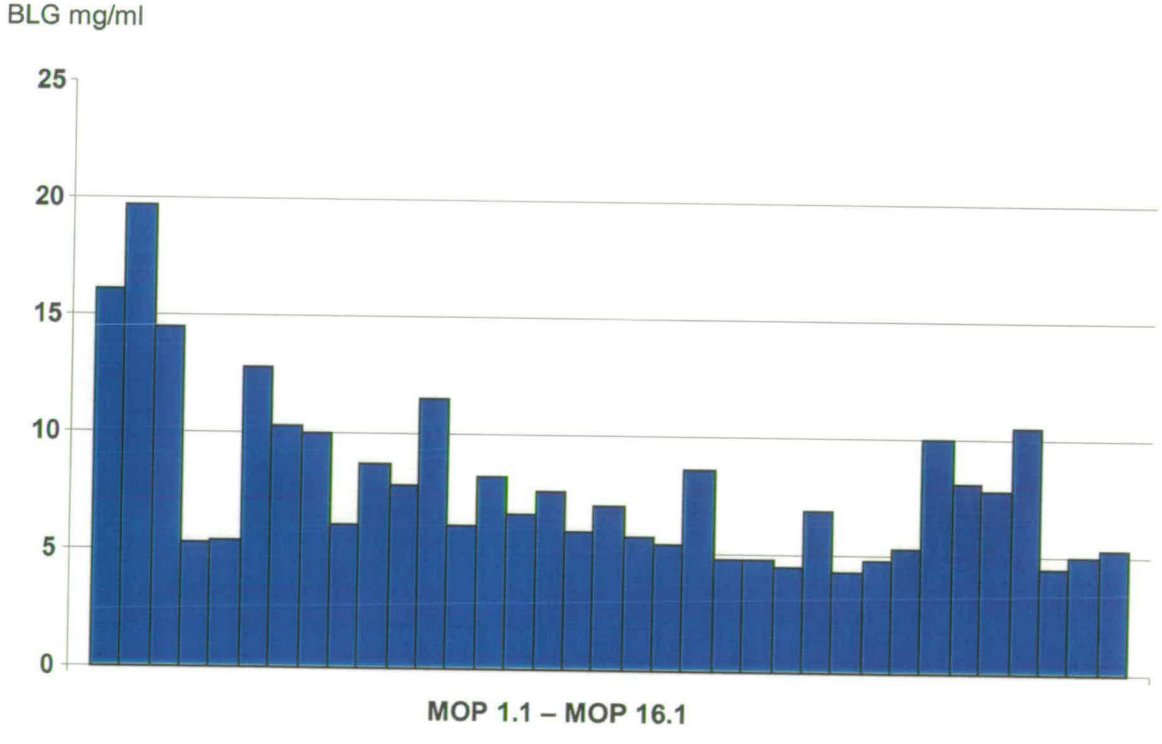
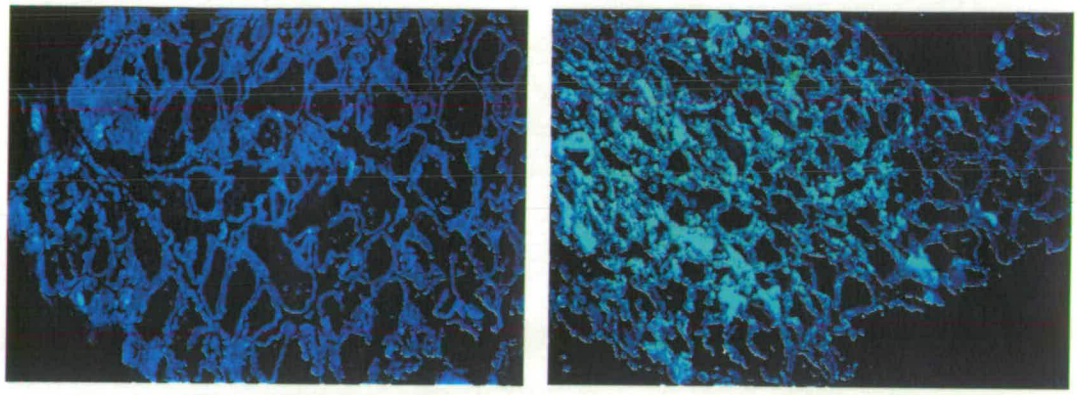


Figure 5.5 B) Bar chart representing individual BLG milk protein yields in mg/ml

5.6.3 MOP mRNA FISH

The MOP mammary tissues were fixed, wax embedded, sectioned and prehybridised (2.4.1 - 3). Initial FISH using oligo cocktails for BLG (Cy3) and LacZ (Fluorescein) mRNA (2.4.8) failed to pick up any LacZ signal even at 80 ng/ml (BLG is clearly visible at 40 ng/ml). These results and reports from other investigators using the PGK mice (C. Neil, personal communication) suggested that LacZ transcription was too low in the mammary gland for primary detection and in need of amplification steps.

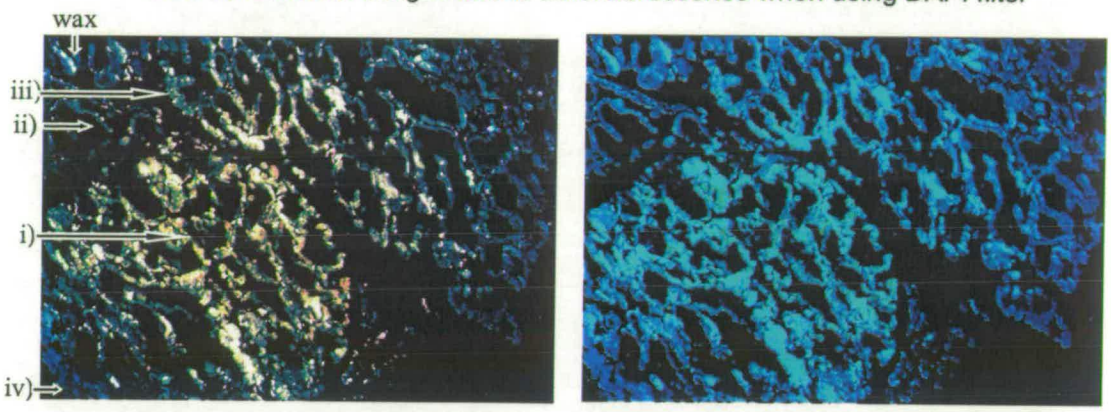
An antibody to fluorescein, conjugated to fluorescein, was used to amplify the signal (2.4.4, 2.4.5, 2.4.7). It took 4 amplification rounds to generate even a weak LacZ signal (Figure 5.5), which I considered to be of too poor quality to use to in an attempt to eludicate the clonal and temporal determination of BLG/7 silencing.



MOP mammary tissue probed with antibody to Fluorescein without primary probe to LacZ

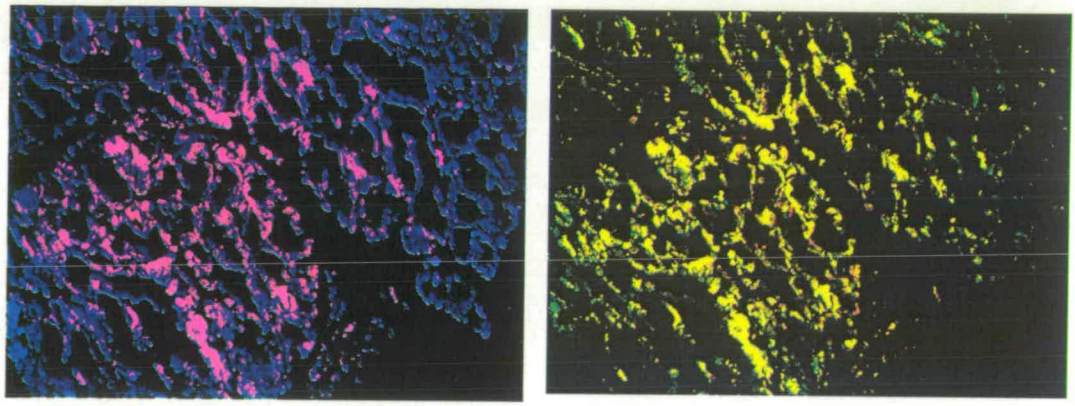
MOP mammary tissue probed for LacZ mRNA

Blue cells in both images due to autofluorescence when using DAPI filter



MOP mammary tissue probed for both LacZ & BLG expression with DAPI filter used.

MOP mammary showing LacZ signal only



MOP mammary showing BLG signal only

MOP mammary showing BLG & LacZ signal Without DAPI background

Figure 5.6 **MOP mammary tissue probed for BLG and LacZ expression. Scale: 1cm = 100 μm.**

Probes used were 5 30'mer (BLG) and 10 30'mer (LacZ) cocktails labelled with either Cy3 (BLG = red signal) or Fluorescein (LacZ = green signal, amplified using antibodies to fluorescein conjugated to fluorescein). Magnification is 10 X. Probe concentration used was 40 ng/ml. Some epithelial cells express LacZ (green), some express BLG (red). Note that alveoli can comprise of cells expressing both genes (**i**), BLG alone (**ii**) LacZ alone (**iii**) or neither (**iv**). Bright yellow fluorescence is caused by autofluorescence of wax residues, easily distinguished from mRNA signals.

Chapter 6

General Discussion

Future experimental strategies & implications of current results

Each of the experimental chapters (3-5) contains its own discussion section where implications of the results are discussed in the context of the experimental findings of the particular chapter. This chapter will therefore contain just a brief summary of the findings of the project. The remainder of the content will be divided between future experimental approaches to expand on the knowledge gained, and putting the results into context with the driving force behind much of the transgenic research taking place today, understanding human gene function, regulation, disease and devising treatment strategies for genetic disorders.

6.1 BLG/7 mice are an ideal model system for studying PEV

BLG/7 is the first line of transgenic mice where the variegated expression was shown to coincide with integration close to centromeric heterochromatin and high copy number, and suggested to be caused by a mechanism analogous to PEV in *Drosophila* (Dobie *et al* 1996). My project has extended the knowledge of how this particular gene array is regulated at this particular site:

Variegation is shown to be irrevocably linked to the transgene array at the site of insertion, with the extent of variegation open to modification by inbreeding factors and multiple minor background-specific loci that probably act in a cumulative manner (Chapter 3).

- The modifier loci operating in an inbred background do not convey any heritable epigenetic mark, as demonstrated through outcrossing

Introduction of a second variegating locus, in the form of generating mice homozygous for the BLG/7 locus does not affect the variegating properties of the locus inasmuch as variable expression is exhibited by homozygous mice. Curiously,

whilst the expected doubling of expression is seen for the minimum expression level, maximum expression remains identical for both hemi- and homozygous populations (Chapter 4).

- Protein and mRNA expression in the form of the number of expressing cells display the same ratio irrespective of whether the hemizyogus or homozygous population is examined.

- Nuclear FISH reveals that in actively expressing homozygote cells, only one of the two BLG/7 loci is active. This is suggestive of some form of *trans*-homology interaction taking place resulting in only one of the homologous loci remaining active, a novel phenomenon in transgenic mice.

Targeted ES cells generating an X-linked marker for the clonality of mammary epithelial cells have been made and will be used to make transgenic mice (Chapter 5)

- The successful targeting proved that when using HPRT as a target, no positive selection marker is necessary to efficiently generate correctly targeted clones.

- Preliminary investigations using an alternative line of mice confirm that the mammary tissue originates from more than a single progenitor cell. Poor expression levels of marker RNA adds to the desirability of making our own specific reporter line.

6.2 Future work based on the backcross experimental results of Chapter 3

Although this experimental chapter answers the question the backcross samples were intended to address, the work also unexpectedly uncovered a range of new questions. The data confirms conclusively that the transgene variegates irrespective of background or passage number through the germline, and is therefore an inherent

property of the array at the locus of integration. The devastating effects of inbreeding depression on all milk protein expression, and the existence of strain-specific modifiers in at least one of the two backgrounds added a whole new dimension to the investigation. These discoveries prompted further investigations, which provided some answers and many new questions. I have outlined how this project could be taken forward.

6.2.1 Continuation of outcrossing experiment

The loss of the CBA BLG/7 backcross population due to poor breeding performance was unfortunate as it meant that the outcrossing experiment was only performed using C57BL/6 BLG/7 backcross population. The assumption has therefore had to be made that any epigenetic modification in the CBA background would behave in exactly the same fashion as in the C57BL/6 background. This assumption needs to be addressed experimentally by breeding back into the CBA background until the 13^{th} generation, and then outcrossing to a C57BL/6 colony to generate first generation outcrossed F1 BLG/7 mice derived from the CBA line.

The current C57BL/6 derived outcrossed F1 BLG/7 mice need to be maintained by mating to the mixed F1 colony, and sampled at each generation to determine how fast the coefficient of variance returns to the level seen in the control outbred F1 BLG/7 mice.

6.2.2 Dissection of multiple modifier loci responsible for inbreeding depression and strain-specific modification

The backcross results suggest that there are multiple modifier loci responsible for inbreeding depression and strain-specific modification. The mechanism through which the modifiers act upon the transgene, and the identities of the loci are currently unknown. Based on the known integration site of the transgene, its persistent variegated phenotype, and the overall depression of inbred endogenous milk proteins, I could speculate that some of the modifiers may be involved in higher order chromatin structure and gene regulation. The dissection of at least some of the

different modifiers could be attempted by utilising three distinct but complementary approaches.

Genomics

The creation of genomic libraries for C57BL/6, CBA and F1 mixed background mice could be used to investigate differences at the DNA level. These libraries could be screened for restriction fragment length polymorphisms (FLP's) identifying regions that differ between the populations. Identification of corresponding gene loci and sequencing may reveal candidate modifier genes.

There are several distinct disadvantages to this approach for my purposes. Huge amount of data would be generated, and the probability of generating false positive results is high. An additional drawback is that this method cannot distinguish between differentially expressed or modified genes and proteins.

Differential Display and Associated Techniques

My favoured technique for rapid dissection of differently expressed genes in the three genotypes would be through differential display (DD). This powerful tool allows detection of both qualitative and quantitative differences in mRNA expression (Miele *et al* 2001). As I am unsure as to what developmental time point the different modifiers are active in, sampling of mammary and other tissues would have to take place at several different stages (virgin, pregnant, early lactation, mid-lactation etc).

The method involves RT PCR of the RNA samples, using anchored and random primers. Even polymorphisms in expressed transcripts may be detected using this technique. The PCR products are radioactively labelled and run on a polyacrylamide gel, the resulting bands can then be quantified as well as qualitatively analysed. Differentially expressed bands are then excised from the gel and either sequenced directly or cloned and then sequenced for identification. A possible additional project may be to try to improve the technique by substitution of the radioactive labels with fluorescent ones.

A complementary approach, although more laborious, would be the use of subtractive hybridisation to identify differentially expressed genes. As DD can generate quite a high false positive rate, and subtractive hybridisation and DD appear to identify positives from two different types of mRNA populations, using both systems when tissue material is plentiful makes sense.

Proteomics

The final approach would be to look at differences in expression at the protein level. A combination of 2D electrophoresis using immobilised pH gradients and subsequent use of mass spectrometry to identify excised protein spots would complement the approaches above and enable identification of differentially modified proteins. Differences in post-translational modification at the protein level would not be apparent in any of the previous techniques. Quantification of protein products would also be possible, enabling identification of differences in protein degradation between the genotypes.

The combination of the techniques described above should reveal a lot of information about differences between the strains of mice at the genomic, cellular and protein level. Such a holistic approach may in the future benefit human medicine, where there are many inherited diseases whose age of onset and/or symptom severity appears to depend on multiple factors. Although there is some evidence that lack of correlation between phenotypes and genotypes could be due to somatic mutation (Gottlieb *et al* 2001), I am confident that the genetic variability between individuals can have a large effect in determining the phenotypic manifestation of an inherited or acquired illness. Variable phenotypic expression is documented for facioscapulohumeral muscular dystrophy (FSHD) (van Deutekom *et al* 1996) and ICF (immunodeficiency, centromeric instability, facial anomalities) (Gimelli *et al* 1993). Muscular dystrophy and cardiomyopathy also display variable phenotypes, and a candidate modifier gene (Myoferlin) has been identified for these conditions (Davis *et al* 2000). If we could dissect out the various modifier loci and their effect on disease progress, then it may be possible to devise ways of alleviating, if not curing, currently incurable and untreatable conditions.

6.3 Future work based on the results of the effects of two homologous BLG/7 loci on transgene expression in Chapter 4

Although the experimental work in this chapter has revealed that BLG/7 homozygotes variegate and exhibit the same maximum expression potential as the hemizygote population due to allelic exclusion, there are many questions left unanswered as to the mechanisms responsible for this phenomenon. The existence of variegating BLG genes capable of doubling output in the homozygous state (McClenaghan *et al* 1995) argues against variegation *per se* being the causative agent of monoallelic expression in BLG/7 mice. I think that several different mechanistic pathways could be involved, whose components may or may not overlap. The future work would be to dissect out the components of these pathways.

6.3.1 Is the epigenetic component leading to monoallelic expression inherited?

Breeding the homozygous BLG/7 mice to heterozygocity would separate the homologous loci and enable investigation into whether the silencing of the one allele is heritable (Matzke *et al* 2001). If the silencing is heritable, only 50% of the offspring will express the transgene, and we would know that the silencing event takes place very early in development, before the segregation of the germline. If the silencing were not heritable, we would know that the process of inactivation takes place later in development. The results will enable estimation of when inactivation takes place, and the temporal determination may be able to provide a few clues as to what mechanistic pathways may be involved.

As the mammalian germline is segregated much earlier in development than the plant germline, it is considered unlikely that any epigenetic marks not involved in imprinting would remain intact, although a recent paper suggests that the murine germline may not be as effective as previously thought in erasure of epigenetic marks (Kearns *et al* 2000). Differences in male and female germline transmission of epigenetic features necessitates that any outcrossing of homozygous BLG/7 mice must be duplicated to include passage through both male and female transgenic parents.

6.3.2 Timing of allelic exclusion in BLG/7 homozygous mice

Variegation in BLG/7 hemizygous mice is maintained through several lactations (Dobie *et al* 1996). As I speculate that allelic exclusion in homozygous BLG/7 mice occurs separately and probably later than the process leading to the underlying variegation of the locus, this maintenance of expression in hemizygotes through several rounds of lactation can be used. I would investigate the timing of the allelic exclusion decision (the *trans*-interaction event) by examining BLG expression in serial lactations from homozygous BLG/7 individuals. Variations in BLG/7 protein content between lactations would indicate that allelic exclusion occurs during the proliferative phase of secretory epithelial cells in early pregnancy and that the decision is not maintained between pregnancies. Maintenance of expression levels between pregnancies would indicate that allelic exclusion takes place earlier in development.

