



Systems for the Analysis of Developmental Control Genes of the Mouse

A thesis submitted to the University of Adelaide
for the degree of Doctor of Philosophy

by

Linda Michelle Whyatt, B.Sc. (Hons.)

Department of Biochemistry
University of Adelaide
Adelaide, South Australia

March 1996

Table of Contents

Table of Contents	ii
Summary	ix
Statement	xii
Acknowledgments	xiii
1 Introduction	1
1.1 The Study of Development.....	1
1.2 The Mouse as a Model for Mammalian Development.....	1
1.3 Early Embryogenesis of the Mouse.....	2
1.3.1 Pre-implantation Development.....	2
1.3.2 Early Post-implantation Development.....	3
1.4 Genetic Control of Mammalian Embryogenesis.....	3
1.4.1 Co-ordination of Mammalian Embryogenesis.....	3
1.4.2 Control of Development at the Extracellular Level.....	4
1.4.3 Control of Development at the Intracellular Level.....	4
1.5 Analysis of Developmental Control Genes in vivo.....	5
1.5.1 Expression Mapping in vivo.....	6
1.5.2 Analysis of Loss or Gain of Function Mutations in vivo.....	6
1.6 A Model System for the Study of Early Mammalian Embryogenesis.....	7
1.6.1 Pluripotent Stem Cells of the Mouse.....	7
1.6.2 Embryonal Carcinoma (EC) Cells.....	8
1.6.3 Embryonic Stem (ES) Cells.....	9
1.6.4 Growth of ES Cells in vitro as Undifferentiated Monocultures.....	10
1.6.5 Differentiation of ES Cells in vitro.....	10
1.6.6 ES Cells as a Model for Pluripotent Stem Cells of the Mouse.....	11
1.6.7 ES Cells as a Route to the Germline.....	12
1.6.8 Gene Targeting in ES cells.....	12
1.7 Analysis of Developmental Control Genes using the ES cell system.....	13
1.7.1 In vitro Expression Mapping.....	13
1.7.2 Analysis of in vitro Gain of Function Mutations.....	14
1.7.3 Analysis of in vitro Loss of Function Mutations.....	15
1.7.4 In vivo Expression Mapping.....	15
1.7.5 Analysis of in vivo Gain of Function Mutations.....	15
1.7.6 Analysis of in vivo Loss of Function Mutations.....	16
1.8 Aims and Approaches.....	18

2	Inducible ES Cell Expression Vectors Suitable for the Expression of Developmental Control Genes.....	20
2.1	Introduction	20
2.2	IFN-inducible ES Cell Expression Vectors.....	24
2.3	Further Discussion	30
3	An Efficient Gene Targeting System for the Repeated Introduction of Precise, Subtle Alterations into the Mouse Genome	33
3.1	Gene Targeting Systems for the Production of Null Alleles.....	33
3.2	Targeting Systems for the Introduction of Independent Subtle Alterations into the Mouse Genome	35
3.2.1	The Need For Targeting Systems Which Permit the Introduction of Independent, Subtle Alterations into the Mouse Genome	35
3.2.2	Targeting Systems for the Introduction of Subtle Alterations into the Mouse Genome.....	37
3.2.3	Gene Targeting Systems Which Permit the Repeated Introduction of Precise, Subtle Alterations into the Mouse Genome.....	38
3.3	An Efficient Gene Targeting System for the Repeated Introduction of Non-Selectable Alterations Into the Mouse Genome	39
3.3.1	The Stable Tag-Exchange Gene Targeting System.....	39
3.3.2	The Selection Systems of the Stable Tag-Exchange Approach	40
3.4	Leukaemia Inhibitory Factor (LIF)	41
3.4.1	Identification and Characterisation of LIF.....	41
3.4.2	The LIF Gene.....	41
3.5	Functional Analysis of LIF.....	42
3.5.1	In vitro Activities of LIF.....	42
3.5.2	Effects of Gain of Function LIF Mutations in the Mouse	43
3.5.3	Effects of Loss of Function LIF Mutations in the Mouse	45
3.6	LIF Expression	46
3.6.1	Control of LIF Expression	46
3.6.2	Sites of LIF Expression	46
3.7	A Gene Targeting Approach to Detect the Cellular sites of LIF Expression	49
3.7.1	The Gene Targeting Approach.....	49
3.7.2	The LIF Reporter System	50
4	Analysis of Stable Tag Step Gene Targeting at the LIF Locus of ES Cells.....	53
4.1	Introduction	53
4.2	Construction of the Tag Step Targeting Vector	53
4.2.1	Vector Design.....	53

4.2.2	The Genomic Mouse LIF Clone.....	55
4.2.3	Construction of a LIF Targeting Vector Carrying the HPRT Gene	56
4.2.4	Construction of the First Round Targeting Vector Carrying the HPRT and <i>neo</i> genes	56
4.3	Gene Targeting at the LIF locus with the tag step vector pHNLX.....	58
4.3.1	Determination of Tag Step Gene Targeting Selection Conditions.....	58
4.3.2	Tag Step Targeting	58
4.3.3	Southern Analysis of Resistant Lines.....	59
4.4	Discussion.....	61
4.5	Recombination Models.....	63
4.5.1	Existing Recombination Models.....	63
4.5.2	Proposed Model for Homologous Recombination of Replacement Vectors.....	66
5	Exchange Step Gene Targeting at the LIF Locus	71
5.1	Introduction	71
5.2	PCR-based Alteration of the <i>lacZ</i> Reporter Gene.....	71
5.2.1	Alteration of Vector pnlacF.....	72
5.2.2	Reconstruction of the <i>lacZ</i> Constructs	72
5.2.3	Construction of the ATG ⁻ <i>lacZ</i> Vector	73
5.2.4	Construction of the ATG ⁺ <i>lacZ</i> Vector	73
5.3	Construction of COS-1 and ES Cell β -Galactosidase Expression Vectors	73
5.4	Analysis of LacZ Expression In COS-1, ES and X Cells	74
5.5	Alteration of the Genomic LIF Clone by PCR.....	76
5.5.1	Introduction of Unique Bgl II Sites into the First and Second Exons of LIF.....	76
5.5.2	Subcloning of the 3 kb Bcl I Fragment of LIF.....	76
5.5.3	Alteration of LIF exon 1D.....	77
5.5.4	Alteration of LIF exon 1M	77
5.5.5	Alteration of LIF exon 1T	77
5.5.6	Alteration of LIF Exon 2	78
5.5.7	Reconstruction of the LIF sequence in pLIFXT7t.....	78
5.6	The Exchange Step Targeting Vectors	78
5.6.1	Construction of the Exchange Step Vectors.....	78
5.7	Investigation of the Conditions Required for 6-TG Selection.....	79
5.7.1	Concentration of 6-TG Required for Cell Toxicity.....	79
5.8	Exchange Step Gene Targeting	79

5.8.1	6-TG Selection and Metabolic Co-operation	79
5.8.2	Transfection of Tag Targeted ES cell line HG60 with the Exchange Step Targeting Vectors	80
5.8.3	Cell Density Required to Prevent Metabolic Co-operation.....	82
5.8.4	Investigation of the Efficiency of Exchange Step Gene Targeting	83
5.9	Discussion	85
6	Analysis of ES cell Lines Resulting from Exchange Step Gene Targeting.....	92
6.1	Introduction	92
6.2	Investigation of the Mechanism of Tag Gene Loss.....	92
6.2.1	Karyotypic Analysis of 6-TG Resistant ES Cell Lines Resulting from Physical Loss of the Tag Gene.....	93
6.2.2	Southern Analysis of 6-TG Resistant ES Cell Lines Resulting from Physical Loss of the Tag Gene.....	94
6.3	Analysis of Reporter Gene Expression in Exchange Step Targeted Lines.....	95
6.3.1	RT-PCR Detection of LIF D Transcription in Exchange Targeted Lines carrying the Reporter Gene in the Second Exon of LIF	96
6.3.2	In vitro Expression of β -galactosidase Activity in Exchange Step Targeted Lines	96
6.5	Discussion.....	99
7	Assessment of Germline Competence of the E14TG2a ES Cell Line and Targeted Derivatives.....	103
7.1	Introduction	103
7.2	Assessment of Germline Competence of Unmanipulated E14TG2a ES Cell Line.....	104
7.2.1	Return of E14TG2a ES Cells to the Mouse	104
7.2.2	Investigation of Germline Competence of E14TG2a ES Cells.....	105
7.3	Assessment of Germline Competence of Tag Step Targeted ES Cells	106
7.3.1	Return of Tag Step Targeted ES Cells to the Mouse	106
7.3.2	Investigation of Germline Competence of Tag Step Targeted ES Cell Lines.....	107
7.4	Assessment of Germline Competence of Exchange Step Targeted ES cells	108
7.4.1	Return of Exchange Step Targeted ES Cell Lines Carrying pE2LacLIF to the Mouse	108
7.4.2	Investigation of Germline Competence of Exchange Step Targeted ES Cell Lines Carrying pE2LacLIF.....	108
7.5	Discussion.....	108
8	General Discussion.....	112
8.1	112	

8.2	Future Work.....	117
8.2.1	6-16-Based Expression Vectors	117
8.2.2	Stable Tag-Exchange Gene Targeting.....	119
9	Materials and Methods.....	123
9.1	Abbreviations	123
9.2	Materials	125
9.2.1	Chemicals and Reagents.....	125
9.2.2	Enzymes.....	126
9.2.3	Nucleic Acid Molecular Weight Standards	127
9.2.4	Radionucleotides	127
9.2.5	Kits.....	127
9.2.6	Miscellaneous Materials	127
9.2.7	Plasmid Vectors	128
9.2.8	Oligonucleotides.....	129
9.2.9	Water and Sterilisation.....	130
9.2.10	Bacterial Strains	130
9.2.11	Bacterial Growth Media.....	131
9.3	Recombinant DNA Techniques	132
9.3.1	Restriction Endonuclease Digestions.....	132
9.3.2	Endfill Reaction.....	132
9.3.3	Endchew Reaction.....	133
9.3.4	CIP Reaction.....	133
9.3.5	Agarose Gel Electrophoresis.....	133
9.3.6	Isolation of DNA from Agarose Gels.....	134
9.3.7	Ligation Reactions.....	134
9.3.8	Preparation of Competent Bacterial Cells.....	135
9.3.9	Transformation of Competent Bacterial Cells.....	135
9.3.10	Bacterial Colony Lift Procedure	136
9.3.11	[alpha- ³² P]-labelled DNA Probes.....	136
9.3.12	Small Scale Isolation of Plasmid DNA	137
9.3.13	Polymerase Chain Reaction Technique	137
9.3.14	Non-denaturing Polyacrylamide Gel Electrophoresis.....	138
9.3.15	Denaturing Polyacrylamide Gel Electrophoresis.....	138
9.3.16	Isolation of DNA Fragments from Polyacrylamide Gels	139
9.3.17	Denaturation of Double Stranded Plasmid DNA for Sequencing....	139
9.3.18	Sequencing of Double Stranded Plasmid DNA.....	139

9.3.19	Large Scale Isolation of Pure Plasmid DNA	139
9.3.20	Small Scale Preparation of High Molecular Weight DNA from Mammalian Cells	140
9.3.21	Southern Blot Analysis	141
9.3.22	Hybridisation of Radioactive Probes to Hybond-N ⁺ Membranes ...	141
9.3.23	Hybridisation of Radioactive Probes to Nitrocellulose Membranes.....	142
9.3.24	Cytoplasmic RNA Preparation.....	142
9.3.25	cDNA Synthesis.....	143
9.4	Tissue Culture Techniques	144
9.4.1	Cell Lines.....	144
9.4.2	Solutions	144
9.4.3	Media	145
9.4.4	ES and COS-1 cell Maintenance.....	146
9.4.5	Transient Transfection of ES cells by Electroporation	146
9.4.6	Transient Transfection of COS-1 Cells by Electroporation.....	147
9.4.7	Stable Transfection of ES Cells by Calcium Phosphate Co-precipitation	148
9.4.8	Screening of Stable ES Cell Lines for CAT Activity.....	148
9.4.9	Time Course of IFN Induced CAT Expression.....	149
9.4.10	Preparation of Crude Cell Extract for the CAT Assay	149
9.4.11	The Chloramphenicol Acetyl-transferase Assay	149
9.4.12	Stable Transfection of ES Cells with Targeting Vector DNA by Electroporation.....	150
9.4.13	Histochemical Staining for β -Galactosidase Activity	151
9.4.14	Differentiation of ES cells to X Cells	152
9.4.15	MBA Differentiation of ES Cells	152
9.4.16	Hoechst Staining of ES cells.....	152
9.4.17	Karyotypic Analysis of ES Cell Lines	152
9.5	Mouse Manipulations and Analysis	153
9.5.1	Chimaerae Production	153
9.5.2	GPI Analysis of Sperm.....	153
9.5.3	Breeding.....	153
9.6	Phosphorimager Analysis and Autoradiograph Scanning	153
9.7	Containment Facilities.....	154
9.7.1	Recombinant DNA Manipulations.....	154
9.7.2	Animal Manipulations.....	154

References.....	155

Summary

The in vitro culture of pluripotential embryonic stem (ES) cells provides a system for the identification and analysis of genes which regulate stem cell self-renewal and differentiation. Since ES cells also provide a vehicle for alteration of the mouse germline, developmental control genes can further be analysed in vivo, in the context of the mouse. The overall aim of work reported in this thesis was to develop systems suitable for use in ES cells, both in vitro and in vivo, for the analysis of genes which regulate embryonic developmental decisions.

The first aim of work presented in this thesis was to characterise further an inducible expression system designed for use in ES cells in vitro. Expression vectors (Whyatt, 1991), based on the interferon- (IFN) inducible promoter of the human 6-16 gene, were designed to permit the controlled expression of genes in ES cells, which regulate stem cell self-renewal, differentiation and death. In this project the efficacy of the vectors was assessed in stable ES cell lines. Treatment of ES cells with IFN did not compromise ES cell pluripotency or germline competence and stable ES cell lines were responsive to IFN treatment. The vectors were able to direct dose dependent, IFN-inducible, heterologous gene expression in the absence of detectable basal expression within the context of the mouse genome. The ability of the vectors to direct precisely controllable levels of transgene expression in the absence of basal expression, makes them ideal for the analysis of genes such as developmental control genes which are likely to be biologically active at very low thresholds and may exert alternate biological effects at different concentrations.

Predetermined genetic alterations can be introduced into the mouse genome and germline by the technique of gene targeting in ES cells (Thomas and Capecchi, 1987). The second aim of work presented in this thesis was to develop a gene targeting system for the efficient, repeated introduction of precise, subtle alterations into the genome of ES cells for the analysis of developmental control genes in vivo. The stable tag-exchange gene targeting approach, developed in this work, is based on the two step targeting approach, known as the tag-and-exchange approach, first described by Reid and co-workers (1990). In the tag-and-exchange

approach the locus to be altered is tagged in an initial round of gene targeting (tag step gene targeting) with a positive/negative selection cassette. Replacement of the tag cassette with the alteration of interest in a subsequent round of gene targeting (exchange step gene targeting) provides direct selection for homologous recombination of the exchange step vector. This theoretically improves the efficiency of exchange step gene targeting by the elimination of non-homologous recombinants and permits the introduction of subtle alterations into the tagged site. However, the efficiency of the tag-and-exchange approach is compromised by a high background of lines which survive exchange step gene targeting due to physical loss of the tag gene by means other than homologous recombination (Askew *et al.*, 1993; Stacey *et al.*, 1994; Wu, 1994). The stable tag-exchange approach was designed to overcome this background by the inclusion of an additional selection cassette on the tag step vector. This positive selection cassette, termed the stability cassette, remains on the tagged allele during exchange step gene targeting, providing selection for the presence of the tagged allele.

The stable tag-exchange system was analysed at the Leukaemia Inhibitory Factor (LIF) locus in E14TG2a (HPRT⁻) ES cells. The LIF gene product has been implicated in diverse developmental processes including haematopoiesis, osteogenesis, tumorigenesis, wound healing and embryonic development. The replacement vector used in tag step gene targeting carried the HPRT and *neo* genes as the tag and stability cassettes respectively, separated by almost five kilobases of homologous DNA. The tag step vector was introduced into the LIF locus with an efficiency of almost seventeen percent. All homologous recombinants contained both the HPRT and *neo* genes correctly integrated at the LIF locus, regardless of whether selection for HPRT or HPRT and *neo* was applied. This finding implies that multiple, separate regions of non-homology, for which selection can not be applied, may be introduced efficiently into the mouse genome on a single replacement vector.

Direct comparison of the unmodified tag-and-exchange approach with that of the stable tag-exchange approach was carried out at the LIF locus. The efficiency of the exchange step of the stable tag-exchange approach was almost ninety percent at the LIF locus compared to only twenty one percent when the tag-and-exchange approach was used. The improved efficiency

of the stable tag-exchange approach makes it suitable for the repeated introduction of alterations into the mouse genome. Selection for the stability cassette during exchange step gene targeting improved the gene targeting efficiency through elimination of lines which had undergone physical loss of the tag gene by means other than homologous recombination. The mechanism responsible for this loss was gene conversion, in which both the tag and stability cassettes were accurately removed from the tagged LIF locus.

Three different isoforms of the LIF gene are known to exist (Rathjen *et al.*, 1990a; B. Haines, P. Rathjen personal communication) which are generated from transcripts carrying alternate first exons. Evidence suggests that the isoforms of LIF may play biologically distinct roles. In the exchange step of the stable tag-exchange approach, a β -galactosidase reporter gene was introduced into the unique exons of LIF or a common exon of LIF, to allow visualisation of the cellular sites of expression of each and all LIF isoforms in mice generated from these cells. Reporter gene expression in the exchange step targeted ES cell lines reflected the known in vitro pattern of LIF expression. In addition, these targeted lines retained germline competence. Identification of the cellular sites of expression of each LIF isoform in the mouse, should lead to a detailed understanding of the normal roles of each of these proteins during developmental processes.

The systems presented in this thesis provide valuable tools with which to exploit the ES cell system for the analysis of genes which regulate developmental decisions both in early and later embryogenesis.

Statement

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED

DATE.....7/3/96.....

Acknowledgments

I thank Professor G. E. Rogers for allowing me to undertake a Ph.D. in the Department of Biochemistry.

I thank Dr. Peter Rathjen for allowing me to undertake a Ph.D. in his lab, for the effort he has put into to my postgraduate education, for providing an excellent environment in which to work, and for his infectious enthusiasm for science which I hope he never loses (and the suit which I hope he does). Thanks also to everyone in the Rathjen lab. Thanks Joy for tissue culture advice and for taking the slogan 'There's no other store....' more seriously than anyone I know, Paul (Paully) for support, friendship and advice particularly in the early years, Julie (Jules) for friendship, Why-Lake counselling and always listening, Bryan (Mad Dog) for quick witted remarks and ability to make up silly names for almost everything, Roger for using words like chagrin and poignant without a hint of embarrassment, Trish for friendship and late night chats on lab retreats (it's almost time I sold a centrifuge, I guess), Jackie for friendship and re-educating me about Americans. To the rest of the lab Jenny, Tom, Mike, Gav, Shiwani thanks also for making the Rathjen lab a cool place in which to work. Thanks to all the rest in the Dept. who have made my life fun over the last few years especially Tim (Timmy) and Rebecca (Reeba) and Kathy and Brett.

Thanks to all who have contributed to work in this thesis. To Dr. Austin Smith whose advice on vector design and gene targeting were invaluable, Dr. Graham Webb for karyotypic analysis by Q banding, Chris Giannakis for assistance with GPI analysis and Jenny Washington for help with getting "samples" from mice. To Jacquie Bealle, Jenny Bennett, Rachel Dempster, Jenny McLean (for the litres, and litres of DMEM) for tissue culture support during my Ph.D. and especially Joe Wrin for always providing help when needed, keeping my mouse colony thriving while I've been writing, and most recently for blastocyst injections. I also owe a great debt to Bresatec and in particular Dr. Ian Lyons, Dr. Mark Nottle, Anita Peura and Steve McIlpatrick for getting my ES cells back into mice and therefore making this whole project possible (see over). Thanks also to Jackie Brinkman, Sheila Jeffries and Judy Whittle for all the litres of FTB and for keeping everything stocked.

Thankyou to Mum and Dad for all your love and support over the years and providing an excellent place in which to live (with Motel-style accommodation especially in the past two months). Without it none of this would have been possible. To Karen, Peter, Jonathan, Darryl and Debra thanks for being family that I would almost certainly pick as friends. To Darryl a special thankyou for formatting this thesis. Thankyou Keith for being clearly, my best friend, making life outside the lab great, for critical reading of this thesis, for always being there, for love and support and for being both indefectible and always gelogenic.



Chapter 1



1 Introduction

1.1 The Study of Development

Development is the process whereby the highly specialised and complex tissues of the mature organism form from a single cell, the fertilised egg. The study of mammalian development has grown from classical beginnings over a century ago into a field in which the application of molecular biological techniques is leading to mechanistic understanding. There are three primary reasons for the study of mammalian development. Firstly, the process of mammalian development is itself intrinsically interesting and different in important respects from the development of other simpler model organisms such as *Caenorhabditis elegans* and *Drosophila melanogaster*. Secondly, a general understanding of mammalian development provides insights into the development of all mammals including the human, and developmentally related pathological conditions. Finally, a detailed understanding of the processes which control mammalian embryogenesis, coupled with molecular techniques, should lead to agricultural and medical advances such as the establishment of animal models for human disease, the generation of genetically altered livestock with improved production characteristics and the development of novel human therapies.

1.2 The Mouse as a Model for Mammalian Development

Extensive genetic analysis of the invertebrates *C. elegans* and *D. melanogaster*, and the vertebrate *Xenopus laevis*, has provided insight into themes underlying development. Parallels between the development of these organisms and mammals can be found and many important mammalian developmental control genes were initially identified in *D. melanogaster*. Despite this, fundamental differences exist between the development of these eukaryotes and that of mammals. For example, the earliest developmental decisions in the mammal are controlled by autonomously produced determinants (Pratt *et al.*, 1983; Flach *et al.*, 1982) rather than by maternally deposited determinants as is the case in some lower eukaryotes (Nusslein-Vollhard *et al.*, 1987; Struhl, 1989). Furthermore, in the case of both *C. elegans* and *D. melanogaster*, relatively rigid cell determination exists even in early development and the fate of a single cell or group of cells can be assigned from very early

stages of development (Rossant, 1988). This does not appear to be the case for mammals in which there is a large amount of cell mixing and developmental flexibility throughout development (Rossant, 1988). These fundamental differences mean that mammalian embryonic development is best studied in a mammalian organism.

The mouse is the model of choice for mammalian development for a number of reasons. Classical embryological study of the mouse has provided a detailed understanding of mouse development at the macroscopic level. Electron microscopy has been used to examine the cellular events that occur during mouse embryogenesis. Extensive genetic research over the past century has provided a more detailed genetic map of the mouse genome than exists for any other experimentally tractable mammalian species. The reduced technical and ethical problems associated with manipulation of the mouse, the ability to culture and genetically manipulate pluripotent stem cells of the mouse embryo, ease of experimental manipulation and relatively short life cycle make it the ideal model for mammalian development.

1.3 Early Embryogenesis of the Mouse

The focus of research in the laboratory in which this work was undertaken is early development of the mouse. A major area of research in the laboratory is the identification and analysis of genes which control early development.

1.3.1 Pre-implantation Development

The fertilised egg divides approximately every 20 hours until the eight cell stage. The four and eight cell stage embryos can be seen in figure 1.1 a and b. Until this stage the cells of the embryo, known as blastomeres (Fig 1.1 a, b), are totipotent, each able to give rise to the entire mouse (Tarkowski, 1959; Hogan *et al.* 1994). Formation of the morula at 2.5 days post coitum (d.p.c.) (Fig. 1.1 f, g) occurs as a result of compaction of the blastomeres (Fig 1.1 c - e). The cells gain polarity at this stage and this may underlie the first partitioning of cellular fate in the embryo (Johnson and Maro, 1988).

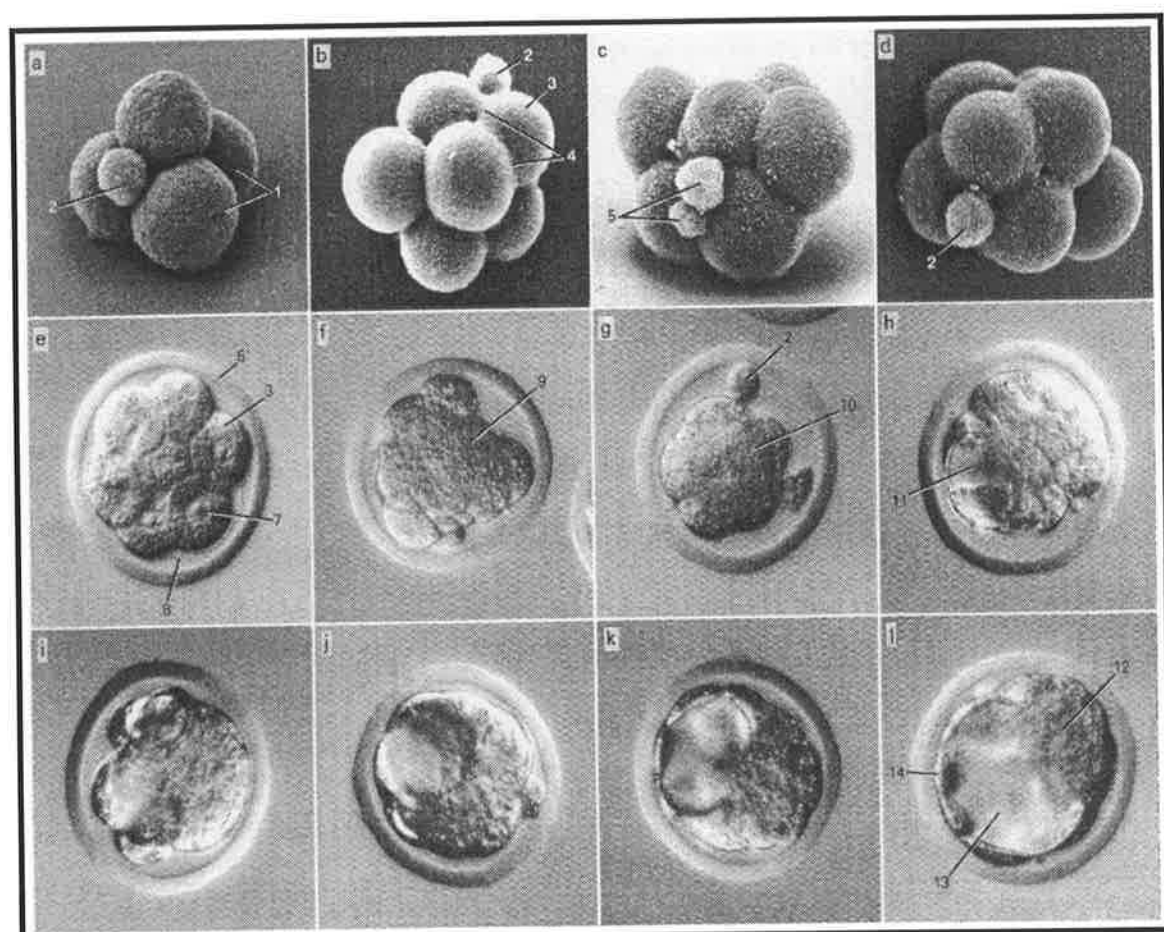
Figure 1.1 Embryonic development during the preimplantation stage

a-d Scanning electron micrographs of 4-and 8- cell embryos. (a) 4-cell stage embryo, (b) 8-cell stage embryo showing the onset of compaction, (c, d) 8-cell stage embryos at early stages of compaction.

e-l Embryos progressing from the 8-cell stage to the blastocyst stage viewed by interference contrast optics.

1. blastomeres of 4-cell stage embryo
2. second polar body
3. blastomere of the 8-cell stage embryo
4. early evidence of compaction in which the blastocyst boundaries become indistinct
5. the two division products of the second polar body after the second meiotic division
6. zona pellucida
7. nucleus of the 8-cell stage embryo
8. perivitelline space
9. early morula stage embryo
10. fully compacted morula stage embryo
11. onset of cavitation
12. location of ICM
13. blastocoelic cavity
14. mural trophectoderm cell

Taken from Kaufman, (1992).



As cleavage continues to and beyond the 16 cell stage, the first differentiation event leads to formation of two distinct cell lineages. The outer blastomeres of the morula give rise to the trophoctoderm, while those positioned internally give rise to the inner cell mass (ICM) (Pederson, 1988; Fig. 1.1 h - l). The blastocyst, formed as a result of this differentiation event and cavitation at 3.5 d.p.c., consists of an epithelial layer of trophoctoderm surrounding a fluid filled space known as the blastocoel and cells of the ICM (Fig. 1.1 l). Cells of the trophoctoderm lineage give rise to extraembryonic tissues required for support of the embryo while cells of the ICM, comprising approximately one third of the 64 cell blastocyst, give rise to the embryo proper and extraembryonic tissues. At 4.5 d.p.c. the blastocyst hatches from the zona pelucida in preparation for implantation.

1.3.2 Early Post-implantation Development

Implantation of the blastocyst into the uterine wall occurs at 4.5 d.p.c. At this time, a second differentiation event occurs, in which the developmental potential of cells of the ICM becomes restricted. Cells on the blastocoel side of the ICM give rise to the primitive endoderm, while the remaining cells of the ICM give rise to the primitive ectoderm. Cells of the primitive endoderm give rise to extraembryonic tissues (Gardner, 1982; 1983), whereas cells of the primitive ectoderm give rise to the three germ layers of the mouse, the germ cells and some extraembryonic components (Gardner and Rossant, 1979).

1.4 Genetic Control of Mammalian Embryogenesis

1.4.1 Co-ordination of Mammalian Embryogenesis

Cell lineage specification, segmentation and regional specification along the body axis are the primary events during early mammalian development (reviewed, Rossant and Joyner, 1989). The gene products which co-ordinate these events have long been sought by the developmental biologist. Application of molecular techniques to this field of research has enabled the identification and analysis of such products. Although fundamental differences exist between the development of mammals and simpler eukaryotic organisms (section 1.2), there is evolutionary conservation between classes of developmental control genes among multicellular eukaryotes.

1.4.2 Control of Development at the Extracellular Level

The fate of a cell within the developing embryo is largely determined by the array of extracellular signalling molecules with which the cell comes into contact (Heath and Smith, 1988). These factors are often polypeptide or chemical growth and differentiation factors.

Evidence from three main sources has implicated the action of these factors in co-ordination of mammalian development. In vivo detection of the sites of expression of these factors, by techniques such as in situ hybridisation, has shown them to be expressed during embryonic development and at sites consistent with them playing co-ordinating roles throughout embryonic and later development (Jackson, 1989). Secondly, the application of exogenously supplied growth and differentiation factors can alter the developmental fate of cells both in vitro and in vivo. For example, both FGF and activin can induce mesoderm formation in the animal cap of *X. laevis* (Green *et al.*, 1990; Sokol *et al.*, 1990) while application of the signalling factor *Sonic hedgehog* or retinoic acid (RA) to the anterior limb bud of the chick leads to mirror image digit duplication (Summerbell, 1983; Riddle *et al.* 1993). Finally, there is accumulating genetic evidence from *D. melanogaster*, *X. laevis* and the mouse which implicates growth and differentiation factors in the co-ordination of development. In *D. melanogaster*, a range of growth factor and growth factor receptor mutants have been analysed. Pattern formation is perturbed by the absence of growth factors such as *wingless* and *hedgehog* (Nusslein-Volhard and Wieschaus, 1980). In *X. laevis*, expression of a dominant inhibitory fibroblast growth factor (FGF) receptor is able to completely block the formation of mesoderm (Amaya *et al.*, 1991). Mice carrying lack of function mutations in signalling molecules also show developmental abnormalities. Examples of this type include null mutations of FGF-4 and Wnt-1 which lead to embryonic lethality (Feldman *et al.*, 1995; Mc Mahon and Bradley, 1990).

1.4.3 Control of Development at the Intracellular Level

Extracellular signalling factors which co-ordinate mammalian development mediate their effects via binding to receptor molecules expressed on the surface of the responding cell. Coupling of ligand and receptor leads to a complex cascade of events which transduce the

signal from the cell surface through the cytoplasm and finally to the nucleus of the responding cell (reviewed in Karin, 1992). Signal transduction molecules activated by coupling of ligand and receptor transmit extracellular signals through the cytoplasm to the nucleus. The genes encoding the signal transduction molecules are the first class of genes involved in the control of development at the intracellular level. Genetic and biochemical evidence has demonstrated direct roles for signal transduction molecules in the transmission of developmentally important extracellular signals to the nucleus of the responding cell. For example, the Src-related tyrosine kinase, Hck has been shown to mediate signal transduction via the signal transducing subunit gp130 of the LIF receptor in ES cells (Ernst *et al.*, 1994). Targeted ES cells which express constitutive Hck activity require approximately fifteen fold less LIF to maintain them in the undifferentiated state than untargeted ES cells.

Transduction of the signal to the nucleus leads to alteration in the array of genes expressed by the responding cell, which ultimately determines cell fate. The effector molecules in the nucleus include DNA binding proteins such as transcriptional regulators (reviewed in Karin 1992). The genes encoding this class of molecules is the final class of genes involved in the control of development at the intracellular level. Genetic evidence has demonstrated an unequivocal role for the transcriptional regulators in control of development at the intracellular level. Prominent among the transcriptional regulators involved in development are the homeobox genes (Scott *et al.*, 1989). Mice generated from targeted ES cells lacking expression of Hox-3.1 show skeletal abnormalities which include transformation of several skeletal segments into more anterior ones (Le Mouellic *et al.*, 1992)

1.5 Analysis of Developmental Control Genes in vivo

A range of techniques are used to identify potential developmental control genes in the mouse (reviewed in Jackson, 1989). Due to the indirect nature of these identification techniques, it is necessary to demonstrate unequivocal developmental roles for the isolated genes. In general, two main approaches have been used. Firstly, mapping the spatial and temporal sites of expression of a gene in vitro and in vivo, by techniques such as RNase protection, RT-PCR and in situ hybridisation, can provide circumstantial evidence that a gene is involved in

developmental processes. Secondly, analysis of gain or loss of function mutants can provide direct evidence that the gene has roles in development.

1.5.1 *Expression Mapping in vivo*

Mapping the spatial and temporal expression pattern of putative developmental control genes in the mouse can provide valuable information about the normal roles of such gene products. This sort of approach requires sensitive detection methods which are able to resolve expression at the cellular level. Mapping approaches are widely used in conjunction with genetic approaches for the analysis of developmental control genes. For example the expression of the signalling molecule Wnt-1 in the developing central nervous system (CNS) of the mouse (Shackleford and Varmus, 1987; Wilkinson *et al.*, 1987) was suggestive of a role for this protein in normal development of the CNS and this has been demonstrated by genetic means (McMahon and Bradley, 1990).

In the case of genes which are expressed at very low levels *in vivo*, conventional mapping techniques may be insufficiently sensitive to detect all sites of expression (Robertson *et al.*, 1993). This is a particular problem for the developmental control genes which can have very low biological thresholds and consequently be expressed at very low levels. A further limitation to expression mapping is encountered in the case of developmental control genes which act at very early stages of development. In particular, the small size and fragility of the mouse embryo prevent use of standard *in situ* hybridisation techniques.

1.5.2 *Analysis of Loss or Gain of Function Mutations in vivo*

Gain of function mutations in the mouse have been generated by a number of approaches. These include exogenous administration of the gene product or over-/inappropriate expression of the gene *in vivo*. For example, the equivalent of gain of function mutations of the LIF gene have been created by administration of LIF by intraperitoneal injection (Metcalf *et al.*, 1990) or by overexpression in a sub-population of T lymphocytes (Shen *et al.*, 1994). Potential roles of LIF in bone formation and haematopoiesis were inferred from these studies.

A limitation of these approaches is that the particular phenotypes seen are primarily determined by the route of administration.

Insertional transgenesis, by pronuclei injection (Gordon *et al.*, 1980) or retroviral infection of preimplantation embryos (van der Putten *et al.*, 1985), has been used to generate gain of function mutations in mice. Expression of the transgene in all tissues of the body allows the consequences of the mutation to be studied in the context of the whole organism. Limitations to this approach are that multiple copies of the transgene randomly integrate into the genome. This can lead to insertional mutagenesis and renders the introduced sequences susceptible to copy number and position effects (Lacy *et al.*, 1983).

A particular limitation of the gain of function approach is that inappropriate application to, or expression of a gene, in cells which do not normally come into contact with the product may lead to a response which is not relevant to normal development and which complicates analysis. These limitations would be overcome by a system which permits direct modification of endogenous loci.

The loss of function mutational approach is preferable to the study of developmental control genes *in vivo* since it can provide direct evidence that a gene has a controlling influence on development. In contrast to other species in which techniques such as expression of antisense transcripts can be used to generate loss of function mutations *in vivo*, no reliable means of directly producing loss of function mutations was possible in the mouse prior to the establishment of ES cell technology.