6.3.3 Spatial distribution of variegated expression pattern

The images generated by mRNA FISH indicate that the variegated expression pattern is highly heterogeneous and remains so throughout the mammary tissue. A more in depth investigation into the expression pattern through serial sections of mammary tissue would give a very accurate picture of the spatial pattern of expression and enable the development of a computer-generated 3D image of the mammary tissue. This would be especially useful for double FISH using markers for clonal propagation of the mammary epithelial cells (6.4), and would serve the dual purpose of determining the clonal history of a single mammary gland as well as the relative expression patterns of BLG/7 and the x-inactivation marker.

6.3.4 Does BLG/7 differ from BLG/45

Although both BLG/7 and BLG/45 variegate (Dobie *et al* 1996), there are distinct differences between the two transgenic lines. BLG/45 exhibits a much higher level of expression, a doubling in homozygote expression levels, and lower levels of variegation when compared to BLG/7 mice. A comparison of expression patterns through mRNA FISH, and using an X-linked marker for clonal expansion, may be

able to tell us something about the mechanistic processes that may or may not link the two variegating constructs.

Crosses between the two lines to generate double transgenics would enable an investigation into whether the two loci remain independent of each other, an ideal test of the Markov model (1.7.7).

Crosses to a non-variegating transgenic BLG line (BLG 14, Dobie *et al* 1996) could also be initiated for both variegating lines in order to investigate the behaviour of these genes in the presence of a non-variegating locus.

6.3.5 Investigation of BLG/7 nuclear localisation

The nuclear localisation of BLG/7 genomic DNA could be highly informative. Double FISH using BLG and probes for various cellular components such as heterochromatic proteins, endogenous milk genes etc could cast some light on to the nature of the nuclear compartment where the locus (active or silent) resides, and thus some of the mechanistic pathways involved in the determination of expression of the locus.

It would also be advantageous for future experimentation to ascertain a more accurate localisation of the genomic integration site of the BLG/7 locus. The recently completed map of the entire mouse genomic sequence should provide the information required to achieve a more specific localisation of the transgene than currently available.

6.3.6 Expression pattern of endogenous milk proteins in nucleus

The absence of good intronic sequence data for the milk protein genes in mouse meant that the ideal controls could not be included in the experiments. A necessary future project would be to clone and sequence the introns of the mouse milk protein genes in order to design probes for nuclear transcription. Again, the recent publication of the mouse genomic sequence may prove useful.

Once the sequence data has been gathered, probes can be designed against intronic sequences for all the milk protein genes. A systematic mapping of nuclear localisation of these endogenous transcripts compared to BLG/7, BLG/45 and BLG/14 transcription would be informative when assessing the effect of nuclear localisation on expression potential. Comparing the nuclear localisation of BLG transcription and/or genomic BLG localisation with heterochromatin-specific proteins, transcription complexes etc would enable detailed speculation into what factors may affect transcription of a variegating locus as opposed to non-variegating loci.

6.3.7 Endogenous neighbouring gene expression patterns

As I speculate on the nature of the expression patterns of several of the genes close to the transgenic integration site (most notably, Ghr, Prlr and IL-7r), it would be natural to investigate their expression patterns, in both mammary tissue and elsewhere. This would naturally involve nuclear expression to determine whether expression is mono- or biallelic. Should any of them exhibit monoallelic expression, this would be a strong candidate for the gene responsible for the monoallelic exclusion seen in BLG/7 homozygous mice. In order to ensure that the observed expression patterns have not been influenced by the BLG/7 integration event (as reported for the Hoxb1 transposition by Kmita and colleagues in 1999), investigations would have to be made using both transgenic and non-transgenic tissue.

Should monoallelic expression be confirmed for any of these candidate genes, a strategy would have to be devised in order to differentiate between the two alleles so that parental effects can be investigated (possibly using crosses between *m. musculus* and *m. spretus*). It would also be interesting to investigate whether the active loci in BLG/7 hemi and homozygous mice are linked to the activational status of one or the other allele of the monoallelic expressing endogenous gene, and if there is a parental preference for the choice of "companion" allele.

6.4 Future work based on the results of the investigation into the clonal origins of BLG/7 variegated expression patterns in Chapter 5

The successful targeting of HPRT proved that positive selection markers are not necessary for efficient HPRT knockout targeting constructs. As only one set of confirmed chimereas were generated, the failure of germline transmission should not prejudice against the remaining clones. These clones could therefore be used in the future. The fact that these cloned cell lines have been kept in storage for over two years, would also justify retargeting into fresh ES cells using the original construct.

MOP mice could also be utilised further, if the mRNA FISH signal could be enhanced. Secondary antibodies have been used to increase signal output, but I feel that the number of amplification steps needed could cause any interpretation of the resulting expression patterns to be questioned. Additional primary oligo probes could be designed and added to the probe cocktail to generate a signal comparable to that seen for BLG/7.

Both of the reporter mice mentioned above have relevant uses, both in addressing the clonal origins of BLG/7 mosaic expression patterns and the clonality of the mammary gland itself. It would be of interest to examine the clonal origins of other variegating lines of mice, such as BLG/45 (Dobie *et al* 1996), and compare their expression patterns with each other and X-inactivation events. Should all have a clonal origin, the patterns of expression may also indicate whether they share the same temporal origin of determination of expression. This would be very informative in elucidating whether all variegation events share a mechanistic link at initiation of variegated expression.

6.5 The position of Position Effect Variegation in human disease

Position effects are increasingly implicated as a factor in a number of human diseases where inappropriate expression or silencing of gene expression cannot be attributed to mutation or deletion of the coding and immediate regulatory sequences.

6.5.1 The different position effects that can cause human disease phenotypes

There are several known examples of disease where position effects are attributed with a great deal of confidence, and these effects can be quite different. One example is the translocation event that positions the MYC oncogene in conjunction with immunoglobulin regulatory sequences. This induces the inappropriate MYC expression that is responsible for Burkitt's Lymphoma (Milot *et al* 1996). Translocations are not the only source of position effects causing disease phenotypes, removal of sequences flanking a gene can also cause position effects due to the effects that the new neighbouring sequences may have. Deletion of upstream sequences has been shown to be responsible for silencing β– and α–globin gene expression in some families with thalassemias (Milot *et al* 1996, Barbour *et al* 2000). PEV in the classical sense is exhibited by some forms of Aniridia, where rearrangements 85 Kb or more 3' to the PAX6 gene causes silencing and thus the haplo-insufficiency that results in the phenotype (Fantes *et al* 1995).

6.5.2 Position effects in transgenic animals, serendipitous model systems

Aberrant transgene behaviour can have several underlying causes (1.1 – 1.7), many of which are not mutually exclusive. PEV is one such phenomenon, originally identified in *Drosophila* (Muller 1930), and now linked to similar variable gene expression in other organisms (Allshire *et al* 1994, Dobie *et al* 1996, Iglesias *et al* 1997). Although the term position effect variegation suggests that site of integration is the sole source of variable gene expression, transgene position effect variegation may be a little more complex. Factors such as copy number, transgene design, integrity of insertion product and genetic background may exert an additional influence on transgene expression in a manner that may or may not be linked to the position effects. These will have to be dissected out before a site can be investigated solely for position effects.

Apart from the obvious links to human disease, understanding PEV is also important for the future of such diverse fields as application of gene therapy, xenotransplantation and generation of bioreactors for "pharming". Without the ability to guarantee stable, heritable and controlled expression of transgenes, neither

treatments of genetic diseases through gene therapy nor production of organs for human transplant purposes will be possible. Understanding the underlying causes of variegated gene expression may help find a way to avoid it. Economic considerations make reliable and heritable expression from a transgene desirable in animals or plants producing pharmacologically important compounds for harvesting, although this is not such a crucial requirement as in the other two applications mentioned above.

6.6 Using knowledge of modifiers of PEV to improve both transgenic livestock and treatment of human diseases

This project has demonstrated that multiple factors can influence BLG/7 expression, but that the extent of variegation is open to modification by recessive background-specific and inbreeding-dependent factors/loci (Chapter 3).

6.6.1 Using knowledge of modifier loci to improve the performance of transgenic animals

The backcrossing results suggest that poorly performing transgenes may benefit from introduction into a mixed genotypic background. Two papers have been published demonstrating that known modifiers of PEV can be experimentally manipulated and appear to exert an effect on a PEV locus without apparently adversely affecting endogenous genes (Festenstein *et al* 1999, McMorrow *et al* 2001). This is a viable strategy for enhancing the expression of a variegating gene in mice, but the difficulty in generating larger transgenic animals renders this strategy unworkable. PEV could be avoided by using targeted insertion of transgenes into permissive sites, but ES cells are not generally available for animals other than mice. This means that transgenic animals are still largely generated through pronuclear injection with random insertion of the transgenes, although recent advances indicate that gene targeting may be feasible using somatic cells (Clark *et al* 2000, Denning *et al* 2001).

A possible strategy to ensure that expression levels are enhanced is to use QTL analysis to identify strains and individuals containing traits associated with high expression in the tissue where the transgene is to be expressed. If the transgenic

protein is to be secreted into the milk, then animals with QTL traits indicative of high milk yields, possibly selected for high yields of a particular fraction of the milk proteins (caseins versus whey), could be chosen in the hope that the transgene will also benefit. Identification of proteins and other factors that generally improve expression from a variegating construct could also be used to identify individuals or strains with the most advantageous qualitative and quantitative combinations of modifying factors.

6.6.2 Exploiting a knowledge of modifier loci in the treatment of human diseases

The multitude of factors shown to exert an effect on gene expression may explain why many human diseases display varied phenotypes despite being caused by the same gene.

Variable expressivity of autosomal dominant diseases is one of the more suggestive examples that different combinations of modifier loci exert a variable effect on disease phenotype. Polydactyly is one example where there is both variable expressivity with differences in phenotypic manifestation and instances of incomplete penetrance. The basic genetic basis for this, especially the incomplete penetrance, is unknown. As limb development is a multifactorial process, some genetic combinations may be able to modify and even prevent the phenotypic manifestation of the mutated gene. Differences in modifier loci may explain a reported incidence of two siblings with the same ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome genotype, only one of which exhibited the full-blown genotype (Gimelli *et al* 1996).

Another example of a phenotypically heterogeneous condition is fasciocapulohumeral muscular dystrophy (FSHD). The actual gene responsible for FSHD has not been identified, but the disease is linked to translocation events occurring between defined regions on chromosomes 4 and 10 involving the telomeres. This suggests that PEV may be responsible for the disease as the translocation breakpoints vary and thus cannot be disrupting the same gene at the

breakage site (van Deutekon *et al* 1996). FSHD is an autosomal dominant disease, with variable clinical severity and age of onset that is evident even within the same family (Fisher & Upadhyaya 1997). The phenotype could be linked to variable spread of heterochromatin from the telomeres into the disease locus causing variable expression of the gene that will differ between individuals. In addition, it is not inconceivable that the amount of expression from the unaffected allele, and/or extent of heterochromatin spread and silencing of the diseased locus, could be influenced by differences in modifier protein combinations that will naturally vary between individuals, even if closely related. It could be that the relative strength of repressive and activating factors able to act at a particular locus will determine the severity of the disease phenotype. A final possibility lies in the identification of one of the candidate FSDH genes as a member of a multigene family with related sequences on other chromosomes. Other family members may, depending on individual allele combinations, be able to substitute for some of the FSHD gene functions and lessen the impact of the disease.

The ability of other gene family members to substitute for each other may explain why holoprosencephaly, caused by sonic hedgehog haploinsufficiency, exhibits variable phenotypic manifestations is not seen in all hemizygous individuals (Kleinjan & van Heyningen 1998).

6.6.3 A role for somatic mutation in human PEV?

Although accepting the presence and involvement of phenotypic modifier genes, Gottlieb and colleagues (2001) suggest that somatic mosaicism may underlie variable expressivity in some human diseases. Androgen insensitivity syndrome (AIS) was quoted as a prime example, with heterogeneity of cancer tissues an accessory example. In view of AIS being an X-linked disease, my personal feeling is that this gene would be more susceptible to somatic-mutation-induced disease phenotypes (as are all sex-linked genes), and should be considered an exception rather than a rule. AIS is a single locus disease. With many diseases, including non-familial cancers, reputed to be multi-locus multi-factorial with environmental conditions impacting on phenotype and disease progression, it is my belief that individual variations in both

qualitative and quantitative aspects of modifier and "accessory" loci undoubtedly contribute to phenotype.

6.6.4 Using the knowledge of modifier loci for future treatment strategies

It is my belief the identification of general modifier genes involved in regulation of PEV, and linked by animal models to defined human PEV medical conditions, could be used to develop alternative treatments. Introduction of proteins that can upregulate expression from the variegating locus, or chemically inducing the relevant cells to produce more of the modifier protein, could be used to alleviate symptoms in conditions where haploinsufficiency in the primary cause of the phenotype. Such intervention could provide a safe alternative to gene therapy at a time when targeted integration needs the harvesting of stem cell populations from the patient and reintroduction of the manipulated stem cells, a process that is lengthy, costly and unable to guarantee long-term success at this moment in time. I have no doubts that targeted gene therapy will one day become almost as routine as treatment of bacterial infections with penicillin and its derivatives, but until that day arrives, alternative treatment options must be sought.

A candidate for testing this approach would be Campomelic dysplacia patients with SOX9 translocations. Patients with translocations usually exhibit much milder clinical symptoms than those with mutations in the SOX9 gene coding region (Kleinjan & van Heyningen 1998). One of the phenotypic symptoms is sex reversal, which could possibly be re-reversed through upregulation of expression in the testis of patients. A more difficult challenge would be targeting of modifier protein expression to skeletal cells in order to alleviate the malformations also associated with this disease. Once the practicalities of the treatment are sorted out, timing of the treatment would be important. I still think that if treatment is initiated early, before completion of development and growth phases, it should be successful and applicable to a wide range of human conditions. The exceptions would be those diseases where development of the affected organs is completed in *utero*. The challenge for these diseases would be early identification of the disease and development of a delivery system that does not affect the mother.

6.7 *Trans*-interactions and PEV loci

The variegating transgene locus of BLG/7 exhibits an unusual property in addition to PEV. This property is only manifested in individuals homozygous for the transgene locus, with only one of the two loci active in any given cell. This is reminiscent of *transvection* and paramutation phenomena in *Drosophila* and plants, but differs significantly inasmuch as maximum expression levels remain on par with hemizygous individuals, as does the number of expressing cells. *Transvection* can lead to up-regulation of expression or, in common with paramutation events, silencing of expression from both paired loci. There is no direct evidence yet for BLG/7 homologous loci interaction, but the mono-allelic expression is highly suggestive of such an event taking place.

The single locus expression seen in BLG/7 has never before been reported in transgenic mammals. Only one similar phenomenon has to my knowledge ever been reported for transgenes, and that involved transgenic plants where the experimental design was to reproduce *transvection* interactions (Matzke *et al* 2001). The infrequent reports of *transvection* in *Drosophila*, suggests that homology interactions affecting transcription are rare events. Monoallelic expression is of course seen in mammals, as the result of imprinting and dosage compensation mechanisms for sex-chromosome linked genes. Recent reports also suggest that some autosomal genes lacking known imprints may be monoallelically expressed (4.5.6), although it is not known if any *trans*-homology interactions are involved in the regulatory process. As none of these genes are subject to PEV, BLG/7 homozygous mice may not be the best model for this documented intended monoallelic expression from endogenous genes.

6.7.1 Why we need to understand the interactions leading to monoallelic expression

Although reported *trans*-homology silencing is rare in animals, especially mammals, elucidation of the mechanisms responsible could not only prove useful in transgenic, but may help explain certain human diseases. Despite a lack of direct evidence, it is speculated that *trans*-homology interactions could play a role in the progression of

some autosomal dominant diseases. Anomalies in linkage maps compiled for several human diseases, including Huntingdon's Chorea, may be explained through trans-inactivation of the healthy homologue that is occasionally maintained through meiosis (Sabl & Laird 1992). There is even a suggestion that *trans*-homology interactions may be involved in the X-chromosome inactivation process (Marahrens 1999), and could be a requirement for monoallelic expression from imprinted genes (LaSalle & Lalande 1996). Should this be the case, then it suggests that homology sensing systems are operational in mammalian systems, and that the monoallelic expression pattern of homozygous BLG/7 mice is the result of activation of this system.

Progressive *trans*-inactivation, where the necessary pairing takes place in a limited period during the cell cycle, could be used as a model to explain progressive diseases that exhibit variable speed of symptom progression and age of onset. It could be that once inactivation takes place the decision is irreversible, such that each round of cell division increases the proportion of inactivated cells until there are not enough active cells remaining to halt the disease. Candidate diseases include many of the autosomal dominant muscular dystrophies, such as fascioscapulohumeral muscular dystrophy (FSHD), as well as the aforementioned Huntingdons Disease. The triplet repeat expansions characteristic of these diseases (Spiegel 1997) may act as foci for heterochromatin formation, and the pairing of the two alleles leading to inactivation of the healthy allele in a manner analogous to transvection of the *brown* gene in *Drosophila* (1.7).

A single paper was found with evidence to support a role for *trans*-homology interactions in human gene expression control and disease. The paper shows that manifestation of type 1 diabetes depends not only on the genes inherited, but also on the allelic identity of the non-inherited homologue from the paternal side (Bennett *et al* 1997). This example combines a *trans*-homology effect with a parent of origin effect, highlighting the complexity of inheritance and the problems that can be encountered when trying to match phenotype with genotype.

6.8 Using BLG/7 as a model system for complex diseases involving both PEV and suspected *trans*-interactions

The experimental evidence, especially that gained in chapter 3, suggests strongly that BLG/7 variegation is a classical position effect analogous to the *Drosophila* PEV phenomenon. This makes BLG/7 an ideal experimental system for investigating general mechanisms affecting PEV in mammals. The monoalleic expression resulting from the presence of two homologous transgene loci in BLG/7 is unique, and could be used to investigate possible common links between rare dominant diseases in mammals. The results could be used to devise a general strategy of combating position effects in human disease and improving the performance of transgenic animals.

6.8.1 BLG/7 as a model system for autosomal dominant diseases

The combination of PEV and possible homologous interactions at the BLG/7 locus in homozygous mice does make them a possible model system for studying common elements that may link certain dominant autosomal human diseases. These diseases could be classified as being caused by a position effect, but instead of translocations causing the disease, expansions of repeat sequences within the area occupied by the gene and its regulatory components are linked with the disease phenotype. The repeat expansions could initiate formation of heterochromatin at the site, which in turn could lead to silencing of the whole gene (as demonstrated by Dorer & Henikoff 1994). Parallels could also be drawn with the dominant *brown*^{D} mutation seen in *Drosophila*, which can *trans*-inactivate homologous genes (Dreesen *et al* 1991, Csink & Henikoff 1996, Dernburg *et al* 1996).

Dominant repeat expansion mutations

A classic example of repeat expansions causing human disease phenotype is the fragile X mental retardation syndrome. Here, mutations of the FMR1 gene commonly consist of a CGG repeat expansion that coincides with methylation and deacetylation of the gene, preventing transcription from the site (Coffee *et al* 1999). At least 12 human diseases are associated with trinucleotide repeat expansions. Some

of these expansions are found in the coding region and the resulting phenotype is attributed to gain/loss of function and/or functional knockout or haploinsufficiency.

6.8.2 Candidate autosomal dominant conditions that may be the result of *trans*-homology interactions.

My particular interest lies in the autosomal dominant conditions where triplet repeat expansions associated with the disease are found outside the coding region.

Myotonic Dystrophy

Myotonic dystrophy (DM) is one example, with the CTG expansion in a non-coding area with a size varying from 35-2000 copies. DM has 3 clinical forms and exhibits huge variations in both phenotypic manifestations and age of onset, with some obligate carriers displaying no clinical symptoms at all (Spiegel 1997). I suggest that the progressive nature of this muscle-wasting disease, coupled with a generally later age of onset in 2/3 clinical forms, indicates the possibility that *trans*-homology interactions between the mutated locus and its normal counterpart during cell division may take place. The repeats could form an initiation site for heterochromatin formation, whose action may be to tether the locus to constitutive heterochromatic sites in the genome and inactivate the gene and/or spread into the whole locus denying access to the transcriptional machinery. Pairing during cell division could result in the irreversible inactivation of the healthy allele in a manner analogous to transvection or paramutation. The outcome of each interaction may depend on the balance of modifying factors in the nuclear location at the time of the interaction taking place. As a person ages, there is increased chance that a given interaction will inactivate the healthy allele, and coupled with the irreversible nature of the inactivation could explain the gradual onset of the disease. Variable phenotypes could thus be a combination of chance and differences in individual modifier genes influencing PEV and/or *trans*-homology pairing. The maternally transmitted neonatal congenital DM could be a combination of maternal germline repeat expansion and/or *in utero* effects that enhance the pairing and inactivation events.

Huntingdon Disease

Huntingdon Disease is another repeat expansion associated disease, but the repeats are in the coding region of the gene, with gain of function the suggested cause of the phenotype (Spiegel 1997). The progressive nature and generally late age of onset of the disease is all the same suggestive of a possible *trans*-homology effect also affecting phenotype. If haplo-insufficiency and/or gain of function was the cause of huntingdons disease, I would expect the symptoms to appear earlier in most patients and the phenotype to remain stable in an individual. An epigenetic modification as one cause of HD is suggested by reports of HD in genotypically normal individuals, indicating the possible inheritance of an epigenetically inactivated normal allele through the germline that is capable of acting as a HD allele, in a manner analogous to paramutation phenomena in plants.

Both HD and DM share the trait of anticipation, associated with increased repeat expansion that may increase heterochromatinisation at the site and strengthen pairing interactions and tip the balance in favour of inactivation. Severity and size of the repeat expansion are linked in all the triplet expansion diseases, which correlates with my theory that the triplet repeats form foci for heterochromatin formation and be central to the directional inactivation events through pairing interaction.

The similarities of the triplet repeat expansion diseases, with variable phenotypes and age of onset with gradual worsening of the condition suggests that they may share some common triggers. BLG/7 could be used to dissect out candidate genes for these triggers and generate a working model for how they act.

6.9 Concluding Remarks

The creation of a transgenic line of mice subject to position effect variegation was serendipitous. The subsequent identification and characterisation of the trait in BLG/7 was due to observation and implementation of scientific methodology. The realisation of the importance of the observations of the initial investigations prompted further investigations of this transgenic line.

My investigations into BLG/7 variegation have revealed this line exhibits a variety of traits that are linked to the variegating transgenic locus. This line can be used as a unique tool to investigate both multiple modifier loci capable of influencing the extent of variegation, and the rare phenomenon of trans-homology interaction resulting in mono-allelic expression.

The first example can be used to devise improvements of product yield from transgenic stock, whilst both can be used to model mechanisms of human diseases and suggest possible therapeutic interventions.

BLG/7 is thus a valuable resource, which with careful experimental design and execution could push the frontiers of both transgenic research and medicine.

REFERENCES

1. Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P. B., Reuter, G., and Jenuwein, T. Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. The European Molecular Biology Organization Journal 18, 1923-1938. 1999.

2. Adam, J. C., Pringle, J. R., and Peifer, M. Evidence for functional differentiation among Drosophila septins in cytokinesis and cellularization. Molecular Biology of The Cell 11, 3123-3135. 2000.

3. Ahmad, K. and Golic, K. G. Somatic reversion of chromosomal position effects in Drosophila melanogaster. Genetics 144, 657-670. 1996.

4. Ahmad, K. and Henikoff, S. Modulation of a transcription factor counteracts heterochromatic gene silencing in Drosophila. Cell 104, 839-847. 2001.

5. Ahuja, N., Li, Q., Mohan, A. L., Baylin, S. B., and Issa, J. P. Aging and DNA methylation in colorectal mucosa and cancer. Cancer Research 58, 5489-5494. 12-1-1998.

6. Aigner, B., Fleischmann, M., Muller, M., and Brem, G. Stable long-term germ-line transmission of transgene integration sites in mice. Transgenic Research 8, 1-8. 1999.

7. Akhmanova, A., Verkerk, T., Langeveld, A., Grosveld, F., and Galjart, N. Characterisation of transcriptionally active and inactive chromatin domains in neurons. Journal of Cell Science 113, 4463-4474. 2000.

8. al-Shawi, R., Kinnaird, J., Burke, J., and Bishop, J. O. Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect. Molecular and Cellular Biology 10, 1192-1198. 1990.

9. al-Shawi, R., Burke, J., Wallace, H., Jones, C., Harrison, S., Buxton, D., Maley, S., Chandley, A., and Bishop, J. O. The herpes simplex virus type 1 thymidine kinase is expressed in the testes of transgenic mice under the control of a cryptic promoter. Molecular and Cellular Biology 11, 4207-4216. 1991.

10. Alami, R., Greally, J. M., Tanimoto, K., Hwang, S., Feng, Y. Q., Engel, J. D., Fiering, S., and Bouhassira, E. E. beta-globin YAC transgenes exhibit uniform expression levels but position effect variegation in mice. Human Molecular Genetics 9, 631-636. 2000.

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

12. Allshire, R. C., Javerzat, J. P., Redhead, N. J., and Cranston, G. Position Effect Variegation at Fission Yeast Centromeres. Cell 76, 157-169. 1994.

13. Amedeo, P., Habu, Y., Afsar, K., Scheid, O. M., and Paszkowski, J. Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. Nature 405, 203-206. 5-11-2000.

14. Amsterdam, A., Lin, S., and Hopkins, N. The Aequorea victoria green fluorescent protein can be used as a reporter in live zebrafish embryos. Developmental Biology 171, 123-129. 1995.

15. Ananiev, E. V. and Gvozdev, V. A. Changed pattern of transcription and replication in polytene chromosomes of Drosophila melanogaster resulting from eu-heterochromatin rearrangement. Chromosoma 45, 173-191. 3-14-1974.

16. Andres, A. C., Schonenberger, C. A., Groner, B., Henninghausen, L., LeMeur, M., and Gerlinger, P. Ha-ras oncogene expression directed by a milk protein gene promoter: tissue specificity, hormonal regulation, and tumor induction in transgenic mice. Proceedings of the National Academy of Sciences U S A 84, 1299-1303. 1987.

17. Antequera, F., Boyes, J., and Bird, A. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. Cell 62, 503-514. 8-10-1990.

18. Aramayo, R. and Metzenberg, R. L. Meiotic transvection in fungi. Cell 86, 103-113. 1996.

19. Archibald, A. L., McClenaghan, M., Hornsey, V., Simons, J. P., and Clark, A. J. High-level expression of biologically active human alpha 1-antitrypsin in the milk of transgenic mice. Proceedings of the National Academy of Sciences U S A 87, 5178-5182. 1990.

20. Aschaffenburg R. and J.Drewry. Improved method for the preparation of crystalline β - lactoglobulin and α – lactalbumin from cow's milk. Biochemical Journal 65, 273-277. 1957.

21. Assaad, F. F., Tucker, K. L., and Signer, E. R. Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis. Plant Molecular Biology 22, 1067-1085. 1993.

22. Aubin-Houzelstein, G., Bernex, F., Elbaz, C., and Panthier, J. J. Survival of patchwork melanoblasts is dependent upon their number in the hair follicle at the end of embryogenesis. Developmental Biology 198, 266-276. 6-15-1998.

23. Bahramian, M. B. and Zarbl, H. Transcriptional and posttranscriptional silencing of rodent alpha-1 (I) collagen by a homologous transcriptionally self-silenced transgene. Molecular and Cellular Biology 19, 274-283. 1999.

24. Balasov, M. L. and Makunin, I. V. Model of continuous and discontinuous chromatin compaction, resulting from position effect variegation. Biologisches Zentralblatt 115, 16-23. 1996.

25. Barash, I, Faerman, A., Richenstein, M., Kari, R., Damary, G. M., Shani, M., and Bissell, M. J. In vivo and in vitro expression of human serum albumin genomic sequences in mammary epithelial cells with beta-lactoglobulin and whey acidic protein promoters. Molecular Reproduction and Development 52, 241-252. 1999.

26. Barbour, V. M., Tufarelli, C., Sharpe, J. A., Smith, Z. E., Ayyub, H., Heinlein, C. A., Sloane-Stanley, J., Indrak, K., Wood, W. G., and Higgs, D. R. alpha-thalassemia resulting from a negative chromosomal position effect. Blood 96, 800-807. 2000.

27. Baulcombe, D. C. Fast forward genetics based on virus-induced gene silencing. Current Opinion in Plant Biology 2, 109-113. 1999.

28. Belyaev, D. K., Ruvinsky, A. O., and Trut, L. N. Inherited activation-inactivation of the star gene in foxes: its bearing on the problem of domestication. Journal of Heredity 72, 267-274. 1981.

29. Belyaeva, E. S. and Zhimulev, I. F. Cytogenetic and molecular aspects of position effect variegation in Drosophila. 3 Continuous and discontinuous compaction of chromosomal material as a result of position effect variegation Chromosoma 100, 453-466. 1991.

30. Belyaeva, E. S., Demakova, O. V., Umbetova, G. H., and Zhimulev, I. F. Cytogenetic and molecular aspects of position effect variegation in Drosophila. melanogaster. Chromosoma 102, 583-590. 1993.

31. Benedict, M. Q., Mcnitt, L. M., Cornell, A. J., and Collins, F. H. Mosaic: A position-effect variegation eye-color mutant in the mosquito Anopheles gambiae. Journal of Heredity 91, 128-133. 2000.

32. Bennett, S. T., Wilson, A. J., Esposito, L., Bouzekri, N., Undlien, D. E., Cucca, F., Nistico, L., Buzzetti, R., Bosi, E., Pociot, F., Nerup, J., CambonThomsen, A., Pugliese, A., Shield, J. P. H., McKinney, P. A., Bain, S. C., Polychronakos, C., Todd, J. A., Pozzilli, P., Visalli, N., Baroni, M., Fioriti, E., Mesturino, C., Signore, A., Cavallo, M., Lucentini, L., Matteoli, M., Crino, A., Teodonio, C., Amoretti, R., Tombesi, A., Ruggeri, M., Pisano, L., Suraci, C., Pennafina, M., Boscherini, B., Stoduto, S., Fonte, M., Mancabitti, M., Multari, G., Suppa, M., DeMattia, G., Faldetta, M. C., Laurenti, O., Marietti, G., Pitocco, D., Ferrazzoli, F., Bizzarri, C., and Ghirlanda, G. Insulin VNTR allele-specific effect in type 1 diabetes depends on identity of untransmitted paternal allele. Nature Genetics 17, 350-352. 1997.

33. Bhadra, U. and Birchler, J. A. Characterization of a sex influenced modifier of gene expression and suppressor of position effect variegation in Drosophila. Molecular & General Genetics 250, 601-613. 1996.

34. Bhadra, U., PalBhadra, M., and Birchler, J. A. A sex-influenced modifier in Drosophila that affects a broad spectrum of target loci included the histone repeats. Genetics 146, 903-917. 1997. **A**

35. Bhadra, U., PalBhadra, M., and Birchler, J. A. A trans-acting modifier causing extensive overexpression of genes in Drosophila melanogaster. Molecular & General Genetics 254, 621-634. 1997. **B**

36. Birchler, J. A., Bhadra, M. P., and Bhadra, U. Making noise about silence: repression of repeated genes in animals. Current Opinion in Genetics and Development 10, 211-216. 2000.

37. Bird, A. P. and Wolffe, A. P. Methylation-induced repression--belts, braces, and chromatin. Cell 99, 451-454. 11-24-1999.

38. Bishop, C. P. and Jackson, C. M. Genomic imprinting of chromatin in Drosophila melanogaster. Genetica 97, 33-37. 1996.

39. Bix, M. and Locksley, R. M. Independent and epigenetic regulation of the interleukin-4 alleles in CD4+ T cells. Science 281, 1352-1354. 1998.

40. Blackwood, E. M. and Kadonaga, J. T. Going the distance: A current view of enhancer action. Science 281, 60-63. 1998.

41. Bleck, G. T. and Bremel, R. D. Variation in expression of a bovine alpha-lactalbumin transgene in milk of transgenic mice. Journal of Dairy Science 77, 1897-1904. 1994.

42. Boerjan, W., Bauw, G., Van Montagu, M., and Inze, D. Distinct phenotypes generated by overexpression and suppression of S-adenosyl-L-methionine synthetase reveal developmental patterns of gene silencing in tobacco. Plant Cell 6, 1401-1414. 1994.

43. Boivin, A. and Dura, J. M. In vivo chromatin accessibility correlates with gene silencing in Drosophila. Genetics 150, 1539-1549. 1998.

44. Bongiorni, S., Cintio, O., and Prantera, G. The relationship between DNA methylation and chromosome imprinting in the Coccid Planococcus citri. Genetics 151, 1471-1478. 1999.

45. Bonifer, C., Huber, M. C., Faust, N., and Sippel, A. E. Regulation of the chicken lysozyme locus in transgenic mice. Critical Reviews in Eukaryotic Gene Expression 6, 285-297. 1996.

46. Boyes, J. and Bird, A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell 64, 1123-1134. 3-22-1991.

47. Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M., and Broach, J. R. Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. Molecular and Cellular Biology 16, 4349-4356. 1996.

48. Brock, D. J. H. Molecular Genetics for the Clinician. 1993.

49. Bronson, S. K., Plaehn, E. G., Kluckman, K. D., Hagaman, J. R., Maeda, N., and Smithies, O. Single-copy transgenic mice with chosen-site integration. Proceedings of the National Academy of Sciences U S A 93, 9067-9072. 8-20-1996.

50. Brown, K. Nuclear structure, gene expression and development. Critical Reviews in Eukaryotic Gene Expression 9, 203-212. 1999.

51. Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M., and Fisher, A. G. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell 91, 845-854. 1997.

52. Brown, K. E., Baxter, J., Graf, D., Merkenschlager, M., and Fisher, A. G. Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. Molecular Cell 3, 207-217. 1999.

53. Brown, K. E., Amoils, S., Horn, J. M., Buckle, V. J., Higgs, D. R., Merkenschlager, M., and Fisher, A. G. Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. Nature Cell Biology 3, 602-606. 2001.

54. Buchenau, P., Hodgson, J., Strutt, H., and ArndtJovin, D. J. The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: Impact on models for silencing. Journal of Cell Biology 141, 469-481. 1998.

55. Bucherna, N., Okkels, F. T., and Palmgren, G. Developmental timing of transgene expression is dosage dependent. Physiologia Plantarum 107, 90-97. 1999.

56. Bulger, M. and Groudine, M. Looping versus linking: toward a model for long-distance gene activation. Genes & Development 13, 2465-2477. 1999.

57. Burdon, T. G., Maitland, K. A., Clark, A. J., Wallace, R., and Watson, C. J. Regulation of the sheep beta-lactoglobulin gene by lactogenic hormes is mediated by a transcription factor that binds an interferon gamma activation site related element. Molecular Endocrinology 8, 1528-1536. 1994.

58. Busturia, A. and Bienz, M. Silencers in abdominal-B, a homeotic Drosophila gene. The European Molecular Biology Organization Journal 12, 1415-1425. 1993.

59. Caldovic, L., Agalliu, D., and Hackett, P. B. Position-independent expression of transgenes in zebrafish. Transgenic Research 8, 321-334. 1999.

60. Cambareri, E. B., Singer, M. J., and Selker, E. U. Recurrence of repeat-induced point mutation (RIP) in Neurospora crassa. Genetics 127, 699-710. 1991.

61. Carrington, E. A. and Jones, R. S. The Drosophila Enhancer of zeste gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. Development 122, 4073-4083. 1996.

62. Carver, A. S., Dalrymple, M. A., Wright, G., Cottom, D. S., Reeves, D. B., Gibson, Y. H., Keenan, J. L., Barrass, J. D., Scott, A. R., and Colman, A. Transgenic livestock as bioreactors: stable expression of human alpha-1-antitrypsin by a flock of sheep. Biotechnology (N Y) 11, 1263-1270. 1993.

63. Caterina, J. J., Ciavatta, D. J., Donze, D., Behringer, R. R., and Townes, T. M. Multiple elements in human beta-globin locus control region 5' HS 2 are involved in enhancer activity and position-independent, transgene expression. Nucleic Acids Research 22, 1006-1011. 1994.