1.6 A Model System for the Study of Early Mammalian Embryogenesis

1.6.1 *Pluripotential Stem Cells of the Mouse*

Although mouse embryogenesis is a highly ordered process, there is a high degree of plasticity during early mouse development. For example, over 80% of cells in the gastrulating egg cylinder can be ablated before organogenesis with no effect on embryogenesis (Snow

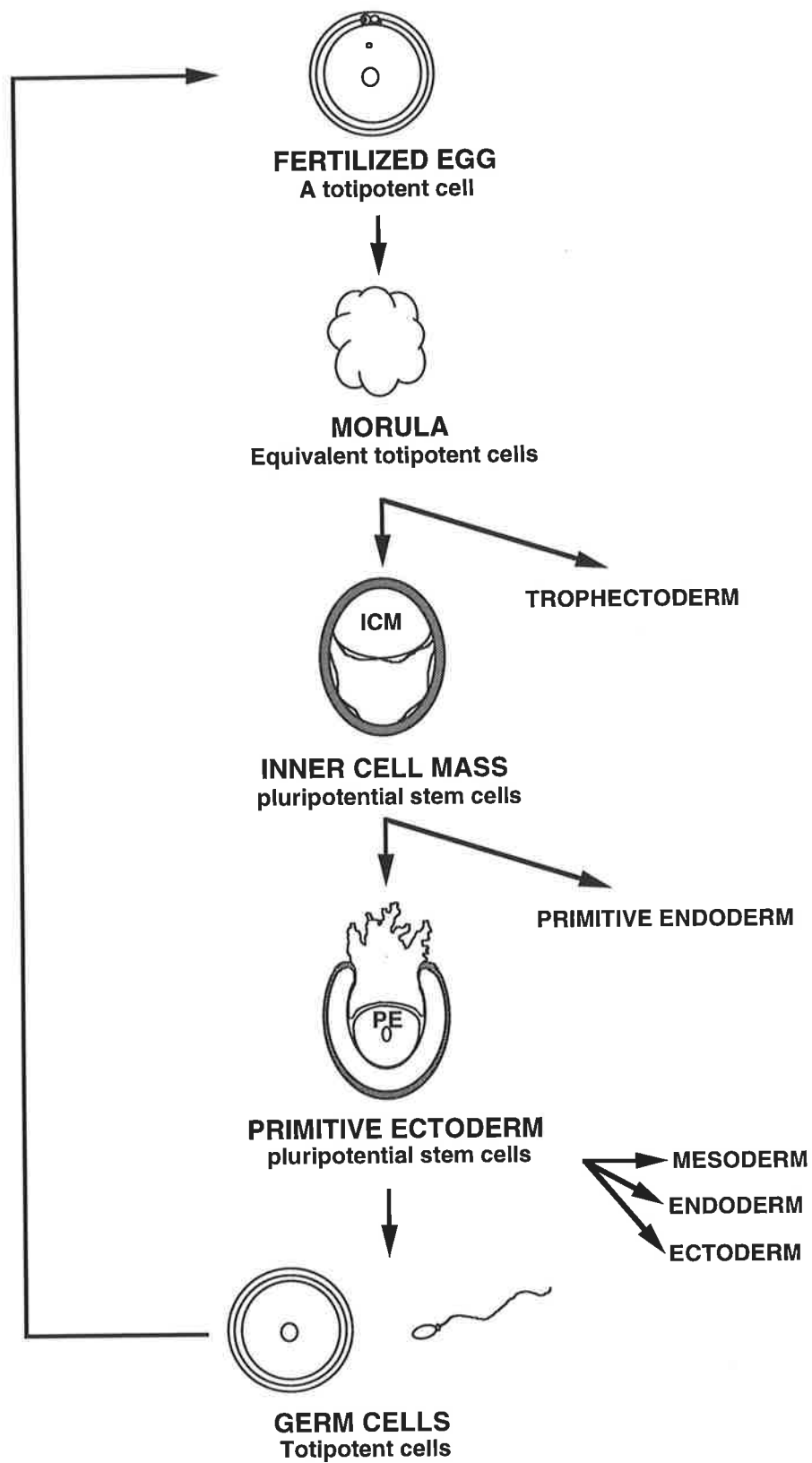
and Tam, 1979) while a normal mouse embryo can be produced by aggregation of more than one preimplantation mouse embryo (Lewis and Rossant, 1982). This plasticity is attributable to the presence of pluripotent cells within the embryo. The pluripotent cells of the mammalian embryo (Fig. 1.2) are established in the ICM, persist in the primitive ectoderm and are maintained throughout adult life as the germ cells. During normal embryogenesis stringent controls ensure that the correct balance between self-renewal and differentiation of the pluripotent stem cells is maintained (Smith, 1992). When this control breaks down, such as in experimental situations, pluripotent stem cells have an inherent ability to respond appropriately to the altered signals they receive. These findings highlight the importance of pluripotent stem cells in early embryogenesis of the mouse. To gain an understanding of the events which co-ordinate embryogenesis at the molecular level, an *in vitro* model system is required in which self-renewal and differentiation of pluripotent stem cells of the mouse can be studied.

1.6.2 Embryonal Carcinoma (EC) Cells

Grafting early stage embryos to ectopic sites within host mice (Hogan *et al.*, 1994) leads to breakdown of the normal controls that regulate stem cell differentiation. In such experiments, tumours derived from the grafted tissue are formed in the host (reviewed by Damjanov *et al.*, 1987). These tumours, known as teratocarcinomas, are characterised by the formation of a plethora of differentiated cell types surrounding a pool of stem cells that have escaped differentiation and can be malignant due to inappropriate proliferation of the remaining stem cells. The stem cells persisting in such tumours have been isolated and cultured *in vitro*, and clonal lines, termed embryonal carcinoma (EC) cells, established (Evans, 1972; Bernstine, 1973). The pluripotent nature of EC cells is evidenced by their ability to give rise to a range of differentiated cell types *in vitro*. Some EC cell lines are also able to contribute to the primary germ layers of the mouse upon reintroduction into the blastocyst (Brinster, 1974; Mintz and Illmensee; 1975; Papaioannou *et al.*, 1975; Stewart, 1982; Fuji and Martin, 1983). Germline transmission of one EC line has also been reported (Bradley, 1984), however the pluripotency of EC lines varies greatly. Some contribute extensively to chimaerae, others contribute with low efficiency, while still others give rise to tumours or developmental

Figure 1.2 Totipotent and pluripotent cells in murine development

The single totipotent cell, the fertilised egg gives rise to the morula which consists of equivalent totipotent cells termed blastomeres. The pluripotential cell lineage is established in the ICM and continues in the primitive ectoderm. The germ cells maintain the pluripotent cell lineage for the next generation. The differentiated tissues arising from each of the totipotent or pluripotent lineages are shown.



abnormalities when reintroduced into the mouse (reviewed in Smith, 1992). The inability of most EC cell lines to pass through the germline has been attributed to the abnormal karyotype of these cell lines (Robertson, 1986). These limitations make EC cells a poor in vitro model for pluripotent stem cells of the mouse.

1.6.3 Embryonic Stem (ES) Cells

The isolation and culture of pluripotent stem cells directly from the ICM of the preimplantation blastocyst was achieved independently by two groups in 1981 (Evans and Kaufman, 1981; Martin, 1981). The cells, now termed embryonic stem (ES) cells, were originally co-cultured with mitotically inactivated mouse embryo fibroblasts which were required to prevent differentiation of the stem cells. Subsequently, it was shown that a cytokine secreted by the feeder-layer, and other cultured cell lines, was sufficient for maintenance of undifferentiated stem cells. This factor was purified and termed Differentiation Inhibiting Activity (DIA) (Smith *et al.*, 1988), and was shown subsequently to be identical to the cytokine Leukaemia Inhibitory Factor (LIF) (Gearing *et al.*, 1987; Hilton *et al.*, 1988).

ES cells can be cultured and genetically manipulated in vitro when grown in the presence of LIF. Even after prolonged periods in culture, ES cells are able to contribute differentiated progeny to all germ layers and the germline of the mouse upon reintroduction into the developing mouse embryo (Bradley *et al.*, 1984; Robertson and Bradley, 1986). ES cells are therefore capable of responding to and generating all the normal developmental signals when reintroduced into the mouse. ES cells cultured in the presence of LIF express the pluripotent stem cell specific homeobox gene Oct-4 (Rosner *et al.*, 1991) and the EC cell specific surface antigen SSEA-1 (Solter and Knowles, 1978), and lack expression of differentiation antigens (Smith, 1992). The marker gene expression profile and the analogous behaviour of ES cells to cells of the ICM in the environment of the mouse blastocyst suggests that ES cells are the in vitro equivalent of the ICM.

1.6.4 Growth of ES Cells *in vitro* as Undifferentiated Monocultures

Propagation of undifferentiated ES cells in the absence of a feeder layer can be achieved by growth in medium supplemented with LIF. This secreted factor acts on target cells which express a cell surface LIF receptor and signal transduction molecule, gp130. Upon binding of LIF to its receptor, heterodimerization of the LIF receptor and gp130 occurs and leads to transduction of the LIF signal (Taga *et al.*, 1992). A growing family of related cytokines which are known to utilise the gp130 signal transduction molecule includes oncostatin M (OSM) (Gearing and Bruce, 1992; Gearing *et al.*, 1992), ciliary neurotrophic factor (CNTF) (Ip *et al.*, 1992), cardiotrophin-1 (CT-1) (Pennica *et al.*, 1995), IL-6 and IL-11 (Murakami *et al.*, 1993; Hilton *et al.*, 1994). CNTF, OSM and CT-1 are also able to inhibit the differentiation of ES cells grown *in vitro* (Conover *et al.*, 1993; Rose *et al.*, 1994, Pennica *et al.*, 1995).

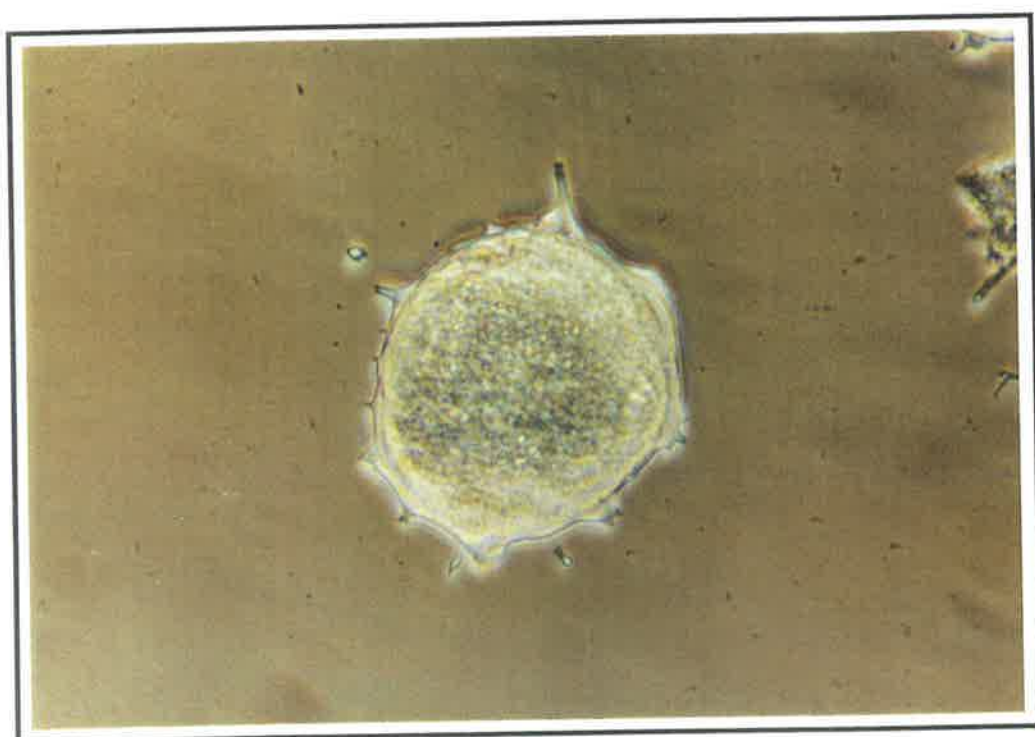
1.6.5 Differentiation of ES Cells *in vitro*

Undifferentiated ES cells grow in dome-shaped, adherent colonies, displaying prominent nuclei and negligible cytoplasm (Fig. 1.3). This distinctive morphology is lost upon differentiation of ES cells by a variety of techniques. Spontaneous and chemical induction of ES cell differentiation gives rise to a variety of cell types which share characteristics with differentiated tissues within the embryo (Smith, 1992). Spontaneous differentiation of ES cells, achieved by withdrawal of LIF, leads to the formation of a range of poorly defined terminally differentiated cell types. ES cells which escape differentiation also remain, and are probably maintained via feedback expression of LIF by differentiated cell types (Rathjen *et al.*, 1990b). Relatively homogeneous differentiation of ES cells can be achieved in the absence of LIF with the chemical inducing agents RA (Smith and Hooper, 1987; Heath and Smith, 1988) and 3-methoxybenzamide (MBA) (Smith, 1991) and leads to formation of poorly defined cell types (Smith, 1991).

ES cells may also be differentiated by growth in suspension cultures. Grown in this way, ES cells form spherical aggregates known as embryoid bodies which share structural similarities with the early mouse blastocyst (Robertson, 1987). Simple embryoid bodies develop after

Figure 1.3 Embryonic stem cell colony morphology

Photograph showing E14 embryonic stem (ES) cells grown as a monoculture in the presence of LIF. ES cells growing as a compact dome shaped colony three days after passaging are shown (photographed under phase contrast optics at 50 x magnification). Photograph was taken using a Nikon Diaphot microscope.



culture for two to four days in the absence of LIF and consist of a central mass of undifferentiated cells, with similarities to the ICM, and an outer layer of endoderm similar to the primitive endoderm. Culture for longer periods of time results in formation of cystic embryoid bodies which contain a central cavity and inner ectoderm-like layer with similarities to the proamniotic cavity and primitive ectoderm, respectively (Doetschman *et al.*, 1985; Shen and Leder, 1992). With further differentiation, a variety of tissues form in embryoid bodies including neurons, muscle, blood and cartilage (Evans and Kaufman, 1981; Martin, 1981; Doetschman *et al.*, 1985; Robertson, 1987). The developmental potential of stem cells in this system becomes progressively more restricted and culminates in the formation of terminally differentiated cell types. Thus, embryoid body formation in vitro reflects the progressive nature of embryonic development.

Differentiation of ES cells can also be achieved in the presence or absence of LIF by growth in medium conditioned by a transformed human hepatic cell line (J. Rathjen, personal communication). The cells produced by this treatment, termed X cells, are a homogeneous monolayer of flattened cells. Analysis of gene expression and differentiation potential of these cells (Bettess, 1993; Lake, 1996) indicates that the cells share similarities with the primitive ectoderm of the mouse embryo. The differentiation of ES cells into X cells in vitro may therefore provide a model for the equivalent differentiation event during mouse embryogenesis, and may also provide the first step toward specific terminal differentiation of ES cells in vitro. With the application of appropriate factors to ES cells and their differentiated derivatives in vitro it may be possible to generate model systems for later developmental events such as haematopoiesis, osteogenesis, neurogenesis and even organogenesis.

1.6.6 *ES Cells as a Model for Pluripotent Stem Cells of the Mouse*

A model system for the analysis of genes with potential roles in development would ideally consist of a homogeneous population of cells in which effects on proliferation, differentiation, death and cell migration can be assessed easily. The ability of ES cells to respond to and generate the signals required for normal development once returned to the mouse blastocyst, and to differentiate in vitro in an analogous manner to that seen during

normal embryogenesis, suggests that the ES cell system is a suitable model system for the identification and analysis of genes that regulate pluripotent stem cell renewal and differentiation.

1.6.7 *ES Cells as a Route to the Germline*

The ability of ES cells to contribute differentiated progeny to the germlayers of the mouse including the germline, even after genetic manipulation (Bradley *et al.*, 1984) means that ES cells can be used as a vehicle for the introduction of genetic alterations into the mouse. Thus, genetic alterations generated and studied in vitro, in ES cell system, can subsequently be analysed in vivo in the mouse.

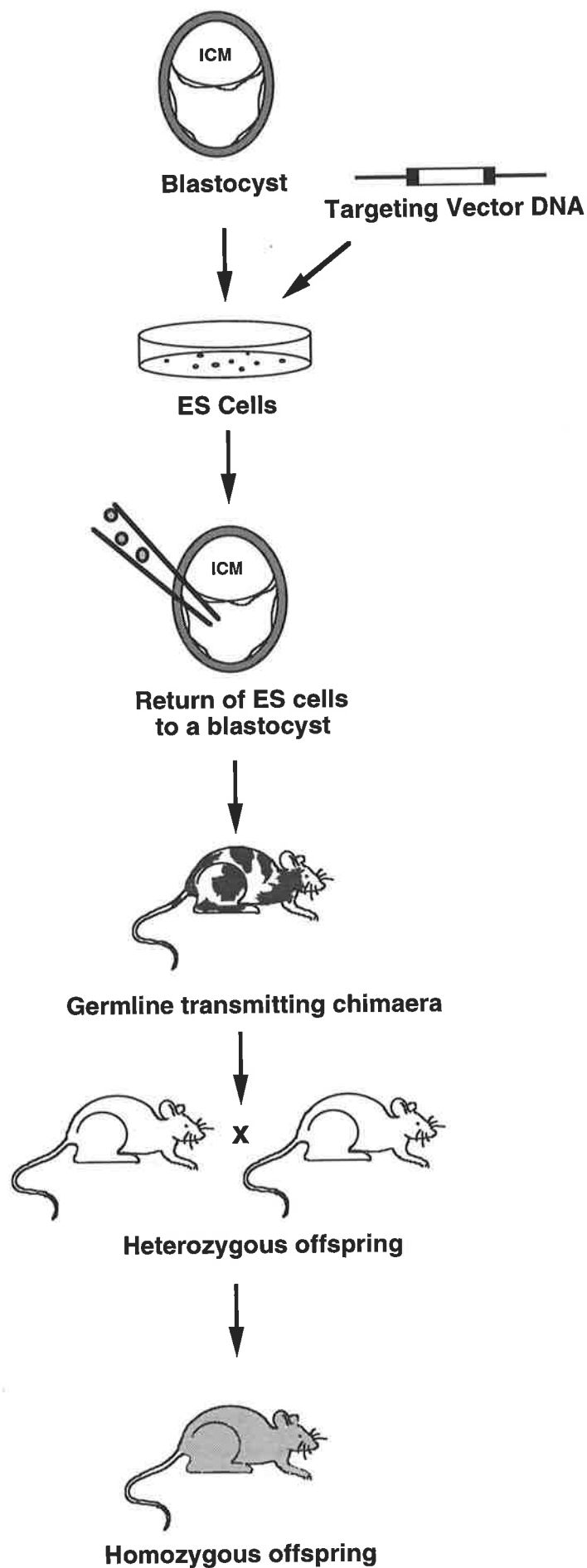
1.6.8 *Gene Targeting in ES cells*

Gene targeting is based on the finding that homologous DNA sequences recombine with their genomic counterparts at low frequency in mammalian cells (Smithies *et al.*, 1985). This allows the introduction of predetermined genetic alterations into any site within the mouse genome for which a genomic clone exists. This technique was first used in mammalian cell lines in the mid-nineteen eighties (Smithies *et al.*, 1985; Thomas *et al.*, 1986; Thomas and Capecchi, 1986), and was rapidly applied to ES cells (Thomas and Capecchi, 1987; Doetschman *et al.*, 1987). ES cells modified by gene targeting retain germline competence (Schwartzberg *et al.*, 1989; Thompson *et al.*, 1989) and can therefore be used for the generation of mice carrying predetermined genetic alterations.

Gene targeting has most commonly been used for the production of loss of function mutations in the mouse. The most common means of achieving this is to replace or disrupt an essential region of the coding region of one allele of the gene of interest with a selectable cassette (reviewed in Capecchi, 1989). Transgenic mice produced from the ES cells altered by gene targeting are bred to homozygosity to produce loss of function mutations in the mouse. This is represented diagrammatically in figure 1.4.

Figure 1.4 Generation of genetically altered mice from cultured ES cells

Representation of the procedure for the generation of genetically altered mice from ES cells. ES cells are isolated from the ICM of the mouse blastocyst, cultured and genetically manipulated by gene targeting and returned to the mouse blastocyst. The blastocysts are returned to foster mothers and chimaeric offspring result. Breeding of heterozygous offspring of chimaerae which transmit progeny of the ES cells through the germline, is carried out to produce homozygous, genetically altered mice.



It is also possible to produce ES cells carrying loss of function mutations by creating a loss of function mutation in both alleles of the gene in ES cells. Two general approaches have been used to achieve this. In the first approach, the first allele is altered in the usual fashion and the second allele is altered via a second round of homologous recombination in which an alternative selection cassette is used. In the second approach, the selection pressure for the selection cassette used to alter the first allele is increased. Lines which have undergone a gene conversion event, in which the wildtype allele is replaced with the targeted allele, are selected by this technique (Mortensen *et al.* 1988).

1.7 Analysis of Developmental Control Genes Using the ES Cell System

The ES cell system provides a model system in which the earliest developmental events of pluripotent cells can be dissected at the molecular level. In addition, the fact that ES cells provide a means to alter the mouse germline, means that genetic alterations studied in vitro can also be assessed in the context of the whole mouse.

1.7.1 *In vitro* Expression Mapping

Determining the sites and stages of expression of a putative developmental control gene during differentiation of ES cells in vitro can give the first indication that the gene under investigation may have effects on early developmental events. An alteration in expression level during differentiation, or expression in distinct differentiated cell types, would implicate the gene product in embryonic development.

In addition to conventional mapping techniques, it is also possible to use gene targeting to introduce a reporter gene into the gene under investigation (Mansour *et al.*, 1990; Le Mouellic *et al.*, 1992; Mountford *et al.*, 1994). Reporter gene expression can then be visualised during ES cell differentiation in vitro (section 1.6.5). This approach was used by Mansour and co-workers (1990) to map expression of the *int-2* protein in ES cells and embryoid bodies. The pattern of reporter gene expression reflected the known pattern of *int-2* RNA expression in ES cells and embryoid bodies.

1.7.2 Analysis of *in vitro* Gain of Function Mutations

In the case of gene products which act at the extracellular level, the equivalent of gain of function mutations can be assessed after exogenous application of such factors to ES cells. When purified factors are unavailable, or factors which act at the intracellular level are under investigation, the equivalent of gain of function mutations can be generated by over/inappropriate expression in ES cells. This is dependent on the availability of expression systems suitable for use in ES cells

The nature of developmental control genes necessitates the use of specialised systems for over/inappropriate expression. Firstly, developmental control genes may induce differentiation, cell death or inhibit proliferation of pluripotent stem cells. This necessitates the use of inducible expression systems for analysis. Gene expression in parallel, clonal cultures in which heterologous gene expression is present or absent can be studied. Secondly, developmental control genes often exert biological effects at extremely low thresholds. This necessitates the use of inducible expression systems which direct minimal or absent basal expression in the absence of induction. Finally, the precise level of expression of a developmental control gene product can have different effects on responding cells. For example, low concentrations of activin (a transforming growth factor family member) can induce ventrolateral mesoderm such as mesenchyme and muscle in *X. laevis* animal cap tissue, whereas high concentrations of activin induce dorsal mesoderm such as notochord (Slack *et al.*, 1987; Sokol *et al.*, 1990). The concentration of activin received by a responding cell determines the subset of genes activated and therefore the fate of the responding cell. This fact necessitates the use of expression systems which direct precisely controllable levels of heterologous gene expression.

Expression systems for use in ES cells fulfilling these requirements had not been described at the commencement of work presented in this thesis. Modified constitutive or inducible eukaryotic promoters have been used to direct expression in ES cells but the expression systems based on such promoters have been unsuitable for the analysis of developmental

control genes (section 2.1). The second chapter of this thesis describes the analysis of an expression system suitable for the analysis of developmental control genes in ES cells.

1.7.3 *Analysis of in vitro Loss of Function Mutations*

Loss of function mutations can be generated in ES cells by the gene targeting approach described in section 1.6.8. The phenotypic consequences of the mutation can then be assessed in vitro in ES cells and their differentiated progeny. For example, both alleles of the GAP-43 gene, which encodes a neuronal growth cone protein, were disrupted in ES cells using this approach (Strittmatter *et al.*, 1995). The phenotypic consequences were assessed in heterozygous and homozygous ES cells and in embryoid bodies produced from these cells. This approach will be particularly valuable to the study of developmental control genes which act at very early stages of development and are therefore likely to exert effects on stem cell renewal, differentiation and death.

1.7.4 *In vivo Expression Mapping*

Reporter genes introduced into an endogenous gene by gene targeting can also be used to map the expression of a gene in vivo in the context of the mouse. This is a superior method to that of insertional transgenesis (section 1.5.2) for expression mapping since the reporter gene, positioned in the endogenous locus, is subject to all the normal controls regulating expression of the gene under investigation. This approach is particularly useful for the analysis of developmental control genes which can be expressed at very low levels in vivo. This approach was used by Le Moulellic and co-workers (1992) to analyse expression of the *Hox-3.1* gene at the protein level in vivo. The expression pattern of the reporter gene reflected that of the *Hox-3.1* transcript in the mouse.

1.7.5 *Analysis of in vivo Gain of Function Mutations*

The generation of gain of function mutations in vivo using the ES cell system can be achieved by two approaches. Stably transfected ES cells over-expressing a gene can be introduced into the mouse embryo in order to generate chimaeric and subsequently transgenic mice expressing the gene of interest. This approach is reliant upon the availability of suitable

expression vectors and suffers from the same limitations as insertional transgenesis. Alternatively, a gene targeting approach, termed gap-repair targeting, can be used to create gene duplications at the endogenous locus using an insertion vector (reviewed in Bronson and Smithies, 1994). This has been carried out at the gene coding for angiotensinogen to produce mice carrying three or four copies of this gene (Smithies and Kim, 1994). The resulting mice displayed elevated levels of plasma angiotensinogen corresponding to the gene copy number. This is a superior method for the generation of *in vivo* gain of function mutations, since the duplicated genes are subject to the normal regulatory controls.

1.7.6 Analysis of *in vivo* Loss of Function Mutations

A large number of loss of function mutant mice have been produced using the technique of gene targeting in ES cells (Brandon *et al.*, 1995 a, b, c), and roles assigned to many genes based on the mutant phenotype. For example, mice lacking Wnt-1 expression die before or shortly after birth lacking a large region of the brain (McMahon and Bradley, 1990), providing definitive evidence that this signalling molecule is essential for embryonic neural development. Similarly, an essential role for the homeobox gene *Hox11* in development of the spleen has been demonstrated by the lack of a spleen in mice carrying a loss of function mutation in the *Hox11* gene (Roberts *et al.*, 1994).

Production of loss of function mutant mice is likely to yield valuable information about a large range of genes although limitations of this approach exist. One limitation has been the lack of a discernible phenotype in the case of some mutants (Brandon *et al.*, 1995 a, b, c). This has been a particular problem with growth factor genes many of which share common signalling pathways. Lack of a discernible phenotype in some mutants may reflect compensatory function by related molecules. For example, the transcription factor product of the MyoD gene was thought to be the key regulator involved in muscle development. Mice lacking expression of this gene did not display an overt phenotype. This has been attributed to elevated levels of Myf-5 gene expression in the mutant mice which is thought to provide compensation for the missing regulator (Rudnicki *et al.*, 1992).

In some cases, the lack of a discernible phenotype may simply reflect limitations of the techniques used to analyse the mutant mice. Subtle phenotypes may be overlooked in initial screening. For example, it has been reported that LIF⁻/LIF⁻ mice possess reduced levels of haematopoietic stem cells (Escary *et al.*, 1993). However, this defect was not observed in the initial report (Stewart *et al.*, 1992). Overt phenotypes for such mutations may be generated by cross-breeding mice lacking related genes or genes which act elsewhere within the developmental hierarchy.

A second limitation is that many null mutations are embryonic lethals. This is a particular problem in the case of developmental control genes since this complicates analysis of the mutant phenotype and prevents the analysis of roles later in development. For example, the growth factor FGF-4 is expressed in the ICM of the blastocyst stage embryo during gastrulation and also during myogenesis, limb and tooth development, suggestive of roles in both early and later developmental events (Niswander and Martin, 1992). Consistent with expression in the ICM, loss of function mutations in FGF-4 are embryonic lethals. Blastocysts lacking expression of this gene do not progress beyond the early stages of postimplantation development due to atrophy of the ICM (Feldman *et al.*, 1995). Although the essential role of this protein in promoting proliferation of the ICM has been demonstrated, roles for this protein in later development can not be assessed using this approach.

To reach a detailed understanding of the control and function of genes in the context of the mouse will require gene targeting approaches which permit the repeated introduction of precise, subtle mutations into the mouse genome designed to probe the structure, function and control of genes. A number of systems which permit such alterations have been published and are described in chapter 3 (sections 3.2.2-3). Alterations at the level of single amino acid, structural motifs and whole protein may be achieved with such approaches. The consequences can then be studied *in vitro* in the ES cell system (Ernst *et al.*, 1994), in which the response of the altered ES cells to known developmental stimuli is monitored, or in the mouse.

In the case of developmental control genes for which loss of function mutations are embryonic lethals, alterations to the expression control elements may permit roles during later stages of development to be assessed. Multiple alterations to the control elements and the gene product itself can be made to dissect the regulatory hierarchies that control expression of a developmental control gene, and probe the structure function relationship of the protein.

The inefficiency of homologous recombination compared to that of non-homologous recombination in the process of gene targeting necessitates the screening of large numbers of transfected cells for homologous recombinants. This is time consuming, labour intensive and laborious. Rigorous analysis at the molecular level in the mouse will therefore require the establishment of efficient gene targeting systems suitable for the repeated introduction of precise, subtle alterations to the mouse genome. Currently, systems fulfilling these requirements are not available. The second part of this thesis addresses this problem. A gene targeting system for the efficient introduction of non-selectable, subtle alterations repeatedly into any site within the mouse genome is described and characterised.

1.8 Aims and Approaches

ES cell technology has the potential to revolutionise our understanding of embryonic development of the mammal. The *in vitro* culture of ES cells provides a system in which the earliest developmental events can be assessed at the molecular level and is particularly amenable to the study of gain or loss of function mutations. In addition, the fact that ES provide a route to the mouse germline, means that genetic alterations generated and studied in ES cells can also be assessed in the context of the whole organism.

To exploit fully the ES cell system for the analysis of developmental control genes, specialised systems are required. Systems for the precisely controllable inducible expression of heterologous genes in ES cells are required for generation of gain of function mutations *in vitro*. Systems which permit the efficient introduction of subtle, predetermined alterations into the mouse germline are required to permit genetic analysis of developmental control genes in

the context of the mammalian organism. The aim of work presented in this thesis was to develop systems fulfilling these requirements.

The first aim of this work was to characterise further an inducible expression system suitable for use in ES cells. The expression system, based on the promoter of the human 6-16 gene, had been shown to direct IFN-inducible gene expression in transiently transfected ES cells (Whyatt, 1991). The approach used in this work was to produce stably transfected ES cell lines in which expression vector activity could be assessed in the context of the mouse genome. The uninduced levels of expression directed by the vectors, dose response and time over which induced expression could be maintained were all investigated.

The second aim of this work was to develop a gene targeting system which could permit the efficient, repeated introduction of precise, subtle alterations into the mouse genome. The system developed, termed the stable tag-exchange approach, is based on the two step targeting approach first described by Reid and co-workers (1990). In this approach, a tag and stability gene are introduced into the locus to be altered. The tag gene provides direct selection for homologous integration at the tagged site and permits the introduction of subtle alterations. The stability gene prevents physical loss of the tag gene by means other than homologous recombination, improving the efficiency of the system. The stable tag-exchange approach was characterised at the LIF locus in ES cells with the introduction of four independent alterations into this gene.

Chapter 2

2 Inducible ES Cell Expression Vectors Suitable for the Expression of Developmental Control Genes.

2.1 Introduction

The response of cultured ES cells to a developmental control gene product which acts at early stages of embryogenesis is likely to reflect that of pluripotent stem cells of the mouse embryo (section 1.6.6). The ES cell system therefore provides a model system for the identification and analysis of developmental control genes which affect pluripotent stem cell behaviour. Such genes are likely to exert effects on pluripotential stem cell renewal, differentiation and death, all of which can be assessed in the ES cell system. The main approach to the analysis of such genes in vitro is the generation of gain or loss of function mutations. In the case of developmental control genes which act at the intracellular level or where purified factors are unavailable, the gain of function mutations can be generated by over-/inappropriate expression of the gene in ES cells.

The ability to carry out such analysis relies on the availability of expression systems suitable for this purpose. As described in section 1.6.6, an ideal expression system for the analysis of developmental control genes in ES cells would direct minimal or no basal expression and direct precisely controllable levels of heterologous gene expression under conditions of induction in stable ES cell cultures.

A range of constitutive viral and eukaryotic promoters can be used to direct high level constitutive heterologous gene expression in ES cells. Although unmodified retroviral promoters are unable to direct gene expression in undifferentiated EC or ES cell lines, possibly due to lack of enhancer function, de novo methylation, intragenic sequences or negative trans-acting factors (Gre $\acute{z}$ *et al.*, 1990), Gre $\acute{z}$ and co-workers (1990) have produced a recombinant mouse retroviral-based system which is able to direct constitutive gene expression in ES cells. The utility of retroviral vectors is compromised by the fact that they act unpredictably in the differentiated cell types of the mouse (Soriano *et al.*, 1991) and are

unable to direct inducible gene expression in ES cells. Eukaryotic promoters such as that of the *Pgk-1* gene (Mortensen *et al.*, 1991) are able to direct constitutive gene expression in ES cells and behave more predictably in mammalian cells.

Two approaches to the construction of inducible ES cell expression systems using eukaryotic promoters have been used. Constitutive eukaryotic promoters can be modified such that they direct inducible gene expression in ES cells. For example, an inducible system based on an engineered *Pgk-1* promoter and lac operator-repressor system (Hannan *et al.*, 1993) is able to direct inducible gene expression in African monkey CV1 cells. The presence of detectable basal expression, resulting from incomplete repression, precludes use of this system for the expression of developmental control genes which are likely to have low biological thresholds. An alternative approach is use of inducible eukaryotic promoters. A number of inducible expression systems for use in mammalian cells based on such promoters have been described, including steroid (Ko *et al.*, 1989; Israel and Kaufman, 1989; Mader and White, 1993), heat shock (Schweinfest *et al.*, 1988) and heavy metal-inducible systems (Smith, unpublished). The heavy metal-inducible promoter of the metallothionein (MT) is able to direct metal-inducible gene expression in ES cells, it also directs appreciable levels of basal expression in these cells (Dr. Austin Smith, personal communication).

We have constructed an ES cell inducible expression system using the promoter of the human 6-16 gene. In 1985, Kelly and co-workers demonstrated that treatment of the human cell line HeLa with type I (α and β) or type II (γ) interferons (IFNs) could induce a range of mRNAs. The inducibility, levels and kinetics of induction by type I or II IFNs varied. Transcription of the human 6-16 gene was undetectable, by Northern analysis using 20 μ g total RNA, in uninduced cell cultures but was induced to high levels by treatment with α and β -IFNs. High level induction of the 6-16 transcript has been seen in all human cell lines tested including T98G neuroblastoma, HeLa, Daudi lymphoblastoid, U937 macrophage and F2001 and MRC5 fibroblasts (Kelly *et al.*, 1986) and expression can be induced in the mouse *Ltk*-cell line after transient or stable transfection with 6-16 based expression vectors (Porter *et al.*, 1988). In contrast to other IFN inducible transcripts, the basal level of 6-16 transcription in

the absence of induction is low or undetectable. Transcription of 6-16 mRNA in response to IFN-induction is a primary response which does not require protein synthesis (Friedman *et al.*, 1984) and is maintained for an extended period. The levels of 6-16 mRNA in T98G neuroblastoma and HeLa cells increased rapidly between six to twelve hours after induction and could be maintained for up to forty hours after IFN treatment (Friedman *et al.*, 1984; Kelly *et al.*, 1985). The level of IFN-induced 6-16 transcription increased in a dose dependent manner in human cells lines including T98G and HeLa cells (Friedman *et al.*, 1984; Kelly *et al.*, 1986).

High induction ratios and very low or absent basal expression in a range of cell lines, dose responsive transcription, maintenance of transcription for over 40 hours and the ability to confer IFN-inducibility on heterologous genes in both mouse and human cell lines made the 6-16 promoter a suitable choice for use in inducible ES cell expression vectors. The following chapter describing the construction and analysis of 6-16-based ES cell vectors is taken from a publication entitled "The Responsiveness of Embryonic Stem Cells to Alpha and Beta Interferon Provides the Basis of an Inducible Expression System for Analysis of Developmental Control Genes" (Whyatt *et al.*, 1993).

The full length and truncated 6-16 promoters (Porter *et al.*, 1988) used in the construction of inducible ES cell expression vectors, were cloned into two different promoterless expression vectors, namely pOT and pOX, in order to compare their efficacy in the context of two vector sequences. Dr. Peter Rathjen provided the vectors p1OT and p3OT carrying the full length and truncated 6-16 promoters, respectively in the pOT based vector. The 6-16 promoters from the pOT based vectors were subcloned into pOX as a part of an honours project (Whyatt, 1991). In the course of the honours project the CAT reporter gene was cloned into the expression site of the pOT and pOX based 6-16 expression vectors and expression analysis was carried out in transiently transfected MBL-5 ES cells.

It was particularly important to determine the efficacy of the 6-16-based expression vectors in stable ES cell lines. The ability of inducible expression vectors to function in stably

transfected ES cell lines would make them suitable for use in duplicate, clonal cultures in which expression is induced or uninduced. By comparison of the uninduced and induced cultures, effects on ES cell behaviour can be attributed directly to expression of the gene under investigation.