64. Cavalli, G. and Paro, R. The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell 93, 505-518. 1998.

65. Cavalli, G. and Paro, R. Epigenetic inheritance of active chromatin after removal of the main transactivator. Science 286, 955-958. 1999.

66. Cerutti, H., Johnson, A. M., Gillham, N. W., and Boynton, J. E. Epigenetic silencing of a foreign gene in nuclear transformants of Chlamydomonas. Plant Cell 9, 925-945. 1997.

67. Chan, C. S., Rastelli, L., and Pirrotta, V. A polycomb response element in the ubx gene that determines and epigenetically inherited state of repression. The European Molecular Biology Organization Journal 13, 2553-2564. 1994.

68. Chareonpornwattana, S., Thara, K. V., Wang, L., Datta, S. K., and Panbangred, W. Inheritance, expression, and silencing of a chitinase transgene in. Theoretical and Applied Genetics 98, 371-378. 1999.

69. Chen, Z. J., Comai, L., and Pikaard, C. S. Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in Arabidopsis allopolyploids. Proceedings of the National Academy of Sciences U S A 95, 14891-14896. 1998.

70. Chess, A., Simon, I., Cedar, H., and Axel, R. Allelic inactivation regulates olfactory receptor gene expression. Cell 78, 823-834. 1994.

71. Chess, A. Expansion of the allelic exclusion principle? Science 279, 2067-2068. 1998.

72. Chiang, A., O'Connor, M. B., Paro, R., Simon, J., and Bender, W. Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex. Development 121, 1681-1689. 1995.

73. Clark, A. J., Archibald, A. L., McClenaghan, M., Simons, J. P., Wallace, R., and Whitelaw, C. B. Enhancing the efficiency of transgene expression. Philosophical Transactions of the Royal Society of London, series B, Biological Sciences 339, 225-232. 1993.

74. Clark, A. J., Harold, G., and Yull, F. E. Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice. Nucleic Acids Research 25, 1009-1014. 1997.

75. Clark, A. J., Burl, S., Denning, C., and Dickinson, P. Gene targeting in livestock: a preview. Transgenic Research 9, 263-275. 2000.

76. Clark, R. F. and Elgin, S. C. R. Heterochromatin protein-1, a known suppressor of position effect variegation, is highly conserved in Drosophila. Nucleic Acids Research 20, 6067-6074. 1992.

77. Clarke, B., McKay, I, Grigliatti, T., Lloyd, V, and Yuan, A. A Markov model for the assembly of heterochromatic regions in position effect variegation. Journal of Theoretical Biology 181, 137-155. 1996.

78. Cleard, F., Delattre, M., and Spierer, P. SU(VAR)3-7, a Drosophila heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. The European Molecular Biology Organization Journal 16, 5280-5288. 1997.

79. Coffee, B., Zhang, F. P., Warren, S. T., and Reines, D. Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nature Genetics 22, 98-101. 1999.

80. Cogoni, C., Irelan, J. T., Schumacher, M., Schmidhauser, T. J., Selker, E. U., and Macino, G. Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. The European Molecular Biology Organization Journal 15, 3153-3163. 1996.

81. Cohen-Tannoudji, M., Babinet, C., and Morello, D. lacZ and ubiquitously expressed genes: should divorce be pronounced? Transgenic Research 9, 233-235. 2000.**A**

82. Cohen-Tannoudji, M., Vandormael-Pournin, S., Drezen, J., Mercier, P., Babinet, C., and Morello, D. lacZ sequences prevent regulated expression of housekeeping genes. Mechanisms of Development 90, 29-39. 2000. **B**

83. Cohn, S. M., Roth, K. A., Birkenmeier, E. H., and Gordon, J. I. Temporal and spatial patterns of transgene expression in aging adult mice provide insights about the origins, organization, and differentiation of the intestinal epithelium. Proceedings of the National Academy of Sciences U S A 88, 1034-1038. 1991.

84. Collas, P. Modulation of plasmid DNA methylation and expression in zebrafish embryos. Nucleic Acids Research 26, 4454-4461. 1998.

85. Collas, P., Liang, M. R., Vincent, M., and Alestrom, P. Active transgenes in zebrafish are enriched in acetylated histone H4. Journal of Cell Science 112, 1045-1054. 1999.

86. Corcoran, A. E., Riddell, A., Krooshoop, D., and Venkitaraman, A. R. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature 391, 904-907. 1998.

87. Cortes, A., Huertas, D., Fanti, L., Pimpinelli, S., Marsellach, F. X., Pina, B., and Azorin, F. DDP1, a single-stranded nucleic acid-binding protein of Drosophila, associates with pericentric heterochromatin and is functionally homologous to the yeast Scp160p, which is involved in the control of cell ploidy. The European Molecular Biology Organization Journal 18, 3820-3833. 1999.

88. Cowell, I. G. and Austin, C. A. Self-association of chromo domain peptides. Biochimica et Biophysica Acta 1337, 198-206. 1997.

89. Cryderman, D. E., Tang, H. B., Bell, C., Gilmour, D. S., and Wallrath, L. L. Heterochromatic silencing of Drosophila heat shock genes acts at the level of promoter potentiation. Nucleic Acids Research 27, 3364-3370. 1999.**A**

90. Cryderman, D. E., Morris, E. J., Biessmann, H., Elgin, S. C. R., and Wallrath, L. L. Silencing at Drosophila telomeres: nuclear organization and chromatin structure play critical roles. The European Molecular Biology Organization Journal 18, 3724-3735. 1999.**B**

91. Csink, A. K. and Henikoff, S. Genetic modification of heterochromatic association and nuclear organization in Drosophila. Nature 381, 529-531. 1996.

92. Custodio, N., Carmo-Fonseca, M., Geraghty, F., Pereira, H. S., Grosveld, F., and Antoniou, M. Inefficient processing impairs release of RNA from the site of transcription. The European Molecular Biology Organization Journal 18, 2855-2866. 1999.

93. Davies, G. J., Sheikh, M. A., Ratcliffe, O. J., Coupland, G., and Furner, I. J. Genetics of homology-dependent gene silencing in Arabidopsis; a role. Plant Journal 12, 791-804. 1997.

94. Davis, D. B., Delmonte, A. J., Ly, C. T., and McNally, E. M. Myoferlin, a candidate gene and potential modifier of muscular dystrophy. Human Molecular Genetics 9, 217-226. 2000.

95. De Angioletti, M., Rovira, A., Notaro, R., Camacho, Vanegas O., Sadelain, M., and Luzzatto, L. Glucose 6-phosphate dehydrogenase expression is less prone to variegation when driven by its own promoter. Gene 267, 221-231. 2001.

96. de Carvalho, F., Gheysen, G., Kushnir, S., Van Montagu, M., Inze, D., and Castresana, C. Suppression of beta-1,3-glucanase transgene expression in homozygous plants. The European Molecular Biology Organization Journal 11, 2595-2602. 1992.

97. Denning, C., Burl, S., Ainslie, A., Bracken, J., Dinnyes, A., Fletcher, J., King, T., Ritchie, M., Ritchie, W. A., Rollo, M., de Sousa, P., Travers, A., Wilmut, I., and Clark, A. J. Deletion of the alpha(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep. Nature Biotechnology 19, 559-562. 2001.

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

99. DeRubertis, F., Kadosh, D., Henchoz, S., Pauli, D., Reuter, G., Struhl, K., and Spierer, P. The histone deacetylase RPD3 counteracts genomic silencing in Drosophila and yeast. Nature 384, 589-591. 1996.

100. Dieguez, M. J., Vaucheret, H., Paszkowski, J., and Scheid, O. M. Cytosine methylation at CG and CNG sites is not a prerequisite for the initiation of transcriptional gene silencing in plants, but it is required for its maintenance. Molecular and General Genetics 259, 207-215. 1998.

101. Dietz, A., Kay, V., Schlake, T., Landsmann, J., and Bode, J. A plant scaffold attached region detected close to a T-DNA integration site is active in mammalian cells. Nucleic Acids Research 22[14], 2744-2751. 1994.

102. Dimitri, P. and Junakovic, N. Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. Trends in Genetics 15, 123-124. 1999.

103. Dobie, K. W., Lee, M., Fantes, J. A., Graham, E., Clark, A. J., Springbett, A., Lathe, R., and McClenaghan, M. Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus. Proceedings of The National Academy of Sciences USA 93, 6659-6664. 1996.

104. Dobie, K., Mehtali, M., McClenaghan, M., and Lathe, R. Variegated gene expression in mice. Trends in Genetics 13, 127-130. 1997.

105. Donaldson, K. M. and Karpen, G. H. Trans-suppression of terminal deficiency-associated position effect variegation in a Drosophila minichromosome. Genetics 145, 325-337. 1997.

106. Dorer, D. R. and Henikoff, S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in drosophila. Cell 77, 993-1002. 1994.

107. Dorer, D. R. and Henikoff, S. Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. Genetics 147, 1181-1190. 1997.

108. Dorer, D. R. Do transgene arrays form heterochromatin in vertebrates? Transgenic Research 6, 3-10. 1997.

109. Dreesen, T. D., Henikoff, S., and Loughney, K. A pairing sensitive element that mediates trans-inactivation is associated with the drosophila brown gene. Genes & Development 5, 331-340. 1991.

110. Duvillie, B., Bucchini, D., Tang, T., Jami, J., and Paldi, A. Imprinting at the mouse Ins2 locus: evidence for cis- and trans-allelic interactions. Genomics 47, 52-57. 1-1-1998.

111. Eberl, D. F., Lorenz, L. J., Melnick, M. B., Sood, V, Lasko, P., and Perrimon, N. A new enhancer of position-effect variegation in Drosophila melanogaster encodes a putative RNA helicase that binds chromosomes and is regulated by the cell cycle. Genetics 146, 951-963. 1997.

112. Eissenberg, J. C., Morris, G. D., Reuter, G., and Hartnett, T. The heterochromatin-associated protein HP-1 is an essential protein in drosophila with dosage-dependent effects on position effect variegation. Genetics 131, 345-352. 1992.

113. Ekwall, K., Olsson, T., Turner, B. M., Cranston, G., and Allshire, R. C. Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. Cell 91, 1021-1032. 1997.

114. Elgin, S. C. R. Heterochromatin and gene regulation in Drosophila. Current Opinion in Genetics & Development 6, 193-202. 1996.

115. Elliot J.I., Festenstein, R., Toliani, M., and Kioussis, D. Random inactivation of a transgene under the control of a hybrid hCD2 locus control region/Ig enhancer regulatory element. The European Molecular Biology Organization Journal 14, 575-584. 1995.

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

117. Engler, P., Doglio, L. T., Bozek, G., and Storb, U. A cis-acting element that directs the activity of the murine methylation modifier locus Ssm1. Proceedings of The National Academy of Sciences USA 95, 10763-10768. 1998.

118. English, J. J., Davenport, G. F., Elmayan, T., Vaucheret, H., and Baulcombe, D. C. Requirement of sense transcription for homology-dependent virus resistance and trans-inactivation. Plant Journal 12, 597-603. 1997.

119. Epstein, H. Polycomb and friends. BioEssays 14, 411-413. 1992.

120. Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., and Mannens, M. Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. Human Molecular Genetics 4, 415-422. 1995.

121. Feldman, M., Ruan, W., Cunningham, B. C., Wells, J. A., and Kleinberg, D. L. Evidence that the growth hormone receptor mediates differentiation and development of the mammary gland. Endocrinology 133, 1602-1608. 1993.

122. Feng, Y. Q., Lorincz, M. C., Fiering, S., Greally, J. M., and Bouhassira, E. E. Position effects are influenced by the orientation of a transgene with respect to flanking chromatin. Molecular and Cellular Biology 21, 298-309. 2001.

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

124.Festenstein, R., SharghiNamini, S., Fox, M., Roderick, K., Tolaini, M., Norton, T., Saveliev, A., Kioussis, D., and Singh, P. Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. Nature Genetics 23, 457-461. 1999.

125.Fisher, J. and Upadhyaya, M. Molecular genetics of facioscapulohumeral muscular dystrophy (FSHD). Neuromuscular Disorders 7, 55-62. 1997.

126.Fitzmaurice, W. P., Nguyen, L. V., Wernsman, E. A., Thompson, W. F., and Conkling, M. A. Transposon tagging of the Sulfur gene of tobacco using engineered maize Ac/Ds elements. Genetics 153, 1919-1928. 1999.

127.Flavell, R. B. Inactivation of gene expression in plants as a consequence of specific sequence duplication. Proceedings of the National Academy of Sciences U S A 91, 3490-3496. 4-26-1994.

128.Foss, E. J., Garrett, P. W., Kinsey, J. A., and Selker, E. U. Specificity of repeat-induced point mutation (RIP) in Neurospora: sensitivity of non-Neurospora sequences, a natural diverged tandem duplication, and unique DNA adjacent to a duplicated region. Genetics 127, 711-717. 1991.

129.Fox, C. A. and Rine, J. Influences of the cell cycle on silencing. Current Opinion in Cell Biology 8, 354-357. 1996.

130.Francastel, C., Walters, M. C., Groudine, M., and Martin, D. I. K. A functional enhancer suppresses silencing of a transgene and prevents its localization close to centromeric heterochromatin. Cell 99, 259-269. 1999.

131.Francastel, C., Schubeler, D., Martin, D. I. K., and Groudine, M. Nuclear compartmentalization and gene activity. Nature Reviews Molecular Cell Biology 1, 137-143. 2000.

132.Frolov, M. V., Benevolenskaya, E. V., and Birchler, J. A. Regena (Rga), a Drosophila homolog of the global negative transcriptional regulator CDC36 (NOT2) from yeast, modifies gene expression and suppresses position effect variegation. Genetics 148, 317-329. 1998.

133.Fuks, F., Burgers, W. A., Brehm, A., HughesDavies, L., and Kouzarides, T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nature Genetics 24, 88-91. 2000.

134.Garrick, D., Sutherland, H., Robertson, G., and Whitelaw, E. Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. Nucleic Acids Research 24, 4902-4909. 1996.

135.Garrick, D., Fiering, S., Martin, D. I. K., and Whitelaw, E. Repeat-induced gene silencing in mammals. Nature Genetics 18, 56-59. 1998.

136.Gdula, D. A., Gerasimova, T. I., and Corces, V. G. Genetic and molecular analysis of the gypsy chromatin insulator of Drosophila. Proceedings of The National Academy of Sciences USA 93, 9378-9383. 1996.

137.Georgiev, P. G. and Corces, V. G. The Su(HW) protein bound to gypsy sequences in one chromosome can repress enhancer promoter interactions in the paired gene located in the other homolog. Proceedings of The National Academy of Sciences USA 92, 5184-5188. 1995.

138. Gerasimova, T. I. and Corces, V. G. Polycomb and trithorax group proteins mediate the function of a chromatin insulator. Cell 92, 511-521. 1998.

139. Gerasimova, T. I., Byrd, K., and Corces, V. G. A chromatin insulator determines the nuclear localization of DNA. Molecular Cell 6, 1025-1035. 2000.

140. Gimelli, G., Varone, P., Pezzolo, A., Lerone, M., and Pistoia, V. ICF syndrome with variable expression in sibs. Journal of Medical Genetics 30, 429-432. 1993.

141. Gindhart, J. G. Jr and Kaufman, T. C. Identification of Polycomb and trithorax group responsive elements in the regulatory region of the Drosophila homeotic gene Sex combs reduced. Genetics 139, 797-814. 1995.

142. Giraldo, P., Gimenez, E., and Montoliu, L. The use of yeast artificial chromosomes in transgenic animals: expression studies of the tyrosinase gene in transgenic mice. Genetic Analysis-Biomolecular Engineering 15, 175-178. 1999.

143. Giraldo, P. and Montoliu, L. Size matters: use of YACs, BACs and PACs in transgenic animals. Transgenic Research 10, 83-103. 2001.

144. Girard, F., Bello, B., Laemmli, U. K., and Gehring, W. J. In vivo analysis of scaffold-associated regions in Drosophila: A synthetic high-affinity SAR binding protein suppresses position effect variegation. The European Molecular Biology Organization Journal 17, 2079-2085. 1998.

145. Glimm, D. R., Baracos, V. E., and Kennelly, J. J. Molecular evidence for the presence of growth hormone receptors in the bovine mammary gland. Journal of Endocrinology 126, R5-R8. 1990.

146. Golic, M. M. and Golic, K. G. A quantitative measure of the mitotic pairing of alleles in Drosophila melanogaster and the influence of structural heterozygosity. Genetics 143, 385-400. 1996.

147. Gottlieb, B., Beitel, L. K., and Trifiro, M. A. Somatic mosaicism and variable expressivity. Trends Genet 17, 79-82. 2001.

148. Graubert, T. A., Hug, B. A., Wesselschmidt, R., Hsieh, C. L., Ryan, T. M., Townes, T. M., and Ley, T. J. Stochastic, stage-specific mechanisms account for the variegation of a human globin transgene. Nucleic Acids Research 26, 2849-2858. 1998.

149. Grewal, S. I. S. Transcriptional silencing in fission yeast. Journal of Cellular Physiology 184, 311-318. 2000.

150. Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R., and Fraser, P. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. Molecular Cell 5, 377-386. 2000.

151. Grosveld, F., deBoer, E., Dillon, N., Fraser, P., Gribnau, J., Milot, E., Trimborn, T., and Wijgerde, M. The dynamics of globin gene expression and gene therapy vectors. Seminars in Hematology 35, 105-111. 1998.

152. Gubb, D., Roote, J., Trenear, J., Coulson, D., and Ashburner, M. Topological constraints on transvection between white genes within the transposing element TE35B in Drosophila melanogaster. Genetics 146, 919-937. 1997.

153. Guo, H. S., Cervera, M. T., and Garcia, J. A. Plum pox potyvirus resistance associated to transgene silencing that can be stabilized after different number of plant generations. Gene 206, 263-272. 1998.

154. Gutierrez, A., Meade, H. M., Ditullio, P., Pollock, D., Harvey, M., Jimenez-Flores, R., Anderson, G. B., Murray, J. D., and Medrano, J. F. Expression of a bovine kappa-CN cDNA in the mammary gland of transgenic mice utilizing a genomic milk protein gene as an expression cassette. Transgenic Research 5, 271-279. 1996.

155. Guy, L. G., Kothary, R., DeRepentigny, Y., Delvoye, N., Ellis, J., and Wall, L. The beta-globin locus control region enhances transcription of but does not confer position-independent expression onto the lacZ gene in transgenic mice. The European Molecular Biology Organization Journal 15, 3713-3721. 1996.