The production of stable MBL-5 ES cell lines and subsequent analysis was carried out as a part of work presented in this thesis. Construction of pn6-16DIA, subsequent analysis, and production of chimaerae were carried out by Annette Düwel and Jenny Nichols in the laboratory of Dr. Austin Smith (CGR, University of Edinburgh, Edinburgh, United Kingdom). Methods carried out as a part of this project are included in chapter 9.

2.2 IFN-inducible ES Cell Expression Vectors

MOLECULAR AND CELLULAR BIOLOGY, Dec. 1993, p. 7971-7976
0270-7306/93/127971-06\$02.00/0
Copyright © 1993, American Society for Microbiology

Vol. 13, No. 12

The Responsiveness of Embryonic Stem Cells to Alpha and Beta Interferons Provides the Basis of an Inducible Expression System for Analysis of Developmental Control Genes

LINDA M. WHYATT,¹ ANNETTE DÜWEL,² AUSTIN G. SMITH,² AND PETER D. RATHJEN^{1*}

Department of Biochemistry University of Adelaide, Adelaide, South Australia 5005, Australia,¹ and AFRC Centre for Genome Research, University of Edinburgh, Edinburgh EH9 3JQ, United Kingdom²

Received 28 April 1993/Returned for modification 16 June 1993/Accepted 17 September 1993

Embryonic stem (ES) cells, derived from the inner cell mass of the preimplantation mouse embryo, are used increasingly as an experimental tool for the investigation of early mammalian development. The differentiation of these cells *in vitro* can be used as an assay for factors that regulate early developmental decisions in the embryo, while the effects of altered gene expression during early embryogenesis can be analyzed in chimeric mice generated from modified ES cells. The experimental versatility of ES cells would be significantly increased by the development of systems which allow precise control of heterologous gene expression. In this paper, we report that ES cells are responsive to alpha and beta interferons (IFNs). This property has been exploited for the development of inducible ES cell expression vectors, using the promoter of the human IFN-inducible gene, 6-16. The properties of these vectors have been analyzed in both transiently and stably transfected ES cells. Expression was minimal or absent in unstimulated ES cells, could be stimulated up to 100-fold by treatment of the cells with IFN, and increased in linear fashion with increasing levels of IFN. High levels of induced expression were maintained for extended periods of time in the continuous presence of the inducing signal or following a 12-h pulse with IFN. Treatment of ES cells with IFN did not affect their growth or differentiation *in vitro* or compromise their developmental potential. This combination of features makes the 6-16-based expression vectors suitable for the functional analysis of developmental control genes in ES cells.

Murine embryonic stem (ES) cells are pluripotential cells derived from the inner cell mass of the preimplantation mouse embryo (5, 16). These cells, which can be cultured in the undifferentiated state *in vitro* for extended periods of time, retain the ability to participate normally in mouse development such that upon reintroduction into the mouse blastocyst, they can contribute differentiated progeny to all tissues of the mouse, including the germ line (3, 22). This implies that ES cells can both generate and respond to all signals that regulate normal murine development. The ability of ES cells to differentiate into a variety of cell types *in vitro* (23) can therefore be exploited as a model system for the identification of inter- and intracellular signals that regulate self-renewal and differentiation during early development.

ES cells can also be used as a vehicle for the introduction of transgenes into the early murine embryo. The consequences of ectopic gene expression on early developmental events can be assessed in chimeric animals generated by blastocyst injection of genetically modified ES cells (2, 29).

The development of improved systems for controlled and reproducible expression of transgenes in ES cells would offer new opportunities for the identification and characterization of gene products that regulate developmental decisions in mammals. In particular, an ability to establish stable transfectants in the absence of transgene expression would allow access to factors which promote differentiation or cell death or which otherwise inhibit stem cell proliferation. The requirement for an inducible expression system which is essentially inactive in the absence of the specific inducer precludes the use of commonly available systems, such as

those based on the metallothionein promoter (9), which have a significant level of uninduced expression in ES cells (unpublished observations). In this report, we demonstrate that ES cells are responsive to alpha and beta interferons (IFN- α and - β) and that this property can be exploited for the construction of an inducible expression system based on the regulatory elements of the human IFN-inducible gene 6-16.

MATERIALS AND METHODS

Vector construction. Plasmid pXMT2 is a modified version of the COS cell expression vector pMT2 (12) containing the restriction enzyme sites *Pst*I, *Kpn*I, *Xho*I, *Xho*I, *Kpn*I, *Pst*I, and *Eco*RI inserted between the unique *Pst*I and *Eco*RI sites in pMT2. Plasmid pOT was constructed by ligation of the 2.6-kb *Eco*RI-*Bgl*II fragment of pBLCAT3 (15) and the 1.6-kb *Eco*RI-*Bgl*II fragment of pSV2- β globin (7). This vector therefore consists of a polylinker upstream of the simian virus 40 small-t intron and polyadenylation sequences.

Human 6-16 promoter fragments were obtained from pSV0CAT-based plasmids described by Porter et al. (19). 6-16 fragment 1 (1,040 bp) extends from the *Bgl*II site at -603 to the *Bam*HI site at +437 in intron 1. This was obtained as a *Bgl*II-*Sma*I fragment (19), end filled, and ligated into *Xba*I-digested, end-filled pOT to produce p1OT. 6-16 fragment 3 (645 bp), which extends from the *Bgl*II site at -603 to the *Nci*II site at +42 in exon 1, was obtained as a *Hind*III-*Nci*II fragment (19), end filled, and ligated to *Xba*I-digested, end-filled pOT to produce p3OT. 6-16 fragments 1 and 3 were isolated from p1OT and p3OT by *Hind*III-*Bgl*II digestion and ligated into *Hind*III-*Bgl*II-digested pXMT2 to produce the vectors p1OX and p3OX, respectively.

The promoterless bacterial chloramphenicol acetyltrans-

* Corresponding author.

7972 WHYATT ET AL.

MOL. CELL. BIOL.

ferase (CAT) gene from pBLCAT3 (15) was isolated as a *Bst*YI fragment and ligated into *Bgl*II-digested expression vectors to produce the CAT expression vectors p1CX, p1CT, p3CX, and p3CT.

Plasmid pn6-16DIA was constructed by insertion of the coding sequence for the diffusible form of human differentiation-inhibiting activity/leukemia-inhibitory factor (DIA/LIF) from plasmid pC10-6R (27) into the *Bgl*II site downstream of the 6-16 sequences in p1OT. The vector also carries the neomycin phosphotransferase resistance (*neo*) cassette p_{gk}neopa, derived from pPNT (28), inserted into the unique *Eco*RI site in p1OT in the opposite transcriptional orientation to the DIA/LIF transcription unit. Digestion with *Eco*RI linearizes the plasmid and leaves the DIA/LIF and *neo* expression cassettes separated by the plasmid backbone.

Plasmid DNA was purified by CsCl density centrifugation. All DNA manipulations were carried out as described by Sambrook et al. (24), using enzymes from Boehringer Mannheim and Bresatec.

Cell culture and assays. CP1 (3), E14TG2a (10), EFC-1 (17), and MBL-5 (18) ES cells have been described elsewhere. CGR-2 ES cells were established from strain 129 mice by established procedures (23). ES cells were maintained in the absence of feeder cells essentially as described previously (26).

Transfections of the CAT expression plasmids were carried out by a modification of the method described by Chen and Okayama (4). Exponentially growing ES cells were seeded at a density of 5×10^5 cells per 9.5-cm² well and grown for 24 h at 37°C under 10% CO₂. Precipitates were formed with 5 µg of DNA mixed with 125 µl of 0.25 M CaCl₂ and 125 µl of 2× BBS (pH 6.99). Cells were transfected for 16 h under 3% CO₂ in a 50:50 mix of Dulbecco's modified Eagle's medium and Ham's F12 (26) containing 10% fetal calf serum. The precipitate was washed from the cells, and the cells were incubated under standard culture conditions. Cells were treated with a mixture of mouse IFN-α and -β (4.7 × 10⁵ IU/ml; Sigma) 2 h after the precipitate was removed from the cells. Transiently transfected cells were cotransfected with a β-galactosidase expression plasmid pCH110 (Pharmacia) at a ratio of 4:1. β-Galactosidase expression was determined by the method of Hall et al. (8) and used to normalize transient transfection efficiencies. Stable transfectants were produced by cotransfection of the expression vector and the Neo^r expression plasmid pPGKneoβ (20) at a ratio of 5:1 and were selected with 200 µg of G418 (GIBCO) per ml. CAT assays were carried out by the method of Gorman et al. (7) and quantitated by liquid scintillation counting. Induction ratios were calculated as percent chloramphenicol conversion in the presence of IFN/percent chloramphenicol conversion in the absence of IFN.

Transfection of pn6-16DIA DNA was carried out by electroporation at 800 mV and 3 µF in a Bio-Rad gene pulser, using 50 µg of linearized DNA and 10⁷ EFC-1 ES cells. Preparation of cytoplasmic mRNA and Northern (RNA) hybridization analysis were carried out as described previously (21). The human DIA/LIF antisense RNA probe was prepared by in vitro transcription of HpGEM1 (20) with T7 polymerase.

Production of chimeric mice. Chimeric mice were produced by established procedures (23), using CGR-2 and E14TG2a ES cells and C57/B16 host blastocysts. Both CGR-2 and E14TG2a ES cells are homozygous *c^{ch}/c^{ch}* so that chimerism was readily detected by the presence of sandy coat color pigmentation. Male chimeras were mated

with outbred MF1 albino females to test for germ line transmission of the ES genotype revealed by the production of chinchilla pups.

RESULTS

Dose-responsive inducible gene expression in transiently transfected ES cells. Two fragments of the human 6-16 promoter were used to construct IFN-inducible expression vectors (Fig. 1). The activities of these promoter fragments in ES cells and their responsiveness to IFN-α and -β were initially investigated in transient assays using CAT as a reporter gene in two different plasmid backbones. Each of the 6-16-based expression vectors directed IFN-inducible expression in transiently transfected ES cells (Fig. 2A). With IFN at 500 IU/ml, levels of induction between 3.1-fold (p3CT) and 46.6-fold (p1CX) were achieved, although absolute induction ratios are difficult to assess when uninduced levels of expression are extremely low. This result establishes for the first time that ES cells are responsive to IFN-α and -β and must therefore express the receptors and signaling molecules that direct the transcriptional response to IFN. Vectors containing the longer 6-16 fragment 1 (p1CX and p1CT) directed higher levels of induced expression than did equivalent vectors based on fragment 3 (p3CX and p3CT). This finding is consistent with previous observations on differentiated cell lines, including T98G neuroblastoma cells (6), HeLa cells (13), U937 macrophages, F2001 and MRC5 fibroblasts (14), and Ltk⁻ cells (13). The levels of expression directed by both 6-16 promoters in the vectors based on pXMT2 (p1CX and p3CX) were higher than the levels directed by the promoters in the vectors based on pOT (p1CT and p3CT). This resulted in 7.5- to 9-fold-higher absolute levels of induced expression and in higher induction ratios. p1CX had the highest absolute levels of induced expression, but this was offset by the marginally higher basal-level expression directed by this vector in the absence of IFN.

All plasmids exhibited a linear increase in the levels of CAT expression in response to increasing levels of IFN between 0 and 1,000 IU/ml. Representative data are shown for p1CX in Fig. 2B and C. Precise control over the level of heterologous gene expression in transiently transfected ES cells could therefore be achieved by induction with the appropriate dose of IFN.

Dose-responsive inducible expression in stably transfected ES cells. Stably transfected MBL-5 ES cell lines containing the 6-16 expression vectors p1CX (Fig. 3) and p3CT (results not shown) were screened for CAT expression before and after treatment with IFN. IFN-inducible expression was observed in nine of nine independent p1CX and nine of nine independent p3CT cell lines. In each case, CAT expression could not be detected in the absence of IFN. This points to a very low level of basal activity of the 6-16 promoter in stably transfected ES cells, since the CAT assay can detect as little as 5 to 10 pg of CAT protein (25). The undetectable basal expression directed by p1CX in stably transfected ES cells contrasts with the low but detectable level observed in transiently transfected ES cells.

CAT expression in stably transfected cell lines was strongly stimulated by treatment of cells with IFN. Induction ratios of up to 100-fold could be achieved by stimulation of the cells with 500 IU of IFN per ml for 48 h (Fig. 3B). Induced expression in stable cell lines containing p1CX (Fig. 4) and p3CT (results not shown) increased linearly with increasing concentrations of IFN, as previously observed

VOL. 13, 1993

INDUCIBLE EXPRESSION SYSTEM FOR USE IN ES CELLS

7973

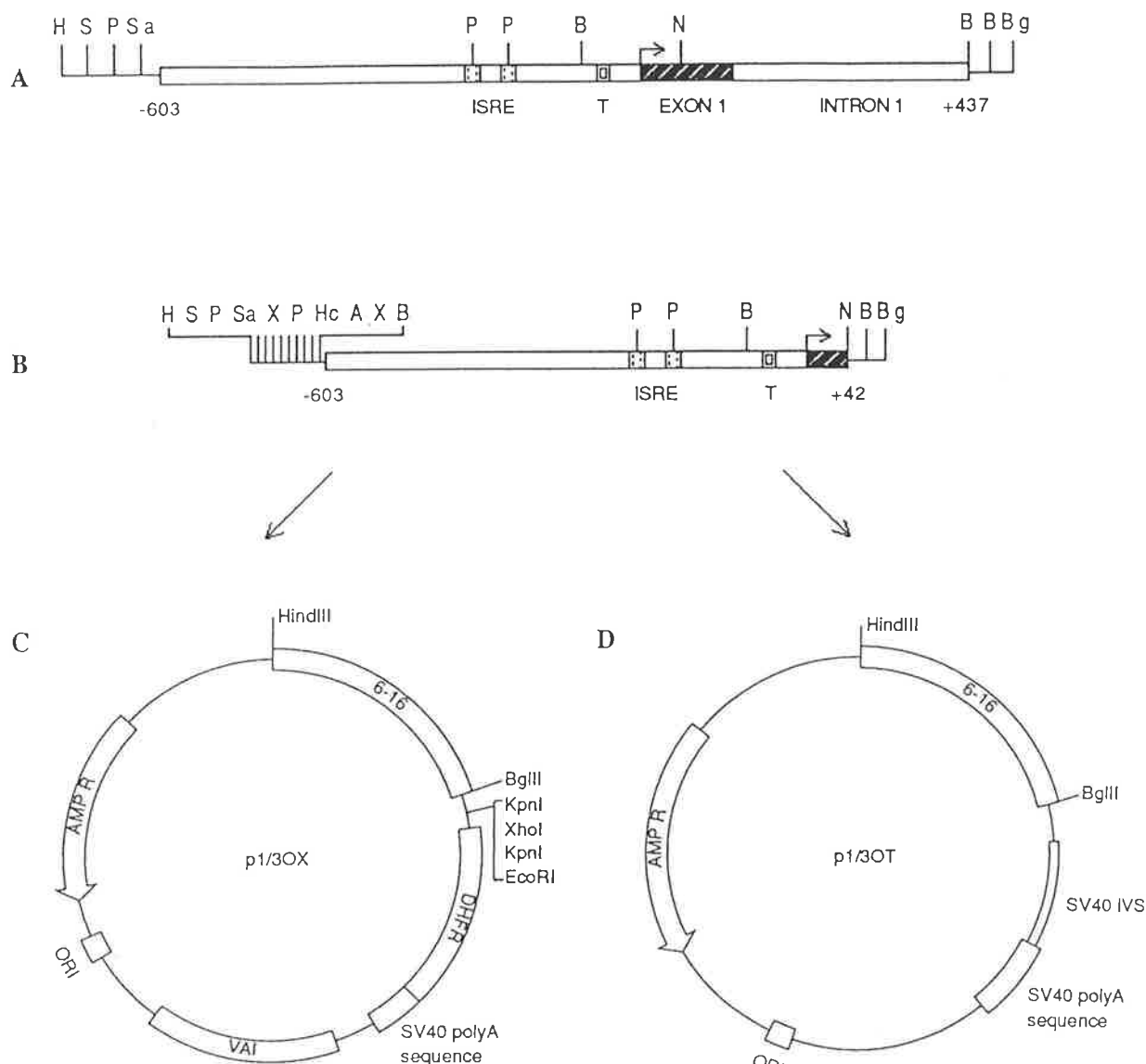


FIG. 1. Schematic diagram of the construction of the IFN-inducible ES cell expression vectors. The 1,040-bp (A) and 645-bp (B) 6-16 promoters were inserted into the pXMT2 and pOT vectors to produce four IFN-inducible ES cell expression vectors (C and D). The CAT reporter gene was then inserted into each of the four vectors at the unique *Bgl*II site. ISRE, IFN-stimulable response element; T, TATA box; DHFR, murine dihydrofolate reductase gene; VAI, adenovirus VAI gene, SV40 IVS, simian virus 40 small-t antigen intron; AMP^R, β -lactamase gene; ORI, *Escherichia coli* origin of replication; horizontal arrow, transcription start site. Restriction sites: A, *Acc*I; B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; Hc, *Hinc*II; N, *Nci*I; P, *Pst*I; S, *Sph*I; Sa, *Sal*I; X, *Xba*I.

with transiently transfected cells. Maximal induced expression was observed with about 1,000 IU of IFN per ml (Fig. 4C).

Time course of induced expression. The time course of IFN-induced expression was analyzed by using a stable MBL-5 ES cell line containing plasmid p1CX. CAT expression was monitored 0, 2, 6, 12, 24, and 48 h after cells had been treated with IFN at 1,000 IU/ml (Fig. 5). High levels of induced CAT activity were detectable 6 h after exposure of ES cells to IFN and were maintained for a further 42 h in the continuous presence of IFN (Fig. 5, filled squares). Removal of IFN from the culture medium after 8 h (Fig. 5, filled diamonds) or 10 h (Fig. 5, open squares) had little effect on the level of CAT expression at 24 h but resulted in a modest

decrease in the levels of expression seen at the 48-h time point.

IFN induction acts at the level of mRNA expression. Plasmid pn6-16DIA was constructed to investigate whether the 6-16 sequences could direct IFN-inducible expression of heterologous genes other than CAT. pn6-16DIA contains the coding sequence for the diffusible form of human DIA/LIF (27) inserted downstream of the 6-16 sequences in p1OT. The plasmid also contains a selectable *neo* cassette under the control of the murine phosphoglycerate kinase 1 (*pgk-1*) promoter (1). *Eco*RI-linearized pn6-16DIA was introduced into EFC-1 ES cells by electroporation, and G418-resistant colonies were isolated. After expansion of individual clones, integration events were analyzed by Southern hybridization

7974 WHYATT ET AL.

MOL. CELL. BIOL.

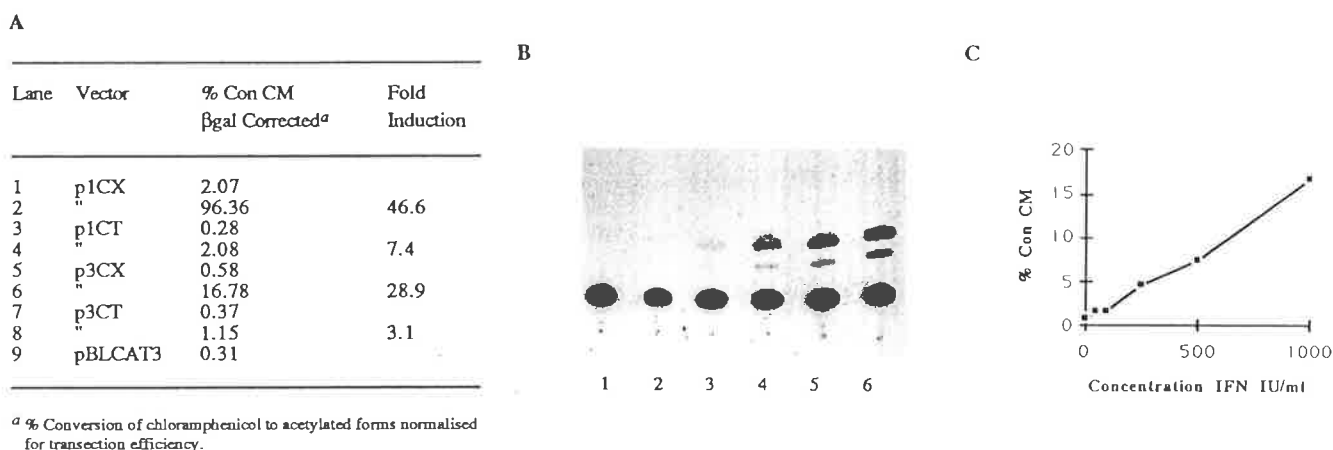


FIG. 2. (A) IFN-inducible expression in transiently transfected MBL-5 ES cells. ES cells were transiently cotransfected with the β -galactosidase expression vector pCH110 and expression vectors p1CX, p1CT, p3CX, p3CT, or pBLCAT3. Cells were grown in the continuous presence of 0 (lanes 1, 3, 5, 7, and 9) or 500 (lanes 2, 4, 6, and 8) IU of IFN per ml for 48 h. CAT assays were carried out with 0.075 U of β -galactosidase (β gal) activity. (B and C) Dose response of the IFN-inducible expression vector p1CX in transiently transfected MBL-5 ES cells. (B) CAT assays were carried out with 30 μ g of protein obtained from ES cells transiently cotransfected with p1CX and pCH110 after growth for 48 h in the continuous presence of 0 (lane 1), 50 (lane 2), 100 (lane 3), 250 (lane 4), 500 (lane 5), or 1,000 (lane 6) IU of IFN per ml. (C) Graphical representation of data shown in panel B.

(not shown). Four independent stable cell lines containing single inserts were subjected to mRNA analysis. Parallel cultures were incubated for 24 h in the presence or absence of 1000 IU of IFN per ml, and cytoplasmic RNA preparations were assayed for expression of the human DIA transgene by Northern hybridization (Fig. 6). In the absence of IFN, human DIA mRNA was not detectable in cell lines 1 and 4 and was present at extremely low levels in clones 2 and 3. All four clones showed similar, very high levels of transgene mRNA in response to IFN. This result indicates that the 6-16 system can be used to direct inducible expression of a range of heterologous genes, including mammalian cDNAs in ES cells, and that IFN induction is mediated largely, if not wholly, at the level of mRNA expression.

IFN treatment does not affect the developmental potential of ES cells. The response of MBL-5 and CP1 ES cells to IFN was assessed in clonal density cultures which were treated with increasing amounts of IFN. Concentrations of IFN up to 5,000 IU/ml had no discernible effect on colony number or colony size. Moreover, inhibition of differentiation was maintained in cultures supplemented with DIA/LIF,

whereas differentiation proceeded normally in cultures withdrawn from DIA/LIF. This finding indicates that IFN does not interfere with the processes of ES cell renewal or differentiation *in vitro*.

ES cells were also injected into blastocysts to investigate whether exposure to IFN had impaired their ability to contribute to chimeras. CGR-2 and E14TG2a ES cells were cultured in the presence of IFN at 500 and 1,000 IU/ml, respectively, for 24 h immediately prior to microinjection. In both cases, several overt coat color chimeras were obtained. These animals were healthy with no signs of any abnormality. The E14TG2a chimeras exhibited particularly extensive ES cell contributions, and one of these has now transmitted the ES genotype to viable offspring. This result demonstrates that pretreatment with IFN is compatible with chimera production and does not compromise the developmental potential of ES cells.

DISCUSSION

At the cellular level, developmental events such as self-renewal, differentiation, and programmed cell death are often critically dependent on the precise timing and level of gene expression. The analysis of these processes would therefore be significantly enhanced by the development of systems that allow the precise manipulation of gene expression. An ideal expression system for use in ES cells would exhibit low levels of uninduced expression (below the threshold required for biological activity), give a controllable range of induced expression levels, and use an inducing agent that does not compromise the pluripotential character of the cells. In this report, we describe the construction and characterization of an inducible expression system which fulfills these requirements in both transiently and stably transfected ES cells.

We have shown that ES cells are responsive to IFN- α and - β . Levels of expression directed by the IFN-inducible 6-16 promoter were extremely low or undetectable in the absence of IFN, could be stimulated up to 100-fold by treatment of cells with IFN, and increased in linear fashion in response to

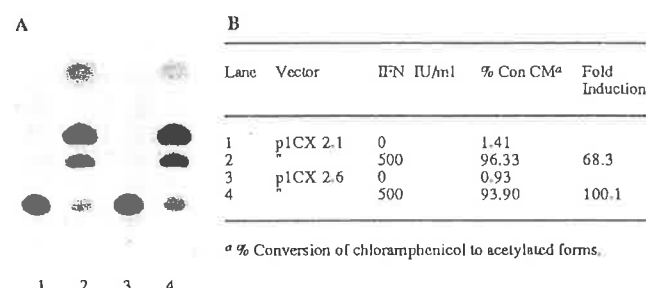


FIG. 3. IFN-induced expression directed by p1CX in stable MBL-5 ES cell lines. (A) CAT assays were carried out with 35 μ g of protein obtained from two independent ES cell lines containing p1CX, i.e., p1CX 2.1 (lanes 1 and 2) and p1CX 2.6 (lanes 3 and 4) after growth for 48 h in the presence of 0 (lanes 1 and 3) or 500 (lanes 2 and 4) IU of IFN per ml. (B) Quantitation of results shown in panel A.

VOL. 13, 1993

INDUCIBLE EXPRESSION SYSTEM FOR USE IN ES CELLS

7975

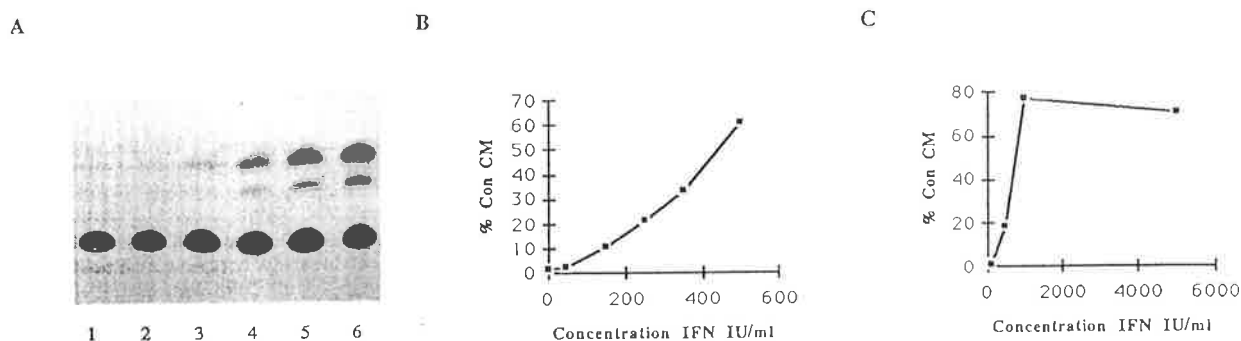


FIG. 4. Dose response of p1CX in stable MBL-5 ES cell line 3.9. CAT assays were carried out with 50 μ g of protein obtained from a stable ES cell line (p1CX3.9) containing p1CX after growth in the presence of 0 (lane 1), 50 (lane 2), 150 (lane 3), 250 (lane 4), 350 (lane 5), or 500 (lane 6) IU of IFN per ml for 48 h (A) or 0, 50, 100, 500, 1,000, or 5,000 IU of IFN per ml for 48 h (C). (B) Graphical representation of results presented in panel A. % Con CM, percent conversion of chloramphenicol to acetylated forms.

increasing levels of IFN. A single 6-16 vector could therefore be used to direct any required level of gene expression by manipulation of the dose of IFN used for induction. This precise refinement of expression levels is not available with other ES cell expression systems.

Although our analysis has been carried out primarily in MBL-5 ES cells, the 6-16 vectors also direct IFN-induced expression in independently isolated ES cell lines such as EFC-1, E14, and CP1. The 6-16 vectors are therefore likely to be generally useful for the regulated expression of transgenes in ES cell lines.

Two derivatives of the 6-16 promoter were tested. Higher levels of induced expression and higher induction ratios were directed by 6-16 fragment 1 (Fig. 1A). This is consistent with results obtained in terminally differentiated cells (19) and points to the existence of a positive signal for gene expression in the first intron of the 6-16 gene. Induced levels of expression for both promoters were higher in vectors derived from pXMT2 than in vectors derived from pOT. This influence of vector context on expression is significant, although its basis remains unclear. It may reflect a positive effect of VAI sequences on IFN-induced gene expression directed by pXMT2-based vectors (12) and/or aberrant pro-

cessing of the simian virus 40 small-t antigen intron (11) in the pOT-based vectors.

High levels of induced heterologous gene expression were detectable 6 h after IFN treatment and were maintained for at least 48 h both in the continuous presence of IFN and following IFN withdrawal after 8 or 10 h of induction. Induction occurred primarily at the level of RNA expression.

Treatment of ES cells with the inducing agent IFN did not induce differentiation or appreciable growth arrest even at concentrations of 5,000 IU/ml. ES cells treated with 1,000 IU of IFN per ml for 24 h retained the ability to participate normally in murine embryogenesis, indicating that the developmental potential of the cells had not been compromised.

The development of an effective inducible expression vector for use in murine ES cells has important implications for the use of these cells as an experimental system for the analysis of mammalian development. Factors that regulate developmental decisions such as differentiation or programmed cell death are difficult to study in pluripotential stem cells, since the consequence of deregulated expression

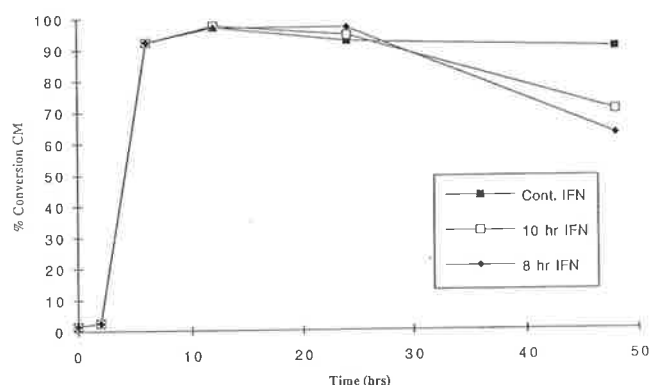


FIG. 5. Time course of IFN-induced expression in stable ES cell lines. CAT assays were carried out with 25 μ g of protein obtained from MBL-5 ES cell line p1CX 2.1 at 0, 2, 6, 12, 24, or 48 h after treatment with 1,000 IU of IFN per ml. Closed squares, continuous presence of IFN; open squares, removal of IFN from culture medium after 10 h; closed diamonds, removal of IFN from culture medium after 8 h. % Conversion CM; percent conversion of chloramphenicol to acetylated forms.

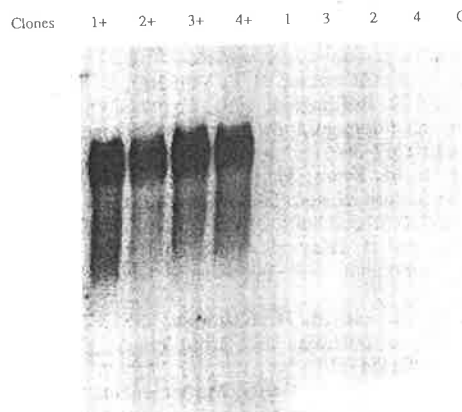


FIG. 6. IFN induction of RNA transcripts. Cytoplasmic RNA was isolated from four independent stable EFC-1 ES cell transfectants carrying the vector pn6-16DIA grown in the continuous presence of 0 (lanes 1 to 4) or 1,000 (lanes 1+ to 4+) IU of IFN per ml for 24 h. Ten micrograms of RNA was analyzed for the expression of human DIA by Northern hybridization. Lane C, nontransfected parental ES cell RNA.

7976 WHYATT ET AL.

MOL. CELL. BIOL.

is terminal alteration of cell identity. The ability to establish stable cell lines in the absence of transgene expression, and subsequently to express heterologous genes in a precisely controlled manner, should help to overcome this difficulty. The extremely low levels of basal expression directed by the 6-16-based expression vectors means that this approach should be applicable to molecules that are biologically active at low concentrations. Inducible expression systems could also be used for the controlled inactivation of stem cell functions by the inducible expression of antisense transcripts. The versatility of the system is extended by the ability to alter expression levels by manipulation of the IFN concentration. This should allow the analysis of factors that are toxic to stem cells when overexpressed and the analysis of factors that can have variable biological activities depending on the level of expression.

Comparison of clonal cell lines under induced and noninduced conditions would allow the developmental response of cells or embryo to be correlated directly with expression of the transgene. The developmental consequences of heterologous gene expression could readily be analyzed in vitro. Moreover, it may be possible to investigate the developmental consequences of transgene expression in vivo, by blastocyst injection following induction in vitro, or by culturing chimeric embryos in the presence of IFN.

The general properties of the 6-16 expression system identify it as an appropriate system for use in ES cells. The availability of these vectors increases the experimental versatility of ES cells and in particular affords new possibilities for the identification and characterization of genes which regulate growth, survival, and differentiation processes in the early mammalian embryo.

ACKNOWLEDGMENTS

This work was supported by grants from the University of Adelaide and the National Health and Medical Research Council (920087).

We thank A. C. G. Porter for valuable discussions and advice, L. R. Whyatt for photography, and J. Rathjen and J. A. Beale for assistance with tissue culture. The contribution of J. Nichols to production of ES cell chimeras is gratefully acknowledged.

REFERENCES

- Adra, C. N., P. H. Boer, and M. W. McBurney. 1987. Cloning and expression of the mouse *pgk-1* gene and nucleotide sequence of its promoter. *Gene* **60**:65-74.
- Boulter, C. A., A. Aguzzi, R. L. Williams, E. F. Wagner, M. J. Evans, and R. Beddington. 1991. Expression of *v-src* induces aberrant development and twinning in chimaeric mice. *Development* **111**:357-366.
- Bradley, I., M. Evans, M. H. Kaufman, and E. J. Robertson. 1984. Formation of germline chimaeras from embryo-derived teratocarcinoma cell lines. *Nature (London)* **309**:255-256.
- Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**:2745-2752.
- Evans, M. J., and M. H. Kaufman. 1981. Establishment in culture of pluripotent cells from mouse embryos. *Nature (London)* **292**:154-156.
- Freidman, R. L., S. P. Manly, M. McMahon, I. M. Kerr, and G. R. Stark. 1984. Transcriptional and posttranscriptional regulation of interferon-induced gene expression in human cells. *Cell* **38**:745-755.
- Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
- Hall, C. V., E. P. Jacob, G. M. Ringold, and F. Lee. 1983. Expression and regulation of *Escherichia coli lacZ* gene fusions in mammalian cells. *J. Mol. Appl. Genet.* **2**:101-109.
- Hamer, D. H., and M. Walling. 1982. Regulation in vivo of a mammalian gene in SV40 vectors. *J. Mol. Appl. Genet.* **2**:1044-1051.
- Hooper, M., K. Hardy, A. Handyside, S. Hunter, and M. Monk. 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature (London)* **326**:292-295.
- Huang, M. T. F., and C. M. Gorman. 1990. The simian virus 40 small-t intron, present in many common expression vectors, leads to aberrant splicing. *Mol. Cell. Biol.* **10**:1805-1810.
- Kaufman, R. J., M. Davies, V. Pathak, and J. W. B. Hershey. 1989. The phosphorylation state of eukaryotic initiation factor 2 alters translational efficiency of specific mRNA. *Mol. Cell. Biol.* **9**:946-958.
- Kelly, J. M., C. S. Gilbert, G. R. Stark, and I. M. Kerr. 1985. Differential regulation of interferon-induced mRNAs and c-myc mRNA by α - and γ -interferons. *Eur. J. Biochem.* **153**:367-371.
- Kelly, J. M., A. C. G. Porter, Y. Chernajovsky, C. S. Gilbert, G. R. Stark, and I. M. Kerr. 1986. Characterization of a human gene inducible by α - and β -interferons and its expression in mouse cells. *EMBO J.* **7**:1601-1606.
- Luckow, B., and G. Schutz. 1987. CAT constructions with multiple cloning sites for the functional analysis of promoters and regulatory sequences. *Nucleic Acids Res.* **15**:5490.
- Martin, G. R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**:7634-7638.
- Nichols, J., E. P. Evans, and A. G. Smith. 1990. Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development* **110**:1341-1348.
- Pease, S., P. Braghetta, D. Gearing, D. Grail, and R. L. Williams. 1990. Isolation of embryonic stem (ES) cells in media supplemented with recombinant leukaemia inhibitory factor (LIF). *Dev. Biol.* **141**:344-352.
- Porter, A. C. J., Y. Chernajovsky, C. Gilbert, G. R. Stark, and I. M. Kerr. 1988. Interferon response element of the human gene 6-16. *EMBO J.* **7**:85-92.
- Rathjen, P. D., J. Nichols, S. Toth, D. R. Edwards, J. K. Heath, and A. G. Smith. 1990. Developmentally programmed induction of differentiation inhibiting activity and the control of stem cell populations. *Genes Dev.* **4**:2304-2318.
- Rathjen, P. D., S. Toth, A. Willis, J. K. Heath, and A. G. Smith. 1990. Differentiation inhibiting activity is produced in matrix-associated and diffusible forms that are generated by alternate promoter usage. *Cell* **62**:1105-1114.
- Robertson, E. J. 1986. Pluripotent stem cell lines as a route into the mouse germline. *Trends Genet.* **2**:9-13.
- Robertson, E. J. 1987. Embryo-derived stem cell lines, p. 71-112. In E. J. Robertson (ed.), *Teratocarcinomas and embryonic stem cells: a practical approach*. IRL Press, Oxford.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Schenborn, E., and D. Titus. 1989. Detection of eukaryotic transcriptional regulatory sequences with the CAT reporter gene assay. *Promega Notes* **21**:1-3.
- Smith, A. G. 1991. Culture and differentiation of embryonic stem cells. *J. Tissue Cult. Methods* **13**:89-94.
- Smith, A. G., J. K. Heath, D. D. Donaldson, G. G. Wong, J. Moreau, M. Stahl, and D. Rogers. 1988. Inhibition of pluripotent embryonic stem cell differentiation by purified polypeptides. *Nature (London)* **336**:688-690.
- Tybulewicz, V. L. J., C. E. Crawford, P. K. Jackson, R. T. Bronson, and R. C. Mulligan. 1991. Neonatal lethality and lymphopenia in mice with a homozygous disruption of the *c-abl* proto-oncogene. *Cell* **65**:1153-1163.
- Williams, R. L., S. A. Courtneidge, and E. F. Wagner. 1988. Embryonic lethality and endothelial tumors in chimaeric mice expressing polyoma virus middle T oncogene. *Cell* **52**:121-131.