156. Guy, L. G., Kothary, R., and Wall, L. Position effects in mice carrying a lacZ transgene in cis with the beta-globin LCR can be explained by a graded model. Nucleic Acids Research 25, 4400-4407. 1997.

157. Haaf, T. and Schmid, M. Chromosome topology in mammalian interphase nuclei. Experimental Cell Research 192, 325-332. 1991.

158. Hagstrom, K., Muller, M., and Schedl, P. Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the Drosophila bithorax complex. Genes & Development 10, 3202-3215. 1996.

159. Hagstrom, K., Muller, M., and Schedl, P. A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the Drosophila bithorax complex. Genetics 146, 1365-1380. 1997.

160. Haller, B. S. and Woodruff, R. C. Varied expression of a Y-linked P[w(+)] insert due to imprinting in Drosophila melanogaster. Genome 43, 285-292. 2000.

161. Han, S. J., Cho, H. S., You, J. S., Nam, Y. W., Park, E. K., Shin, J. S., Park, Y. I., Park, W. M., and Paek, K. H. Gene silencing-mediated resistance in transgenic tobacco plants carrying potato virus Y coat protein gene. Molecules and Cells 9, 376-383. 1999.

162. Hark, A. T. and Tilghman, S. M. Chromatin conformation of the H19 epigenetic mark. Human Molecular Genetics 7, 1979-1985. 1998.

163. Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. Proceedings of the National Academy of Sciences U S A 94, 2122-2127. 1997.

164. Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Sasaki, N., Hirose, K., Imoto, H., Okuizumi, H., and Muramatsu, M. Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method. Nat Genet 6, 33-40. 1994.

165. Hazelrigg, T., Levis, R., and Rubin, G. M. Transformation of white locus DNA in drosophila: dosage compensation, zeste interaction, and position effects. Cell 36, 469-481. 1984.

166. Henchoz, S., De Rubertis, F., Pauli, D., and Spierer, P. The dose of a putative ubiquitin-specific protease affects position-effect variegation in Drosophila melanogaster. Molecular and Cellular Biology 16, 5717-5725. 1996.

167. Henikoff, S. and Dreesen, T. D. Trans- inactivation of the drosophila brown gene – evidence for transcriptional repression and somatic pairing dependence. Proceedings of The National Academy of Sciences USA 86, 6704-6708. 1989.

168. Henikoff, S. Position-effect variegation after 60 years. Trends in Genetics 6, 422-426. 1990.

169. Henikoff, S. Dosage-dependent modification of position-effect variegation in Drosophila. BioEssays 18, 401-409. 1996.

170. Henikoff, S. Nuclear organization and gene expression: Homologous pairing and long-range interactions. Current Opinion in Cell Biology 9, 388-395. 1997.

171. Henikoff, S., Greene, E. A., Pietrokovski, S., Bork, P., Attwood, T. K., and Hood, L. Gene families: the taxonomy of protein paralogs and chimeras. Science 278, 609-614. 1997.

172. Henikoff, S. and Matzke, M. A. Exploring and explaining epigenetic effects. Trends in Genetics 13, 293-295. 1997.

173. Henikoff, S. and Comai, L. Trans-sensing effects: the ups and downs of being together. Cell 93, 329-332. 1998.

174. Henikoff, S. Heterochromatin function in complex genomes. Biochimica Et Biophysica Acta-Reviews On Cancer 1470, 2000.

175. Hollander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, A. K., and Burakoff, S. J. Monoallelic expression of the interleukin-2 locus. Science 279, 2118-2121. 1998.

176. Hollick, J. B., Dorweiler, J. E., and Chandler, V. L. Paramutation and related allelic interactions. Trends in Genetics 13, 302-308. 1997.

177. Holliday, R. The inheritance of epigenetic defects. Science 238, 163-170. 1987.

178. Holliday, R. Epigenetics: an overview. Developmental Genetics 15, 453-457. 1994.

179. Holmquist, G. P. Role of replication time in the control of tissue-specific gene expression. American Journal of Human Genetics 40, 151-173. 1987.

180. Hsieh, J. and Fire, A. Recognition and silencing of repeated DNA. Annual Review of Genetics 34, 187-204. 2000.

181. Huber, M. C., Kruger, G., and Bonifer, C. Genomic position effects lead to an inefficient reorganization of nucleosomes in the 5'-regulatory region of the chicken lysozyme locus in transgenic mice. Nucleic Acids Research 24, 1443-1452. 1996.

182. Iglesias, V. A., Moscone, E. A., Papp, I., Neuhuber, F., Michalowski, S., Phelan, T., Spiker, S., Matzke, M., and Matzke, A. J. Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. Plant Cell 9, 1251-1264. 1997.

183. Ingelbrecht, I., Van Houdt, H., Van Montagu, M., and Depicker, A. Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. Proceedings of the National Academy of Sciences U S A 91, 10502-10506. 1994.

184. Itoh, K., Nakajima, M., and Shimamoto, K. Silencing of waxy genes in rice containing Wx transgenes. Molecular and General Genetics 255, 351-358. 1997.

185. Iyer, L. M. and Hall, T. C. Virus recovery is induced in Brome mosaic virus p2 transgenic plants showing synchronous complementation and RNA-2-specific silencing. Molecular Plant Microbe Interactions 13, 247-258. 2000.

186. Jablonka, E., Lachmann, M., and Lamb, M. J. Evidence, mechanisms and models for the inheritance of acquired characters. Journal of Theoretical Biology 158, 245-268. 1992.

187. Jaenisch, R. DNA methylation and imprinting: why bother? Trends in Genetics 13, 323-329. 1997.

188. Jakowitsch, J., Papp, I., Moscone, E. A., van der Winden, J., Matzke, M., and Matzke, A. J. Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans. Plant Journal 17, 131-140. 1999.

189. Jazwinski, S. M. Longevity, genes, and aging: a view provided by a genetic model system. Experimental Gerontology 34, 1-6. 1999.

190. Jeddeloh, J. A., Bender, J., and Richards, E. J. The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. Genes & Development 12, 1714-1725. 1998.

191. Jensen, S., Gassama, M. P., and Heidmann, T. Cosuppression of I transposon activity in drosophila by I-containing sense and antisense transgenes. Genetics 153, 1767-1774. 1999.

192. Jenuwein, T. Re-SET-ting heterochromatin by histone methyltransferases. Trends in Cell Biology 11, 266-273. 2001.

193. Jeong, S., Lee, Y. J., Jang, J. S., Park, C. W., Chung, J. H., Seong, J. K., Lee, K. K., and Yu, D. Y. A novel epigenetic control operating on Vme1(+) locus leads to variegated monoallelic expression. Biochemical and Biophysical Research Communications 279, 884-890. 2000.

194. John, R. M. and Surani, M. A. Agouti germ line gets acquisitive. Nature Genetics 23, 254-256. 1999.

195. Jorgensen, R. Altered gene expression in plants due to trans interactions between homologous genes. Trends in Biotechnology 8, 340-344. 1990.

196. Jorgensen, R. A. Cosuppression, flower color patterns, and metastable gene-expression states. Science 268, 686-691. 1995.

197. Jorgensen, R. A., Atkinson, R. G., Forster, R. L., and Lucas, W. J. An RNA-based information superhighway in plants. Science 279, 1486-1487. 1998.

198. Judd, B. H. Transvection: allelic cross talk. Cell 53, 841-843. 6-17-1988.

199. Kabra, N. H., Kang, C., Hsing, L. C., Zhang, J., and Winoto, A. T cell-specific FADD-deficient mice: FADD is required for early T cell development. Proceedings of the National Academy of Sciences U S A 98, 6307-6312. 2001.

200. Karpen, G. H. and Spradling, A. C. Reduced DNA polytenization of a minichromosome region undergoing position-effect variegation in drosophila. Cell 63, 97-107. 1990.

201. Kearns, M., Preis, J., McDonald, M., Morris, C., and Whitelaw, E. Complex patterns of inheritance of an imprinted murine transgene suggest incomplete germline erasure. Nucleic Acids Research 28, 3301-3309. 2000.

202. Kellum, R. and Alberts, B. M. Heterochromatin protein-1 is required for correct chromosome segregation in drosophila embryoes. Journal of Cell Science 108, 1419-1431. 1995.

203. Kennison, J. A. The Polycomb and trithorax group proteins of Drosophila: Trans- regulators of homeotic gene function. Annual Review of Genetics 29, 289-303. 1995.

204. Kieffer, L. J., Yan, L., Hanke, J. H., and Kavathas, P. B. Appropriate developmental expression of human CD8 beta in transgenic mice. Journal of Immunology 159, 4907-4912. 1997.

205. Kilby, N. J., Leyser, H. M., and Furner, I. J. Promoter methylation and progressive transgene inactivation in Arabidopsis. Plant Molecular Biology 20, 103-112. 1992.

206. Kim, S., Benguria, A., Lai, C. Y., and Jazwinski, S. M. Modulation of life-span by histone deacetylase genes in Saccharomyces cerevisiae. Molecular Biology of The Cell 10, 3125-3136. 1999.

207. Kim, S. K. and Wang, J. C. Gene silencing via protein-mediated subcellular localization of DNA. Proceedings of The National Academy of Sciences USA 96, 8557-8561. 1999.

208. Kioussis, D. and Festenstein, R. Locus control regions: overcoming heterochromatin-induced gene inactivation in mammals. Current Opinion in Genetics & Development 7, 614-619. 1997.

209. Kleinjan, D. J. and vanHeyningen, V. Position effect in human genetic disease. Human Molecular Genetics 7, 1611-1618. 1998.

210. Kmita, M., vanderHoeven, F., Zakany, J., Krumlauf, R., and Duboule, D. Mechanisms of Hox gene colinearity: transposition of the anterior Hoxb1 gene into the posterior HoxD complex. Genes & Development 14, 198-211. 2000.

211. Koetsier, P. A., Mangel, L., Schmitz, B., and Doerfler, W. Stability of transgene methylation patterns in mice: position effects, strain specificity and cellular mosaicism. Transgenic Research 5, 235-244. 1996.

212. Kooter, J. M., Matzke, M. A., and Meyer, P. Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends in Plant Science 4, 340-347. 1999.

213. Kordon, E. C. and Smith, G. H. An entire functional mammary gland may comprise the progeny from a single cell. Development 125, 1921-1930. 1998.

214. Kricker, M. C., Drake, J. W., and Radman, M. Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. Proceedings of the National Academy of Sciences U S A 89, 1075-1079. 2-1-1992.

215. Kunz, C., Schob, H., Stam, M., Kooter, J. M., and Meins, F. Developmentally regulated silencing and reactivation of tobacco chitinase transgene expression. Plant Journal 10, 437-450. 1996.

216. Laible, G., Wolf, A., Dorn, R., Reuter, G., Nislow, C., Lebersorger, A., Popkin, D., Pillus, L., and Jenuwein, T. Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at S-cerevisiae telomeres. The European Molecular Biology Organization Journal 16, 3219-3232. 1997.

217. Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R., and Berns, A. Simplified mammalian DNA isolation procedure. Nucleic Acids Research 19, 4293. 1991.

218. Lardon, A., Georgiev, S., Aghmir, A., LeMerrer, G., and Negrutiu, I. Sexual dimorphism in white campion: Complex control of carpel number. Genetics 151, 1173-1185. 1999.

219. Larsson, J., Zhang, J. P., and RasmusonLestander, A. Mutations in the Drosophila melanogaster gene encoding S- adenosylmethionine suppress position-effect variegation. Genetics 143, 887-896. 1996.

220. LaSalle, J. M. and Lalande, M. Homologous association of oppositely imprinted chromosomal domains. Science 272, 725-728. 1996.

221. Lewis, E. B. The theory and application of a new method of detecting chromosomal rearrangements in Drosophila melanogaster. American Naturalist 88, 225-239. 1954.

222. Lien, L. L., Lee, Y., and Orkin, S. H. Regulation of the myeloid-cell-expressed human gp91-phox gene as studied by transfer of yeast artificial chromosome clones into embryonic stem cells: suppression of a variegated cellular pattern of expression requires a full complement of distant cis elements. Molecular and Cellular Biology 17, 2279-2290. 1997.

223. Lloyd, V. K., Sinclair, D. A., and Grigliatti, T. A. Competition between different variegating rearrangements for limited heterochromatic factors in Drosophila melanogaster. Genetics 145, 945-959. 1997.

224. Lloyd, V. K., Sinclair, D. A., and Grigliatti, T. A. Genomic imprinting and position-effect variegation in Drosophila melanogaster. Genetics 151, 1503-1516. 1999.

225. Lu, B. Y., Bishop, C. P., and Eissenberg, J. C. Developmental timing and tissue specificity of heterochromatin- mediated silencing. The European Molecular Biology Organization Journal 15, 1323-1332. 1996.

226. Lu, B. Y., Ma, J. Y., and Eissenberg, J. C. Developmental regulation of heterochromatin-mediated gene silencing in Drosophila. Development 125, 2223-2234. 1998.

227. Lu, B. Y., Emtage, P. C. R., Duyf, B. J., Hilliker, A. J., and Eissenberg, J. C. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. Genetics 155, 699-708. 2000.

228. Lundgren, M., Chow, C. M., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell 103, 733-743. 2000.

229. Lyko, F., Ramsahoye, B. H., Kashevsky, H., Tudor, M., Mastrangelo, M. A., Orr-Weaver, T. L., and Jaenisch, R. Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. Nature Genetics 23, 363-366. 1999.

230. Maessen, G. D. F. Genomic Stability and stability of expression in genetically modified plants. Acta Botanica Neerlandica 46, 3-24. 1997.

231. Mahmoudi, T. and Verrijzer, C. P. Chromatin silencing and activation by Polycomb and trithorax group proteins. Oncogene 20, 3055-3066. 2001.

232. Marahrens, Y. X-inactivation by chromosomal pairing events. Genes & Development 13, 2624-2632. 1999.

233. Martin, C. C. and McGowan, R. Genotype-specific modifiers of transgenic methylation and expression in the zebrafish, Danio-Rerio. Genetical Research 65, 21-28. 1995.

234. MartinMorris, L. E. and Henikoff, S. Conservation of brown gene trans-inactivation in drosophila. Genetics 140, 193-199. 1995.

235. MartinMorris, L. E., Csink, A. K., Dorer, D. R., Talbert, P. B., and Henikoff, S. Heterochromatic trans-inactivation of Drosophila white transgenes. Genetics 147, 671-677. 1997.

236. Matsuoka, T., Hirata, M., Tanaka, H., Takahashi, Y., Murata, T., Kabashima, K., Sugimoto, Y., Kobayashi, T., Ushikubi, F., Aze, Y., Eguchi, N., Urade, Y., Yoshida, N., Kimura, K., Mizoguchi, A., Honda, Y., Nagai, H., and Narumiya, S. Prostaglandin D2 as a mediator of allergic asthma. Science 287, 2013-2017. 2000.

237. Matzke, A. J. M. and Matzke, M. A. Trans-inactivation of homologous sequences in Nicotiana tabacum. Current topics in microbiology and immunology: "Gene silencing in higher plants and related phenomena" pages 1-14. 1995.

238. Matzke, A. J. M. and Matzke, M. A. Position effects and epigenetic silencing of plant transgenes. Current Opinion in Plant Biology 1, 142-148. 1998.**A**

239. Matzke, M. A. and Matzke, A. J. M. Epigenetic silencing of plant transgenes as a consequence of diverse cellular defence responses. Cellular and Molecular Life Sciences 54, 94-103. 1998.**B**

240. Matzke, M. A., Neuhuber, F., and Matzke, A. J. M. A variety of epistatic interactions can occur between partially homologus transgene loci brought together by sexual crossing. Molecular & General Genetics 236, 379-386. 1993.

241. Matzke, M. A., Mette, M. F., and Matzke, A. J. M. Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Molecular Biology 43, 401-415. 2000.

242. Matzke, M., Mette, M. F., Jakowitsch, J., Kanno, T., Moscone, E. A., van der Winden, J., and Matzke, A. J. M. A test for transvection in plants: DNA pairing may lead to trans-activation or silencing of complex heteroalleles in tobacco. Genetics 158, 451-461. 2001.

243. Mayeux - Portas, V., File, S. E., Stewart, C. L., and Morris, R. J. Mice lacking the cell adhesion molecule Thy-1 fail to use socially transmitted cues to direct their choice of food. Current Biology 10, 68-75. 2000.

244. McCabe, M. S., Schepers, F., vanderArend, A., Mohapatra, U., deLaat, A. M. M., Power, J. B., and Davey, M. R. Increased stable inheritance of herbicide resistance in transgenic lettuce carrying a petE promoter-bar gene compared with a CaMV 35S- bar gene. Theoretical and Applied Genetics 99, 587-592. 1999.

245. McCall, K. and Bender, W. Probes for chromatin accessibility in the Drosophila bithorax complex respond differently to Polycomb-mediated repression. The European Molecular Biology Organization Journal 15, 569-580. 1996.

246. McClenaghan, M., Springbett, A., Wallace, R. M., Wilde, C. J., and Clark, A. J. Secretory proteins compete for production in the mammary gland of transgenic mice. Biochemical Journal 310, 637-641. 1995.

247. McFarlane, M. and Wilson, J. B. A model for the mechanism of precise integration of a microinjected transgene. Transgenic Research 5, 171-177. 1996.

248. McGowan, R., Campbell, B., Peterson, A., and Sapienza, C. Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse. Genes & Development 3 , 1669-1676. 3-11-1989.

249. McMorrow, T., van den Wijngaard, A., Wollenschlaeger, A., van de Coreut, M., Monkhorst, M., Trimborn, T., Fraser, P., van Lohuizen, M., Jenuwein, T., Djabali, M., Philipsen, S., Grosveld, F., and Milot, E. Activation of the beta globin locus by transcription factors and chromatin modifiers. The European Molecular Biology Organization Journal 19, 4986-4996. 2000.

250. Mehtali, M., LeMeur, M., and Lathe, R. The methylation-free status of a housekeeping transgene is lost at high copy number. Gene 91, 179-184. 1990.

251. Meyer, P., Heidmann, I, and Niedenhof, I. Differences in DNA-methylation are associated with a paramutation phenomenon in transgenic petunia. Plant Journal 4, 89-100. 1993.