2.3 Further Discussion

A limitation to the use of ES cells as a system for the analysis of developmental control genes has been the lack of expression systems suitable for the expression of such genes in these cells. We sought to address this problem by exploiting a naturally occurring, IFN-inducible promoter which displayed characteristics suitable for the expression of genes which exert biological effects at low thresholds such as developmental control genes. Features of an ideal expression system for use in the ES cell system would include use of an induction system which does not compromise the pluripotency of ES cells, and which would provide the means to direct precisely controllable levels of gene expression which could be maintained for extended periods of time. At the commencement of this work it had been demonstrated that ES cells were responsive to type I IFNs, indicating that ES cells possessed all the factors required for IFN-responsiveness (Whyatt, 1991). In addition, treatment of ES cells with IFN, even at high doses, did not have any measurable effect on growth rate or differentiation, nor did it compromise their pluripotency.

Construction and analysis of stable ES cell lines carrying the vectors p1CX and p3CT was carried out in work presented in this thesis. Expression directed by the vectors in stable ES cell lines was undetectable in the absence of induction, increased rapidly between two and six hours after treatment with IFN and reached almost maximal expression levels after six hours. High level expression was maintained for at least forty two hours after induction. The period of time over which a gene must be expressed to cause alteration in cell fate is not clear and may vary with the type and amount of protein expressed. However, treatment of ES cells with the chemical inducer MBA for forty eight hours is sufficient to induce terminal differentiation of these cells *in vitro*, suggesting that induced expression for up to forty hours should be sufficient to exert biological effects on ES cells. The characteristics displayed by the 6-16-based expression vectors suggest that they will be suitable for the analysis of developmental control genes which exert effects on pluripotent stem cells of the mouse.

Using the 6-16-based expression vectors it should therefore be possible to analyse directly the consequences of gene expression in stably transfected, duplicate clonal cultures of ES cells in which expression is uninduced or induced. Effects on ES cell proliferation, differentiation and cell death can be monitored readily. The effects of various levels of induced gene expression could be assessed after induction with appropriate concentrations of IFN and effects on differentiation assessed after treatment with chemical inducing agents, LIF withdrawal or growth as embryoid bodies (section 1.6.5).

Since this work was carried out the 6-16-based vectors have been used to express a range of other genes in transiently and stably transfected ES cell lines. The vector has been used to express inducible *lacZ* gene expression in transiently transfected ES and X cells (sections 1.6.5, 5.4). Inducible expression of the homeobox gene *Hesx1* (Thomas, 1995) and variants of the LIF gene (B. Haines, personal communication) have been achieved with these vectors. The level of transcription, as assessed by RNase protection, in uninduced lines varied according to the vector copy number, although lines with very low or absent basal expression were obtained in each case. Inducible expression of the FLP recombinase protein (Vetter *et al.*, 1983) in transiently transfected ES cells has also been achieved using the 6-16-based vectors (Harrison, 1993; Brown, 1994). These results suggest that the vectors will be suitable for the expression of a wide range of genes in ES cells.

The ability of the 6-16-based vectors to function in X cells indicates that they are active in a differentiated derivative of ES cells. In addition, the ubiquitous nature of the IFN response (reviewed in Williams, 1991) suggests that the vectors may be suitable for use in a wide range of mammalian cell types. This would extend the versatility of the vectors and may make them suitable for the use in cell types generated by lineage specific differentiation of ES cells.

In 1995, a polyoma virus-based, constitutive gene expression vector for use in ES cells was described by Gassmann and co-workers. This vector can be maintained in ES cells as an autonomously replicating extrachromosomal element or episome. This feature provides two major advantages over conventional vectors which require integration into the host genome

for stable maintenance. The expression driven by such vectors is not subject to position effects due to chromosomal location, and mutagenesis is not caused by random integration into the genome. High superinfection efficiencies can be achieved with cell lines carrying episomal-based vectors (Gassmann *et al.*, 1995). This feature of the episomal-based system makes possible functional screening of cDNA libraries in ES cells. Such technology would facilitate the identification of genes which exert effects on pluripotent stem cells of the mouse. An inducible, precisely controllable, episomal expression system for use in ES cells, which lacks appreciable basal expression, would provide the best possible means to assess the function of developmental control genes in this system. It may be possible to produce such a system with use of the 6-16 promoter in conjunction with the polyoma virus-based vectors.

Chapter 3

3 An Efficient Gene Targeting System for the Repeated Introduction of Precise, Subtle Alterations into the Mouse Genome

3.1 Gene Targeting Systems for the Production of Null Alleles

Gene targeting in ES cells has most commonly been used to produce loss of function mutant mice (Brandon *et al.*, 1995a). In the case of genes for which there is no direct selection for loss of function in vitro, this has generally been achieved by inactivation with a selectable gene. A selectable gene such as the bacterial *neo* gene, which provides resistance to the aminoglycoside G418, is introduced into the target locus by homologous recombination such that it disrupts or replaces an essential region of the coding sequence. Cells carrying the targeting vector are selected in G418 and homologous recombinants are identified by PCR or Southern blot analysis (reviewed in Capecchi, 1989).

Two types of vectors are used for gene targeting: replacement or insertion vectors (Thomas and Capecchi, 1987). Replacement vectors are linearised at the extremities or outside the homologous sequence and the vectors remain co-linear with the target sequence (Fig. 3.1a). To produce null mutations a critical section of the coding region is replaced or disrupted with the selection cassette. In contrast to this, insertion vectors are linearised within the region of homology and homologous recombination of these vectors leads to duplication of the genomic sequence (Fig. 3.1b). To produce null mutations with insertion vectors, the coding sequence is disrupted with the vector sequence which also carries a selection cassette within an essential region of the coding sequence.

Although the mechanisms by which homologous recombination of both types of vectors occurs in mammalian cells is still poorly understood, the parameters which govern the process have been investigated. The efficiency of homologous recombination can be increased by increasing the total amount of homology contained in a targeting vector (Thomas and Capecchi, 1987; Hasty *et al.*, 1991b; Deng and Capecchi, 1992). The relationship between length of homology and gene targeting efficiency is exponential between 2 and 10

Figure 3.1 Types of Gene Targeting Vectors

a Representation of gene targeting using a replacement vector with (i) the genomic locus, (ii) replacement targeting vector and (iii) targeted locus shown. Thin line, genomic DNA; numbered black boxes, exons; hatched box, *neo* cassette.

b Representation of gene targeting using an insertion vector with (i) the genomic locus, (ii) insertion targeting vector and (iii) targeted locus shown. Thin line, genomic DNA; thick line, non-genomic DNA; numbered black boxes, exons; hatched box, *neo* cassette.

Adapted from Soriano (1995).

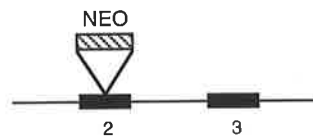
a

Replacement Vector

i Genomic Locus



ii Targeting Vector



iii Targeted Locus



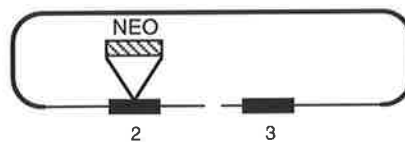
b

Insertion Vector

i Genomic Locus



ii Targeting Vector



iii Targeted Locus



kb, whereas increases in homology beyond 14 kb have little effect. The efficiency of homologous recombination can also be increased by using DNA isogenic with the strain of mice from which the ES cells were derived for vector construction (te Riele *et al.*, 1992; van Deursen and Wieringa, 1992). The importance of using isogenic DNA is reflected by the finding that the frequency of homologous recombination at the creatine kinase M locus was twenty five fold higher when isogenic DNA was used compared to non-isogenic DNA (van Deursen and Wieringa, 1992). The demonstration that genes not expressed in ES cells could be efficiently modified by gene targeting (Johnson *et al.*, 1989), made the technique applicable to all genes of the mouse genome. In addition to this, the finding that the frequency of homologous recombination is unaffected by the length of non-homologous DNA introduced into the target locus (Mansour *et al.*, 1990) vastly increased the flexibility of vector design.

An enrichment system, known as positive-negative selection (Mansour *et al.*, 1988; Fig. 3.2) has been based on the finding that non-homologous regions at the extremities of replacement vectors are integrated during non-homologous but not homologous recombination. Incorporation of a negative selection cassette, such as the Herpes simplex virus thymidine kinase (HSV-*k*) gene, positioned outside the homologous sequence on a replacement vector can therefore be used to reduce the background of non-homologous recombinants which survive selection.

In initial gene targeting experiments, in which non-isogenic DNA was used for the construction of targeting vectors, the vast majority of clones surviving selection were non-homologous recombinants. The positive-negative enrichment system was able to reduce the high background of non-homologous recombinants and led to enrichments in the order of 1000 to 2000-fold (Mansour *et al.*, 1988; Mansour *et al.*, 1990). With the use of isogenic targeting vectors however, a much higher proportion of transfected cells are homologous recombinants (in the range of 1/10 to 1/500), reducing the number of non-homologous recombinants which must be screened. When isogenic DNA vectors are used, enrichments

Figure 3.2 Positive-Negative gene targeting approach

Representation of gene targeting using the positive-negative approach with (i) the genomic locus, (ii) targeting vector, (iii) targeted locus and (iv) random insertion shown. Thin line, genomic DNA; numbered black boxes, exons; hatched box, *neo* cassette; stippled box, HSV-*tk* cassette.

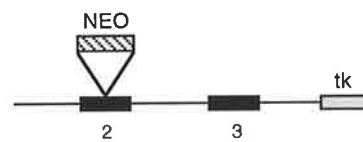
Adapted from Soriano (1995).

Positive-Negative System

i Genomic Locus



ii Targeting Vector



iii Targeted Locus



iv Random Insertion



seen with the positive-negative selection system are much more modest and generally fall in the range of 5 to 10 fold.

3.2 Targeting Systems for the Introduction of Independent Subtle Alterations into the Mouse Genome

3.2.1 *The Need For Targeting Systems Which Permit the Introduction of Independent, Subtle Alterations into the Mouse Genome*

To gain a complete understanding of the *in vivo* roles of a gene product and the control of its expression, analysis of mice carrying independent precise, subtle alterations in the gene of interest will be required. These might include alterations to the control elements and the protein itself, such as to functional domains, structural motifs, and single amino acids. Using this sort of approach, promoter and/or enhancer elements that regulate expression throughout development and/or in a tissue specific manner could be systematically identified and inactivated. This would help to define the regulatory hierarchies controlling expression of the gene of interest.

Up to 35% of the null mutations produced by gene targeting to date have been embryonic or neonatal lethal (Brandon *et al.*, 1995a) and is a particular problem in the case of the developmental control genes. This complicates analysis of the mutant phenotype and can make functional analysis of a gene product difficult or impossible. Using sophisticated gene targeting approaches, the elements required for expression at later stages of development and in the adult could be inactivated whilst those responsible for expression during embryonic development are left intact. This would overcome the problems associated with embryonic lethality of the developmental control genes and thereby permit later developmental roles to be analysed.

In addition to these goals, the ability to introduce independent, precise alterations into the mouse genome would have medical applications, particularly to the study of genetic disease. Gene targeting in ES cells has been used for the creation of mouse models for human genetic

disease (reviewed in Melton, 1994). For example, mouse models for cystic fibrosis (CF) have been generated using by gene targeting. Mice carrying completely or partially inactivated cystic fibrosis transmembrane conductance regulator (CFTR) genes have been produced, which vary in the severity of symptoms (reviewed in Melton, 1994). Many of the symptoms of human CF are exhibited by these mice, but in most cases the severity of the symptoms is greater than of the human disease. This is due to the fact that the most common defect in the human CFTR gene is a single amino acid substitution rather than a partial or complete absence of the protein. Thus, mice carrying the analogous amino acid substitution are likely to make the best models for the human disease. A number of other important human genetic diseases are also most commonly caused by minor changes to proteins at the amino acid level (reviewed in Melton, 1994). Thus, effective production of mouse models which closely resemble human genetic disease requires the development of targeting systems which permit the introduction of subtle changes into the mouse genome. The fact that a single human genetic disease in the population can be caused by a range of subtle mutations at the protein level, or that different mutations in a single gene can cause a variety of symptoms (Detloff *et al.*, 1994), also suggests that introduction of multiple alterations to disease associated loci will be required to produce accurate models for human disease.

The ability to repeatedly introduce alterations to the mammalian genome by gene targeting may also have applications for the improvement of pharmaceutical and agricultural products. For example, Stacey and co-workers (1994) have exchanged the mouse α -lactalbumin gene, a protein found in milk, for its human counterpart by gene targeting. This has been carried out with the commercial aim of altering the composition of milk for therapeutic and pharmaceutical applications.

The main obstacle to the repeated introduction of alterations to a single site within the genome by gene targeting is the inefficiency of homologous recombination compared to non-homologous recombination. When alterations are to be made to genes for which direct selection can not be applied, the majority of transfected lines are non-homologous recombinants. The ratio of homologous to non-homologous recombinants is generally in the

range of 1 to 10 percent but can be lower depending on the locus being altered. Thus, a large number of transfected ES cell lines must be grown up and screened in order to detect the relatively rare homologous recombinants. This is a time consuming and laborious process which hampers or precludes the repeated introduction of alterations into a single locus. To realise the full potential of gene targeting, targeting systems with improved efficiency are required to permit the repeated introduction of precise, non-selectable alterations to a single site of the genome.

3.2.2 Targeting Systems for the Introduction of Subtle Alterations into the Mouse Genome

A number of gene targeting systems have been described which permit the introduction of subtle alterations into genes of interest. In the hit-and-run system, also known as the in-and-out system, (Hasty *et al.*, 1991a; Valancius and Smithies, 1991; Fig. 3.3) the genetic alteration is introduced into the target gene by homologous recombination using an insertion vector which also carries a positive/negative selection cassette. This event leads to duplication of the target sequence and is positively selected. This is followed by negative selection for loss of the positive/negative selection cassette by an intrachromosomal recombination event in which the duplicated sequence is removed while the subtle alteration remains on the targeted allele (Fig. 3.3 iv). A frequent problem with this approach, is loss of the subtle alteration during the intrachromosomal recombination event. The approach is further limited by the fact that, although it provides a means to introduce subtle alterations into the genome, it does not afford any advantage when multiple independent alterations at a single locus are required.

An alternative approach has been to introduce the subtle alteration and a selection cassette, which remains on the altered allele, using a single replacement vector (Rubinstein *et al.*, 1993). A precise, subtle alteration has been introduced into the coding region of the mouse pro-opiomelanocortin (POMC) gene on a replacement vector which contained a *neo* positive selection cassette immediately 3' of the coding region (Rubinstein *et al.*, 1993). Although the *neo* selection cassette remained on the altered allele it did not activate transcription of the normally non-expressed POMC gene in ES cells in vitro. However, as in the case of the hit

Figure 3.3 Hit-and-Run targeting approach

Representation of the hit-and-run gene targeting approach with (i) the genomic locus, (ii) targeting vector, (iii) targeted locus and (iv) targeted locus after intrachromosomal recombination shown. Thin line, genomic DNA; thick line, non-genomic DNA; numbered black boxes, exons; hatched box, *neo* cassette; stippled box, HSV-*tk* cassette; white box, subtle alteration.

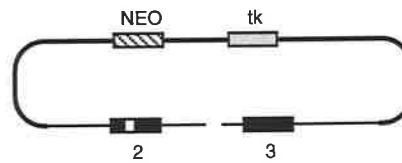
Adapted from Soriano (1995).

Hit-and-Run

i Genomic Locus



ii Targeting Vector



iii Targeted Locus



Intrachromosomal
Recombination



iv Targeted Locus



and run strategy, the system does not afford any advantage when repeated alterations to single locus are required.

3.2.3 *Gene Targeting Systems Which Permit the Repeated Introduction of Precise, Subtle Alterations into the Mouse Genome*

In the targeting system first suggested by Reid and co-workers (1990) known as double-replacement (Stacey *et al.*, 1994 ; Wu *et al.*, 1994) or tag-and-exchange (Askew *et al.*, 1993; Fig. 3.4; referred to as tag-and-exchange in this thesis) two successive gene targeting steps are required to introduce subtle alterations into the site of interest. A positive/negative selection cassette or tag gene is introduced into the locus to be altered in a first round of gene targeting (the tagging step) in which null alleles can be produced (Fig. 3.4 ii, iii). This step, for which direct selection can not be applied, is carried out once at any given locus and the tag gene provides direct selection for subsequent homologous recombination at this site. The tag gene is replaced by the alteration of interest in a subsequent round of gene targeting, known as the exchange step (Fig. 3.4 iv), in which selection for loss of the tag gene is applied. This feature of the system prevents survival of non-homologous recombinants in the exchange step and therefore increases the targeting efficiency at this site. This system obviates the need for exogenous selection cassettes on exchange step vectors permitting the introduction of subtle alterations into the locus under investigation. It was hoped that the improved efficiency of the exchange step would permit the repeated introduction of alterations into the tagged site.

In practise this system has been shown to be ineffective due to a high background of resistant lines surviving the exchange step selection due to physical loss of the tag gene by means other than homologous recombination (Stacey *et al.*, 1994 ; Wu *et al.*, 1994). Likely mechanisms responsible for physical loss of the tag gene include chromosome loss, chromosome deletion, or gene conversion in which the tagged allele is restored to wildtype. As a result of this, exchange step targeting efficiencies are much lower than expected, making the repeated introduction of alterations to a single gene laborious. In fact in several published examples of the tag-and-exchange approach the efficiency of the exchange step has been lower than that of the tag step in which direct selection was not applied (Askew *et al.*, 1993;

Figure 3.4 Tag-and-exchange gene targeting approach

Representation of the tag-and-exchange gene targeting approach with (i) the genomic locus, (ii) tag step targeting vector, (iii) tag step targeted locus and exchange step targeting vector and (iv) exchange step targeted locus shown. Thin line, genomic DNA; numbered black boxes, exons; stippled box, HPRT cassette; white box, subtle alteration.

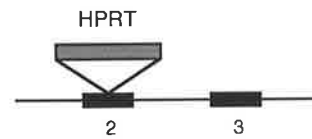
Adapted from Soriano (1995).

Tag-and-Exchange

i Genomic Locus



ii Tag Targeting Vector



iii Tagged Locus



Negative
Selection



iv Targeted Locus



Stacey *et al.*, 1994; Wu, 1994). This strategy has therefore afforded no real advantage over the hit-and-run approach.

3.3 An Efficient Gene Targeting System for the Repeated Introduction of Non-Selectable Alterations Into the Mouse Genome

3.3.1 *The Stable Tag-Exchange Gene Targeting System*

The second aim of this thesis was to design and characterise an efficient gene targeting system which permits the repeated introduction of precise, subtle alterations into the mouse genome. The system described and characterised in this thesis is based on the tag-and-exchange system and was designed to overcome the high background of non-homologous recombinants which survive during exchange step targeting due to physical loss of the tag cassette by means other than homologous recombination. The modified tag-and-exchange system was termed the stable tag-exchange approach.

The system incorporates an additional positive selection cassette, termed the stability cassette, positioned downstream of the coding region on the tag step vector (Fig. 3.5). This cassette, which is not replaced during the exchange step, provides a means to select for the presence of the tagged allele during exchange step gene targeting. The stable tag-exchange strategy is based on the premise that physical loss of the tag cassette by means other than homologous recombination would also lead to loss of the stability cassette. Thus, selection for the presence of this cassette and consequently the tagged allele, can be maintained during the exchange step. Lines which have physically lost both the tag cassette and the stability cassette by means other than homologous recombination would be eliminated by dual selection while cell lines which have lost the tag cassette due to homologous recombination but have retained the stability cassette, would survive. It was therefore envisaged that close to one hundred percent of cells surviving in the exchange step would be homologous recombinants (Fig. 3.5 iv). Direct selection during exchange step gene targeting obviates the need for exogenous selection cassettes on exchange step targeting vectors, permitting the introduction of precise,

Figure 3.5 Stable Tag-Exchange gene targeting approach

Representation of the stable tag-exchange gene targeting approach with (i) the genomic locus, (ii) tag step targeting (iii) tag step targeted locus, (iv) exchange step targeting and (v) exchange step targeted locus and expected gene targeting frequencies shown. Thin line, genomic DNA; numbered black boxes, exons; stippled box, HPRT cassette; hatched box, *neo*; white box, subtle alteration.

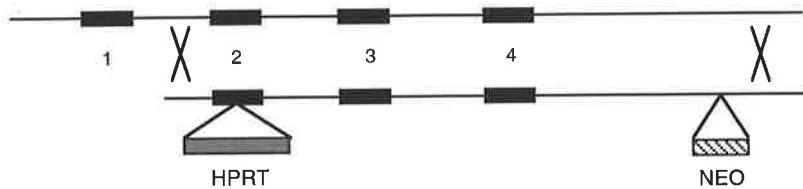
Stable Tag-Exchange

i Genomic Locus



ii Tag Targeting

Frequency 0.01 - 50 %



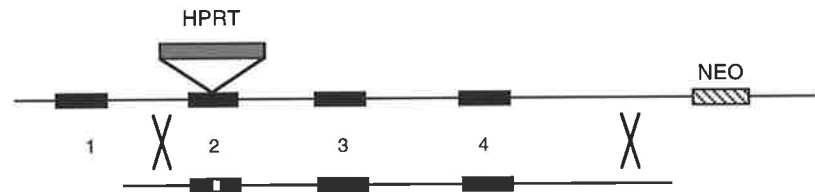
HAT/G418

iii Tagged Locus



iv Exchange Targeting

Frequency ~100 %



6-TG/G418

v Targeted Locus



subtle alterations to the tagged site while the improved efficiency of the system would permit the repeated introduction of alterations to the genome.

3.3.2 *The Selection Systems of the Stable Tag-Exchange Approach*

We have used a tag vector carrying an hypoxanthine phosphoribosyl transferase (HPRT) expression cassette driven by the 5' thymidine kinase (TK) promoter and mutant polyoma enhancer (Reid *et al.*, 1990) as the tag gene, and the *neo* gene, driven by the mouse phosphoglycerate kinase-1 (*Pgk-1*) promoter as the stability cassette (Fig. 3.5).

In order to use the selection system based on the HPRT gene, ES cells which lack a functional copy of the X-linked HPRT gene are required. The product of the HPRT gene catalyses an early step in the salvation pathway for purine synthesis. HPRT gene expression in mammalian cells provides resistance to the drug amethopterin present in hypoxanthine amethopterin thymidine (HAT) medium, while it's absence provides resistance to the base analogue 6-thioguanine (6-TG).

Hooper and co-workers (1987) isolated ES cell lines carrying spontaneous mutations in the HPRT gene. The HPRT deficient ES cells derivatives retained germline competence making the HPRT selection system suitable for the production of mice. A germline competent HPRT⁻ derivative of the E14 ES cell line termed E14TG2a containing a deletion in the HPRT gene (Hooper *et al.*, 1987) was used in this project. The deletion begins at least ten kilobases 5' to the HPRT gene promoter and extends into the second intron of the gene making reversion events unlikely (Doetschman *et al.*, 1987). Expression of the *neo* gene provides resistance to the aminoglycoside G418 in mammalian cells and dual selection for these genes can be applied to ES cells without loss of pluripotence.

Using the HPRT⁻ E14TG2a ES cell line, first round targeted (tagged) lines are selected in medium containing HAT and G418. The tag gene, but not the stability cassette, is then replaced by the alteration of interest in a subsequent round of targeting (exchange step), and is selected in 6-TG and G418. Homologous recombinants which have lost expression of the

tag cassette but not the stability cassette will survive selection in 6-TG and G418, and should carry the desired alteration.

3.4 Leukaemia Inhibitory Factor (LIF)

3.4.1 Identification and Characterisation of LIF

Maintenance of undifferentiated ES cells in vitro can be achieved by growth in media conditioned by a number of cell lines including the Buffalo rat liver (BRL) epithelial cell line (Smith and Hooper, 1987). The secreted factor responsible for this activity was purified from BRL conditioned medium and termed differentiation inhibiting activity (DIA) (Smith *et al.*, 1988). Subsequent functional and sequence analysis of DIA revealed it to be identical to the mouse cytokine previously known as D-factor (Tomida *et al.*, 1984), now known as LIF (Gearing *et al.*, 1987; Hilton *et al.*, 1988).

The active molecule is a basic, hydrophobic, single amino acid chain glycoprotein of 43,000 M_r with a core polypeptide of 180 amino acids of M_r 20,000. LIF is a member of a family of cell signalling molecules that are characterised by a four α -helix bundle topology (Robinson *et al.*, 1994) typified by the prototype member of this group: growth hormone (GH). LIF has an up-up-down-down four-helix bundle conformation which is shared by the haematopoietic growth factors (Robinson *et al.*, 1994). Signal transduction by LIF requires interaction of the ligand with a transmembrane, cell surface receptor subunit (LIFR), a member of the cytokine-binding family of receptor subunits (Cosman, 1993) and association of this complex with the transmembrane signalling molecule, gp-130 (Gearing *et al.*, 1992; Gearing and Bruce, 1992). Signal transduction by LIF is complex and involves activation of protein kinases including Hck (Ernst *et al.*, 1994).

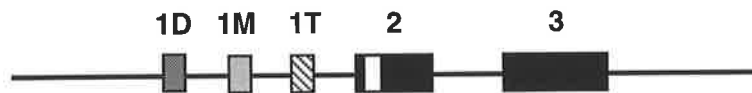
3.4.2 The LIF Gene

The genomic organisation of the mouse LIF gene is shown in figure 3.6 i. Two distinct LIF transcripts which vary at the 5' end have been identified in the mouse and have been termed D and M for diffusible and matrix associated, respectively (Rathjen *et al.*, 1990a). These

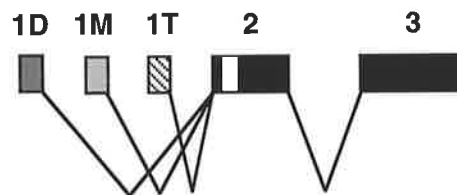
Figure 3.6 Genomic organisation of the mouse LIF locus

Representation of the mouse LIF locus showing (i) the genomic organisation (Stahl *et al.*, 1990; P. Rathjen, personal communication) of the mouse LIF gene, (ii) the alternate LIF transcripts and (iii) the spliced LIF transcripts (ATG, translation initiation codon; arrow, cleavage point of the mature protein). Line, genomic sequence; dark grey box, first exon of LIF D; light grey box, first exon of LIF M; hatched box, first exon of LIF T; black boxes, exons two and three of LIF; white box, signal sequence.

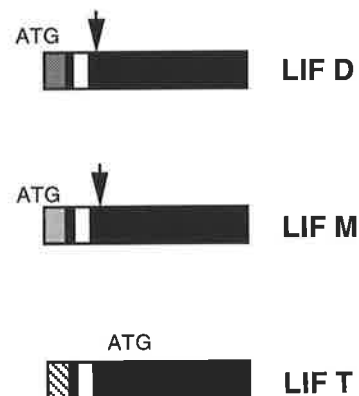
i LIF gene



ii Alternate Transcripts



iii Spliced Transcripts



transcripts have unique first exons which are spliced to common second and third exons of the LIF gene (Fig. 3.6 ii). This leads to alternative forms of the LIF protein with variant amino terminal sequences, but with an unchanged hydrophobic secretory signal and mature protein. The alternative first exons determine the cellular localisation of the protein. LIF translated from the D transcript is secreted from the cell in a diffusible form (LIF D), whereas that translated from the M transcript remains associated with the extracellular matrix (ECM) (LIF M) (Rathjen *et al.*, 1990a).

A third LIF transcript in the mouse has recently been identified and characterised (P. Rathjen, B. Haines personal communication). This transcript carries an alternate first exon which is also spliced to the common second and third exons of the LIF gene and encodes a LIF protein referred to as LIF T. Lack of an initiation codon in the first exon of this transcript causes translation to initiate at the first ATG within the second exon of LIF (Fig. 3.6 iii). Translation of the LIF T transcript begins downstream of the secretory signal in exon two of LIF (Fig. 3.6, open box) and leads to the generation of a truncated form of LIF (17,000 M_r core polypeptide) (B. Haines, P. Rathjen, personal communication). The lack of a secretory signal on this form of LIF suggests that this protein may have alternative cell autonomous roles (B. Haines, P. Rathjen, personal communication).

Evidence suggests that expression of the LIF D, M and T transcripts are controlled by distinct promoters (Rathjen *et al.*, 1990a and b; B. Haines, P. Rathjen, personal communication), which permit differential regulation of the LIF isoforms both in in vitro cell lines (Rathjen *et al.*, 1990a; 1990b) and in the mouse (Rathjen *et al.*, 1990a, Robertson, 1993).

3.5 Functional Analysis of LIF

3.5.1 *In vitro* Activities of LIF

Extracellular LIF is a multi-functional cytokine which can cause distinct biological effects on a wide range of cells in vitro. For example LIF inhibits the differentiation of ES cells, renal cells/kidney rudiments and Da-1a myeloid cells, increases the survival and/or proliferation of haematopoietic progenitors, primordial germ cells and sensory neurons, induces terminal

differentiation of the mouse leukaemic cell line M1, induces cholinergic differentiation of sympathetic neurons, stimulates bone resorption in foetal calvaria, inhibits bone resorption in foetal long bones, induces the acute phase response in hepatocytes (reviewed in Hilton, 1992; Metcalf, 1992) and improves the survival and development of the eight-cell mouse embryo (Lavranos *et al.*, 1995). LIF T has distinct biological roles compared to those of LIF D and M in a terminally differentiated cell type in vitro. Overexpression of LIF T in COS-1 cells in vitro causes apoptosis of these cells and this effect is not seen with either LIF D or LIF M (B. Haines, personal communication). This suggests that the LIF T protein may display biologically distinct roles in vivo. Elucidation of the normal in vivo roles of this LIF isoform will therefore be of particular interest.

This diverse range of activities in vitro suggests that, similar to other cytokines, LIF may be used repeatedly throughout embryonic and adult life to mediate a range of distinct biological activities that are determined by the responding cell. Many of LIF's actions in vitro on a range of cells including haematopoietic, osteogenic, primordial germ cells and ES cells are consistent with a role in maintenance of the stem cell pool both during embryogenesis and in regenerating tissues of the adult. The proposal that LIF may act as a generalised stem cell factor (Smith and Rathjen, 1991; Smith *et al.*, 1992) however, requires further investigation.

3.5.2 *Effects of Gain of Function LIF Mutations in the Mouse*

The equivalent of gain of function mutations in the LIF gene have been created by overexpression of LIF in mice or exogenous application of LIF to mice. Two independent studies have been carried out in which mice were engrafted with haematopoietic cell lines overexpressing LIF. Metcalf and Gearing (1989) saw a range of symptoms that led to death within 2 to 3 months when mice were engrafted with the haematopoietic cell line FD. The engrafted mice were hyperexcitable, had reduced body fat, calcium depositions in the heart, skeletal muscle and liver, osteoblast formation in the marrow cavity, loss of corpora lutea formation and spermatogenesis, elevated acute phase response proteins and showed atrophy of the thymus in addition to effects on the pancreas and adrenal cortex. The primary cause of death was due to effects on calcium metabolism and osteogenesis which led to calcification of

essential organs such as the heart. These phenotypes show that LIF can have effects on calcium and lipid metabolism, osteogenesis, acute phase response and the autonomic nervous system. Major effects on the haematopoietic system were not seen in this study. In contrast to this, Shen and co-workers (1994) observed primarily haematopoietic effects when they engrafted mice with T cells overexpressing LIF. These included increased megakaryocyte formation, extramedullary haematopoiesis, thymic and lymph node abnormalities, in addition to induction of the acute phase response seen in both studies. Reduced body fat or calcium deposits in cardiac or skeletal muscle were not observed.

Elevated levels of LIF *in vivo* have also been achieved by intraperitoneal injection with high doses of LIF (Metcalf and Gearing, 1989). The treated mice showed some symptoms in common with both of these studies. The mice were hyperexcitable, had reduced body fat, increased megakaryocyte and platelet formation and elevated levels of the acute phase response proteins.

To assess the effect of LIF overexpression on embryogenesis, Conquet and co-workers (1992) returned ES cells which overexpressed either the diffusible or matrix associated form of LIF to the blastocyst. The phenotypic consequences on the host blastocysts were assessed at various stages of gastrulation. Embryos overexpressing diffusible LIF were phenotypically normal whereas those overexpressing matrix associated LIF died by 9.5 d.p.c. as a result of aberrant development of the primitive ectoderm which normally gives rise to all embryonic tissues. In these embryos, development of mesoderm from the primitive ectoderm did not occur. This result suggests that matrix associated LIF may play a role in the maintenance of the stem cell pool of the embryo.

Limitations of these analyses are twofold. Firstly, the particular phenotype seen is dependent on the route of administration and quantity of LIF used in each study. Secondly, although this sort of approach gives an indication of the potential *in vivo* roles of LIF, it does not give any information about the actual roles LIF is playing during normal development and adult life.

3.5.3 Effects of Loss of Function LIF Mutations in the Mouse

To investigate the roles of LIF in normal mouse embryonic development and in the adult, LIF null mutant mice were generated by gene targeting by two independent groups. Mice lacking expression of all forms of LIF are grossly normal (Stewart *et al.*, 1992; Escary *et al.*, 1993). Homozygous LIF⁻/LIF⁻ mice displayed a retarded postnatal growth rate and adults were 25-35% smaller in body weight than their heterozygous or wild type siblings (Stewart *et al.*, 1992). However, maternal LIF expression was absolutely essential for uterine implantation of the mouse blastocyst (Stewart *et al.*, 1992). Although homozygous mutant female are fertile, lack of LIF expression in the uterine environment prevented blastocyst implantation in these mice. In addition to this finding, Escary and co-workers (1993) have reported that mice lacking LIF expression also have vastly reduced numbers of haematopoietic stem cells in the spleen and bone marrow. The observed effects on mouse size and the haematopoietic system are consistent with other findings (Lavranos *et al.*, 1995; Metcalf and Gearing, 1989; Shen *et al.*, 1994).

From these results, an essential role for LIF in implantation of the mouse blastocyst has been established, however other essential roles for LIF were not identified. In the privileged environment of the laboratory, LIF expression in the mouse is not essential for normal embryonic or adult survival. At least three possible explanations for this exist. Firstly, LIF may not play essential roles in the mouse. Secondly, essential roles for LIF may only become apparent when the LIF⁻/LIF⁻ mice are placed in a wildtype environment where they are subjected to normal environmental challenges. Thirdly, LIF may play essential roles in the mouse which can be reproduced by other related cytokines (Robertson *et al.*, 1993). The family of cytokines that utilise the signal transduction molecule gp130 continues to increase and there is a great deal of overlap in activities of many of these molecules. This suggests that compensation by other factors in vivo may override defects caused by a lack of LIF expression. Subtle defects in the LIF⁻/LIF⁻ mice may only become evident upon much closer examination of the mutant mice or upon cross breeding with other null mutant mice.

The effects seen in mice lacking expression of the transmembrane LIFR are more severe than those seen in the LIF null mutant mice (Ware *et al.*, 1995; Li *et al.*, 1995). These mice die at or soon after birth from a range of placental, skeletal, neural and metabolic defects. The contribution to these phenotypes by LIF is unknown because the LIF-related cytokines CNTF and CT-1, and possibly as yet unidentified ligands, utilise the LIFR to mediate signal transduction. It was found in this study that LIF expression in the uterus at the time of implantation does not act directly on the blastocysts via the cell surface receptor (Ware *et al.*, 1995).

3.6 LIF Expression

3.6.1 Control of LIF Expression

Little is known about the sequences that control expression of the various forms of the LIF protein (Kaspar *et al.*, 1993). TATA-like elements positioned upstream of the transcription start site of LIF D are likely to have regulatory functions on expression of this form of the protein (Kaspar *et al.*, 1993). Additional control elements for the LIF gene are likely to be contained in a conserved, non-methylated CpG island at the 5' end of the LIF gene. This region, which starts within the untranslated region of the first exon of LIF D and ends between the first exons of LIF M and T, is able to bind nuclear proteins suggestive of a role in transcriptional regulation (Kaspar *et al.*, 1993). A distal enhancer positioned upstream of the LIF gene, between nucleotides -1200 and -3200, which is required for transcription of the D and M forms of LIF has also been identified (Hsu and Heath, 1994). This sequence contains minimal steroid response elements, putative DNase I hypersensitive sites and a region of alternating purine/pyrimidine nucleotides.