252. Meyer, P. and Saedler, H. Homology-dependent gene silencing in plants. Annual Review of Plant Physiology and Plant Molecular Biology 47, 23-48. 1996.

253. Meyer, P. The role of chromatin remodeling in transgene silencing and plant development. In Vitro Cellular & Developmental Biology-Plant ITRO, 29-36. 1999.

254. Michelotti, E. F., Sanford, S., and Levens, D. Marking of active genes on mitotic chromosomes. Nature 388, 895-899. 1997.

255. Miele, G., Manson, J., and Clinton, M. A novel erythroid-specific marker of transmissible spongiform encephalopathies. Nature Medicine 7, 361-364. 2001.

256. Milot, E., Strouboulis, J., Trimborn, T., Wijgerde, M., deBoer, E., Langeveld, A., TanUn, K., Vergeer, W., Yannoutsos, N., Grosveld, F., and Fraser, P. Heterochromatin effects on the frequency and duration of LCR- mediated gene transcription. Cell 87, 105-114. 1996.

257. Montoliu, L., Chavez, S., and Vidal, M. Variegation associated with lacZ in transgenic animals: a warning note. Transgenic Research 9, 237-239. 2000.

258. Morgan, H. D., Sutherland, H. G., Martin, D. I., and Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. Nature Genetics 23, 314-318. 1999.

259. Morjan, W. E., Obrycki, J. J., and Krafsur, E. S. Inbreeding effects on Propylea quatuordecimpunctata (Coleoptera: Coccinellidae). Annals of The Entomological Society of America 92, 260-268. 1999.

260. Morley, S. D., Viard, I, Chung, B. C., Ikeda, Y., Parker, K. L., and Mullins, J. J. Variegated expression of a mouse steroid 21-hydroxylase/beta- galactosidase transgene suggests centripetal migration of adrenocortical cells. Molecular Endocrinology 10, 585-598. 1996.

261. Morris, J. R., Chen, J. L., Filandrinos, S. T., Dunn, R. C., Fisk, R., Geyer, P. K., and Wu, C. T. An analysis of transvection at the yellow locus of Drosophila melanogaster. Genetics 151, 633-651. 1999.

262. Mottus, R., Sobel, R. E., and Grigliatti, T. A. Mutational analysis of a histone deacetylase in Drosophila melanogaster: Missense mutations suppress gene silencing associated with position effect variegation. Genetics 154, 657-668. 2000.

263. Mueller, E., Gilbert, J., Davenport, G., Brigneti, G., and Baulcombe, D. C. Homology-dependent resistance – transgenic virus-resistance in plants related to homology-dependent gene silencing. Plant Journal 7, 1001-1013. 1995.

264. Muller, H. J. Types of visible variations induced by X-rays in *Drosophila*. Journal ff Genetics 22, 299-334. 1930.

265. Muller, H. J. Further studies on the nature and causes of gene mutations. Proceedings of the .International.Congress on Genetics 1, 213-225. 1932.

266. Muller, J. and Bienz, M. Long range repression conferring boundaries of Ultrabithorax expression in the Drosophila embryo. The European Molecular Biology Organization Journal 10, 3147-3155. 1991.

267. Muller, M., Hagstrom, K., Gyurkovics, H., Pirrotta, V, and Schedl, P. The Mcp element from the Drosophila melanogaster bithorax complex mediates long-distance regulatory interactions. Genetics 153, 1333-1356. 1999.

268. Munakata, T., Adachi, N., Yokoyama, N., Kuzuhara, T., and Horikoshi, M. A human homologue of yeast anti-silencing factor has histone chaperone activity. Genes To Cells 5, 221-233. 2000.

269. Netzer, C., Rieger, L., Brero, A., Zhang, C. D., Hinzke, M., Kohlhase, J., Bohlander, S. K. SALL1, the gene mutated in Townes-Brocks syndrome, encodes a transcriptional repressor which interacts with TRF1/PIN2 and localizes to pericentromeric heterochromatin. Human Molecular Genetics 10 (26), 3017 – 3024. 2001.

270. Neuhuber, F., Park, Y. D., Matzke, A. J. M., and Matzke, M. A. Susceptibility of transgene loci to homology-dependent gene silencing. Molecular & General Genetics 244, 230-241. 1994.

271. Ng, H. H. and Bird, A. DNA methylation and chromatin modification. Current Opinion in Genetics & Development 9, 158-163. 1999.

272. Ogilvy, S., Elefanty, A. G., Visvader, J., Bath, M. L., Harris, A. W., and Adams, J. M. Transcriptional regulation of vav, a gene expressed throughout the hematopoietic compartment. Blood 91, 419-430. 1998.

273. Okano, M., Bell, D. W., Haber, D. A., and Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99, 247-257. 1999.

274. Orlando, V. and Paro, R. Chromatin multiprotein complexes involved in the maintenance of transcription patterns. Current Opinion in Genetics and Development 5, 174-179. 1995.

275. Pal-Bhadra, M., Bhadra, U., and Birchler, J. A. Cosuppression in Drosophila: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent. Cell 90, 479-490. 1997.

276. Palmiter, R. D., Chen, H. Y., and Brinster, R. L. Differential regulation of metallothionein-thymidine kinase fusion genes in transgenic mice and their offspring. Cell 29, 701-710. 1982.

277. Palmiter, R. D., Wilkie, T. M., Chen, H. Y., and Brinster, R. L. Transmission distortion and mosaicism in an unusual transgenic mouse pedigree. Cell 36, 869-877. 1984.

278. Palmiter, R. D., Sandgren, E. P., Koeller, D. M., and Brinster, R. L. Distal regulatory elements from the mouse metallothionein locus stimulate gene expression in transgenic mice . Molecular and Cellular Biology 13, 5266-5275. 1993.

279. Panning, B. and Jaenisch, R. RNA and the epigenetic regulation of X chromosome inactivation. Cell 93, 305-308. 5-1-1998.

280. Park, Y. D., Ronis, D. H., and Lorenzen, J. H. Suppression of marker genes in sequentially transformed potato (Solanium tuberosum L) American Potato Journal 72, 599-604. 1995

281. Park, Y.-D., Papp, I, Moscone, E. A., Iglesias, V. A., Vaucheret, H., Matzke, A. J., and Matzke, M. A. Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. The Plant Journal 9, 183-194. 1996.

282. Paro, R. and Hogness, D. S. The polycomb protein shares a homologous domain with a heterochromatin-associated protein of drosophila. Proceedings of The National Academy of Sciences USA 88, 263-267. 1991.

283. Persuy, M. A., Stinnakre, M. G., Printz, C., Mahe, M. F., and Mercier, J. C. High expression of the caprine beta-casein gene in transgenic mice. European Journal of Biochemistry 205, 887-893. 1992.

284. Philbrick, W. M., Dreyer, B. E., Nakchbandi, I. A., and Karaplis, A. C. Parathyroid hormone-related protein is required for tooth eruption. Proceedings of the National Academy of Sciences U S A 95, 11846-11851. 1998.

285. Pinto, Y. M., Kok, R. A., and Baulcombe, D. C. Resistance to rice yellow mottle virus (RYMV) in cultivated African rice varieties containing RYMV transgenes. Nature Biotechnology 17, 702-707. 1999.

286. Pittius, C. W., Henninghausen, L., Lee, E., Westphal, H., Nicols, E., Vitale, J., and Gordon, K. A Milk protein gene promoter directs the expression of human tissue plasminogen-activator cDNA to the mammary-gland in transgenic mice. Proceedings of The National Academy of Sciences USA 85, 5874-5878. 1988.

287. Platero, J. S., Sharp, E. J., Adler, P. N., and Eissenberg, J. C. In vivo assay for protein-protein interactions using Drosophila chromosomes. Chromosoma 104, 393-404. 1996.

288. Porter, S. D. and Meyer, C. J. 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. 1994.

289. Ramirez, A., Milot, E., Ponsa, I, Marcos-Gutierrez, C., Page, A., Santos, M., Jorcano, J., and Vidal, M. Sequence and chromosomal context effects on variegated expression of keratin 5/lacZ constructs in stratified epithelia of transgenic mice. Genetics 158, 341-350. 2001.

290. Ratcliff, F., Harrison, B. D., and Baulcombe, D. C. A similarity between viral defense and gene silencing in plants. Science 276, 1558-1560. 1997.

291. Reeve, A. E. Role of genomic imprinting in Wilms' tumour and overgrowth disorders. Medical and Pediatric Oncology 27, 470-475. 1996.

292. Renauld, H. and Gasser, S. M. Heterochromatin: A meiotic matchmaker? Trends in Cell Biology 7, 201-205. 1997.

293. Reuter, G., Giarre, M., Fara, J., Gausz, J., Spierer, A., and Spierer, P. Dependence of position-effect variegation in drosophila on dose of a gene encoding an unusual zinc-finger protein. Nature 344, 219-223. 1990.

294. Rhoades, K. L., Singh, N., Simon, I., Glidden, B., Cedar, H., and Chess, A. Allele-specific expression patterns of interleukin-2 and Pax-5 revealed by a sensitive single-cell RT-PCR analysis. Current Biology 10, 789-792. 6-29-2000.

295. Rijkers, T., Peetz, A., and Ruther, U. Insertional mutagenesis in transgenic mice. Transgenic Research 3, 203-215. 1994.

296. Robertson, G., Garrick, D., Wilson, M., Martin, D. I. K., and Whitelaw, E. Age-dependent silencing of globin transgenes in the mouse. Nucleic Acids Research 24, 1465-1471. 1996.

297. Roemer, I., Reik, W., Dean, W., and Klose, J. Epigenetic inheritance in the mouse. Current Biology 7, 277-280. 1997.

298. Ronsseray, S., Marin, L., Lehmann, M., and Anxolabehere, D. Repression of hybrid dysgenesis in Drosophila melanogaster by combinations of telomeric P-element reporters and naturally occurring P elements. Genetics 149, 1857-1866. 1998.

299. Sabl, J. F. and Laird, C. D. Epigene conversion - a proposal with implications for gene mapping in humans. American Journal of Human Genetics 50, 1171-1177. 1992.

300. Sabl, J. F. and Henikoff, S. Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in Drosophila. Genetics 142, 447-458. 1996.

301. Santoso, B., Ortiz, B. D., and Winoto, A. Control of organ-specific demethylation by an element of the T-cell receptor-alpha locus control region. Journal of Biological Chemistry 275, 1952-1958. 2000.

302. Sass, G. L. and Henikoff, S. Comparative analysis of position-effect variegation mutations in Drosophila melanogaster delineates the targets of modifiers. Genetics 148, 733-741. 1998.

303. Sass, G. L. and Henikoff, S. Pairing-dependent mislocalization of a Drosophila brown gene reporter to a heterochromatic environment. Genetics 152, 595-604. 1999.

304. Scheid, O. M., Afsar, K., and Paszkowski, J. Release of epigenetic gene silencing by trans-acting mutations in Arabidopsis. Proceedings of The National Academy of Sciences USA 95, 632-637. 1998.

305. Schmulling, T. and Rohrig, H. Gene silencing in transgenic tobacco hybrids: frequency of the event and visualization of somatic inactivation pattern. Molecular & General Genetics 249, 375-390. 1995.

306. Scholz, H., Deatrick, J., Klaes, A., and Klambt, C. Genetic dissection of pointed, a drosophila gene encoding 2 ETS-related proteins. Genetics 135, 455-468. 1993.

307. Schultz and Dobzhansky. The relation of a dominant eye colour in Drosophila melanogaster to the associated chromosome rearrangement. Genetics 19, 344-364. 1934.

308. Schweizer, J., ValenzaSchaerly, P., Goret, F., and Pourcel, C. Control of expression and methylation of a hepatitis B virus transgene by strain-specific modifiers. DNA and Cell Biology 17, 427-435. 1998.

309. Scrable, H. and Stambrook, P. J. A genetic program for deletion of foreign DNA from the mammalian genome. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 429, 225-237. 1999.

310. Selker, E. U. Epigenetic phenomena in filamentous fungi: useful paradigms or repeat-induced confusion? Trends in Genetics 13, 296-301. 1997.

311. Selker, E. U. Gene silencing: Repeats that count. Cell 97, 157-160. 1999.

312. Seo, S., Okamoto, N., Seto, H., Ishizuka, K., Sano, H., and Ohashi, Y. Tobacco MAP-kinase – a possible mediator in wound signal-transduction pathways. Science 270, 1988-1992. 1995.

313.Seum, C., Spierer, A., Pauli, D., Szidonya, J., Reuter, G., and Spierer, P. Position-effect variegation in Drosophila depends on the dose of the gene encoding the E2F transcriptional activator and cell cycle regulator. Development 122, 1949-1956. 1996.

314.Shani, M., Barash, I., Nathan, M., Ricca, G., Searfoss, G. H., Dekel, I., Faerman, A., Givol, D., and Hurwitz, D. R. Expression of human serum albumin in the milk of transgenic mice. Transgenic Research 1, 195-208. 1992.

315.Sharp, P. A. RNAi and double-strand RNA. Genes & Development 13, 139-141. 1999.

316.Sheardown, S., Norris, D., Fisher, A., and Brockdorff, N. The mouse Smcx gene exhibits developmental and tissue specific variation in degree of escape from X inactivation. Human Molecular Genetics 5, 1355-1360. 1996.

317.Sherman, J. M. and Pillus, L. An uncertain silence. Trends in Genetics 13, 308-313. 1997.

318.Silva, A. J. and White, R. Inheritance of allelic blueprints for methylation patterns. Cell 54, 145-152. 1988.

319.Simon, D. and Knowles, B. B. Newly acquired peri-telomeric heterochromatin in a transgenic mouse lineage. Cytogenetics and Cell Genetics 62, 211-213. 1993.

320.Simon, J., Peifer, M., Bender, W., and O'Connor, M. Regulatory elements of the bithorax complex that control expression along the anterior-posterior axis. The European Molecular Biology Organization Journal 9, 3945-3956. 1990.

321.Simon, J., Chiang, A., Bender, W., Shimell, M. J., and O'Connor, M. Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products. Developmental Biology 158, 131-144. 1993.

322.Simons, H. and Tucker, G. A. Simultaneous co-suppression of polygalacturonase and pectinesterase in tomato fruit: inheritance and effect on isoform profiles. Phytochemistry 52, 1017-1022. 1999.

323.Simons, J. P., McClenaghan, M., and Clark, A. J. Alteration of the quality of milk by expression of sheep b - lactoglobulin in transgenic mice. Nature 328, 530-532. 1987.

324.Sinclair, D. A. R., Schulze, S., Silva, E., Fitzpatrick, K. A., and Honda, B. M. Essential genes in autosomal heterochromatin of Drosophila melanogaster. Genetica 109, 9-18. 2000.

325.Sinclair, K. D., Young, L. E., Wilmut, I., and McEvoy, T. G. In-utero overgrowth in ruminants following embryo culture: lessons from mice and a warning to men. Human Reproduction 15 Supplement 5, 68-86. 2000.

326.Singer, M. J., Marcotte, B. A., and Selker, E. U. DNA methylation associated with repeat-induced point mutation in Neurospora crassa. Molecular and Cellular Biology 15, 5586-5597. 1995.

327.Singh, P. B., Miller, J. R., Pearce, J., Kothary, R., Burton, R. D., Paro, R., James, T. C., and Gaunt, S. J. A sequence motif found in a drosophila heterochromatin protein is conserved in animals and plants. Nucleic Acids Research 19, 789-794. 1991.

328.Singh, P. B. Molecular mechanisms of cellular determination – their relation to chromatin structure and parental imprinting. Journal of Cell Science 107, 2653-2668. 1994.

329.Singh, P. B. and Huskisson, N.S. Chromatin complexes as aperiodic microcrystalline arrays that regulate genome organisation and expression. Developmental Genetics 22, 85 – 99. 1998

330. Smit, A. F. The origin of interspersed repeats in the human genome. Current Opinion in Genetics and Development 6, 743-748. 1996.

331. Smith, C. J., Watson, C. F., Morris, P. C., Bird, C. R., Seymour, G. B., Gray, J. E., Arnold, C., Tucker, G. A., Schuch, W., and Harding, S. Inheritance and effect on ripening of antisense polygalacturonase genes in transgenic tomatoes. Plant Molecular Biology 14, 369-379. 1990.

332. Spiegel, R. Molecular genetics of triplet repeat expansions causing hereditary neurodegenerative disorders. Developmental Brain Dysfunction 10, 293-305. 1997.

333. Spoffford, J. Position effect variegation in Drosophila. 955-1018. 1976.

334. Stam, M., Viterbo, A., Mol, J. N., and Kooter, J. M. Position-dependent methylation and transcriptional silencing of transgenes in inverted T-DNA repeats: implications for posttranscriptional silencing of homologous host genes in plants. Molecular and Cellular Biology 18, 6165-6177. 1998.

335. Stankunas, K., Berger, J., Ruse, C., Sinclair, D. A. R., Randazzo, F., and Brock, H. W. The Enhancer of Polycomb gene of Drosophila encodes a chromatin. Development 125, 4055-4066. 1998.

336. Stuart, G. W., Vielkind, J. R., McMurray, J. V., and Westerfield, M. Stable lines of transgenic zebrafish exhibit reproducible patterns of transgene expression. Development 109, 577-584. 1990.

337. Suckow, M. A., Danneman, P., and Brayton, C. The Laboratory Mouse. 2001.

338. Sun, F. L. and Elgin, S. C. R. Putting boundaries on silence. Cell 99, 459-462. 1999.

339. Sun, Q. X., Ni, Z. F., and Liu, Z. Y. Differential gene expression between wheat hybrids and their parental inbreds in seedling leaves. Euphytica 106, 117-123. 1999.

340. Sutherland, H. G., Martin, D. I., and Whitelaw, E. A globin enhancer acts by increasing the proportion of erythrocytes expressing a linked transgene. Molecular and Cellular Biology 17, 1607-1614. 1997.

341. Sutherland, H. G. E., Kearns, M., Morgan, H. D., Headley, A. P., Morris, C., Martin, D. I. K., and Whitelaw, E. Reactivation of heritably silenced gene expression in mice. Mammalian Genome 11, 347-355. 2000.

342. Svensson, K., Mattsson, R., James, T. C., Wentzel, P., Pilartz, M., MacLaughlin, J., Miller, S. J., Olsson, T., Eriksson, U. J., and Ohlsson, R. The paternal allele of the H19 gene is progressively silenced during early mouse development: the acetylation status of histones may be involved in the generation of variegated expression patterns. Development 125, 61-69. 1998.

343. Talbert, P. B., Leciel, C. D. S., and Henikoff, S. Modification of the drosophila heterochromatic mutation brown (dominant) by linkage alterations. Genetics 136, 559-571. 1994.

344. Tam, P. P., Williams, E. A., and Tan, S. S. Expression of an X-linked HMG-lacZ transgene in mouse embryos: implication of chromosomal imprinting and lineage-specific X-chromosome activity. Developmental Genetics 15, 491-503. 1994.

345. Tartof, K. D., Hobbs, C., and Jones, M. A structural basis for variegating position effects. Cell 37, 869-878. 1984.

346. Tartof, K. D. and Henikoff, S. Trans-sensing effects from Drosophila to humans. Cell 65, 201-203. 1991.

347. Tartof, K. D. Position effect variegation in yeast. BioEssays 16, 713-714. 1994.

348. Tenllado, F. and DiazRuiz, J. R. Complete resistance to pepper mild mottle tobamovirus mediated by viral replicase sequences partially depends on transgene homozygosity and is based on a gene silencing mechanism. Transgenic Research 8, 83-93. 1999.

349. tenLohuis, M., Muller, A., Heidmann, I, Niedenhof, I, and Meyer, P. A repetitive DNA fragment carrying a hot spot for de novo DNA methylation enhances expression variegation in tobacco and petunia. Plant Journal 8, 919-932. 1995.

350. Thomas, G. H. and Elgin, S. C. Protein/DNA architecture of the DNase I hypersensitive region of the Drosophila hsp26 promoter. The European Molecular Biology Organization Journal 7, 2191-2201. 1988.

351. Thomas, K. R. and Capecchi, M. R. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51, 503-512. 1987.

352. Thomas, K. R., Deng, C., and Capecchi, M. R. High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. Molecular and Cellular Biology 12, 2919-2923. 1992.

353. Tilghman, S. M. The sins of the fathers and mothers: genomic imprinting in mammalian development. Cell 96, 185-193. 1999.

354. Toyota, M. and Issa, J. P. The role of DNA hypermethylation in human neoplasia. Electrophoresis 21, 329-333. 2000.

355. Ulker, B., Allen, G. C., Thompson, W. F., Spiker, S., and Weissinger, A. K. A tobacco matrix attachment region reduces the loss of transgene expression in the progeny of transgenic tobacco plants. Plant Journal 18, 253-263. 1999.

356. Urbanek, P., Fetka, I., Meisler, M. H., and Busslinger, M. Cooperation of Pax2 and Pax5 in midbrain and cerebellum development. Proceedings of the National Academy of Sciences U S A 94, 5703-5708. 1997.

357. Vain, P., Worland, B., Kohli, A., Snape, J. W., Christou, P., Allen, G. C., and Thompson, W. F. Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny Plant Journal 18, 233-242. 1999.

358. van der Vlag, J., den Blaauwen, J. L., Sewalt, R. G. A. B., van Driel, R., and Otte, A. P. Transcriptional repression mediated by polycomb group proteins and other chromatin-associated repressors is selectively blocked by insulators. Journal of Biological Chemistry 275, 697-704. 2000.

359. vanBlokland, R., van der Geest, N., Mol, J. N. M., and Kooter, J. M. Transgene-mediated suppression of chalcone synthase expression in petunia hybrida results from an increase in RNA turnover. Plant Journal 6, 861-877. 1994.

360. vanBlokland, R., tenLohuis, M., and Meyer, P. Condensation of chromatin in transcriptional regions of an inactivated plant transgene: evidence for an active role of transcription in gene silencing. Molecular and General Genetics 257, 1-13. 1997.

361. vanDeutekom, J. C. T., Bakker, E., Lemmers, R. J. L. F., vanderWielen, M. J. R., Bik, E., Hofker, M. H., Padberg, G. W., and Frants, R. R. Evidence for subtelomeric exchange of 3.3 kb tandemly

repeated units between chromosomes 4q35 and 10q26: Implications for genetic counselling and etiology of FSHD1. Human Molecular Genetics 5, 1997-2003. 1996.

362. VanHoudt, H., Kovarik, A., VanMontagu, M., and Depicker, A. Cross-talk between posttranscriptionally silenced neomycin phosphotransferase II transgenes. Federation of European Biochemical Societies Letters 467, 41-46. 2000.

363. Vaucheret, H. Identification of a general silencer for 19S and 35S promoters in transgenic tobacco plant – 90 bp of homology in the promoter sequence are sufficient for trans-inactivation. Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences 316, 1471-1483. 1993.

364. Vaucheret, H. Promoter-dependent trans-inactivation in transgenic tobacco plants – kinetic aspects of gene silencing and gene reactivation. Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences 317, 310-323. 1994.

365. Villeponteau, B. The heterochromatin loss model of aging. Experimental Gerontology 32, 383-394. 1997.

366. Voinnet, O. and Baulcombe, D. C. Systemic signalling in gene silencing. Nature 389, 553 1997.

367. Waddington, C. H. Endeavour 1 , 18. 1942.

368. Wakimoto, B. T. and Hearn, M. G. The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of Drosophila melanogaster. Genetics 125, 141-154. 1990.

369. Wallrath, L. L. and Elgin, S. C. R. Position-effect variegation in drosophila is associated with an altered chromatin structure. Genes & Development 9, 1263-1277. 1995.

370. Wallrath, L. L., Guntur, V. P., Rosman, L. E., and Elgin, S. C. R. DNA representation of variegating heterochromatic P-element inserts in diploid and polytene tissues of Drosophila melanogaster. Chromosoma 104, 519-527. 1996.

371. Wallrath, L. L. Drosophila telomeric transgenes provide insights on mechanisms of gene silencing. Genetica 109, 25-33. 2000.

372. Walters, M. C., Fiering, S., Eidemiller, J., Magis, W., Groudine, M., and Martin, D. I. Enhancers increase the probability but not the level of gene expression. Proceedings of the National Academy of Sciences U S A 92, 7125-7129. 1995.

373. Wang, G. Z., Ma, A., Chow, C. M., Horsley, D., Brown, N. R., Cowell, I. G., and Singh, P. B. Conservation of heterochromatin protein 1 function. Molecular and Cellular Biology 20, 6970-6983. 2000.

374. Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H. L. RNA-directed de novo methylation of genomic sequences in plants. Cell 76, 567-576. 1994.

375. Waterhouse, P. M., Smith, N. A., and Wang, M. B. Virus resistance and gene silencing: killing the messenger. Trends in Plant Science 4, 452-457. 1999.

376. Weiler, K. S. and Wakimoto, B. T. Heterochromatin and gene expression in Drosophila. Annual Review of Genetics 29, 577-605. 1995.

377. Whitelaw, C. B., Archibald, A. L., Harris, S., McClenaghan, M., Simons, J. P., and Clark, A. J. Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice. Transgenic Research 1, 3-13. 1991.

378. Whitelaw, C. B., Harris, S., McClenaghan, M., Simons, J. P., and Clark, A. J. Position-independent expression of the ovine beta-lactoglobulin gene in transgenic mice. Biochemical Journal 286, 31-39. 1992.

379. Wijgerde, M., Grosveld, F., and Fraser, P. Transcription complex stability and chromatin dynamics in vivo. Nature 377, 209-213. 1995.

380. Wilkie, G. S., Shermoen, A. W., OFarrell, P. H., and Davis, I. Transcribed genes are localized according to chromosomal position within polarized Drosophila embryonic nuclei. Current Biology 9, 1263-1266. 1999.

381. Willard, H. F. and Salz, H. K. Remodelling chromatin with RNA. Nature 386, 228-229. 3-20-1997.

382. Winkler, H. Die Konversion der Gene. 1930.

383. Wolffe, A. P. Transcription control: repressed repeats express themselves. Current Biology 7, R796-R798. 1997.

384. Wolffe, A. P. When more is less. Nature Genetics 18, 5-6. 1998.

385. Woloshuk, C. P., Yousibova, G. L., Rollins, J. A., Bhatnagar, D., and Payne, G. A. Molecular characterization of the AFL-1 locus in aspergillus flavus. Applied and Environmental Microbiology 61, 3019-3023. 1995.

386. Wright, G., Carver, A., Cottom, D., Reeves, D., Scott, A., Simons, P., Wilmut, I., Garner, I., and Colman, A. High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. Biotechnology (N Y) 9, 830-834. 1991.

387. Wu, C. T. and Morris, J. R. Transvection and other homology effects. Current Opinion in Genetics & Development 9, 237-246. 1999.

388. Wu, C. T. and Morris, J. R. Genes, genetics, and epigenetics: A correspondence. Science 293, 1103-1105. 2001.

389. Yoder, J. A., Walsh, C. P., and Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. Trends in Genetics 13, 335-340. 1997.

390. Young, L. E., Fernandes, K., McEvoy, T. G., Butterwith, S. C., Gutierrez, C. G., Carolan, C., Broadbent, P. J., Robinson, J. J., Wilmut, I, and Sinclair, K. D. Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nature Genetics 27, 153-154. 2001.

391. Yull, F., Harold, G., Wallace, R., Cowper, A., Percy, J., Cottingham, I., and Clark, A. J. Fixing human factor IX (fIX): correction of a cryptic RNA splice enables the production of biologically active fIX in the mammary gland of transgenic mice. Proceedings of the National Academy of Sciences U S A 92, 10899-10903. 1995.

392. Zentilin, L., Qin, G., Tafuro, S., Dinauer, M. C., Baum, C., and Giacca, M. Variegation of retroviral vector gene expression in myeloid cells. Gene Therapy 7, 153-166. 2000.

393. Zhu, X. X., Deng, C., Kuick, R., Yung, R., Lamb, B., Neel, J. V., Richardson, B., and Hanash, S. Analysis of human peripheral blood T cells and single-cell-derived T cell clones uncovers extensive

clonal CpG island methylation heterogeneity throughout the genome. Proceedings of The National Academy of Sciences USA 96, 8058-8063. 1999.

394. Zinovieva, N., Lassnig, C., Schams, D., Besenfelder, U., Wolf, E., Muller, S., Frenyo, L., Seregi, J., Muller, M., and Brem, G. Stable production of human insulin-like growth factor 1 (IGF-1) in the milk of hemi- and homozygous transgenic rabbits over several generations. Transgenic Research 7, 437-447. 1998.

395. Zolotukhin, S., Potter, M., Hauswirth, W. W., Guy, J., and Muzyczka, N. A "humanized" green fluorescent protein cDNA adapted for high-level expression in mammalian cells. Journal of Virology 70, 4646-4654. 1996.

Appendix

A.1) Raw data for inbred and outcrossed milk samples

Label	SD	BLG ug	Bcasein ug	Adi.BLG mg/ml	Label	SD	BLG ug	Bcasein ug	Adi.BLG mg/ml
C57 7I 1	0.02	1.11	1.44	6.24	C57 1	0.03	1.08	1.44	6.16
C57 7I 2	0.01	0.75	1.48	4.10	C57 2	0.02	0.76	1.43	4.16
C57 7I 3	0.02	1.08	1.5	5.83	C57 3	0.04	1.05	1.49	4.86
C57 7I 4	0.02	0.91	1.79	4.12	C57 4	0.03	0.94	1.76	5.43
C57 7I 5	0.01	0.65	1.4	3.76	C57 5	0.02	0.67	1.41	3.59
C57 7I 6	0.03	1.53	1.66	7.47	C57 6	0.05	1.39	1.52	7.45
C57 7I 7	0.02	0.88	1.51	4.72	C57 7	0.03	0.87	1.52	4.89
C57 7I 8	0.02	0.93	1.44	5.23	C57 8	0.03	0.92	1.45	5.36
C57 7I 9	0.01	0.98	1.35	5.88	C57 9	0.03	0.97	1.4	4.85
C4 Bcas	0.02	0	1.62		C4 Bcas	0.04	0	1.63	
C57 7I 1	0.03	1.15	1.47	6.26	C57 I7 10	0.03	0.94	1.69	4.48
C57 7I 2	0.02	0.75	1.45	4.14	C57 I7 11	0.04	1.47	1.9	6.23
C57 7I 3	0.03	1.07	1.45	5.90	C57 I7 12	0.03	1.12	1.51	5.97
C57 7I 4	0.03	0.96	1.82	4.22	C57 I7 13	0.08	2.45	2.39	8.25
C57 7I 5	0.02	0.71	1.42	4.00	C57 I7 14	0.01	0.23	1.63	1.14
C57 7I 6	0.05	1.53	1.66	7.37	C57 I7 15	0.04	1.75	2.24	6.29
C57 7I 7	0.03	0.88	1.5	4.69	C57 I7 16	0.03	0.88	1.4	5.06
C57 7I 8	0.03	0.93	1.41	5.28	C57 I7 17	0.04	1.27	1.71	5.98
C57 7I 9	0.02	0.87	1.31	5.31	C57 I7 18	0.03	1.22	1.48	6.64
C4 Bcas	0.06	0	1.6		C4 Bcas	0.04	0	1.61	
C57 7I 10	0.01	0.53	1.14	3.35	C57 7I 10	0.03	0.97	1.76	4.63
C57 7I 11	0.02	0.96	1.23	5.62	C57 7I 11	0.05	1.44	1.91	6.33
C57 7I 12	0.02	0.84	1.06	5.71	C57 7I 12	0.04	1.11	1.56	5.98
C57 7I 13	0.05	1.95	1.76	7.98	C57 7I 13	0.08	2.47	2.48	8.37
C57 7I 14	0.01	0.01	0.6	0.12	C57 7I 14	0.01	0.39	1.7	1.93
C57 7I 15	0.03	1.62	1.92	6.08	C57 7I 15	0.07	1.77	2.31	6.44
C57 7I 16	0.02	0.63	1.07	4.24	C57 7I 16	0.02	0.89	1.47	5.09
C57 7I 17	0.02	0.95	1.22	5.61	C57 7I 17	0.04	1.26	1.79	5.91
C57 7I 18	0.02	1.03	1.15	6.45	C57 7I 18	0.04	1.22	1.55	6.61
C4 Bcas	0.03	0	1.44		C4 Bcas	0.05	0	1.68	
C57 I7 19	0.06	1.67	1.38	9.80	CBA 14	0.03	1.04	1.53	5.47
C57 I7 19	0.05	1.66	1.35	9.96	CBA16	0.03	0.83	1.24	5.39
C57 I7 20	0.07	1.05	1.5	5.67	CBA17	0.02	0.8	1.32	4.88
C57 I7 20	0.04	0.97	1.43	5.49	C57 3	0.03	1.15	1.57	5.90
C57 I7 21	0.06	1.47	1.21	9.84	C57 4	0.03	0.91	1.79	4.09
C57 I7 21	0.06	1.45	1.2	9.79	C57 9	0.03	0.97	1.36	5.74
C57 I7 11	0.06	1.37	1.81	6.13	C57 9	0.03	0.98	1.34	5.89
c57 I7 13	0.06	2.25	2.17	8.40	C57 10	0.02	0.94	1.68	4.50
C57 I7 16	0.03	0.9	1.35	5.40	C57 19	0.04	1.66	1.37	9.75
CBA I7 9	0.03	0.82	1.25	5.31	C4 Bcas	0.06	0	1.61	
C4 Bcas	0.05	0	1.62						
					CBA 10	0.04	1.34	2.21	6.46
CBA 1	0.02	0.61	1.33	3.26	CBA 12	0.02	0.76	1.79	4.52
CBA 2	0.01	0.5	1.31	2.71	CBA 15	0.03	0.58	1.41	4.38
CBA 3	0.02	0.56	1.14	3.49	CBA 16	0.03	0.58	1.3	4.75
CBA 4	0.03	1.1	1.63	4.79	CBA 18	0.03	0.56	1.86	3.21
CBA 5	0.03	1.06	1.86	4.05	C57 19	0.05	1.63	2.02	8.59
CBA 6	0.02	0.9	1.85	3.45	C5720	0.03	0.98	1.94	5.38
CBA 8	0.01	0.47	1.35	2.47	C57 21	0.03	1.6	1.85	9.21
C4 Bcas	0.03	0	1.42		C4 Bcas	0.03	0	2.13	

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
CBA 1	0.03	0.73	1.82	3.25
CBA 2	0.02	0.52	1.67	2.52
CBA 7	0.05	1.44	1.77	6.59
CBA 10	0.05	1.45	1.81	6.49
CBA 11	0.02	0.56	1.67	2.72
CBA 13	0.03	0.79	1.64	3.90
CBA 13	0.03	0.79	1.66	3.85
CBA 14	0.03	0.94	1.42	5.36
C4 Bcas	0.06	0	1.62	
CBA 1	0.03	0.75	1.81	3.34
CBA 2	0.02	0.62	1.8	2.77
CBA 3	0.02	0.67	1.49	3.62
CBA 4	0.04	1.06	1.86	4.59
CBA 5	0.04	1.07	2.05	4.20
CBA 6	0.04	1.03	2.11	3.93
CBA 7	0.05	1.43	1.75	6.58
CBA 8	0.02	0.56	1.62	2.78
CBA 9	0.03	0.91	1.37	5.35
CBA 9	0.03	0.92	1.35	5.49
C4 Bcas	0.04	0	1.61	
CBA 10	0.04	1.44	1.81	6.17
CBA 11	0.02	0.66	1.8	2.84
CBA 12	0.02	1.16	1.62	5.55
CBA 13	0.02	0.91	1.77	3.98
CBA 14	0.02	0.93	1.43	5.04
CBA 15	0.02	0.99	1.54	4.98
CBA 16	0.02	0.72	1.06	5.26
CBA 17	0.01	0.68	1.14	4.62
CBA 18	0.01	0.73	1.72	3.29
CBA 18	0.01	0.76	1.73	3.40
C4 Bcas	0.03	0	1.55	
MP 1.3	0.02	1.25	1.35	10.42
MP 1.5	0.02	0.99	1.78	6.26
MP 1.6	0.02	1.23	1.39	9.96
MP 1.7	0.02	0.82	1.57	5.88
MP 1.11	0.03	1.2	1.65	8.18
MP 1.12	0.02	0.9	1.24	8.17
MP 1.14	0.02	0.88	1.46	6.78
MP 1.15	0.01	0.8	1.26	7.14
MP 3.2	0.02	0.98	1.07	10.30
C5 Bcas	0.04	0	1.8	0.00
MP 3.3	0.02	0.74	1.38	6.07
MP 5.3	0.03	1.04	1.59	7.40
MP 6.3	0.03	1.17	1.27	10.42
MP 6.4	0.03	1.11	1.45	8.66
MP 6.5	0.03	1.11	2	6.28
MP 6.6	0.02	0.85	0.89	10.80
MP 6.7	0.02	0.66	1.17	6.38
C5 Bcas	0.04	0	1.81	