3.6.2 Sites of LIF Expression

Extracellular LIF exerts biological effects at picomolar concentrations due to high affinity for its receptor (Robertson *et al.*, 1993) and consequently the level of expression of LIF in most tissues of the mouse is extremely low. This fact has hampered expression mapping of LIF in the mouse. LIF expression has been detected in the preimplantation blastocyst by reverse transcription followed by PCR (RT-PCR) (Conquet and Brûlet, 1990; Murray *et al.*, 1990)

and in extraembryonic tissues of the embryo up to midgestation stage (12.5 d.p.c.) by in situ hybridisation and RT-PCR (Conquet and Brûlet, 1990). Similarly, LIF expression has been detected in the egg cylinder stage embryo at 6.5 d.p.c with higher levels detected in the extraembryonic tissues at 6.5 and 7.5 d.p.c. by RNase protection (Rathjen *et al.*, 1990b).

Limited LIF expression has been detected in the neonate, specifically in skin and calvaria, by RNase protection or Northern analysis (Bhatt *et al.*, 1991). The highest levels of LIF expression in the adult are seen in the uterus. LIF expression has been detected in the uterus of adult virgin female mice and the level of expression fluctuates with the oestrous cycle (Shen and Leder, 1992; Smith *et al.*, 1992). In contrast to this, Bhatt and co-workers (1991) failed to detect LIF expression in virgin mice, but detected LIF expression in the uterus during pregnancy. Uterine LIF expression has been localised to the endometrial glands of the uterus by in situ hybridisation (Bhatt *et al.*, 1991; Smith *et al.*, 1992) and has been shown to peak at the stage of blastocyst implantation (Bhatt *et al.*, 1991). Low level LIF expression has been reported by Bhatt and co-workers (1991) in limited tissues of the adult using RNase protection or Northern analysis.

In contrast to other reports (Bhatt *et al.*, 1991; Shen and Leder, 1992; Hilton, 1992) widespread, low level LIF expression has been detected in many tissues of the adult using a highly sensitive RNase protection assay (Robertson *et al.*, 1993). Relatively high levels of LIF expression in the adult were detected in the brain and the major self-renewing tissues including the skin, intestine and lung in addition to high levels in the female reproductive tract. LIF expression was also detected throughout late gestation and in the neonate in this study. LIF expression was detected during organogenesis of the heart, lung and intestine and in the calvaria. Expression at many of these sites persists in the neonate. Differential expression of the D and M transcripts was seen and the expression of each was regulated throughout development and in a tissue specific manner (Robertson *et al.*, 1993).

The low level LIF expression detected by Robertson and co-workers (1993) represented less than 10 transcripts per cell if equally expressed in all cells of a given tissue. However, it is

possible that a subset of cells within a given tissue express LIF at higher levels. High levels of LIF expression detected in the major self-renewing tissues of the adult led Robertson and co-workers (1993) to conclude that LIF expression in the adult is associated with tissues in which a pool of stem cells remain. This is supporting evidence for the theory that LIF expression is required for maintenance of the stem cell niche (Rathjen *et al.*, 1990b; Smith and Rathjen, 1991). Resolution of LIF expression at the cellular level would be required to resolve this (Robertson *et al.*, 1993).

Limitations of the approaches used for the detection of LIF expression *in vivo* are several fold. Firstly, the methods used detect expression of the LIF transcript(s) which do not necessarily correlate with expression of the LIF protein. Secondly, in general the techniques used have been insufficiently sensitive to detect very low levels of LIF expression. This may be an inherent problem with detection of the LIF transcript(s) which contain(s) several copies of a motif implicated in turnover of short-lived transcripts (Mountford *et al.*, 1994). Finally, the techniques used, with the possible exception of *in situ* hybridisation, are unable to resolve expression at the cellular level.

The information that we now have about the expression of LIF D, M and T indicates that these forms of LIF are regulated throughout development and adulthood and that even in a single tissue there is tight control over the particular form of LIF expressed (Rathjen *et al.*, 1990b; Robertson *et al.*, 1993). This suggests that each LIF isoform plays distinct biological roles. LIF T is also likely to exert biologically distinct roles due to its truncated form and cellular localisation. To gain an accurate understanding of the normal *in vivo* roles of each LIF isoform a mapping technique is required which will enable sensitive resolution at the cellular level of the sites of expression of each LIF isoform at the protein level.

3.7 A Gene Targeting Approach to Detect the Cellular sites of LIF Expression

3.7.1 The Gene Targeting Approach

LIF can not be detected in the circulation (Metcalf, 1992) and the immobilised matrix associated form of the protein can only act in an autocrine manner or affect cells in close proximity to the expressing cell. This suggests that LIF acts over very short ranges. Thus, determination of the cellular sites of LIF expression should give a good indication of the sites of LIF activity in the mouse.

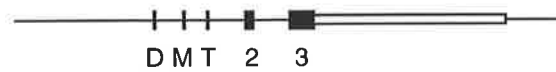
In work presented in this thesis, a gene targeting approach was used to produce mice in which the cellular sites of LIF expression could be identified. ES cell lines have been produced which carry reporter genes in the appropriate exon of the LIF gene to enable expression of all and each transcript of LIF to be mapped in vivo. To achieve this, a reporter gene was introduced into the common second exon of LIF and the unique first exons of LIF D, M and T by gene targeting in ES cells. An example of gene targeting with a vector carrying a reporter gene in exon 1 of LIF D is shown in figure 3.7. Once returned to the mouse, the offspring of germline transmitting chimaerae will be bred to produce heterozygous mice carrying one genetically marked allele of LIF. Development of these mice will proceed normally due to expression of the unaltered LIF allele (Stewart *et al.*, 1992; Escary *et al.*, 1993). The cellular sites of each form of LIF and of all forms together can then be detected during embryogenesis and in the adult.

The requirement for the repeated introduction of precise alterations into the LIF locus justified use of the stable tag-exchange gene targeting approach described in section 3.3.1. The stable tag-exchange system should be suitable for the introduction of reporter genes into the mouse genome since exchange step vectors lack exogenous selection cassettes and the stability cassette introduced during the tag step can be positioned well away from the coding region of the gene under investigation. A *neo* gene under the transcriptional control of the *Pgk-1* promoter positioned immediately 3' of the last exon of POMC (Rubinstein *et al.*, 1993) did not activate POMC transcription in ES cells in which the gene is normally silent. Similarly, in

Figure 3.7 Stable tag-exchange gene targeting at the LIF locus

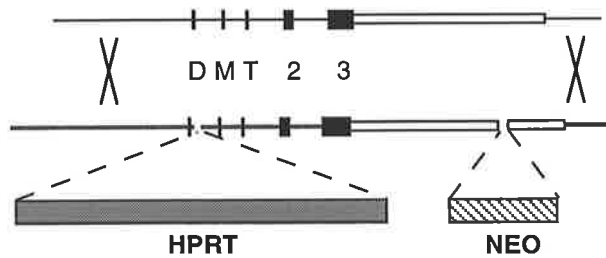
Representation of stable tag-exchange gene targeting at the LIF locus showing the (i) LIF locus, (ii) tag step gene targeting, (iii) tagged LIF locus, (iv) exchange step gene targeting with a vector carrying a reporter in the first exon of LIF D and (v) targeted LIF locus and selection used throughout targeting. Line, genomic sequence; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR, stippled box, HPRT; hatched box, *neo*; light stippled box, reporter.

i LIF Locus



ii Tag Step

Gene Targeting



HAT/G418

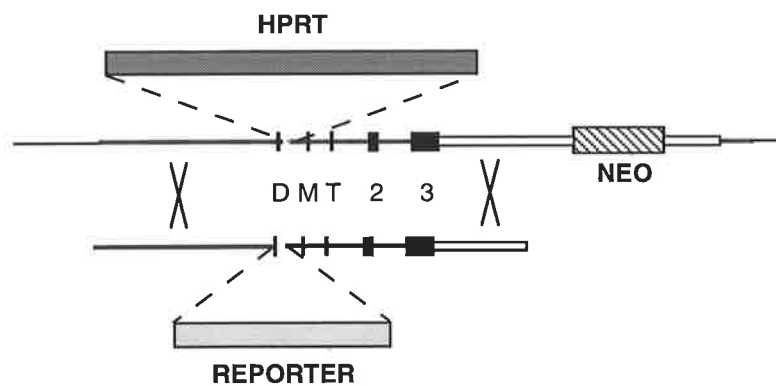
iii Tagged

LIF Locus



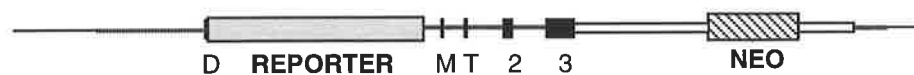
iv Exchange Step

Gene Targeting



6-TG/G418

v Targeted Locus



mice generated from targeted ES cells carrying a *lacZ* reporter gene and downstream *neo* expression cassette in the *Hox-3.1* gene, reporter gene expression was not affected by the presence of the *neo* gene (Le Mouellic *et al.*, 1992). These findings suggest that a *neo* gene positioned well downstream of the LIF coding region should not affect expression of LIF gene.

Using the stable tag-exchange approach, the exchange step gene targeting was carried out with four separate targeting constructs which introduced the reporter gene into the appropriate exon of the LIF gene. The exchange step vectors used to alter the first exons of LIF D and M, carried an ATG⁻ reporter gene fused in frame in the first exon of LIF D and M. Translation of the fusion transcripts should therefore reflect expression of the LIF D and M proteins. The exchange step vectors used to alter the first exon of LIF T (which lacks an initiation codon) and the second exon of LIF carried an ATG⁺ reporter gene. In these cases translation of the reporter gene should therefore reflect expression of the LIF T and all LIF transcripts.

3.7.2 The LIF Reporter System

Due to the extremely low levels of LIF expression in the mouse, the choice of reporter/marker gene for use in this project was critical. Ideally, the reporter gene used would allow detection of LIF protein expression at the cellular level and be highly sensitive. The ability to quantify the levels of LIF expression was not a requirement of this project. A number of reporter genes frequently used in mammalian cells allow the level of expression to be quantified but do not permit detection at the cellular level.

Sensitive detection systems exist for the product of the luciferase gene (Ow *et al.*, 1986) and green fluorescent protein (Chalfie *et al.*, 1994) but limitations of both these reporters make them unsuitable for use in this project. Direct detection of the firefly luciferase gene activity can not be carried out in situ (Promega Corporation, 1995) and the very short halflife of the luminescence produced by this reporter gene precludes it's use in this study. The green fluorescent protein can be detected in situ at the cellular level but was unsuitable for use in this project as the protein readily diffuses through biological membranes, and the

fluorescence produced by this protein has a short half life (Chalfie *et al.*, 1994) making cellular identification difficult. Less commonly used, and consequently less well characterised, reporter genes for use in transgenic mice which permit in situ cellular detection include human placental alkaline phosphatase (Berger *et al.*, 1988) and *D. melanogaster* alcohol dehydrogenase (Lin and Culp, 1991).

The bacterial β -galactosidase (*lacZ*) (Gorman *et al.*, 1982) gene has been used extensively as a reporter gene in the mouse and expression can be visualised in situ at the cellular level (reviewed in Cui *et al.*, 1994). The enzyme product of the *lacZ* gene catalyses hydrolysis of β -galactosides and incubation with the chromogenic substrate 5-bromo-4-chloro-indolyl- β -D-galactoside (X-gal) produces a blue precipitate which can be visualised at the single cell level in cells including mammalian cells (Hughes and Blau, 1990; Weis *et al.*, 1991). The *lacZ* gene was the most suitable reporter gene for use in this project, since it provides a visual marker of gene expression at the cellular level (Cui *et al.*, 1994) and since highly sensitive immunological and fluorescent detection systems for β -galactosidase activity exist.

The *lacZ* gene has been used as a reporter gene in transgenic animals produced by conventional means (reviewed in Cui *et al.*, 1994), as a reporter in promoter trap experiments (Friedrich and Soriano, 1991) and in mice altered by gene targeting (Mansour *et al.*, 1990; Le Mouellic *et al.*, 1992; Mountford *et al.*, 1994). In 1990, Mansour and co-workers inactivated a single copy of the *int-2* locus with a *lacZ* reporter gene introduced by gene targeting. Expression of the *int-2* - β -gal fusion protein reflected that of the *int-2* gene. This approach allowed visualisation of *int-2* expression in ES cells and their differentiated progeny in vitro and in hetero- and homozygous mice in vivo. In a similar approach, Le Mouellic and co-workers (1992) used the *lacZ* gene to inactivate the *Hox-3.1* gene and provide a marker for *Hox-3.1* expression during embryogenesis. The pattern of *lacZ* expression seen in this study reflected the known *Hox-3.1* transcription pattern throughout embryogenesis in heterozygous and homozygous mutant mice. These examples suggest that the *lacZ* gene is a suitable reporter gene for use in mice generated from targeted ES cells.

A nuclear localised form of the *lacZ* was used in the exchange step vectors. Thus, β -galactosidase activity located in the nucleus of an expressing cell should reflect LIF expression. This provides the dual advantage of increasing sensitivity and overcoming the problem of diffusion of β -galactosidase activity to adjacent cells. Several copies of a motif implicated in the turnover of shortlived transcripts are contained in the UTR of the LIF transcript (Mountford *et al.*, 1994) and are likely to increase the turnover of the LIF message and consequently reduce the levels of β -galactosidase activity produced. A *lacZ* gene carrying its own polyadenylation sequence was therefore used to prevent potentially rapid turnover of the β -galactosidase transcript. Use of the reporter gene approach described in this project should overcome the limitations of previous approaches to LIF expression mapping in vivo.

In establishing the stable tag-exchange system at the LIF locus, two outcomes were achieved. By tagging the LIF locus with a positive/negative selection cassette and stability cassette, an ES cell line has been produced in which subsequent homologous recombination at this locus can be directly selected. This should permit the efficient introduction of any desired mutation into the LIF locus of ES cells for analysis in vitro and in vivo. This should facilitate the analysis of the LIF protein itself and the elements which control expression of the protein in the context of the whole organism. Secondly, the production of ES cells and subsequently mice, in which the sites of expression of each and all isoforms of LIF can be visualised will sophisticate our knowledge of this cytokine which is suspected of having roles in a range of developmental processes. Identification of the cells expressing each form of LIF should provide hitherto unobtainable information about the roles of each form of this protein during embryogenesis and in the adult.

Chapter 4

4 Analysis of Stable Tag Step Gene Targeting at the LIF Locus of ES Cells

4.1 Introduction

The stable tag-exchange gene targeting system described in chapter 3 (section 3.3.1) was designed to permit the efficient, repeated introduction of non-selectable, subtle alterations into the mouse genome. The system has been characterised at the LIF locus. In the first step of this targeting approach, the tag and stability cassettes have been introduced into the LIF locus. In so doing, the efficacy of the tag step has been assessed and a system for the efficient introduction of non-selectable alterations into the endogenous LIF gene has been established. Construction of the tag step targeting vector, tag step targeting at the LIF locus in ES cells and analysis of the tag step of the stable tag-exchange targeting approach is reported in this chapter.

4.2 Construction of the Tag Step Targeting Vector

4.2.1 Vector Design

A number of factors must be considered when designing a gene targeting vector including the type of vector, the amount of homologous DNA and choice of selection systems. Design of the LIF tag step targeting vector is discussed below.

For ease of design, replacement rather than insertion targeting vectors have been used in this project since it has been demonstrated that when using vectors constructed from isogenic DNA, the targeting efficiency of replacement and insertion vectors is comparable (Deng and Capecchi, 1992). Because the E14TG2a ES cell line used in this work was derived from the mouse 129 strain, a genomic LIF clone from a 129 mouse strain library was used for the construction of each of the targeting vectors. Using DNA isogenic with the strain of mouse from which the ES cells were derived increases the efficiency of gene targeting (te Riele *et al.*, 1992; van Deursen and Wieringa, 1992), presumably by reducing the number of base-pair mismatches between incoming DNA and cognate locus.

Null mutations can be generated during the tag step of the stable tag-exchange approach. In this project however, the tag step vector was not designed to create a null mutation of the LIF gene, since such mutations were being produced by other researchers at the commencement of the work. The complete coding sequence of the LIF gene could thus be included in the tag step vector, allowing the maximum amount of homologous DNA to be used for the construction of this vector. The total amount of homologous DNA contained in a gene targeting vector is one of the most important factors in determining the efficiency of gene targeting in ES cells. Increases in targeting frequency are seen with increasing vector homology particularly between 2 and 10 kb (Thomas and Capecchi, 1987; Hasty *et al.*, 1991b; Deng and Capecchi, 1992). For this reason, close to 10 kb of homologous DNA was used for the construction of the tag step targeting vector. The presence of less than 1 kb of homology on the shortest arm of an insertion vector has been reported to reduce the targeting frequency and fidelity of recombination (Thomas *et al.*, 1992). For this reason a minimum of 1.8 kb of homology was used on the short arm of the tag step targeting vector.

The amount of non-homologous DNA carried by a targeting vector does not affect significantly the frequency of gene targeting (Mansour *et al.*, 1990). This enabled use of the HPRT minigene as the tag cassette (Reid *et al.*, 1990), which is over 6 kb in length is efficiently expressed in ES cells. The *neo* gene was chosen as the stability cassette since dual selection for this and the HPRT gene can be carried out in ES cells. The *neo* cassette used had also been shown to be expressed efficiently in ES cells (Thonglairoam, 1994).

The stable tag-exchange system requires the introduction of two separated regions of extended non-homology (tag and stability cassettes) into the locus under investigation. The possibility of using a single replacement vector carrying tag and stability cassettes separated by almost 5 kb of DNA was investigated in tag step gene targeting.

To facilitate the introduction of alterations into the LIF coding region, the HPRT tag gene was positioned within this region on the tag step vector. The stable tag-exchange approach

requires that the stability cassette remains on the initially tagged allele during exchange step gene targeting. This is a potential drawback of the stable tag-exchange approach since the stability cassette on the exchange targeted allele may affect expression of the gene under investigation. Careful design of tag step vectors is therefore required to minimise these effects. In this project, the stability cassette was positioned 3' of the translated region of the LIF gene, since regulatory elements are likely to exist within the 5' and intronic regions of a gene. Due to the long 3' UTR of the LIF gene and paucity of homology beyond this region on the genomic clone used for vector construction, the stability cassette was positioned within the 3' UTR of the LIF gene. In the ideal case, the stability cassette would be positioned further downstream of the gene of interest in a non-transcribed region.

4.2.2 *The Genomic Mouse LIF Clone*

An approximately 13.5 kb Hind III fragment of the mouse LIF gene was isolated from a mouse 129 strain library and cloned into pBluescript II KS⁻ forming the vector mDIAG12(2) (Mountford *et al.*, 1994, Fig 4.1 a, b). Thorough restriction mapping was used to characterise the clone and showed that the LIF clone was actually 13.2 kb in length. An example of the restriction analysis is shown in figure 4.1 c. The clone contains the entire coding region of the LIF gene, 6.4 kb of sequence 5' to the coding region and 4.0 kb of sequence 3' of the translated exons which contains the 3' UTR of 3.5 kb.

When screening ES cell lines resulting from gene targeting by Southern analysis, it is preferable to use probes which hybridise to a region outside the homology used in the targeting vectors. This prevents detection of vectors integrated at non-homologous sites. To enable such a probe to be used for the detection of lines targeted at the LIF locus, the 9.6 kb Bgl II/Hind III fragment of the LIF clone in mDIAG12(2) was used for construction of the LIF targeting vectors used in this project. A sequence within the remaining portion of the clone was used as a 5' external probe. Positioning of the stability cassette close to the 5' end of the genomic clone used for vector construction precluded the use of a 3' external probe for Southern analysis.

Figure 4.1 Genomic mLIF clone

a Representation of the 13.2 kb genomic mouse LIF clone obtained from a 129 library (section 4.2.2). The restriction sites for the enzymes Bgl II, Eco RI, Hind III, Mlu I, Nhe I and Sac II are shown. Line, genomic sequence; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR.

b Representation of the vector mDIAG12(2) from which the LIF clone used for targeting vector construction was taken (restriction sites and symbols as above).

c Examples of restriction analysis (section 9.3.1, 9.3.5) of vector mDIAG12(2) digested with the appropriate enzyme(s) and electrophoresed on a 1 % TAE gel (section 9.3.5). B:Bgl II, H:Hind III, M:Mlu I, N:Nhe I, S: Sac II, Spp 1: Eco RI digested SPP-1 size markers: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36 kb.

Expected sizes from the digests are as follows:

S/H: 6.6, 6.5, 2.9 kb

H/N: 6.7, 2.9, 2.9, 1.9, 0.8, 0.7 kb

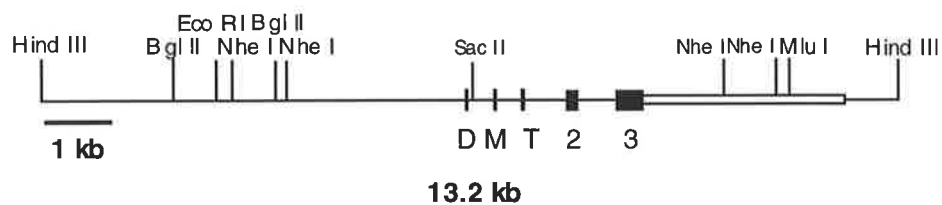
B/N: 6.9, 6.7, 1.0, 0.7, 0.6, 0.2 kb

B/H: 9.6, 2.9, 2.0, 1.6 kb

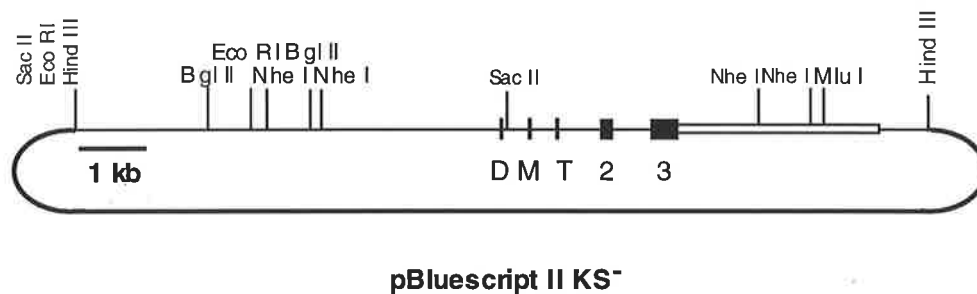
N/S: 4.9, 4.0, 3.0, 2.8, 0.8, 0.7 kb

M/H: 11.4, 2.9, 1.8 kb

a Genomic mLIF clone from 129 Library

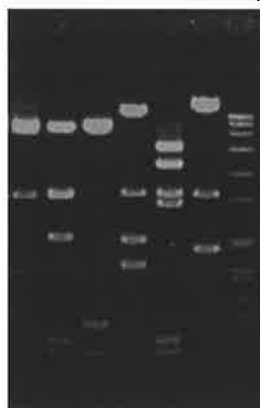


b mDIAG12(2)



c Restriction Enzyme Analysis

S/H H/N B/N B/H N/S M/H Spp1



4.2.3 Construction of a LIF Targeting Vector Carrying the HPRT Gene

The 9.6 kb Bgl II/Hind III fragment of mDIAG12(2), containing 2.8 kb of 5' sequence, the entire LIF coding region and 4 kb of 3' sequence (Fig. 4.2 i), was subcloned into pT7T319U (Stratagene) from which the Xba I site had been removed by digestion with Xba I, endfilling and religation (pXT7T3, Fig. 4.2 ii) to produce the vector pLIFXT7t (Fig. 4.2 iii). The LIF fragment was endfilled and cloned into the Sma I site of pXT7T3 with the 5' end of LIF closest to the Hind III site in the polylinker.

The HPRT selection cassette was taken from the vector pnI2(11s) (Reid *et al.*, 1990). This vector contains a 6.1 kb HPRT minigene consisting of the coding region of the human HPRT gene and intronic sequences required for efficient expression in ES cells (Reid *et al.*, 1990). The minigene carries the 3' simian virus 40 (SV40) polyadenylation signal and expression is driven by the 5' thymidine kinase (TK) promoter with mutant polyoma enhancer (Fig. 4.3 i). The HPRT minigene was isolated as a Bam HI fragment, the ends of which were filled, and ligated into pLIFXT7t which had been digested with Sac II and blunted by exonuclease digestion (Fig. 4.3 ii). The HPRT gene was ligated in the same transcriptional orientation as the LIF gene. The resultant vector, termed pHPRTLIFXt (Fig. 4.3 iii), contains the entire coding region of the LIF gene with the HPRT expression cassette positioned between the first exons of the diffusible and matrix associated forms of LIF.

4.2.4 Construction of the First Round Targeting Vector Carrying the HPRT and neo genes

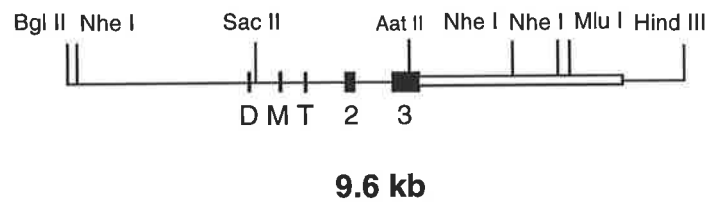
Vector pLIFXT7 was used for construction of the LIF targeting vector carrying both the HPRT and *neo* genes. The *neo* expression cassette (Fig. 4.4 i) taken from pgkNeopA (Thonglairoam, 1994) carries the *neo* gene under the transcriptional control of the mouse *Pgk-1* promoter and polyadenylation sequences of the *Pgk-1* gene. The *neo* expression cassette from pgkNeopA was isolated as an Eco RI/Bgl II fragment, endfilled and ligated, in the same transcriptional orientation as LIF, to pLIFXT7t which had been digested with Mlu I and endfilled (Fig. 4.4 ii). The resultant vector termed pNEOLIFXt (Fig. 4.4 iii) carries the *neo* gene positioned within the 3' noncoding region of mLIF in pLIFXT7t.

Figure 4.2 Construction of pLIFXT7t

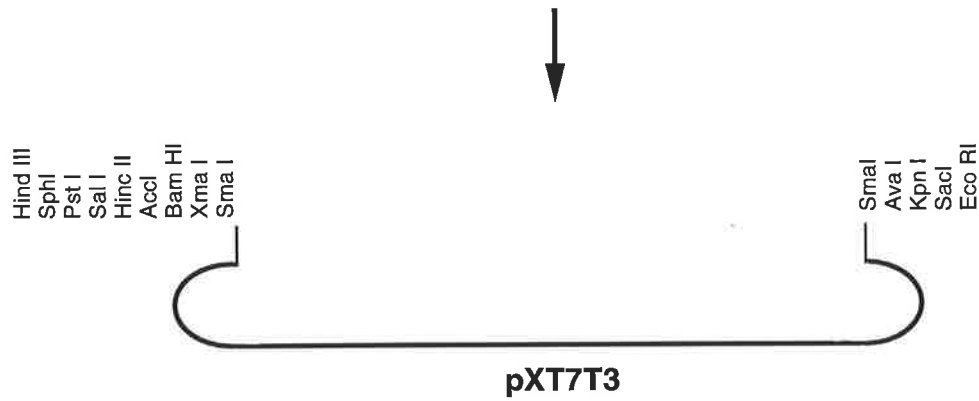
Representation of (i) the Bgl II/Hind III fragment of LIF used in construction of the targeting vectors, (ii) the Sma I linearised vector pXT7T3 into which the LIF fragment was cloned and (iii) the resultant vector pLIFXT7t. The restriction sites for the enzymes Aat II, Bgl II, Hind III, Mlu I, Nhe I and Sac II and additional restriction sites in the polylinker of pXT7T3 are shown. Thin line, genomic DNA; thick line, non-genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR.

Construction of pLIFXT7t

i



ii



iii

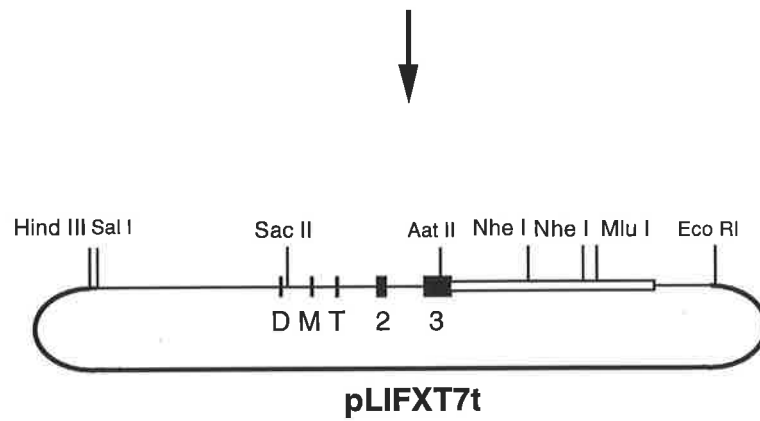


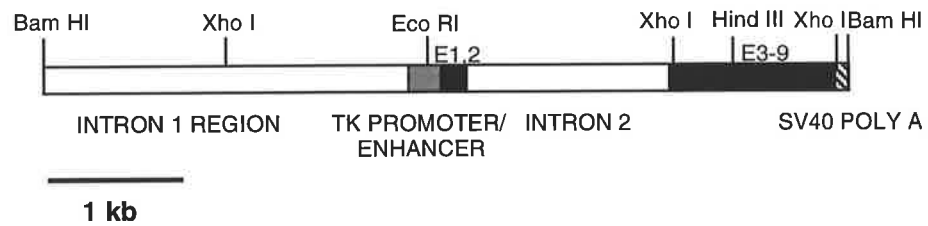
Figure 4.3 Construction of pHPRTLIFXt

Representation of (i) the HPRT minigene taken from pnI2(I1s) showing the restriction sites for Bam HI, Eco RI, Hind III, and Xho I (open boxes, intronic regions; black boxes, exons of the human HPRT gene; stippled box, HSV-tk 5' promoter and mutant polyoma enhancer; hatched box, SV40 polyadenylation signal) (ii) cloning of the HPRT minigene into the Sac II site of the LIF clone in pLIFXT7t and (iii) the vector pHPRTLIFXt. The restriction sites for Aat II, Bam HI, Eco RI, Hind III, Mlu I and Sal I are shown. Thin line, genomic DNA; thick line, non-genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR, stippled box, HPRT cassette.

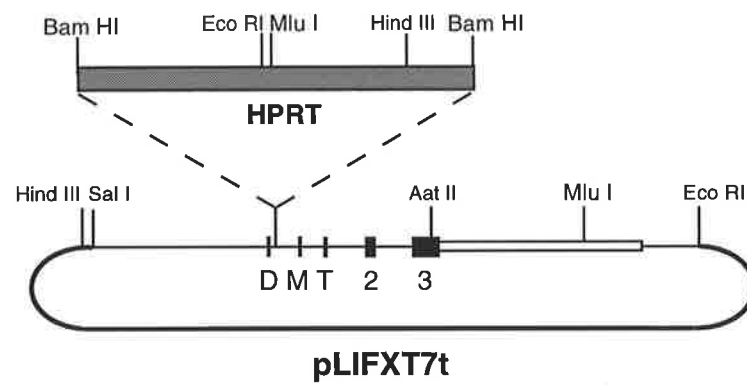
Construction of pHPRTLIFXt

i

HPRT minigene isolated from pnl2(l1s)



ii



iii

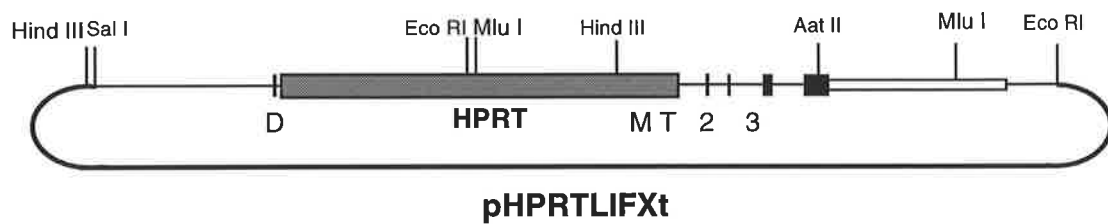


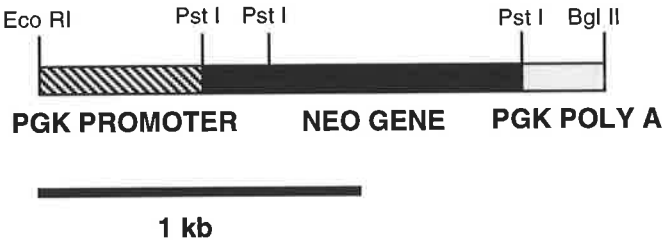
Figure 4.4 Construction of pNEOLIFX

Representation of (i) the *neo* cassette isolated from pgkNEOpA showing the restriction sites for Bgl II, Eco RI and Pst I (dark hatched box, *Pgk-1* promoter; black box, *neo* gene; stippled box, polyadenylation signal of *Pgk-1*) (ii) cloning of the *neo* cassette into the Mlu I site in pLIFXT7t, (iii) the vector pNEOLIFXt, (iv) the sequence of the truncated 3' end of pNEOLIFXt and (v) restoration of the LIF sequence from vector mDIAG12(2). The restriction sites for Aat II, Bgl II, Eco RI, Hind III, Mlu I and Sal I and additional sites in the polylinker of pNEOLIFXt are shown. Thin line, genomic DNA; thick line, non-genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR, light hatched box, *neo* cassette.

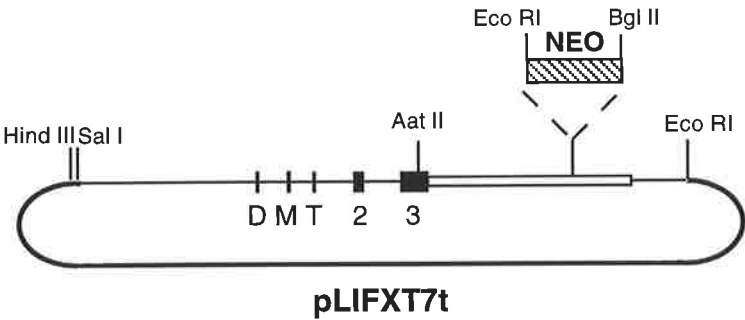
Construction of pNEOLIFX

NEOMYCIN CASSETTE isolated from pgkNeopA

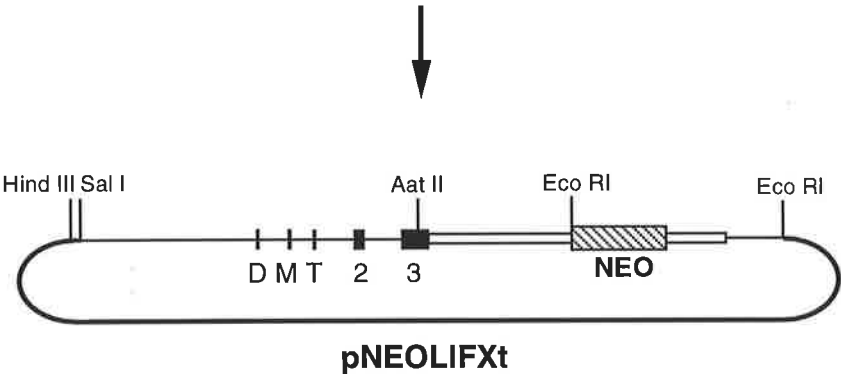
i



ii



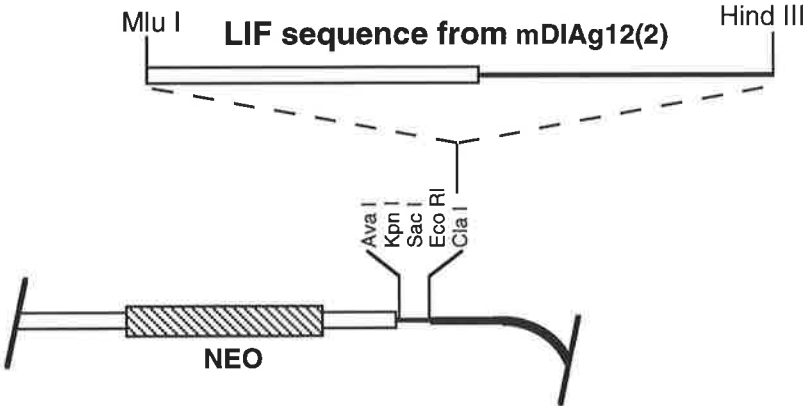
iii



iv

Mlu I
ACGCGTGTCAGCGACAAAGTTACTCCACCGTACTCCTGTTGCTGCGTCAGGCTCATCTCAGGACTG
GCTGCCCTTCTCCAAGCTGAGAGTCAATTTGTCTAAAAGCCAAGATGATGCCACAGCCTGGG GGGA

v



Restriction enzyme analysis of pNEOLIFXt revealed a truncation at the 3' end of the LIF sequence. Sequencing revealed that approximately 1550 bp of the LIF clone was missing and that the truncation was 123 bp 3' of the inserted *neo* gene (Fig. 4.4 iv). The missing homologous sequence 3' to the *neo* gene was required in the targeting vector to permit a crossover event downstream of the *neo* cassette, since less than 1 kb of homology on the short arm of a replacement vector can reduce the targeting frequency and lead to aberrant homologous recombination (Thomas *et al.*, 1992). This portion of the LIF clone was subcloned back into pNEOLIFXt in the following manner. Vector pNEOLIFXt was linearised with Cla I, which cuts the vector within the polylinker 3' of the LIF sequence, and endfilled. The missing LIF sequence was isolated from mDIAG12(2) by digestion with Mlu I and Hind III, endfilled and ligated to linearised pNEOLIFXt in the correct orientation to restore the vector now termed pNEOLIFX (Fig. 4.4 v). Vector pNEOLIFX carries a 123 bp duplication of LIF DNA and part of the polylinker from pXT7T3 contained within the homologous region of LIF DNA at the 3' end. This was not expected to be a problem since the short duplicated LIF sequence was unlikely to participate in homologous base pairing, and due to the polylinker sequence positioned downstream, it was likely to form part of the non-homologous loop-out containing the *neo* cassette during gene targeting.

A vector which contained both the HPRT and *neo* selection cassettes was required for tag step targeting. The fact that a Mlu I site (used for insertion of the *neo* gene into the LIF sequence) was present within the HPRT cassette, necessitated that the HPRT gene be inserted into the vector after insertion of the *neo* gene. The HPRT gene was therefore isolated as a Sal I/Aat II fragment from pHPRTLIFX (Fig. 4.5 i) and ligated into pNEOLIFX that had also been digested with Sal I and Aat II (Fig 4.5 i) to form the targeting vector pHNLX (Fig. 4.5 ii). This vector contains the HPRT expression cassette positioned between the first exons of the diffusible and matrix associated forms of the LIF and the *neo* expression cassette located within the 3' UTR of the LIF gene. Restriction analysis of the vector pHNLX and a restriction map are shown in Fig. 4.6.

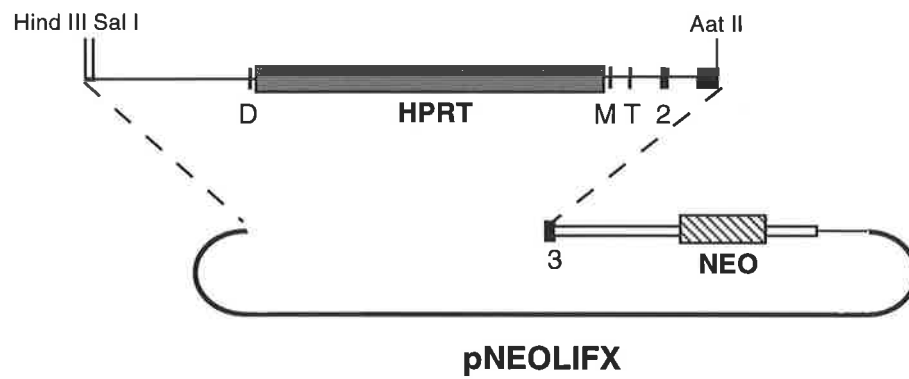
Figure 4.5 Construction of the tag step targeting vector pHNLX

Representation of (i) the cloning of the Hind III/Aat II fragment of pHPRTLIFX, carrying the HPRT cassette, into pNEOLIFX and (ii) the vector pHNLX. The restriction sites for Aat II, Hind III and Sal I are shown. Thin line, genomic DNA; thick line, non-genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR, stippled box HPRT cassette; hatched box, *neo* cassette.

Construction of Tag Step Vector pHNLX

i

Fragment isolated from pHPRTLIFX



ii

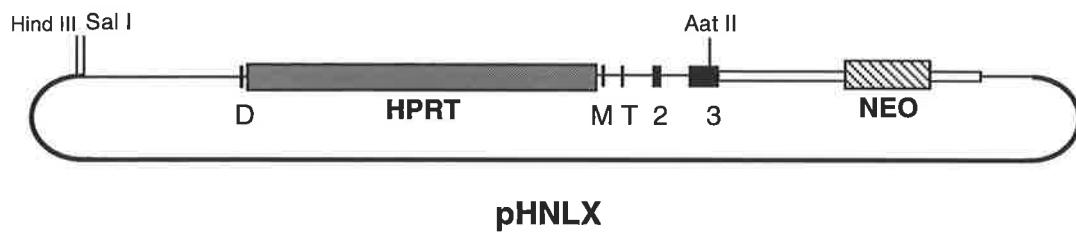


Figure 4.6 The tag step targeting vector pHNLX

a Representation of the tag step targeting vector pHNLX showing the restriction sites for Bam HI, Eco RI, Hind III, and Sal I. Thin line, genomic DNA; thick line, non-genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR, stippled box, HPRT cassette; hatched box, *neo* cassette.

b Restriction enzyme analysis (section 9.3.1) of pHNLX. pHNLX was digested with the appropriate enzyme(s) and electrophoresed on a 1 % TAE gel (section 9.3.5). S, Sal I; E, Eco RI, B, Bam HI; H, Hind III, Spp1: Eco RI digested SPP-1 size markers: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36 kb.

Expected sizes from the digests are as follows:

S: 20.5

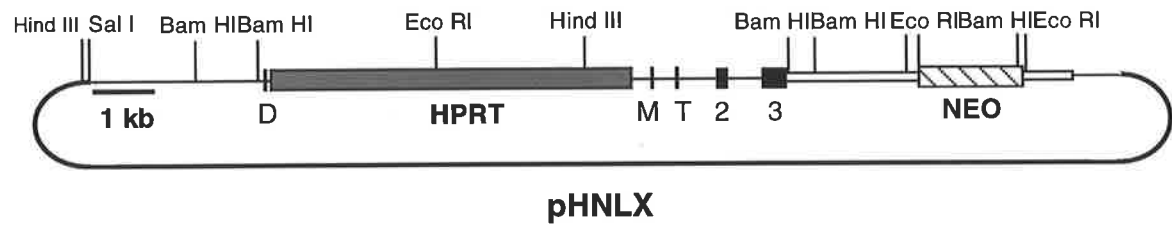
E: 10.4, 8.2, 1.9

B: 9.0, 5.3, 3.1, 1.8, 0.9, 0.4

E/H: 5.7, 5.7, 4.6, 2.5, 1.9 kb

a

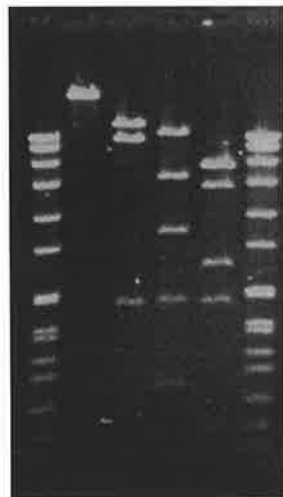
Tag Step Vector pHNLX



b

Restriction Enzyme Analysis of pHNLX

Spp1 S E B E/H Spp1



4.3 Gene Targeting at the LIF locus with the tag step vector pHNLX

4.3.1 Determination of Tag Step Gene Targeting Selection Conditions

The E14 ES cell derivative E14TG2a (Hooper *et al.*, 1987) was used in this project since it lacks a functional HPRT gene, is germline competent and can be grown as a monoculture in the absence of feeder cells. Variation between ES cell lines and batches of selective agents requires that selection conditions are determined for each cell line and selective agent.

To determine the dose of G418 and HAT required for effective selection against E14TG2a, ES cells were grown in the presence of HAT and G418. HAT selection with a final concentration of 100 μ M hypoxanthine, 0.4 μ M amethopterin, 16 μ M thymidine was sufficient to completely kill the ES cells (Table 4.1). HAT selection was used in tag step gene targeting at that concentration. A G418 concentration of 200 μ g/ml was sufficient to cause greater than 97 percent cell death (Table 4.2) and was used in the tag step gene targeting at that concentration.

4.3.2 Tag Step Targeting

E14TG2a ES cells were transfected by electroporation with pHNLX DNA that had been linearised by digestion with Sal I (Fig. 4.7 ii), or with no DNA as a mock control. The linearised targeting vector DNA contained 2.9 kb of heterologous plasmid DNA at the 3' end of homology. It has been shown previously that non-homologous regions at the extremities of targeting vectors are not introduced into the cognate locus (Mansour *et al.*, 1988) and do not have a deleterious effect on the targeting efficiency (Deng and Capecchi, 1992). The electroporation conditions used were determined by workers in the laboratory of Dr. Austin Smith (Centre for Genome Research, University of Edinburgh) and have been shown to minimise cell death caused by electroporation and to prevent the integration of multiple copies of the targeting vector into the genome (A. Smith, personal communication).

The tag step targeting event at the LIF locus is illustrated in figure 4.7. After transfection, cells were grown in the absence of selection to determine cell survival after electroporation, or

Table 4.1 Results of HAT toxicity assay

E14TG2a ES cells were grown in the presence of HAT selective medium at various concentrations for seven days (section 4.3.1), fixed and viewed. H, hypoxanthine; A, amethopterin; T, thymidine.

Table 4.2 Results of G418 toxicity assay

E14TG2a ES cells were grown in the presence of G418 selective medium at various concentrations for seven (section 4.3.1) fixed and viewed.

Table 4.1

Seeding Density	Selection	Concentration	Cell Survival (DAY 7)
5×10^5	No Selection	-	100 %
5×10^5	HAT (1)	100 μ M H, 0.8 μ M A, 20 μ M T	0 %
5×10^5	HAT (2)	100 μ M H, 0.4 μ M A, 16 μ M T	0 %

Table 4.2

Seeding Density	Selection	Concentration	Cell Survival (DAY 7)
1×10^6	No Selection	-	100 %
1×10^6	G418	175 μ g/ml	< 5 %
1×10^6	G418	200 μ g/ml	< 3 %
1×10^6	G418	400 μ g/ml	< 1 %

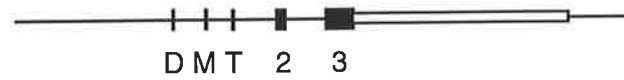
Figure 4.7 Tag step gene targeting at the LIF locus

Representation of (i) the untargeted LIF locus, (ii) tag step gene targeting at the LIF locus and (iii) the tag step targeted LIF locus. The selective agents used during gene targeting and the Sal I site used for vector linearisation (arrow) are shown. Thin line and thick stippled line, genomic DNA; dashed line, plasmid vector DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR; stippled box, HPRT cassette; hatched box, *neo* cassette.

Tag Step Gene Targeting

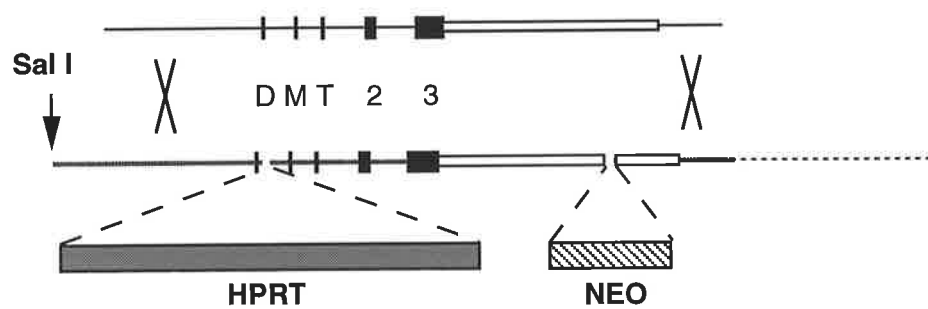
i

LIF Locus ($HAT^s/G418^s$)



ii

Gene Targeting



HAT/G418

iii

Tagged LIF Locus ($HAT^r/G418^r$)



in medium containing HAT, G418, or HAT plus G418. Because pHNLX contains two regions of non-homology separated by a large region of homologous DNA (Fig. 4.7 ii), it was possible that either the HPRT or *neo* cassette or both could be integrated into the LIF locus. Integration of either would occur if a crossover event occurred in the region of homologous DNA between the tag and stability cassettes contained on the tag step vector. This would result in HAT^r/G418^s or HAT^s/G418^r lines. To investigate this possibility, transfected cells were subjected to selection in HAT, G418 or HAT plus G418.

All cells were killed in the mock transfection control samples grown in HAT, G418, or HAT plus G418, whereas resistant colonies appeared in the transfected samples and were picked from day 7 of selection. Resistant colonies were picked, expanded, frozen and used for isolation of genomic DNA. Table 4.3 shows the results from this experiment. Of three hundred and three resistant colonies, sixty were screened by Southern blot analysis in order to identify homologous recombinants.

Fewer colonies would be expected to survive under conditions of dual selection if recombination events resulting in the integration of either of the selection cassettes into the genome occurred with high frequency. The number of colonies appearing under each selection condition suggested that this was not the case. From 9.2×10^6 cells plated in either selective medium, sixty HAT^r, ninety G418^r and 77 HAT^r/G418^r colonies arose. Fewer resistant colonies did not appear under conditions of dual selection, suggesting that the HPRT and *neo* resistance cassettes were both integrated into the genome. Analysis at the DNA level was required to determine the frequency of homologous recombination and to confirm that both selection cassettes were correctly integrated in targeted lines.

4.3.3 Southern Analysis of Resistant Lines

Genomic DNA isolated from HAT^r or HAT^r/G418^r lines was digested with Hind III/Eco RI and subjected to Southern analysis. The restriction sites for Hind III and Eco RI and the probes used for screening are illustrated in figure 4.8 a. Sixty lines were screened by

Southern analysis and eighteen potentially targeted lines were screened in detail (Fig. 4.8 b, c, d). Four probes were used in this analysis to identify correctly targeted lines.

Probe 1 The 5' external LIF specific probe was used to detect targeted lines. The probe hybridises to fragments of 10.5 kb and 2.7 kb from the untargeted LIF allele (Fig. 4.8 a i, 4.8 b all clones), and of 2.7 kb and 7.5 kb from the correctly targeted allele (Fig. 4.8 a i, ii, 4.8 b HAT^r clones 4, 9, 12 (partial digest inconclusive) 14, 20, 24; HAT^r /G418^r clones 26, 38, 60, 61, 62). Non-homologous recombinants are not detected with the external genomic probe which simplifies the screening process.

Probe 2 The *neo* probe was used to detect the *neo* selection cassette in transfected lines which were selected in HAT medium. The probe hybridises to a fragment of 1.9 kb in all lines carrying the tag step targeting vector and was used in conjunction with probe 1 in Southern analysis (Fig. 4.8 a ii, 4.8 b all clones).

Probe 3 The HPRT probe was used to detect targeted lines also carrying the vector at multiple non-homologous sites within the genome. The probe hybridises to a vector fragment of 7.5 kb if integrated at the homologous site (Fig. 4.8 a ii, 4.8 c HAT^r clones 4, 9, 14, 20, 24; HAT^r /G418^r clones 26, 38, 60, 61, 62) and fragments of other sizes if integrated at non-homologous sites (Fig. 4.8 c, clones HAT^r/G418^r 8, 29, 43, 44, 49, 81, 82).

Probe 4 The 3' probe was used to detect correct integration, in the absence of bacterial plasmid DNA, at the 3' end of the vector. The probe hybridises to a fragment from the untargeted LIF allele of 10.5 kb (Fig. 4.8 a i, 4.8 d all clones) and a 1.9 kb fragment if integrated at the homologous site (Fig 4.8 d HAT^r clones 4, 9, 14, 20, 24; HAT^r /G418^r clones 26, 38, 60, 61, 62). The 1.9 kb Eco RI/ Hind III fragment results from digestion at an external Hind III site contained within the LIF locus, but not on the targeting vector itself. Thus, fragments of other sizes are detected by this probe if the vector has integrated into non-homologous sites (Fig. 4.8 d HAT^r/G418^r clones 8, 29, 43, 44, 49, 81, 82).

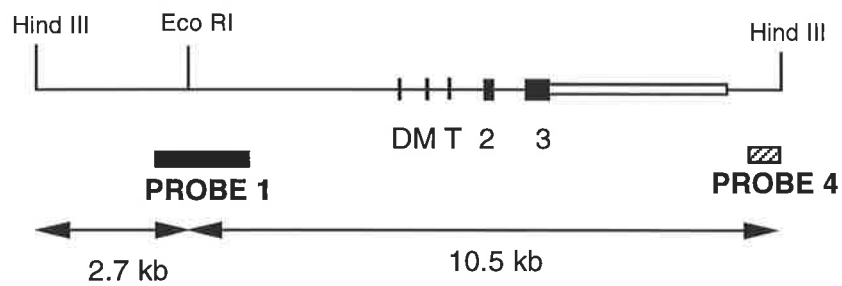
Figure 4.8 a Southern analysis of tag step targeted lines

Representation of the expected sizes of restriction fragments resulting from digestion of (i) the untargeted LIF allele and the (ii) tag step targeted LIF allele with Eco RI/Hind III, showing the probes used for Southern analysis (Figs. 4.8 b-d). Probe 1, 5' external probe; probe 2, *neo* probe; probe 3, HPRT probe; probe 4, 3' internal probe, thin line and thick stippled line, genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR; shaded box, HPRT cassette; hatched box, *neo* cassette.

a Southern Analysis of Tag Step Targeting

i

Untargeted LIF Allele



ii

Tag Step Targeted LIF Allele

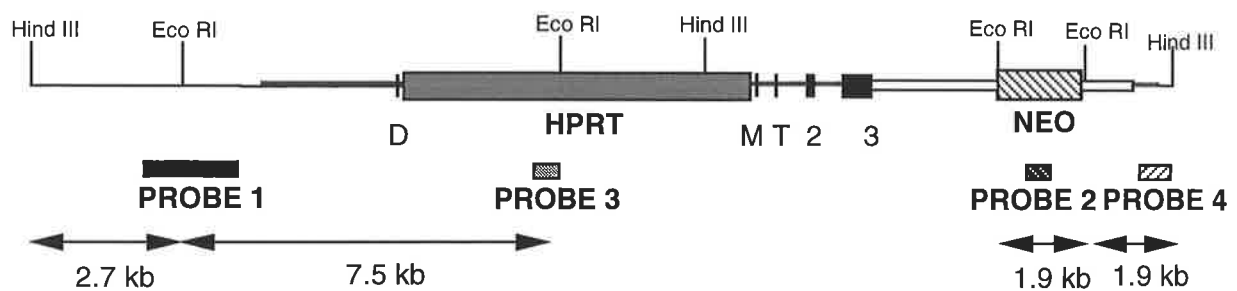


Figure 4.8 b Southern analysis of tag step targeted lines

Approximately 20 µg of genomic DNA from HAT^r and HAT^r/G418^r clones was digested with Eco RI/Hind III, electrophoresed on a 0.8 % TAE agarose gel and blotted (section 9.3.21). The filter was probed with a 5' external Bgl I LIF probe of 1.7 kb (probe 1) and Pst I *neo* probe of 0.7 kb (probe 2) and washed in 0.1 x SSC/0.1 % SDS at 65°C for 30 minutes (section 9.3.11, 9.3.22). The sizes of the fragments to which the LIF probes hybridises are 2.7 and 10.5 kb from the untargeted allele, 2.7 and 7.5 kb from the targeted allele. The size of the fragment to which the *neo* probe hybridises is 1.9 kb. The type of selection (HAT or HAT/G418) is indicated and the numbers above the lanes refer to the cell lines, where HAT^r line 4 is referred to as H4 etc.

b

5' External and *neo* Probes (1 & 2)

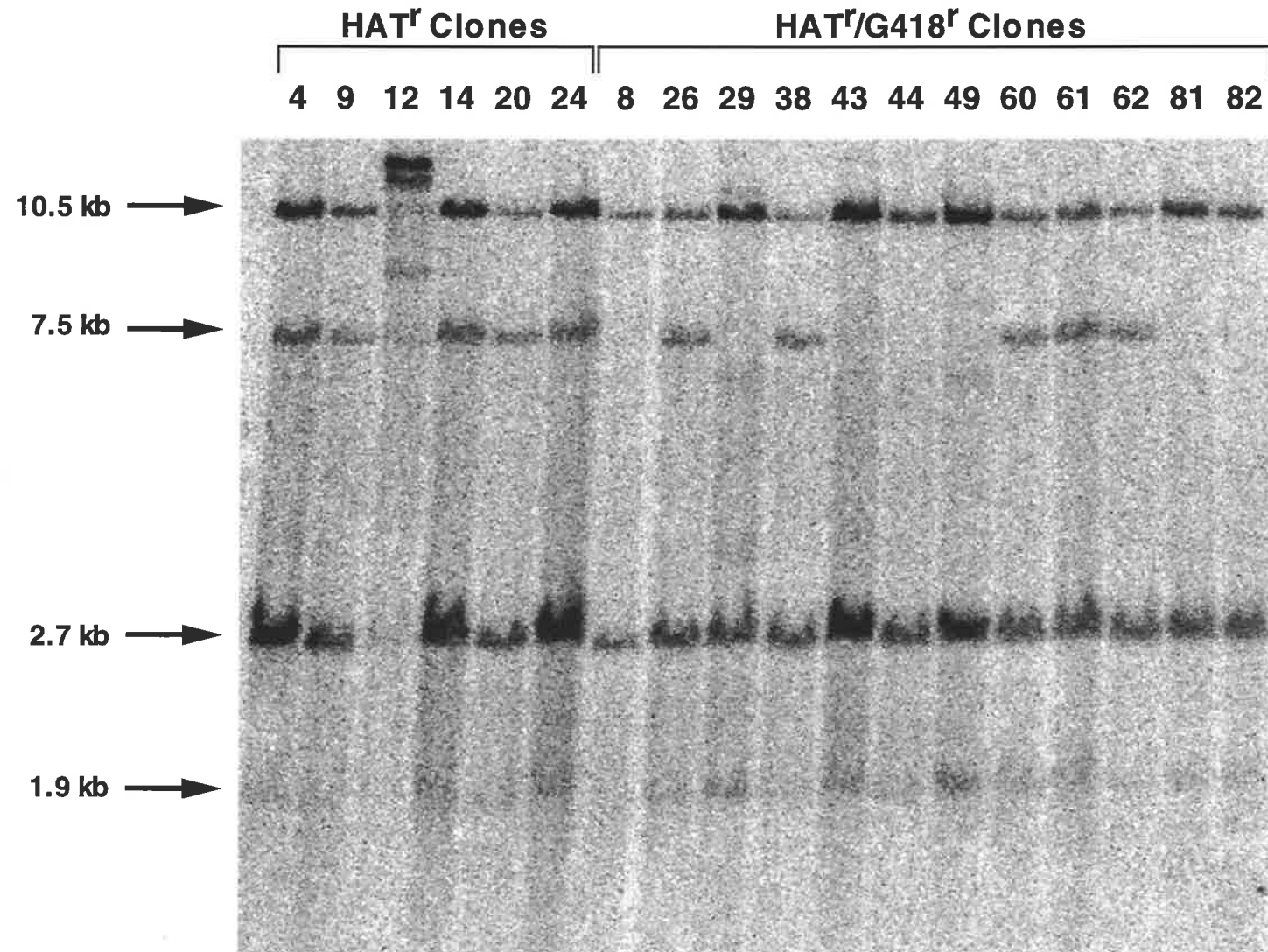


Figure 4.8 c Southern analysis of tag step targeted lines

Approximately 20 µg of genomic DNA from HAT^r and HAT^r/G418^r clones was digested with Eco RI/Hind III, electrophoresed on a 0.8 % TAE agarose gel and blotted (as described in Fig. 4.8 b). The filter was probed with a Cla I/Eco RI HPRT probe of 0.2 kb (probe 3) and washed (as described in Fig. 4.8 b). The size of the fragment to which the HPRT probe hybridises is 7.5 kb from the targeted allele. Fragments of other sizes result in non-homologous recombinant lines. The lanes are labelled with the selection used and cell line names (as described in Fig. 4.8 b).

c

HPRT Probe (3)

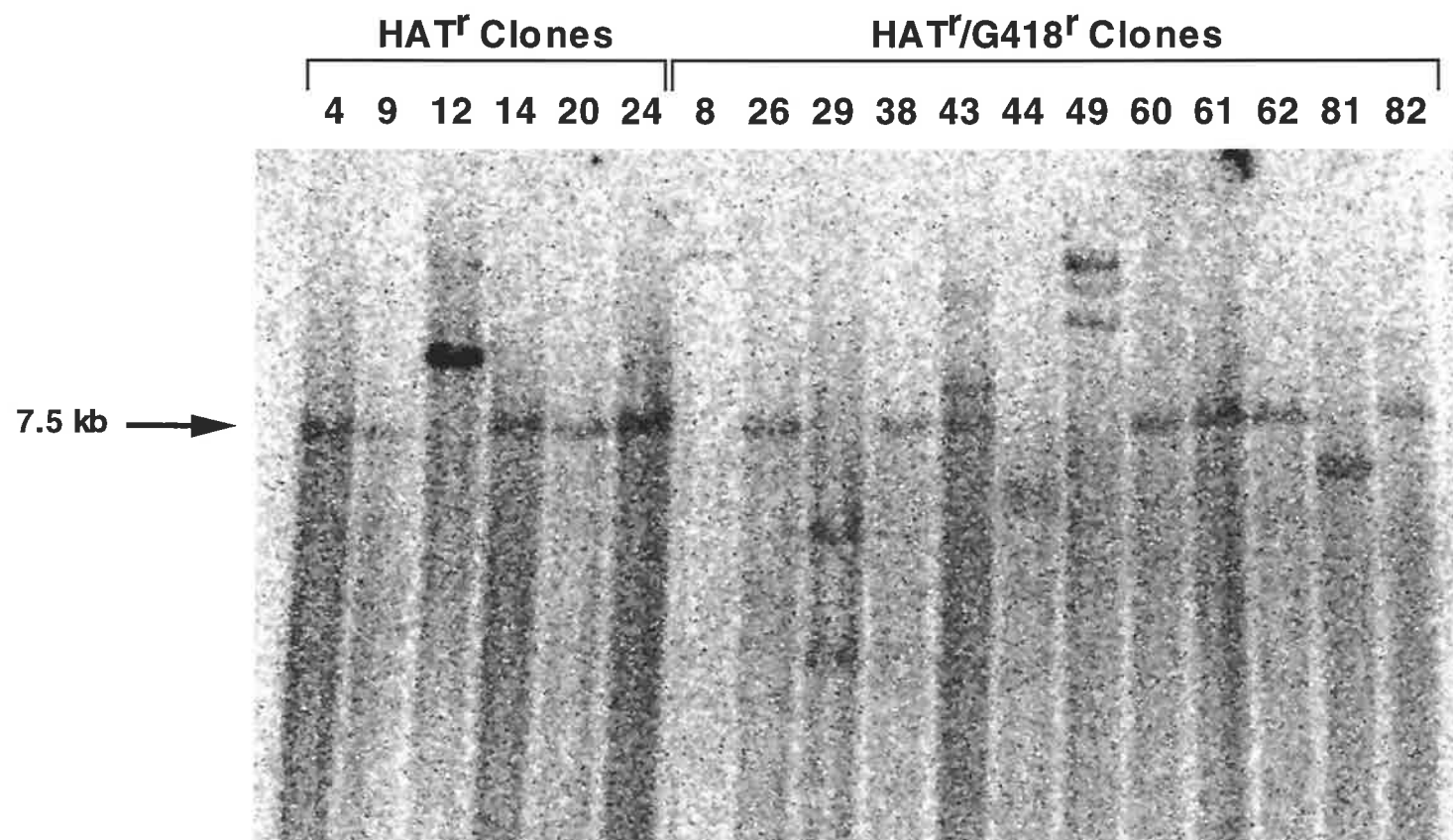


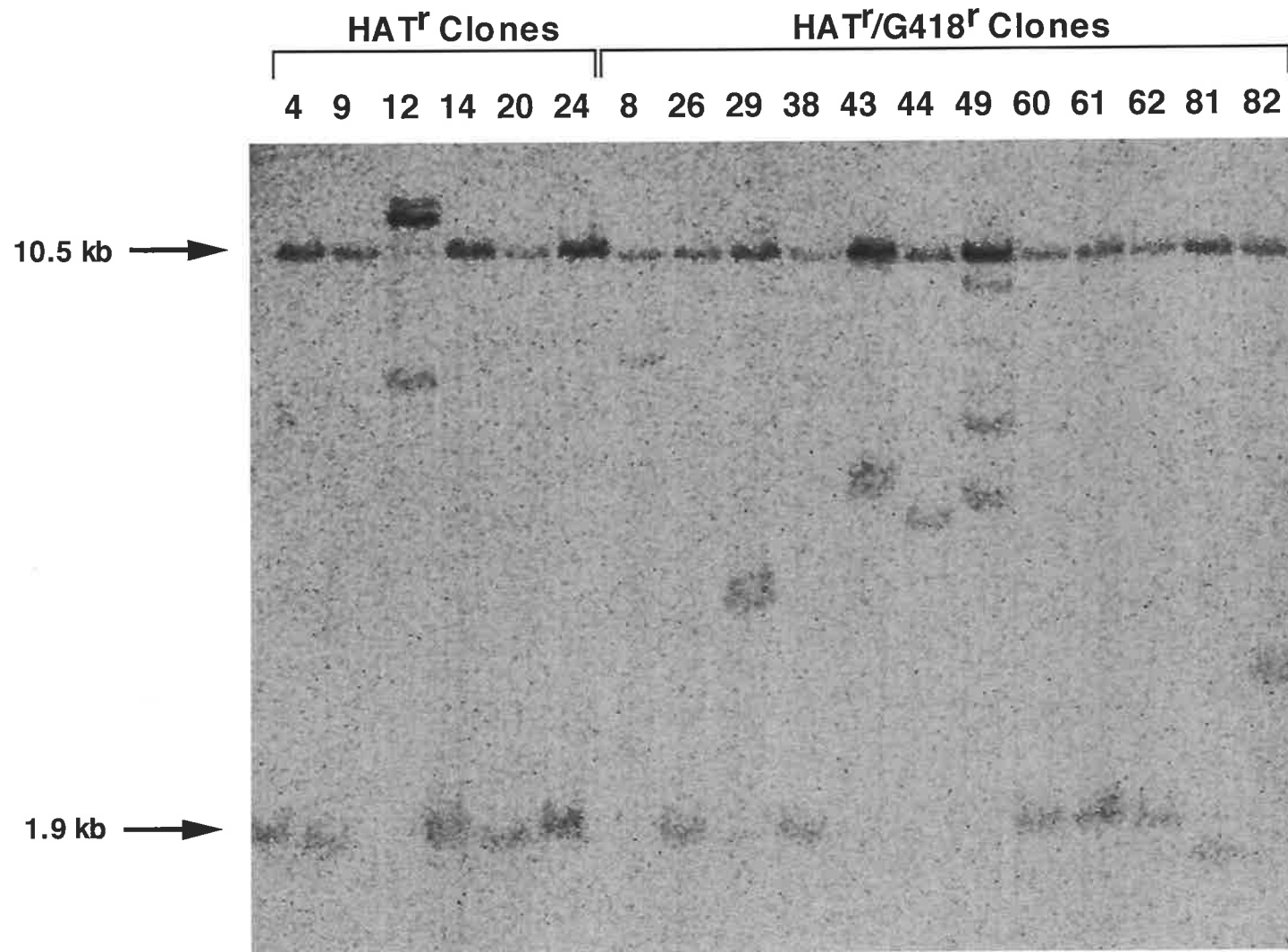
Figure 4.8 d Southern analysis of tag step targeted lines

Approximately 20 µg of genomic DNA from HAT^r and HAT^r/G418^r clones was digested with Eco RI/Hind III, electrophoresed on a 0.8 % TAE agarose gel and blotted (as described in Fig. 4.8 b).

The Southern was probed with a 3' internal Ssp I/Hind III LIF probe of 0.9 kb (probe 4) and washed (as described in Fig. 4.8 b). The sizes of the fragments to which the LIF probes hybridises are 10.5 kb from the untargeted allele, 10.5 and 1.9 kb from the targeted lines. Fragments of other sizes result in non-homologous recombinant lines. The lanes are labelled with the selection used and cell line names (as described in Fig. 4.8 b).

d

3' Internal Probe (4)



Five/fifteen HAT^r (excluding line H12 for which the Southern analysis was not conclusive) and five/forty five HAT^r/G418^r lines were correctly targeted (Table 4.3), representing targeting frequencies of 33.3 and 11.1 percent, respectively. The overall targeting frequency was therefore 16.7 percent (Table 4.3). The difference between the targeting efficiencies of cells grown under HAT and HAT/G418 selection may reflect the small number of HAT^r lines (15 versus 45 HAT^r/G418^r) that were screened. The results of the Southern analysis indicate that the homologous recombinants were correctly targeted, carrying the HPRT and *neo* cassettes integrated into the LIF locus in the absence of bacterial plasmid DNA or integrations at non-homologous sites. All homologous recombinants contained both the HPRT and *neo* selection cassettes correctly integrated even when selection for only the HPRT gene was applied. This was also the case for non-homologous recombinants some of which carried the vector at more than one location within the genome (Fig. 4.8 c HAT^r/G418^r clones 29, 43, 49).

Four correctly tag step targeted lines (lines H9, H20, HG26 and HG60) were returned to the mouse to assess their ability to contribute to chimaeric mice and their germline competence (section 7.3).

4.4 Discussion

The vector pHNLX, used for tag step gene targeting, contains almost 10 kb of homologous DNA and two separate large regions of non-homology. The vector was efficiently introduced into the LIF locus by homologous recombination in ES cells. One in six (16.7 %) resistant colonies screened by Southern blot contained the tag step vector correctly integrated at the LIF locus in the absence of integrations at non-homologous sites. In comparison, a targeting frequency at the LIF locus of 6.25 % was seen when 5.6 kb of non-isogenic DNA was used (Stewart *et al.*, 1992) while a frequency of 26 % was seen when 7 kb of isogenic DNA was used (Mountford *et al.*, 1994). The results of this study indicate that a vector carrying two large, separate regions of non-homology can be efficiently introduced into the mouse genome on a single targeting construct. It would be of interest to compare directly the targeting efficiency of an equivalent vector carrying one region of non-homology with that of pHNLX.

Table 4.3 Results from tag step gene targeting

Results of tag step gene targeting carried out with the vector pHNLX using various selection conditions (as described in sections 4.3.2; 9.4.12). Screening was carried out by Southern blot analysis (sections 9.3.11; 9.3.21-22).

Table 4.3

	G418	HAT	HAT/G418	Total
Cells Plated	9.2×10^6	9.2×10^6	1.8×10^7	3.6×10^7
Cells Surviving	6.4×10^5	6.4×10^5	1.3×10^6	2.5×10^7
Resistant Lines	90	60	153	303
Screened	-	15	45	60
Targeted	-	5	5	10
Efficiency	-	33.3 %	11.1 %	16.7 %

The HPRT (6.1 kb) and *neo* (1.8 kb) selection cassettes contained on pHNLX are separated by almost five kilobases of homologous LIF DNA. No significant decrease in the number of colonies resistant to HAT/G418 selection compared to HAT or G418 selection was seen, suggesting that both the HPRT and *neo* cassettes were always introduced into the genome, regardless of whether homologous or non-homologous recombination occurred. Analysis at the DNA level verified this. The absence of HPRT⁺/*neo*⁻ recombinants appearing under HAT selection indicated that functional crossover events did not occur between the regions of non-homology contained on the targeting vector during either homologous or non-homologous recombination.

This finding suggests that the positive/negative selection (tag) cassette used in a tag step targeting vector could be substituted by a negative selection cassette. Selection for expression of the stability cassette during tag step targeting is all that is required to achieve homologous recombination of the tag step targeting vector. During exchange step targeting, positive selection for the stability cassette and selection for loss of the tag cassette would be used to select for homologous recombinants. Such an approach would increase the versatility of the stable tag-exchange approach since it would increase the choice of genes that could be used as the tag cassette.

The fact that during homologous recombination crossover events did not occur between the two regions of non-homology, despite the fact that close to half the homology carried by the vector was contained in this region, suggests that the DNA ends of replacement vectors are essential for the process of homologous recombination. pHNLX contained almost 3 kb of heterologous plasmid DNA at the 3' end which was not integrated into the homologous site. This indicates that, although involved in the process of recombination, the vector ends do not have to participate in the recombinant product.

The implication that heterologous sequences positioned near the termini of replacement vectors can "force" crossover events to occur at the extremities of such vectors suggests that

multiple heterologous sequences may be introduced into the mouse genome on a single replacement targeting vector. It may therefore be possible to produce gain of function mutations by the introduction of multiple copies of the gene under investigation into the endogenous site using a single replacement vector. An advantage of this sort of approach, compared to conventional transgenesis, is that the multiple copy genes are positioned at the correct site within the genome and thus should be subject to normal regulatory controls.

It may also be possible to introduce multiple different genes into a single site within the genome on a single replacement vector. Many genes can contribute to a complex disease state (Bronson and Smithies, 1994) and thus to produce effective animal models for such diseases, systems are required which permit the introduction of the genes responsible into a site within the genome. Similarly, such systems would also be required to produce mice for example carrying the genes involved in a particular synthetic pathway.

The finding that almost five kilobases of homologous DNA can be integrated efficiently into the genome of ES cells, when positioned between two regions of non-homology, suggests that large regions of DNA may be introduced efficiently into the mouse genome using this approach. The upper limit to the amount of DNA that can be introduced in this way is not known and requires further investigation. Once established, this feature may permit use of stable tag-exchange tag step vectors which carry the stability cassette positioned tens of kilobases from the gene that is to be altered in the exchange targeting step. This should prevent any effects on the expression of the gene of interest by the stability cassette.