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
CBA 3	0.01	0.59	1.58	3.40
CBA 4	0.02	1.08	2.04	4.82
CBA 5	0.02	1.06	2.27	4.25
CBA 6	0.02	0.99	2.31	3.90
CBA 7	0.02	1.44	1.95	6.72
CBA 8	0.01	0.56	1.83	2.78
C4 Bcas	0.03	0	1.82	
CBA 11	0.02	0.56	1.02	3.07
CBA 12	0.03	0.86	1.35	3.57
CBA 15	0.03	0.83	0.98	4.74
CBA 17	0.02	0.64	0.86	4.17
C4 Bcas	0.03	0	1.12	
MP 1.3	0.08	1.33	1.46	10.31
MP 1.5	0.04	0.81	1.5	6.11
MP 1.6	0.07	1.07	1.22	9.92
MP 1.7	0.05	0.78	1.54	5.73
MP 1.11	0.07	1.09	1.56	7.90
MP 1.12	0.07	1.02	1.32	8.74
MP 1.14	0.06	0.98	1.57	7.06
MP 1.15	0.05	0.94	1.5	7.09
MP 3.2	0.06	1.32	1.38	10.82
C5 Bcas	0.1	0	1.81	0.00
MP 1.3	0.03	1.34	1.43	10.13
MP 1.5	0.03	1.17	2.03	6.23
MP 1.6	0.04	1.81	1.96	9.99
MP 1.7	0.03	1.12	2.04	5.94
MP 1.11	0.05	1.91	2.62	7.88
MP 1.12	0.04	1.81	2.05	9.55
MP 1.14	0.03	1.67	2.31	7.82
MP 1.15	0.03	1.49	1.99	8.10
MP 3.2	0.04	1.62	1.66	11.04
C5 Bcas	0.04	0	1.73	0.00
MP 3.3	0.04	0.97	1.7	6.60
MP 5.3	0.03	0.88	1.37	7.43
MP 6.3	0.07	1.62	1.74	10.77
MP 6.4	0.07	1.33	1.7	9.05
MP 6.5	0.07	1.39	2.4	6.70
MP 6.6	0.05	1.06	1.09	11.24
MP 6.7	0.04	0.75	1.27	6.83
C5 Bcas	0.06	0	1.85	
MP 3.3	0.02	1.46	0.81	6.28
MP 5.3	0.02	1.47	0.92	7.08
MP 6.3	0.02	1.29	1.21	10.61
MP 6.4	0.02	1.4	1.03	8.32
MP 6.5	0.03	2.23	1.36	6.90
MP 6.6	0.02	1.34	1.29	10.89
MP 6.7	0.01	1.27	0.76	6.77
C5 Bcas	0.04	0	1.8	

A.2) Raw data for hemi- and homozygous milk samples

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
H1	0.04	1.82	1.27	14.33
H2	0.02	0.91	1.07	8.50
H3	0.03	1.1	1.34	8.21
H4	0.06	2.23	1.23	18.13
H5	0.04	1.35	2.1	6.43
H6	0.05	1.9	1.69	11.24
H7	0.02	0.79	1.51	5.23
H8	0.03	0.99	1.33	7.44
H9	0.02	0.97	1.2	8.08
C5 Bcas	0.05	0	1.6	

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
H1	0.03	1.83	1.29	14.19
H2	0.02	0.98	1.09	8.99
H3	0.02	1.11	1.35	8.22
H4	0.03	2.19	1.23	17.80
H5	0.02	1.32	2.09	6.32
H6	0.04	1.89	1.65	11.45
H7	0.01	0.79	1.53	5.16
H8	0.02	1.04	1.35	7.70
H9	0.02	0.93	1.16	8.02
H10	0.02	0.81	1.44	5.63
C5 Bcas	0.03	0	1.6	

H1	0.02	1.8	1.3	13.76
H2	0.01	0.93	1.03	8.97
H3	0.02	1.1	1.36	8.04
H4	0.03	2.04	1.06	19.13
H5	0.02	0.94	1.7	5.49
H6	0.03	1.68	1.46	11.43
H7	0.02	0.83	1.54	5.36
H8	0.02	1.09	1.36	7.96
H9	0.02	0.94	1.19	7.85
H10	0.01	0.78	1.42	5.46
C5 Bcas	0.02	0	1.59	

H11	0.03	0.72	0.95	7.67
H12	0.03	0.59	1.17	5.11
H13	0.04	0.97	1.57	6.26
H14	0.02	0.49	1.19	4.17
H15	0.04	1.19	1.1	10.95
H16	0.03	0.72	0.99	7.36
H17	0.04	0.95	1.1	8.74
H19	0.01	0.4	1.13	3.58
H20	0.03	0.93	1.26	7.47
C5 Bcas	0.05	0	1.62	

H10	0.03	0.83	1.48	5.68
H11	0.03	0.79	0.99	8.08
H12	0.02	0.69	1.2	5.82
H13	0.03	0.96	1.51	6.44
H14	0.02	0.54	1.25	4.37
H15	0.05	1.21	1.1	11.14
H16	0.03	0.79	1.01	7.92
H17	0.03	0.99	1.02	9.83
H19	0.01	0.51	1.12	4.61
H20	0.03	0.96	1.25	7.78
C5 Bcas	0.05	0	1.62	

H11	0.03	1.02	1.24	8.23
H12	0.02	0.87	1.47	5.92
H13	0.04	1.11	1.75	6.34
H14	0.01	0.63	1.33	4.74
H15	0.04	1.11	1.2	9.25
H16	0.02	0.96	1.27	7.56
H17	0.04	1.22	1.37	8.91
H19	0.01	0.55	1.26	4.37
H20	0.03	0.98	1.33	7.37
C5 Bcas	0.05	0	1.6	

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
HH1	0.05	2.41	1.4	17.21
HH2	0.05	3.1	2.18	14.22
HH3	0.04	1.47	1.49	9.87
HH4	0.05	2.43	2.52	9.64
HH5	0.06	2.72	2.78	9.78
HH6	0.05	2.61	2.68	9.74
HH7	0.04	1.47	1.92	7.66
HH8	0.04	1.91	1.63	11.72
HH9	0.02	0.52	0.49	10.61
C5 Bcas	0.04	0	1.6	
HH1	0.05	3.03	2.31	13.12
HH2	0.06	3.41	2.66	12.82
HH3	0.04	1.47	1.59	9.25
HH4	0.06	2.75	2.82	9.75
HH5	0.06	2.86	2.93	9.76
HH6	0.06	3.33	2.88	11.56
HH7	0.04	1.47	1.92	7.66
HH8	0.05	2.66	2.24	11.88
HH9	0.02	0.52	0.56	9.29
C5 Bcas	0.04	0	1.6	
HH11	0.06	2.28	1.19	22.51
HH12	0.04	1.28	1.17	12.85
HH13	0.06	1.98	1.49	15.61
HH14	0.05	1.47	1.85	9.34
HH15	0.06	1.95	1.98	11.57
HH16	0.06	2.07	2.53	9.61
HH17	0.08	3.29	1.96	19.72
HH18	0.03	1.02	1.05	11.41
HH19	0.06	1.65	1.25	15.51
HH20	0.06	1.97	1.69	13.70
C5 Bcas	0.05	0	1.88	
HH11	0.08	2.94	1.66	17.82
HH12	0.05	1.69	1.95	8.72
HH13	0.07	2.41	2.19	11.07
HH14	0.06	1.84	2.07	8.94
HH15	0.08	2.77	2.4	11.61
HH16	0.07	2.54	3.38	7.56
HH17	0.09	3.45	1.88	18.47
HH18	0.04	1.29	1.11	11.69
HH19	0.06	1.68	1.36	12.43
HH20	0.06	1.97	1.65	12.01
C5 Bcas	0.04	0	1.61	

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
HH1	2.02	1.36	0.44	17.22
HH2	2.56	1.7	0.67	14.33
HH3	1.21	1.33	0.62	7.32
HH4	1.85	1.95	0.4	17.34
HH5	1.86	1.93	0.41	17.01
HH6	2.24	1.77	0.41	20.49
HH7	1.44	1.9	0.52	10.38
HH8	1.84	1.43	0.64	10.78
HH9	0.6	0.65	0.26	8.65
HH11	2.74	1.45	0.65	15.81
C5 Bcas	0	1.58	0	
HH1	0.04	2.73	2.04	13.63
HH2	0.06	2.65	1.9	14.21
HH3	0.03	1.37	1.48	9.43
HH4	0.05	2.42	2.58	9.56
HH5	0.06	2.79	2.86	9.94
HH6	0.06	2.8	2.32	12.30
HH7	0.03	1.49	1.95	7.78
HH8	0.04	2.12	1.69	12.78
HH9	0.01	0.69	0.74	9.50
HH11	0.04	2.95	1.66	18.10
C5 Bcas	0.03	0	1.63	
HH 12	0.05	2.35	2.81	9.41
HH 13	0.05	2.73	2.72	11.29
HH 14	0.04	2.09	2.56	9.18
HH 15	0.05	2.91	2.76	11.86
HH 16	0.05	2.55	3.57	8.04
HH 17	0.07	3.6	2.12	19.10
HH 18	0.04	1.59	1.62	11.04
HH 19	0.03	1.69	1.52	12.51
HH 20	0.04	2	1.83	12.30
C5 Bcas	0.04	0	1.8	

A.3) Raw data for MOP milk samples

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
MOP 1.1	0.07	2.25	1.46	15.80
MOP 1.2	0.08	2.42	1.31	18.94
MOP 1.3	0.07	2	1.47	13.95
MOP 2.1	0.03	0.88	1.66	5.43
MOP 2.3	0.06	1.48	1.15	13.19
MOP 2.4	0.05	1.32	1.32	10.25
MOP 2.5	0.04	1.22	1.21	10.33
MOP 4.3	0.03	0.78	0.91	8.79
MOP 4.4	0.02	0.62	0.83	7.66
C5 Bcas	0.06	0	1.64	
MOP 1.1	0.09	2.26	1.43	15.90
MOP 1.2	0.09	2.4	1.28	18.87
MOP 1.3	0.08	1.7	1.15	14.88
MOP 2.1	0.06	0.88	1.66	5.33
MOP 2.2	0.04	0.73	1.37	5.36
MOP 2.3	0.08	1.58	1.3	12.23
MOP 2.4	0.07	1.34	1.29	10.45
MOP 2.5	0.05	1.16	1.25	9.34
MOP 3.1	0.04	0.92	1.24	7.47
C5 Bcas	0.07	0	1.61	
MOP 4.3	0.04	0.97	1.17	8.71
MOP 4.4	0.03	0.89	1.18	7.92
MOP 5.2	0.06	1.67	1.51	11.61
MOP 6.3	0.03	0.59	1.12	5.53
MOP 6.4	0.05	1.22	1.58	8.11
MOP 6.5	0.05	1.05	1.65	6.68
MOP 7.2	0.04	0.94	1.31	7.53
MOP 7.4	0.03	0.78	1.43	5.73
MOP 9.1	0.03	0.95	1.43	6.98
C5 Bcas	0.07	0	1.68	
MOP 5.2	0.07	1.94	1.76	11.37
MOP 6.3	0.03	1.02	1.55	6.79
MOP 6.4	0.05	1.6	1.94	8.51
MOP 6.5	0.05	1.35	1.91	7.29
MOP 7.2	0.05	1.43	1.77	8.33
MOP 7.4	0.04	1.18	1.86	6.54
MOP 9.1	0.05	1.33	1.76	7.79
MOP 10.1	0.03	1.01	1.56	6.68
MOP 10.2	0.02	0.84	1.49	5.81
C5 Bcas	0.05	0	1.65	

Label	SD	BLG ug	Bcasein ug	Adj.BLG mg/ml
MOP 1.1	0.04	2.01	1.13	16.68
MOP 1.2	0.05	2.15	0.95	21.22
MOP 1.3	0.04	1.74	1.12	14.56
MOP 2.1	0.02	0.87	1.55	5.26
MOP 2.2	0.02	0.76	1.3	5.48
MOP 2.3	0.04	1.46	1.06	12.91
MOP 2.4	0.03	1.42	1.32	10.09
MOP 2.5	0.02	1.03	0.94	10.27
MOP 3.1	0.02	0.75	1.29	5.45
C5 Bcas	0.04	0	1.5	
MOP 2.2	0.02	0.75	1.4	5.39
MOP 3.1	0.02	0.75	1.39	5.43
MOP 4.3	0.02	0.91	1.07	8.56
MOP 4.4	0.02	0.82	1.05	7.86
MOP 10.2	0.02	0.85	1.48	5.78
MOP 11.1	0.02	0.93	1.43	6.54
MOP 13.2	0.02	0.96	1.25	7.73
MOP 15.2	0.01	0.56	1.13	4.99
MOP 16.1	0.02	0.7	1.37	5.14
C5 Bcas	0.04	0	1.61	
MOP 5.2	0.04	1.42	1.24	11.59
MOP 6.3	0.02	0.48	0.99	4.91
MOP 6.4	0.04	1.18	1.5	7.97
MOP 6.5	0.03	0.79	1.36	5.88
MOP 7.2	0.03	0.7	1.04	6.81
MOP 7.4	0.03	0.74	1.41	5.31
MOP 9.1	0.03	0.71	1.15	6.25
MOP 10.1	0.02	0.61	1.17	5.28
MOP 10.2	0.02	0.54	1.17	4.67
C5 Bcas	0.06	0	1.62	
MOP 10.1	0.02	0.6	1.22	5.01
MOP 10.3	0.05	1.1	1.31	8.55
MOP 10.8	0.03	0.6	1.36	4.49
MOP 10.9	0.02	0.47	1.17	4.09
MOP 10.12	0.01	0.29	1.02	2.90
MOP 11.1	0.05	1.34	1.87	7.30
MOP 11.5	0.02	0.55	1.44	3.89
MOP 11.7	0.02	0.6	1.42	4.30
MOP 11.8	0.02	0.72	1.5	4.89
C5 Bcas	0.06	0	1.63	

Label	SD	BLG ug	Bcasein ug	Adj./Bcas mg/ml
MOP 10.3	0.05	1.38	1.6	8.89
MOP 10.8	0.03	0.77	1.58	5.03
MOP 10.9	0.03	0.7	1.45	4.98
MOP10.12	0.04	0.76	1.55	5.06
MOP 11.5	0.03	0.77	1.62	4.90
MOP 11.7	0.03	0.72	1.54	4.82
MOP 11.8	0.03	0.94	1.75	5.54
MOP 12.4	0.05	1.57	1.62	9.99
MOP 12.5	0.04	1.46	1.82	8.27
C5 Bcas	0.05	0	1.65	
MOP 12.4	0.05	1.43	1.42	10.01
MOP 12.5	0.05	1.5	1.83	8.15
MOP 13.2	0.05	1.36	1.68	8.04
MOP 13.7	0.05	1.44	1.36	10.52
MOP 13.8	0.03	0.61	1.4	4.33
MOP 13.7	0.04	1.45	1.38	10.44
MOP 13.8	0.03	0.63	1.37	4.57
MOP 15.2	0.02	0.5	1.09	4.56
MOP 16.1	0.03	0.76	1.44	5.24
C5 Bcas	0.05	0	1.59	

Label	SD	BLG ug	Bcasein ug	Adj./Bcas mg/ml
MOP 10.3	0.03	1.29	1.53	8.48
MOP 10.8	0.02	0.78	1.58	4.97
MOP 10.9	0.02	0.81	1.5	5.43
MOP10.12	0.02	0.91	1.64	5.58
MOP 11.1	0.03	1.22	1.79	6.86
MOP 11.5	0.02	0.61	1.51	4.06
MOP 11.7	0.02	0.88	1.68	5.27
MOP 11.8	0.02	0.9	1.69	5.36
MOP 12.4	0.03	1.53	1.53	10.06
C5 Bcas	0.03	0	1.61	
MOP 12.5	0.03	1.2	1.54	7.89
MOP 13.2	0.03	1.05	1.39	7.65
MOP 13.7	0.03	1.31	1.25	10.61
MOP 13.8	0.01	0.61	1.37	4.51
MOP 15.2	0.02	0.68	1.27	5.42
MOP 16.1	0.03	0.82	1.51	5.50
MOP 5.2	0.04	1.37	1.2	11.56
MOP 6.3	0.02	0.68	1.18	5.83
MOP 6.4	0.02	1.01	1.29	7.93
C5 Bcas	0.04	0	1.62	

B) Statistical Terms

Mean: the sum of all the values divided by the number of values

Variance: average of the squared distance of observations (i.e. sample values) from the mean

Standard Deviation (*s.d*): the square root of the variance

Coefficient of variation: (*s.d.*/mean) x 100

C) Mating regimes for BLG/7 line maintenance

BLG/7 was made by microinjection of the 16.2 Kb Sal I fragment of the ovine β-lactoglobulin genomic sequence into F1 (CBA: C57BL/6) fertilised eggs (Simons *et al* 1987).

G_0 BLG/7 founder was bred to F1 (CBA: C57BL/6).

F1 mice are generated by crossing a CBA with a C57BL/6, both from inbred colonies maintained by brother-sister matings.

C.1) Mating regimes for hemi- and homozygous BLG/7 populations

Hemizygotes:

BLG/7 hemi (F2) male x F1 (CBA: C57BL/6) non- transgenic female

BLG/7 hemi (F2)

Homozygotes:

a) BLG/ 7 hemi (F2) x BLG/ 7 hemi (F2)

b) BLG/7 homozygote confirmed by Southern Blotting

c) Maintained as closed homozygous colony

C.2) Mating regimes for backcrossed BLG/7 populations

CBA backcross:

a) BLG/7 hemi (F2) male x CBA purebred inbred female

b) Backcrossed male BLG/7 hemi progeny x CBA purebred inbred female

CBA backcross stopped at generation 16 due to poor breeding performance.

C57BL/6 backcross:

c) BLG/7 hemi (F2) male x C57BL/6 purebred inbred female

d) Backcrossed male BLG/7 hemi progeny x C57BL/6 purebred inbred female

C57BL/6 backcross currently at generation 20.