4.5 Recombination Models

4.5.1 Existing Recombination Models

The mechanism(s) by which replacement vectors undergo homologous recombination in mammalian cells are still poorly understood (Hasty *et al.*, 1992), although parallels with simpler organisms can be made. The Holliday model was proposed in 1964 to describe meiotic homologous recombination in fungi. In this model, it was proposed that a single

stranded DNA break permits invasion of homologous DNA by the single stranded region. Ligation to the complementary homologous strand then follows (Fig. 4.9 a, i). Heteroduplex DNA is thus formed on both chromatids with consequent formation of a structure known as the Holliday junction (Fig. 4.9 a, i). Branch point migration, which is stabilised by base pairing between the heteroduplex DNA strands, leads to lateral movement of the Holliday junction. In this model, DNase mediated resolution of the Holliday junction is proposed. Resolution of the Holliday junction by joining of the crossed strands (Fig. 4.9 a, i, horizontal arrow) or the uncrossed strands (Fig. 4.9 a, i, vertical arrow) leads to formation of two possible products (Fig. 4.9 a, ii).

The Holliday model forms the basis of models proposed to describe homologous recombination in eukaryotes including mammals. The primary models which can be applied to homologous recombination of replacement targeting vectors, are the double-strand break repair (Szostak *et al.*, 1983) and single strand annealing (Lin *et al.*, 1984) models. In the double-strand-break repair model (Szostak *et al.*, 1983), a double stranded gap within the homology is required to initiate the process of strand invasion (Fig. 4.9 b, i). Exonuclease action may be required to produce single stranded regions required for strand invasion in which the 3' end invades the homologous duplex DNA displacing a D loop (Fig. 4.9 b, ii). The D loop is enlarged by repair synthesis (Fig. 4.9 b, iii) until the invading 3' end can anneal to the complementary single stranded end. Repair synthesis in the second duplex results in completion of the repair process (Fig. 4.9 b, iv). Branch point migration and resolution of the Holliday junctions occurs as in the Holliday model (Fig. 4.9 b, iv) leading to two possible products in each of the duplexes (Fig. 4.9 b, v). In the single stranded annealing model (Lin *et al.*, 1984) double stranded breaks in the DNA close to, but not necessarily within the homology, in both the donor and recipient strand are required to initiate recombination (Fig. 4.9 c, i). The DNA ends are then specifically degraded exposing complementary single stranded regions which can form heteroduplex DNA (Fig. 4.9 c, ii). The gaps are repaired to complete the recombinant product (Fig. 4.9 c, iii).

Figure 4.9 Models of recombination

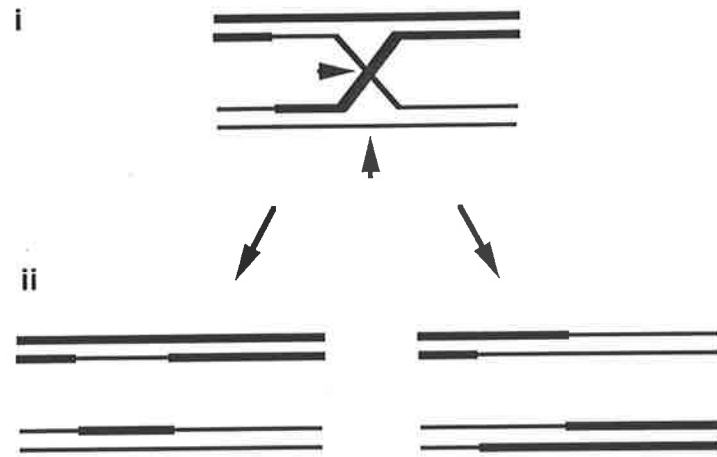
a Representation of the Holliday model (Holliday, 1964). (i) A single stranded break has initiated strand invasion and ligation to the complementary homologous strand. The structure formed, consisting of heteroduplex DNA and a branch point is known as the Holliday structure. Resolution of the Holliday junction can occur by joining of the crossed strands (horizontal arrow) or the uncrossed strands (vertical arrow). (ii) The resulting structures formed from resolution of the crossed strands (left) and uncrossed strands (right) are illustrated. Adapted from Szostak *et al.*, (1983).

b Representation of the double strand gap repair model. (i) A double stranded cut is made in one duplex, and a gap flanked by 3' single strands is formed by exonuclease activity. (ii) A 3' end invades the homologous duplex DNA displacing a D loop. (iii) The D loop is enlarged by repair synthesis until the invading end can ligate to the complementary homologous strand. (iv) Repair synthesis from the second 3' end completes the process of gap repair and results in the formation of two Holliday junctions. (v) Resolution of the Holliday structures can occur by joining of the crossed or uncrossed strands resulting in two possible crossover configurations. The right hand junction was resolved by joining of the crossed strands. Taken from Szostak *et al.*, (1983).

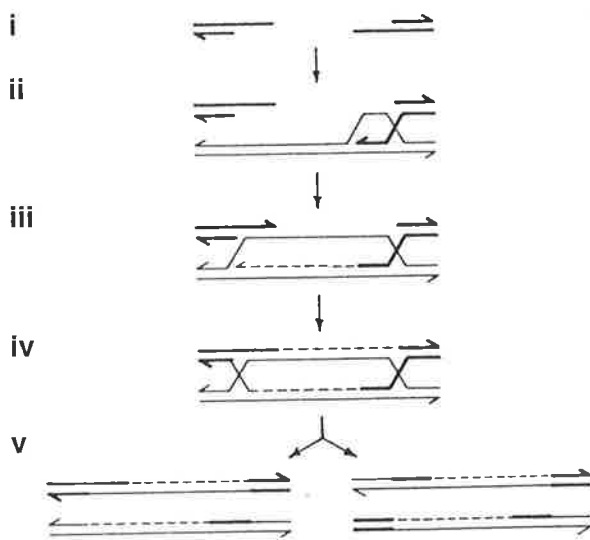
c Representation of the single stranded annealing model (Lin *et al.*, 1984). (i) Double stranded breaks in both duplexes are required to initiate recombination and gaps are formed by strand-specific exonuclease activity (3' exonuclease activity is illustrated), exposing complementary single stranded ends. (ii) The complementary single strands anneal. (iii) Repair synthesis from the 3' ends of both strands and (iv) ligation completes the process. Taken from Lin *et al.*, (1990). The upper duplex represented by the thick line is analogous to homologous targeting vector DNA in the figures **a-c**.

d Representation of the Hasty model (Hasty *et al.*, 1992). (i) Strand invasion is initiated by the formation of D loops without the involvement of DNA ends leading to formation of heteroduplexes. Heteroduplexes can be resolved by joining of the crossed (horizontal line) or (ii) uncrossed strands (vertical line). (iii) Resolution of the Holliday junctions can result in the formation of the illustrated structure in which a gene replacement event has occurred. Thin line, genomic DNA; Thick line, homologous DNA on targeting vector; wavy lines, terminal heterologous sequence; open box, selection cassette. Taken from Hasty *et al.* (1992).

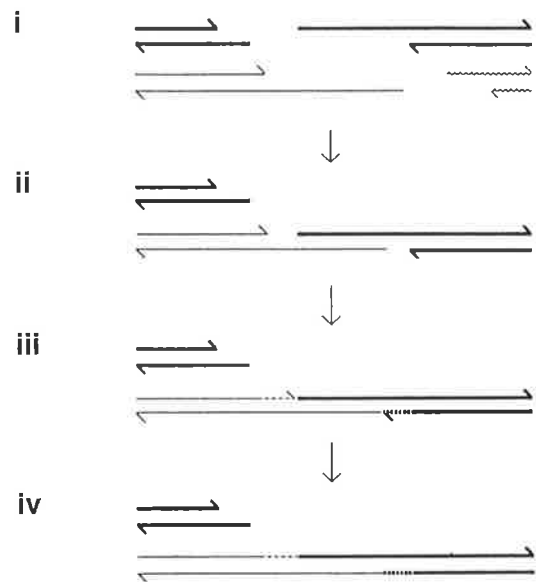
a



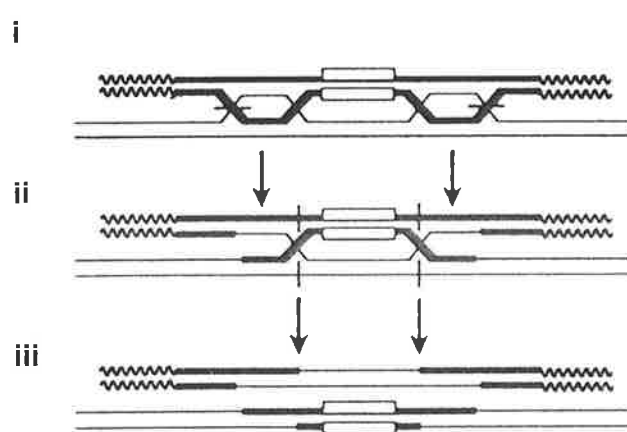
b



c



d



Hasty and co-workers (1992) proposed a modification of the double-strand break repair model to describe homologous recombination of replacement targeting vectors into the mammalian genome. To account for the non participation of heterologous DNA ends of replacement vectors in the recombination products, Hasty and co-workers (1992) proposed that strand invasion and D loop formation would occur close to the extremities of the targeting vector (Fig. 4.9 d, i). Heterologous ends cannot participate in the basepairing required to stabilise the formation of heteroduplex DNA since homologous DNA is not present at the site of integration. In this model, the heterologous ends remain base paired and do not participate in initiation of the recombination event. A double-strand break within the homology contained on the vector or chromosome, or invasion without a break is proposed to initiate the recombination process. Resolution of the Holliday junctions then proceeds as described in the Holliday model.

If DNA breaks or gaps or D loop formation, required to initiated homologous recombination, can occur randomly throughout the genome or targeting vectors, up to fifty percent of functional crossover events would occur between the HPRT or *neo* selection cassettes contained on pHNLX since close to fifty percent of the homology is contained between the HPRT and *neo* cassettes on this vector. This would result in integration of either the HPRT or *neo* genes but not both into the non-homologous site. Under conditions of selection for HPRT, one would expect that up to half of the HAT^r lines would carry the HPRT but not *neo* gene correctly integrated.

The results presented in this chapter suggest that D loops and/or single stranded nicks or double stranded gaps in either the vector or chromosome did not form between the HPRT and *neo* selection cassettes since functional crossover events did not occur in this region of the vector. It can therefore be concluded that either these do not form at all or form very infrequently in ES cells. As a consequence, none of the existing models adequately account for the findings presented in this chapter.

4.5.2 *Proposed Model for Homologous Recombination of Replacement Vectors*

Aspects of the Hasty model (1992) in conjunction with the findings of this work have been used to form a model to account for homologous recombination of replacement vectors in ES cells, which is illustrated in figure 4.10. The model accounts for the integration of replacement vectors carrying homologous or heterologous ends, or as in the case of pHNLX, one homologous and one heterologous end.

i. In order to initiate homologous recombination, the mammalian recombination machinery attaches to the free DNA end at the terminus of the replacement vector (Fig. 4.10 i). Exonuclease activity may be required to create an extended single stranded region to which the machinery can attach and the homology search then ensues.

ii. The recombination machinery “scans” from the end of the targeting vector, displacing the complementary strand and thereby creating a single stranded region of DNA. If homologous DNA is in the vicinity, strand invasion may occur leading to the formation of heteroduplex DNA between the incoming and chromosomal DNA, which displaces the homologous chromosomal strand. A nick induced at the end of the 5' D loop would permit ligation of the vector DNA to the complementary chromosomal end (Fig. 4.10 ii, left).

iii. If heterologous sequences are contained at the end of the targeting vector these will reanneal after the passage of the recombination machinery due to the absence of contiguous homologous sequence within the chromosome. In this case a D loop with two crossovers would form (Fig. 4.10 ii, right) and the terminal heterologies would not participate in the recombinant product as described in the model proposed by Hasty and co-workers (1992). Alternatively, DNA not participating in the recombination process may be removed by degradation.

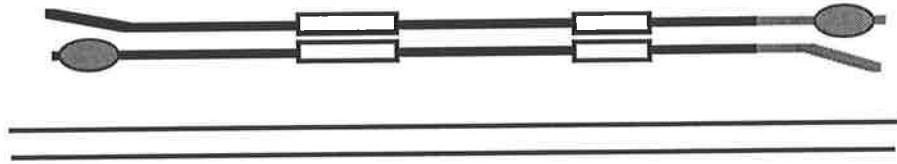
iv. Holliday junctions are formed and branch point migration ensues (Fig. 4.10 iii, horizontal arrow) until a crossover event occurs. Resolution of the Holliday junctions by joining of the appropriate strands may lead to formation of a recombination event in which the

Figure 4.10 Proposed model of homologous recombination

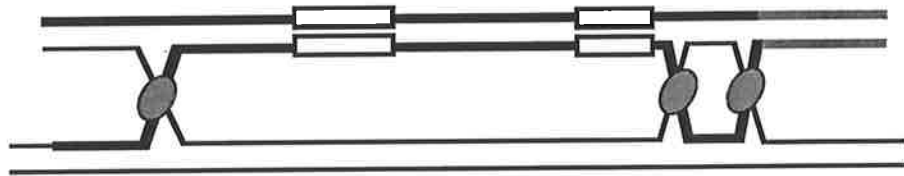
Representation of the model proposed to account for homologous recombination of replacement vectors. (i) The recombination machinery attaches to the single stranded DNAs end of the targeting vector. (ii) Strand invasion and formation of the Holliday structure occurs as proposed in the Holliday model if heterologous DNA is not present at the terminus of the vector (left). Strand invasion and formation of two D loops occurs if heterologous DNA is present at the terminus of the vector (right). (iii) Branch point migration ensues (horizontal arrow) until resolution of the Holliday junction occurs or a large heterologous region is encountered. In the latter case the recombination event may abort (boxed) or the recombination machinery may stall. (iv) Resolution of each Holliday junction can occur by joining of the crossed or uncrossed strands leading to formation of a gene replacement event. In the illustrated example the Holliday junction on the left is resolved by joining of the uncrossed strands and the Holliday junctions on the right are resolved by joining of the crossed strands at the inner junction and the uncrossed strands at the outer junction.

Stippled oval, recombination machinery; thin line, genomic locus; thick line, homologous DNA on targeting vector; stippled line, non-homologous plasmid DNA; open boxes, selection cassettes.

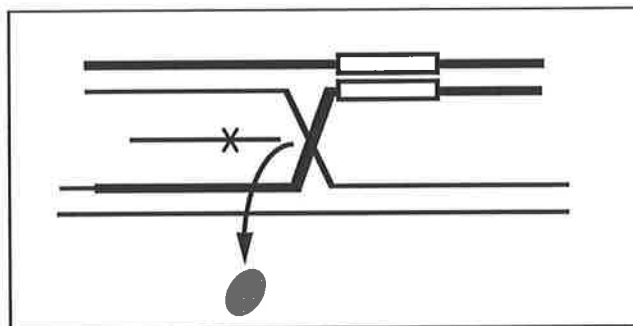
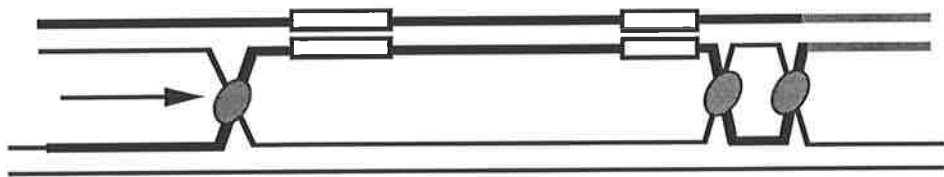
i



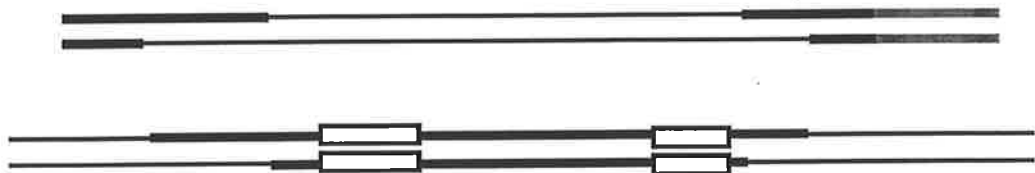
ii



iii



iv



cognate locus is replaced with the vector sequence as has been proposed (Hasty *et al.*, 1992; Fig. 4.10 iii, iv).

v. If a large region of heterologous DNA is encountered prior to formation of a crossover event, branch point migration would be prevented due to the absence of base-pairing which is required to stabilise the interaction. Two outcomes can be envisaged. The recombination machinery may “fall off” and the recombination event would be aborted (Fig. 4.10 boxed), or the recombination machinery may pause, possibly increasing the chance of a crossover event before the non-homologous sequence is encountered.

The novel aspects of this model are (i) the essential involvement of the vector ends in initiation of the recombination event (ii) the scanning mechanism of the recombination machinery and (iii) the role of two separate large regions of non-homology in preventing crossover formation between them.

In addition to the results of this study, a number of lines of evidence provide support for the proposed model. Parallels between the gene products involved in homologous recombination of simpler organisms and those of mammals are now becoming apparent. Single stranded DNA ends are substrates for the RecA protein which carries out the homology search, heteroduplex formation and crossover events in bacteria (Heyer, 1994). In the yeast *Saccharomyces cerevisiae* the protein encoded by *RAD51* is required for repair of double strand breaks in DNA and is homologous to RecA (Shinohara *et al.*, 1993), suggesting that the mechanisms of recombination are conserved between prokaryotes and eukaryotes. The *RAD51* and RecA proteins form very similar structures with double stranded DNA in the presence of ATP (Ogawa *et al.*, 1993). Homologues of the bacterial molecules involved in recombination have been identified in eukaryotes including mammals (Morita *et al.*, 1993; Shinohara *et al.*, 1993; Bezzubova *et al.*, 1994). RAD 51 homologues have been identified in yeast, filamentous fungi, avian and mammalian species (Heyer, 1994) suggesting that the basic mechanisms of recombination have been conserved throughout evolution (Shinohara *et al.*, 1993). By analogy to the bacterial system, DNA ends would be required for the initiation

of homologous recombination in eukaryotes. Other models for homologous recombination in mammals propose the involvement of DNA breaks within incoming DNA to initiate the recombination process. I propose that where double stranded DNA ends at the termini of a replacement vector are available they are used most frequently to initiate the recombination process.

The involvement for DNA ends in mammalian recombination is well established. DNA ends are involved in the extrachromosomal homologous recombination (Kucherlapati *et al.*, 1984), homologous recombination of insertion vectors (Hasty *et al.*, 1992) and non-homologous recombination of targeting vectors (Folger *et al.*, 1982; Thomas *et al.*, 1986) in mammalian cells. The frequency of extrachromosomal recombination between two plasmids containing homologous sequences is enhanced by the inclusion of a double strand gap within the homology contained on one vector (Kucherlapati *et al.*, 1984). This indicates that double stranded DNA ends can act as substrates for the recombination machinery in mammalian cell lines. The double-strand-break repair model is used to describe homologous recombination of insertion vectors which carry two adjacent homologous ends (Hasty *et al.*, 1992). The ends of such vectors are thought to initiate the initial strand invasion event leading to the crossover events and as such would be the substrates for the recombination machinery. The fact that terminal heterologies contained on replacement vectors are included in the products of non-homologous recombination in mammalian cell lines (Folger *et al.*, 1982; Thomas *et al.*, 1986) also indicates that DNA ends mediate non-homologous integration of replacement vectors into the genome. The results presented in this chapter suggest that DNA ends of a replacement vector also mediate homologous recombination of this class of vectors in mammalian cells.

It has been proposed that homologous recombination of insertion and replacement vectors occurs by distinct pathways (Hasty *et al.*, 1991c; Hasty *et al.*, 1992). Homologous recombination of insertion vectors has been proposed to occur by a modified double strand-strand break repair model, initiated by the vector ends (Hasty *et al.*, 1992). The finding by Deng and Capecchi (1992) that insertion and replacement vectors have similar targeting efficiencies when constructed from isogenic DNA and the findings presented in this chapter,

suggest that the mechanisms of homologous recombination utilised by both types of vector may be similar. The results of this chapter suggest that homologous recombination of replacement vectors also occurs by a mechanism which is initiated by vector ends. Such a mechanism is also more closely related to that proposed for extrachromosomal homologous recombination, homologous recombination of insertion vectors and non-homologous recombination of replacement targeting vectors in mammalian cells. This suggests closer alignment between the mechanisms responsible for all types of recombination in mammalian cells.

In several published examples, selection cassettes have been introduced into the mouse genome on targeting vectors which also contain a region of non-homology positioned at a second site within the homologous sequence (Le Mouellic *et al.*, 1992; Gu *et al.*, 1993; Michalska and Choo, 1993). In each of these examples, the additional region of non-homology was less than sixty bp in length. In two such examples the additional regions of non-homology were not introduced into the recombinant locus along with the selection cassettes (Le Mouellic *et al.*, 1992; Gu *et al.*, 1993). In the third example however, all homologous recombinants contained both regions of non-homology (Michalska and Choo, 1993). It is likely that the additional short region of non-homology contained on these vectors would be insufficient to disrupt the base pairing required to stabilise branch point migration of Holliday junctions. Thus, a short region of non-homology would be unable to prevent the formation of functional crossovers between it another region of non-homology. The minimum length of heterologous sequence required to prevent passage of the mammalian recombination machinery and hence crossover events forming between two region of non-homology contained on a targeting vector is not known but falls between 60 bp and 1.9 kb. Defining this parameter will have useful applications for the future construction of replacement and stable tag-exchange targeting vectors.

Introduction of small nucleotide substitutions and a selection cassette into the genome on replacement vectors containing the selection cassettes positioned as far away as three kilobases from the substitution can be achieved with a frequency of close to ninety percent

(Capecchi, 1989). This provides additional evidence that crossover events most frequently occur near the ends of replacement vectors and implicates ends of such vectors in the recombination process.

The results presented in this chapter indicate that the tag step of the stable tag-exchange approach can be carried out efficiently in ES cells. In addition, features of the mechanistic model proposed to account for homologous recombination of replacement vectors suggest that greater flexibility can be applied to the future design of these vectors.

Chapter 5

5 Exchange Step Gene Targeting at the LIF Locus

5.1 Introduction

In work presented in this thesis, a modified tag-and-exchange gene targeting approach, known as the stable tag-exchange system, was designed to enable the efficient introduction of precise, subtle alterations repeatedly into any site of the mouse genome. The efficacy of the system, which utilises two rounds of gene targeting, was assessed at the LIF locus in ES cells. The first of the targeting steps, known as the tag step, was assessed in the preceding chapter. In this chapter, exchange step gene targeting at the LIF locus was investigated. The exchange step targeting vectors described in this chapter were designed to enable visualisation of the cellular sites of expression of each and all LIF isoforms in the mouse using a β -galactosidase reporter system. To this end, ES cells were produced which carry a reporter gene in the first unique exon of LIF D, M, T or in the second common exon of LIF, as described in chapter 3 (sections 3.7.1-2).

This chapter describes construction and characterisation of a reporter system for use in the exchange step vectors, construction of exchange step vectors and exchange step gene targeting at the LIF locus using these vectors. Direct comparison of the efficiency of the unmodified tag-and-exchange gene targeting system and that of the stable tag-exchange system is reported.

5.2 PCR-based Alteration of the *lacZ* Reporter Gene

In order to map the expression of each and all LIF isoforms at the cellular level, a *lacZ* reporter gene was required that would reflect accurately LIF expression in ES cells and the mouse. To achieve this, a *lacZ* reporter gene suitable for expression in mammalian cells and carrying a NLS (section 3.7.2) was cloned into the unique first exons or into the second common exon of LIF (section 5.6.1). Two forms of the *lacZ* cassette were constructed. The first, lacking an initiation codon, was cloned in frame in the first exons of LIF D and M in which case translation of the reporter gene would start at the native initiation codons of LIF D and M. The second, containing an initiation codon, was used in the first exon of LIF T and

the second common exon of LIF. Translation of this reporter gene would start at the initiation codon in the *lacZ* gene to reflect LIF T expression. Both *lacZ* cassettes were altered by a PCR-based approach such that they were contained on a Bgl II fragment. This enabled insertion of the ATG^{+/-} *lacZ* cassettes into unique Bgl II sites (section 5.6.1) which were introduced into the appropriate exons of LIF (section 5.5.1).

5.2.1 Alteration of Vector *pnlacF*

Vector *pnlacF* (J. Peschon, unpublished; Fig. 5.1 a) was used for the construction of a nuclear localised *lacZ* expression cassette suitable for use in exchange step targeting vectors. This vector contains the NLS of SV40 (Lanford *et al.*, 1988) in place of the first eight amino acids of the bacterial *lacZ* gene followed by the mouse protamine I intron, 3' untranslated region and polyadenylation sequence (Peschon *et al.*, 1987). The *lacZ* gene is efficiently expressed in mammalian cells (J. Peschon, personal communication).

PCR amplification of the sequence at the 5' end of the *lacZ* coding sequence of *pnlacF* was used to modify this region of the *lacZ* expression cassette (Fig. 5.1 b). Two 5' primers were used for PCR amplification and are illustrated in figure 5.1 b. The first primer was used to introduce a Bgl II site into the 5' end of the *lacZ* cassette (ATG⁺ clone, lacATG5') and the second was used to introduce an in frame Bgl II site into the *lacZ* cassette and remove the initiation ATG contained in *pnlacF* (ATG⁻ clone, lac5'). The 3' primer (lac3') was used to remove an internal ATG and enabled cloning of the PCR products via the Bam HI site contained within the primer (Fig. 5.1 b). Removal of the internal ATG was carried out to prevent initiation of translation at this ATG which would not reflect accurately expression initiated at the ATG sequences in the first exons of LIF D and M. The sequences of the products amplified by this approach are illustrated in figure 5.1 c.

5.2.2 Reconstruction of the *lacZ* Constructs

Direct cloning of PCR products 14ATG⁻ and 12ATG⁺ into *pnlacF* proved difficult and were therefore initially subcloned. PCR products 14ATG⁻ and 12ATG⁺ were endfilled and ligated into Sma I digested pT7T319U (Stratagene). PCR product 14ATG⁻ was ligated with the

Figure 5.1 Alteration of the *lacZ* cassette in pnlacF

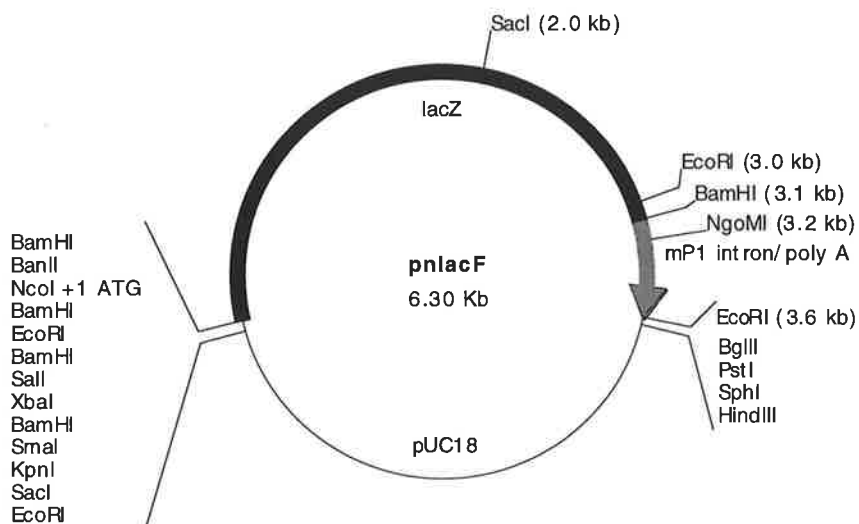
a Representation of the plasmid pnlacF showing the *lacZ* gene, mouse P1 intron and polyadenylation sequence (mP1 intron/poly A). The restriction sites for Bam HI, Eco RI and Sac I sites in the *lacZ* and mouse P1 sequence and additional sites in the polylinker of pnlacF are shown. Thick line, *lacZ* gene; thin line, pUC18 plasmid DNA; stippled line, mP1 intron and polyadenylation sequence.

b Representation of the 5' polylinker and coding sequence of pnlacF. The restriction sites for Bam HI, Eco RI, Kpn I, Sal I, Sma I, Sst I and Xba I (underlined), the NLS (bold) and the primers lac5', lacATG5' and lac3' used for amplification of the ATG⁻ *lacZ* sequence (14ATG⁻) and ATG⁺ *lacZ* sequence (12ATG⁺), respectively are shown.

c Sequence of the 5' regions of the *lacZ* gene amplified by PCR and termed 14ATG⁻ and 12ATG⁺. The restriction sites for Bam HI and Bgl II and the primer sequence (bold) are shown.

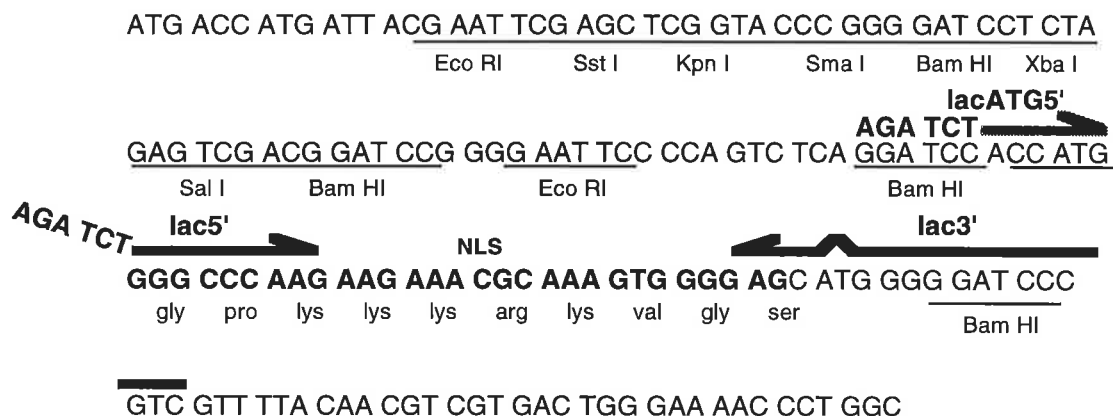
a

Map of pnlacF



b

5' Polylinker and Coding Sequence of pnlacF



c

Sequence of the 5' Region of the lacZ Gene Altered by PCR

Bgl II

14ATG⁻

AGA TCT GGG CCC AAG AAG AAA CGC AAA GTG GGG AGC GTG GGG GAT CCC GTC
TCT AGA CCC GGG TTC TTC TTT GCG TTT CAC CCC TCG CAC CCC CTA GGG CAG

Bam HI

Bgl II

12ATG⁺

AGA TCT ACC ATG GGG CCC AAG AAG AAA CGC AAA GTG GGG AGC GTG GGG GAT CCC GTC
TCT AGA TGG TAC CCC GGG TTC TTC TTT GCG TTT CAC CCC TCG CAC CCC CTA GGG CAG

Bam HI

Bgl II site proximal to the Eco RI site of the polylinker to form p14T7, and 12ATG⁺ was ligated in the opposite orientation to form p12T7.

5.2.3 Construction of the ATG⁻ lacZ Vector

Vector p14T7 was linearised at the Eco RI site within the polylinker of pT7T319U and endfilled. The PCR product 14ATG⁻ was isolated by digestion with Bam HI (Fig. 5.2 i) and ligated into Sma I/Bam HI digested pUC18 containing the 3' fragment of pnlacF (Fig. 5.2 ii) to form p14lacZ. The altered pnlacF vector was reconstructed by reinsertion of the 3.1 kb Bam HI fragment of pnlacF (Fig. 5.2 iii) into Bam HI digested p14lacZ to form pnlacF14 (Fig. 5.2 iv, v). This vector contains an ATG⁻ NLS⁺ lacZ gene, suitable for expression in eukaryotic cells, contained on a Bgl II cassette of 3.6 kb. Sequence analysis at the 5' end of the ATG⁻ lacZ gene was carried out to confirm that the lacZ cassette had been altered correctly for in frame cloning.

5.2.4 Construction of the ATG⁺ lacZ Vector

The PCR product 12ATG⁺ was isolated from p12T7 by digestion with Bam HI (Bam HI cuts at the 3' end of the PCR product and at the 5' end within the polylinker of pT7T319U, Fig. 5.3 i) and ligated into the pUC18 fragment of Bam HI digested pnlacF containing the 3' fragment of pnlacF (Fig. 5.3 ii) to form p12lacZ. p12lacZ was digested with Ban II and Ngo MI and ligated to the 2 kb Ban II/Sac I and 1.1 kb Sac I/Ngo MI fragments of pnlacF (Fig. 5.3 iii, iv) to form pnlacF12 (Fig. 5.3 v). This vector contains a nuclear localised lacZ gene with its own ATG, suitable for expression in eukaryotic cells, contained on a Bgl II cassette of 3.6 kb. Sequence analysis of the 5' end of the ATG⁺ lacZ gene was carried out to confirm that the lacZ cassette had been altered correctly.

5.3 Construction of COS-1 and ES Cell β -Galactosidase Expression Vectors

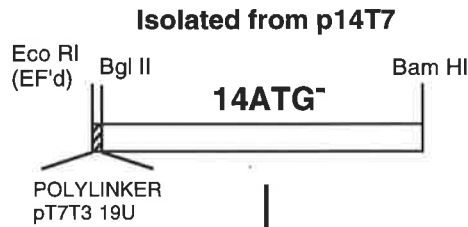
It was important to demonstrate that the ATG⁺ lacZ reporter gene (section 5.2.4) could be expressed efficiently in mammalian cells in vitro before use in the exchange step targeting constructs. Expression was assessed in the context of three different eukaryotic expression vectors. The lacZ cassettes from pnlacF14 and pnlacF12 (Fig. 5.4 i) were isolated as Bgl II

Figure 5.2 Construction of the ATG⁻ *lacZ* cassette

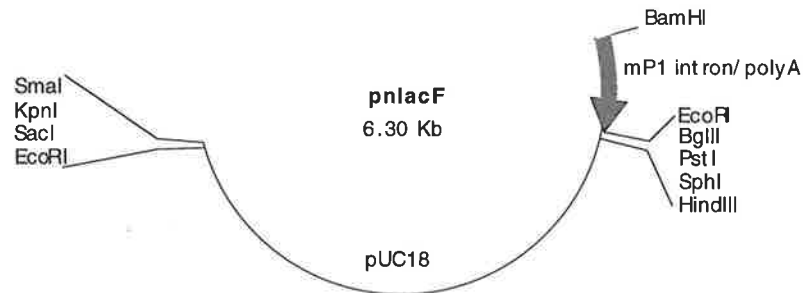
Representation of the construction of pnlacF14 showing (i) the cloning of 14ATG⁻ isolated from p14T7 into (ii) Sma I/Bam HI linearised pnlacF and the restoration of (iii) the Bam HI fragment of pnlacF into (iv) p14lacZ to form (v) pnlacF14. Restriction sites for Bam HI, Bgl II, Eco RI, Hind III, Kpn I, Pst I, Sac I, Sma I, Sph I are shown. EF'd, endfilled; mP1 intron/ poly A, intron and polyadenylation sequence of the mouse mP1 gene; open box, PCR product; thick line, *lacZ* gene; thin line, pUC18 plasmid DNA; stippled line, mP1 intron and poly adenylation sequence.

Construction of the ATG⁻ *lacZ* Cassette

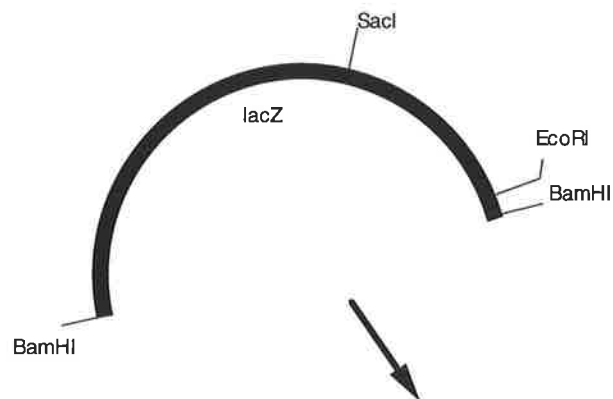
i



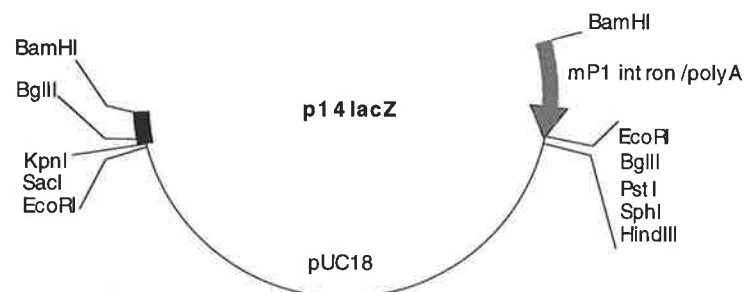
ii



iii



iv



v

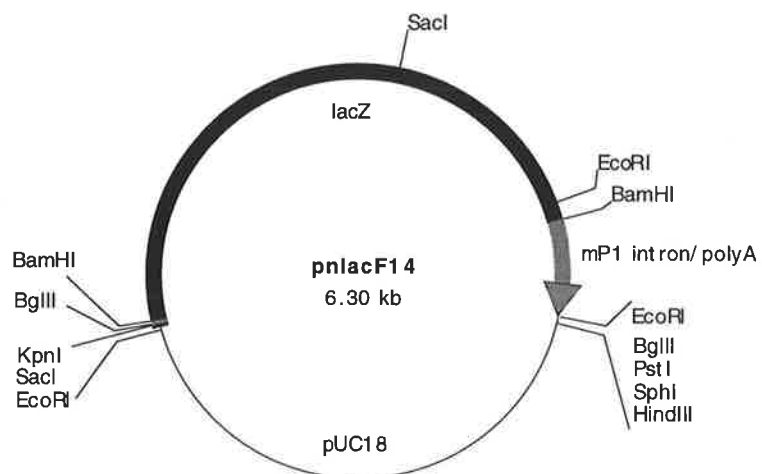
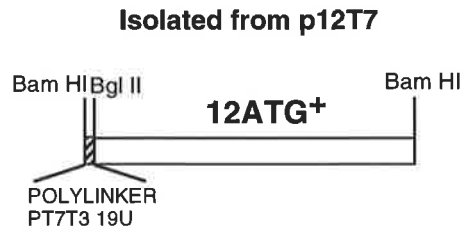


Figure 5.3 Construction of the ATG⁺ *lacZ* cassette

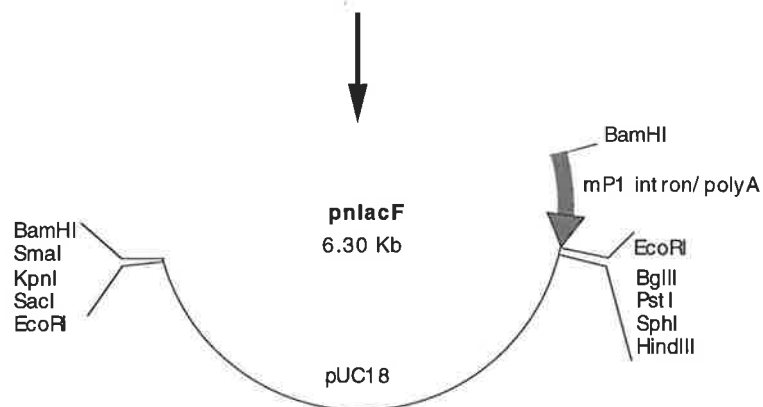
Representation of the construction of pnlacF12 showing (i) the cloning of 12ATG⁺ isolated from p12T7 into (ii) Bam HI linearised pnlacF and the restoration of (iii) the Ban II/Sac I and Sac I/Ngo MI fragments of pnlacF into (iv) p12lacZ to form (v) pnlacF12. Restriction sites for Bam HI, Ban II, Bgl II, Eco RI, Hind III, Kpn I, Ngo MI, Pst I, Sac I, Sma I, Sph I are shown. mP1 intron/poly A, intron and polyadenylation sequence of the mouse mP1 gene; open box, PCR product; hatched box, region of pT7T319U polylinker; thick line, *lacZ* gene; thin line, pUC18 plasmid DNA; stippled line, mP1 intron and poly adenylation sequence.

Construction of the ATG⁺ *lacZ* Cassette

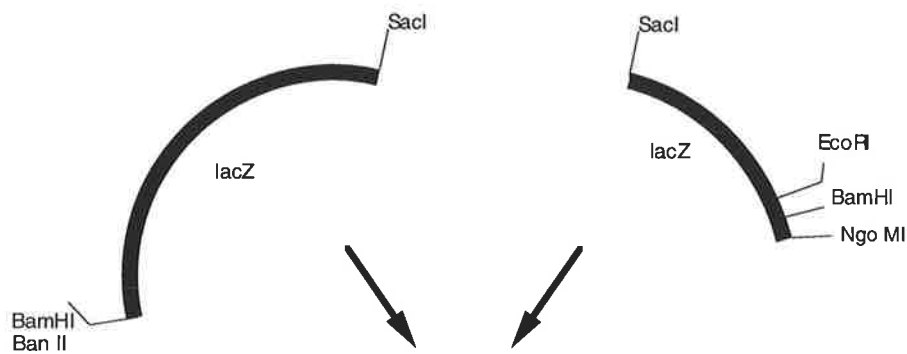
i



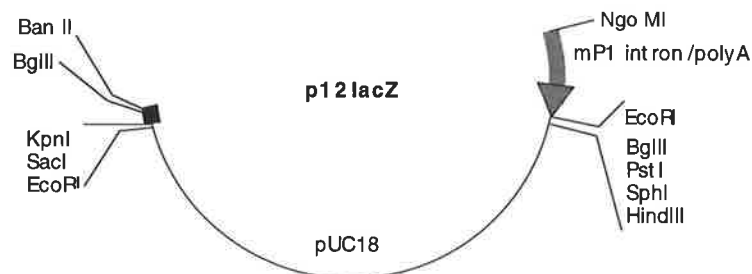
ii



iii



iv



v

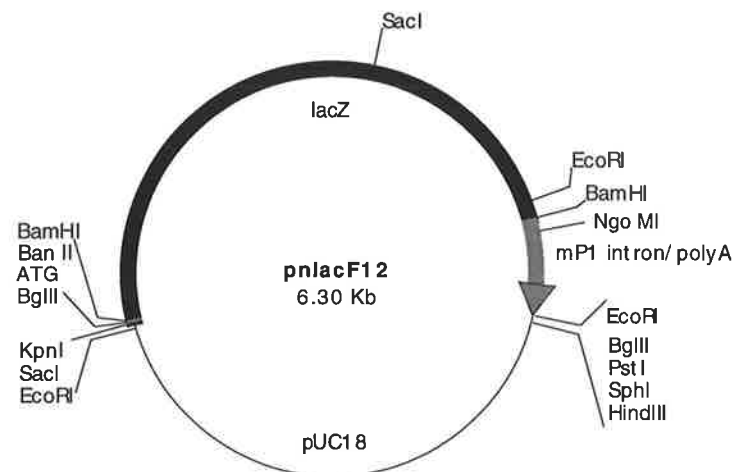
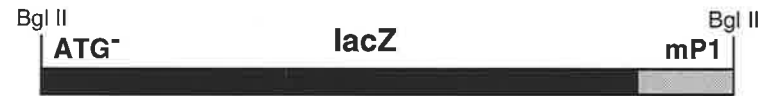


Figure 5.4 Construction of the COS-1 and ES cell *lacZ* Expression Vectors

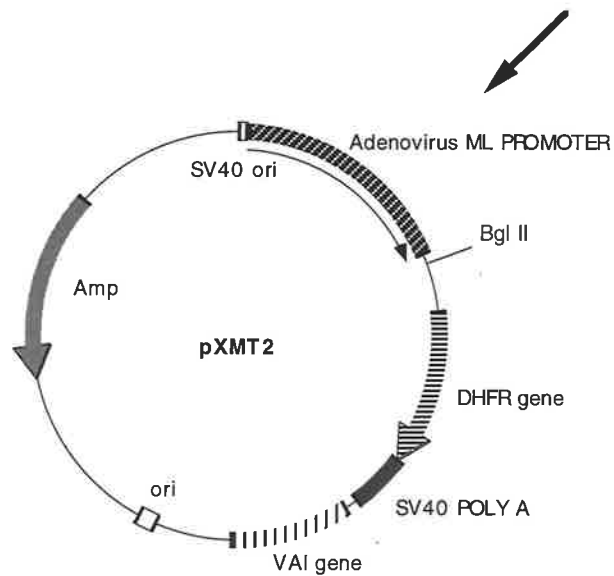
Representation of the construction of constitutive *lacZ* expression vectors for use in COS-1 cells and constitutive and inducible *lacZ* expression vectors for use in ES and X cells showing (i) the ATG⁻ and ATG⁺ *lacZ* cassettes isolated as Bgl II fragments from pnlacF14 and pnlacF12, respectively and (ii) the vectors pXMT2, pPOX/p1OX into which the cassettes were cloned. Dark box, coding sequence of the *lacZ* gene; stippled box, mouse P1 intron and polyadenylation sequence; Amp, ampicillin resistance gene; SV40 ori, SV40 origin of replication; Adenovirus ML promoter, adenovirus major late promoter; DHFR gene, dihydrofolate reductase gene; SV40 POLY A, polyadenylation sequence of SV40; VAI gene, adenovirus VAI gene; ori, bacterial origin of replication; mPGK-1/6-16 (1) promoter, mouse *Pgk-1* promoter or human 6-16 promoter (1) (section 2.2).

Construction of lacZ Expression Vectors

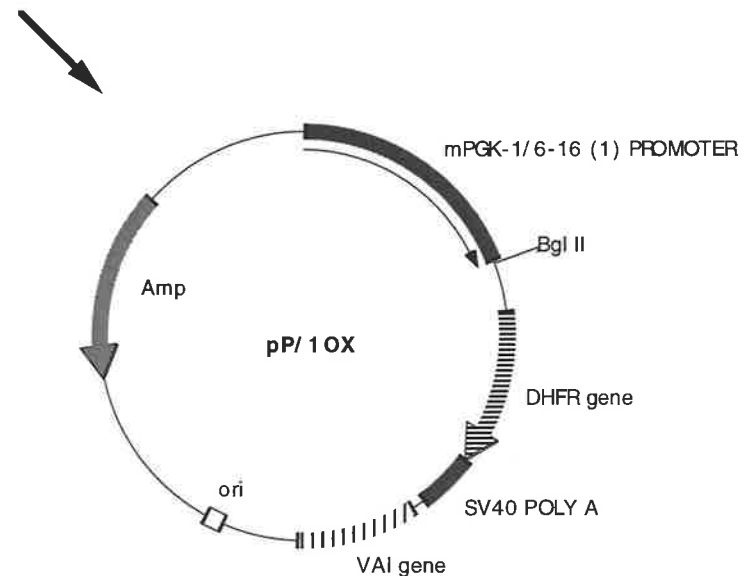
i



ii



**Constitutive COS-1 Cell
Expression Vectors**



**Constitutive and Inducible ES Cell
Expression Vectors**

fragments and ligated into the Bgl II site of pXMT2 (Fig. 5.4 ii, left) to form the β -galactosidase expression vectors plac14MT2 (ATG⁻) and plac12MT2 (ATG⁺), suitable for use in COS-1 cells. The *lacZ* cassettes were also cloned into the Bgl II sites of pPOX (Fig. 5.4 ii, right; pPOX is a constitutive ES cell expression vector containing the promoter of the *Pgk-1* gene in place of the 6-16 promoter in p1OX) and p1OX (section 2.2). The resultant vectors termed pPlac14X and pPlac12X were constitutive *lacZ* ES cell expression vectors, while vectors p1lac14X and p1lac12X were IFN-inducible *lacZ* ES cell expression vectors.

5.4 Analysis of LacZ Expression In COS-1, ES and X Cells

Analysis of *lacZ* expression was carried out in COS-1 cells, ES cells and a differentiated derivative of ES cells known as X cells (section 1.6.5). COS-1 cells were used because they are a differentiated cell type (African green monkey, kidney cell line, ATCC, 1994) and have a large cytoplasm and distinct nucleus, allowing clear visualisation of nuclear β -galactosidase activity. In contrast, undifferentiated ES cells have negligible cytoplasm and it is difficult to distinguish between cytoplasmic and nuclear staining in these cells. To map LIF expression in the mouse, the *lacZ* reporter gene must be functional in the differentiated progeny of ES cells. The ability of the *lacZ* cassette to function in X cells, a differentiated derivative of ES cells, was therefore assessed. Analysis of the *lacZ* expression in the context of an inducible ES cell expression vector was carried out to ensure that the reporter gene reflected accurately promoter activity in mammalian cell lines.

COS-1 or ES cells (tag step gene targeted line HG60; section 4.3.3) were transiently transfected with circular plasmid DNA. COS-1 cells were transfected with pCH110 (Pharmacia, constitutive β -galactosidase expression vector), plac14MT2 or plac12MT2 and histochemical staining for β -galactosidase activity was carried out after 48 hours. ES cells were transfected with pPlac14X, pPlac12X, p1lac14X, p1lac12X, p1OX or pCH110. Cells were seeded into complete ES cell medium or into 50% Med II medium (conditioned medium required for X cell formation) and left for two to three days to permit X cell formation. After transfection with the constitutive expression vectors, cells were grown for 48 hours prior to histochemical staining for β -galactosidase activity. Cells transfected with the inducible

expression vectors were left uninduced or were induced with 1000 U/ml IFN for 24 hours prior to histochemical staining for β -galactosidase activity.

Cytoplasmic β -galactosidase activity was detected in COS-1 cells transfected with pCH110 (Fig. 5.5 a), cells transfected with plac12MT2 displayed nuclear β -galactosidase activity (Fig. 5.5 b, c, d), while no β -galactosidase activity was detected in parallel cultures transfected with plac14MT2 carrying the ATG⁻ *lacZ* cassette (results not shown). The level of β -galactosidase activity varied from cell to cell (Fig. 5.5 c, d) but was similar between cultures transfected with pCH110 or plac12MT2, with high level expressing cells in both cultures (Fig. 5.5 a, c). However, expression could not be directly compared due to the fact that the promoters in the test and control plasmids are different and because pCH110 is able to replicate in COS-1 cells.

Similarly, cytoplasmic β -galactosidase activity was detected in ES and X (results not shown) cells transfected with pCH110 (Fig. 5.6 a), nuclear β -galactosidase activity was detected in both ES and X cells transfected with pPlac12X (Fig. 5.6 b) while no β -galactosidase activity was detected in parallel cultures transfected with pPlac14X which lacks an initiation ATG (not shown). Diffusion of β -galactosidase activity to surrounding cells from cells expressing cytoplasmic β -galactosidase activity can be seen in figure 5.6 a. Using the nuclear localised form of *lacZ*, diffusible β -galactosidase activity is likely to be contained within the cytoplasm of the expressing cell. Nuclear β -galactosidase activity was seen in ES and X cell cultures transfected with the IFN-inducible expression vector p1lac12X only after treatment with IFN (Fig. 5.6 c, d). In contrast, inducible β -galactosidase activity was not seen in ES or X cell cultures transfected with pPlac14X carrying the ATG⁻ *lacZ* cassette (results not shown). Fewer cells in cultures transfected with the pnlacF derived ATG⁺ *lacZ* cassette expressed β -galactosidase activity when compared to ES or X cell cultures transfected with pCH110 (Table 5.1, 5.2), but the level of expression between cells transfected with pCH110 and plac12MT2, pPlac12X or p1lac12X was similar. Direct comparison of the level of expression was not possible due to the fact that the promoters in the test and control plasmids are different.

Figure 5.5 Analysis of β -galactosidase Expression in COS-1 cells

a Photograph of β -galactosidase activity produced by the positive control β -galactosidase expression vector pCH110 in transiently transfected COS-1 cells, showing cytoplasmic staining (photographed under bright field optics at 100 x magnification).

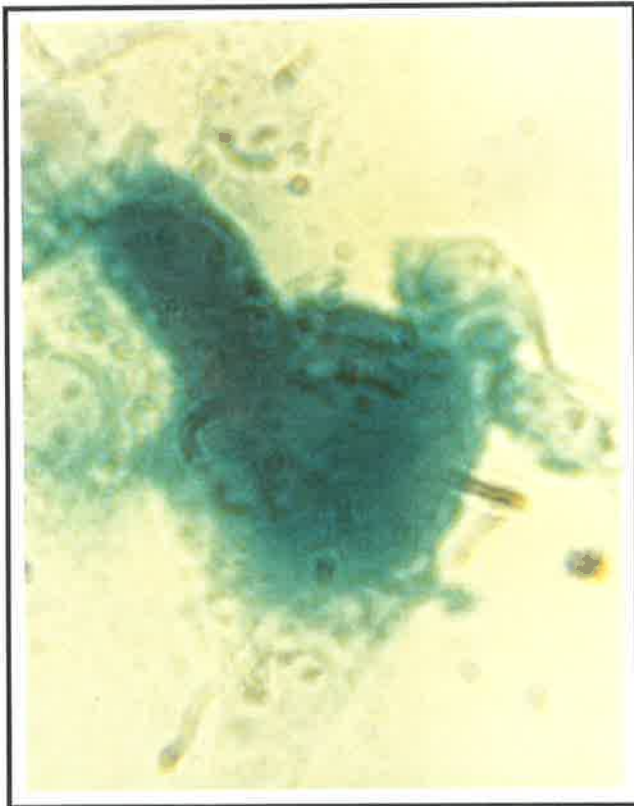
b Photograph of β -galactosidase activity produced by the β -galactosidase expression vector plac12MT2 in transiently transfected COS-1 cells, showing nuclear staining (photographed under bright field optics at 100 x magnification).

c Photograph of β -galactosidase activity produced by the β -galactosidase expression vector plac12MT2 in transiently transfected COS-1 cells, showing high levels of nuclear staining (photographed under bright field optics at 1000 x magnification).

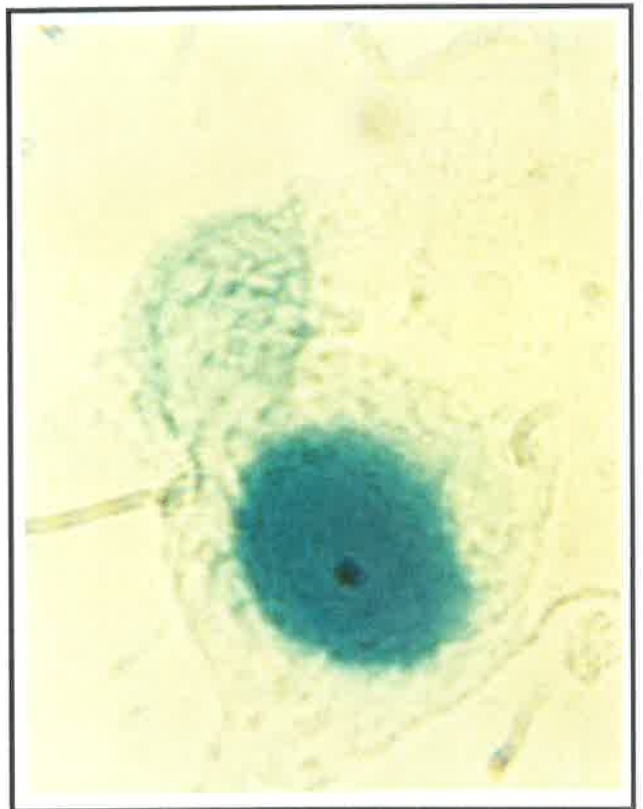
d Photograph of β -galactosidase activity produced by the β -galactosidase expression vector plac12MT2 in transiently transfected COS-1 cells, showing low levels of nuclear staining (photographed under bright field optics at 1000 x magnification).

Histochemical staining for β -galactosidase activity was carried out forty eight hours after transfection using X-gal (40 mg/ml) as the substrate. Photographs were taken on the Zeiss Axioplan Universal microscope.

a



b



c



d

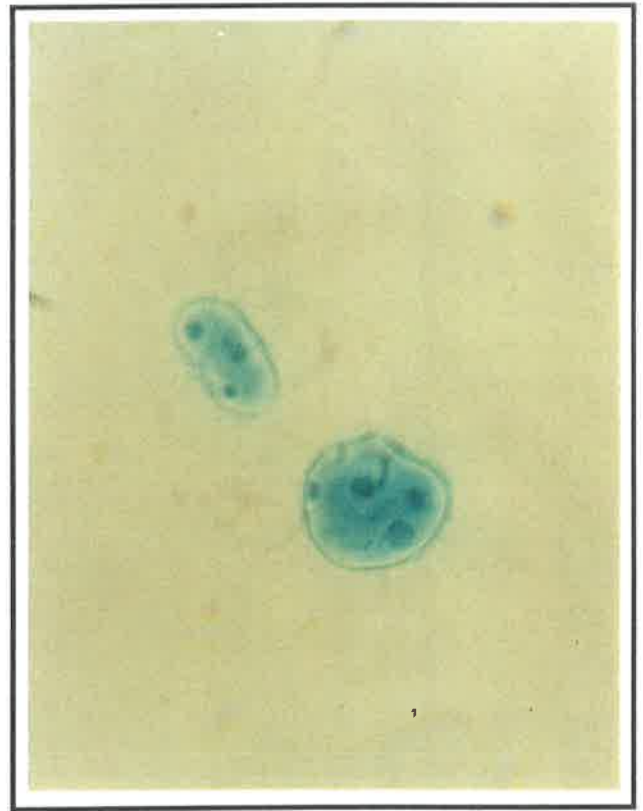


Figure 5.6 Analysis of β -galactosidase Expression in ES cells

a Photograph of β -galactosidase activity produced by the β -galactosidase expression vector pCH110 in transiently transfected ES cells, showing cytoplasmic staining (under bright field optics at 250 x magnification).

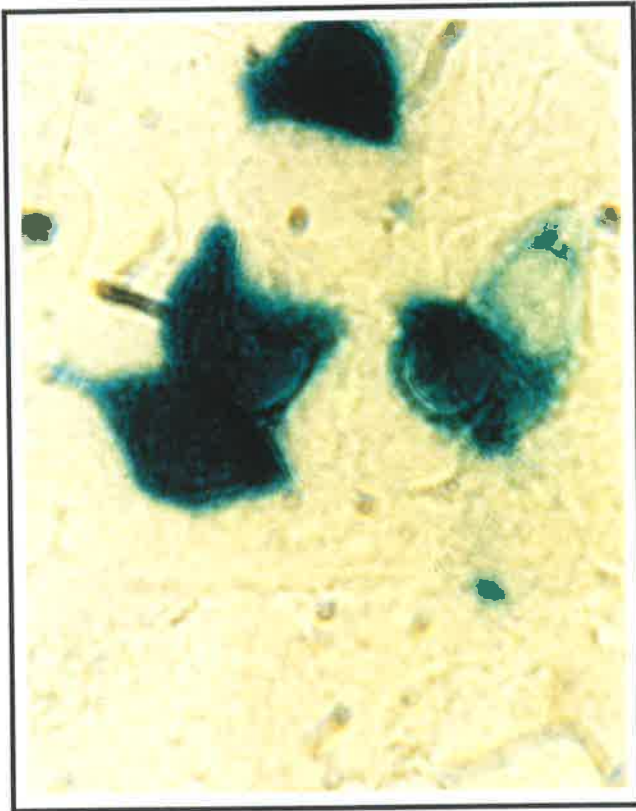
b Photograph of β -galactosidase activity produced by the β -galactosidase expression vector pPlac12X in transiently transfected HG60 ES cells showing nuclear staining (under bright field optics at 100 x magnification).

c Photograph of HG60 cells transiently transfected with p1lac12X and stained for β -galactosidase activity without IFN treatment (under bright field optics 250 x magnification).

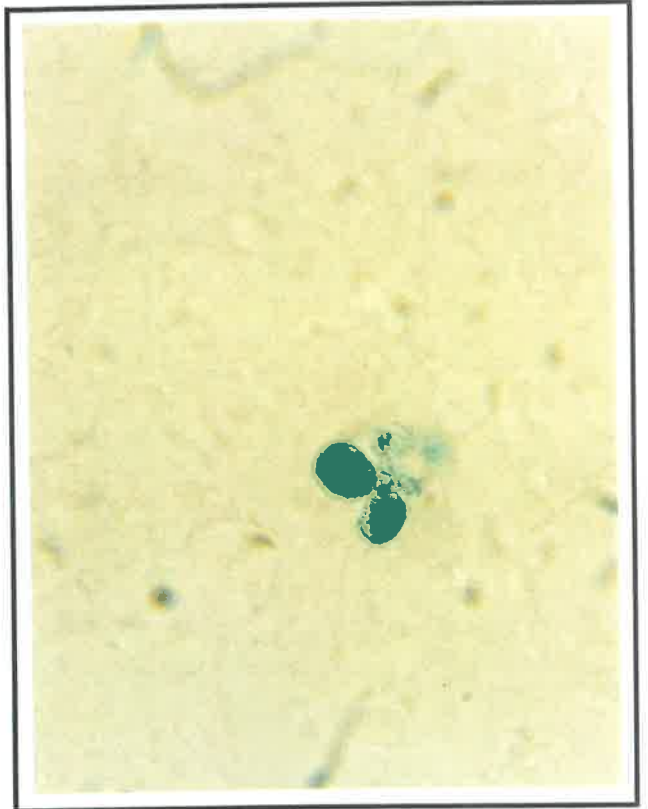
d Photograph of HG60 cells transiently transfected with p1lac12X and stained for β -galactosidase activity after treatment with IFN treatment (under bright field optics 250 x magnification).

Histochemical staining for β -galactosidase activity was carried out forty eight hours after transfection using X-gal (40 mg/ml) as the substrate. Photographs were taken on the Zeiss Axioplan Universal microscope.

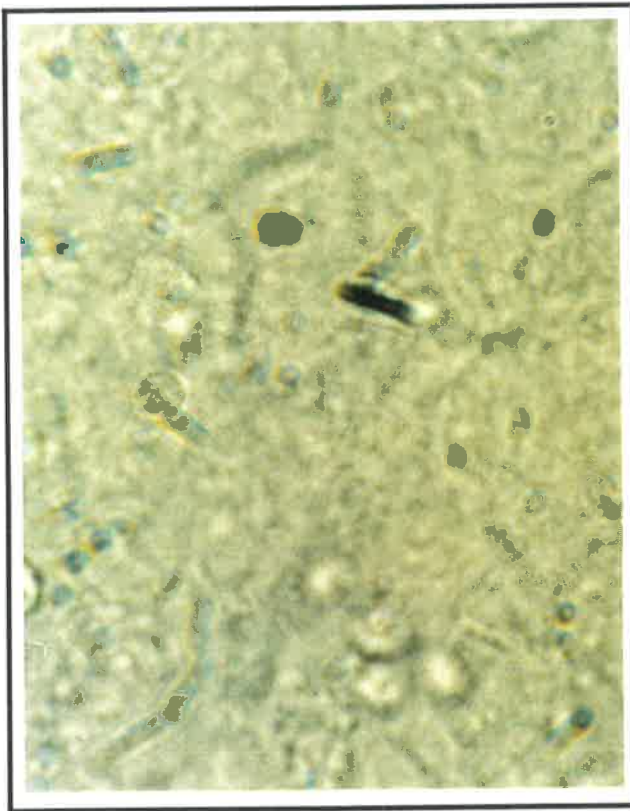
a



b



c



d

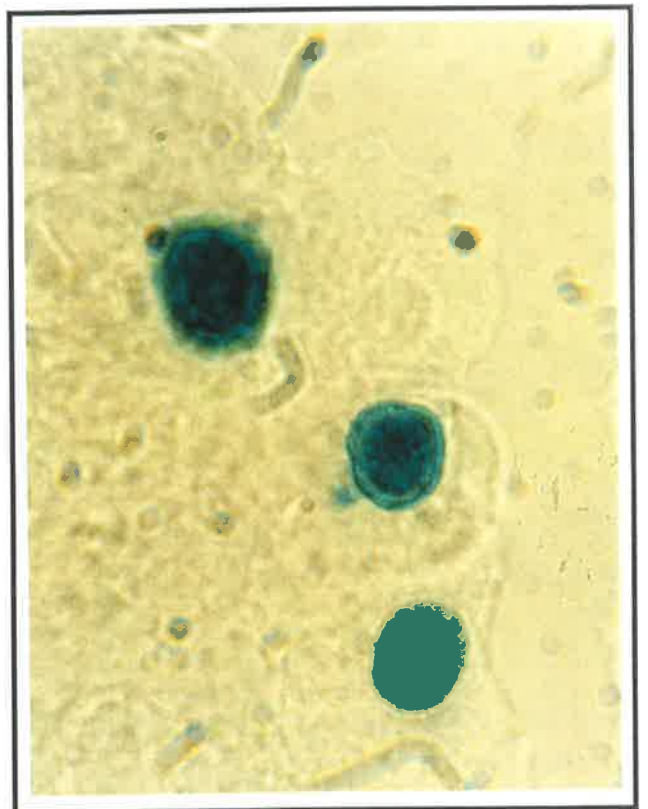


Table 5.1

The number of β -galactosidase positive ES and X cells in cultures transiently transfected with pCH110, pPlac12X or pPlac14X (as described in sections 5.4; 9.4.5; 9.4.13).

Table 5.2

The number of β -galactosidase positive ES and X cells in cultures transiently transfected with pCH110, p1OX, p1lac12X or p1lac14X with or without IFN induction (as described in sections 5.4; 9.4.5; 9.4.13).

Table 5.1

DNA	β-gal Positive ES Cells	β-gal Positive X Cells
pCH110	16	18
pPlac12X	4	3
pPlac14X	0	0

Table 5.2

DNA and Induction	β-gal Positive ES Cells	β-gal Positive X Cells
pCH110 - IFN	99	44
p10X - IFN	0	0
p1lac12X - IFN	0	0
p1lac12X + IFN	24	10
p1lac14X - IFN	0	0
p1lac14X + IFN	0	0

From the results presented in this chapter, it can be seen that the altered ATG⁺ β -galactosidase expression cassette is expressed efficiently, producing nuclear β -galactosidase activity in a range of cell types under the control of both constitutive and inducible promoters. An initiation ATG is required for *lacZ* expression since vectors carrying *lac14* did not direct β -galactosidase expression. Since the ATG⁻ β -galactosidase expression cassette has all the features of its ATG⁺ counterpart with the exception of an initiation codon, these results also suggest that once cloned as a fusion protein the ATG⁻ β -galactosidase expression cassette will also be expressed efficiently in a range of cell types. These features make the β -galactosidase expression cassettes appropriate for use in the exchange step targeting vectors.

5.5 Alteration of the Genomic LIF Clone by PCR

5.5.1 Introduction of Unique Bgl II Sites into the First and Second Exons of LIF

Introduction of a *lacZ* reporter gene into each and all forms of the endogenous LIF gene required creation of four unique restriction sites at an appropriate position on the genomic LIF clone used for vector construction. The restriction site chosen for this purpose was Bgl II since the site is not contained in the genomic LIF clone or *lacZ* coding region. Unique Bgl II sites were introduced into the first exons of LIF D, M, T and the second exon of the genomic LIF clone contained in pLIFXT7t. The Bgl II sites were introduced using a PCR based approach and the LIF sequence in pLIFXT7t was reconstructed using unique restriction sites contained at the ends of each PCR product. The 3 kb Bcl I LIF fragment from pLIFXT7t, which carries each of the first exons and the second exon of LIF, was used as the template for PCR amplification and for initial cloning of the PCR products.

5.5.2 Subcloning of the 3 kb Bcl I Fragment of LIF

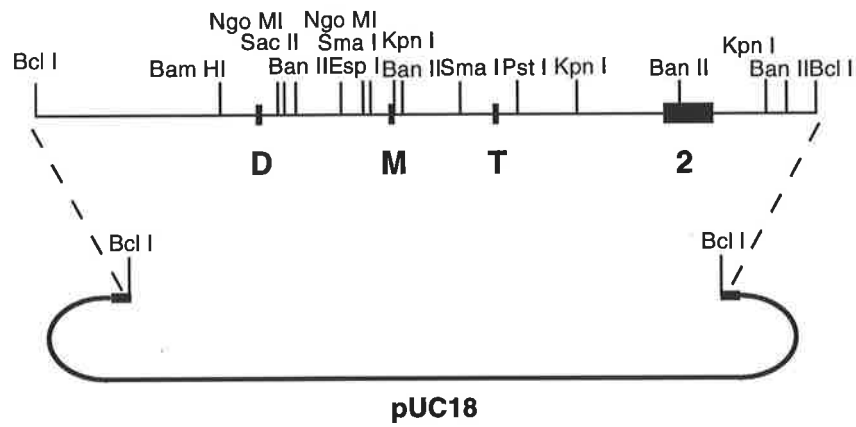
pUC18 (New England Biolabs), was digested with Hind III and Eco RI, endfilled and ligated to Bcl I linkers (CTGATCA). After recircularisation, the vector pUCBcl was linearised by Bcl I digestion and ligated to the 3 kb Bcl I fragment of mLIF containing exons 1D, 1M, 1T and 2 (Fig. 5.7 i) to form pUCBclLIF (Fig. 5.7 ii) with the first exon of LIF D

Figure 5.7 Construction of pUCBclLIF

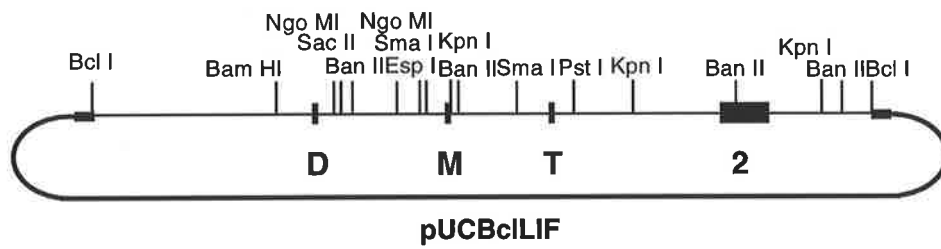
Representation of (i) the cloning of the Bcl I fragment of the genomic LIF clone from pXT7LIFt into pUCBcl and (ii) the resultant vector pUCBclLIF. The restriction sites for Bam HI, Ban II, Bcl I, Bgl II, Esp I, Kpn I, Ngo MI, Pst I, Sac II and Sma I and additional sites in the polylinker of pBLUESCRIPT-Ngo are shown. Thin line, genomic LIF DNA; thick line, plasmid vector DNA with Bcl I linkers shown; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF.

Construction of pUCBclLIF

i



ii



closest to the Eco RI site of the pUC18 polylinker. PCR reactions were carried out using pUCBclLIF as the template.

5.5.3 *Alteration of LIF exon 1D*

The Ngo MI site of pBluescript was removed by digestion with Ngo MI followed by endfilling to form the vector pBluescript-Ngo. The Bam HI/Sac II fragment of LIF (243 bp, Fig. 5.8 i) was cloned into pBluescript-Ngo to form the vector pBLUELIF. PCR product 1D (230 bp, amplified with the primers 1D5' and 1D3' indicated in figure 5.8 ii), retaining the ATG of exon 1D upstream of the Bgl II site created for in frame insertion of the reporter cassette, was digested with Bam HI and Ngo MI and cloned into pBLUELIF digested with Bam HI and Ngo MI, to form pIDBLUELIF (Fig. 5.8 iii). The LIF fragment from pIDBLUELIF, carrying the introduced Bgl II site, was removed by digestion with Bam HI and Sac II (Fig. 5.8 iv) and ligated to pUCBclLIF digested with Bam HI and Sac II to form vector p1DBclLIF.

5.5.4 *Alteration of LIF exon 1M*

The 700 and 735 bp Kpn I fragments of LIF were removed from pUCBclLIF by digestion with Kpn I (Fig. 5.9 i) and the vector recircularised to form pBclLIF-Kpnf. The PCR product 1M (116 bp, amplified with the primers 1M5' and 1M3' indicated in Fig. 5.9 ii) retaining the ATG of exon 1M upstream of the Bgl II site created for in frame insertion of the reporter cassette, was digested with Esp I and Kpn I and cloned into pBclLIF-Kpnf digested with Esp I and Kpn I to form the vector p1MBclLIF-Kpnf (Fig. 5.9 iii). The 1435 bp partial fragment of LIF, consisting of the 700 and 735 bp Kpn I fragments removed previously, was subcloned back into pBclLIF-Kpnf to form p1MBclLIF.

5.5.5 *Alteration of LIF exon 1T*

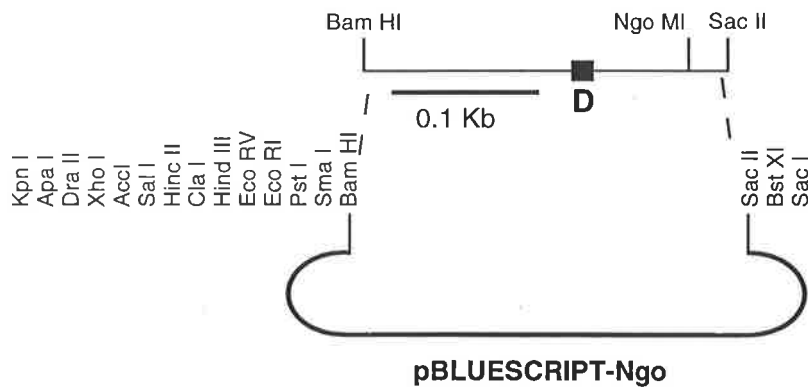
The Ban II site was removed from pUC18 by digestion with Ban II, exonuclease digestion of the overhangs and religation to form pUC18-Ban. Vector pUC-(H/S/B) was created by digestion of pUC18-Ban with Hind III and Sma I, endfilling of the Hind III end, and religation to remove the Pst I site contained within the polylinker of pUC18-Ban. The

Figure 5.8 Construction of pBLUELIF and p1DBclLIF

Representation of (i) the cloning of the Bam HI/ Sac II fragment of the genomic LIF clone from pUCBclLIF into pBLUESCRIPT-Ngo to form pBLUELIF, (ii) the PCR product 1D, (iii) cloning of the Bam HI/Ngo MI cut PCR product 1D into pBLUELIF and (iv) cloning of the Bam HI/Sac II fragment, containing PCR product 1D, into pUCBclLIF to give p1DBclLIF. The primers 1D5' and 1D3' used for amplification of 1D and the restriction sites for Bam HI, Ban II, Bcl I, Bgl II, Esp I, Kpn I, Ngo MI, Pst I, Sac II and Sma I are shown. Thin line, genomic LIF DNA; thick line, plasmid vector DNA with Bcl I linkers shown; thick stippled line, PCR product 1D; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; boxed codons, translated codons of LIF exon 1D.

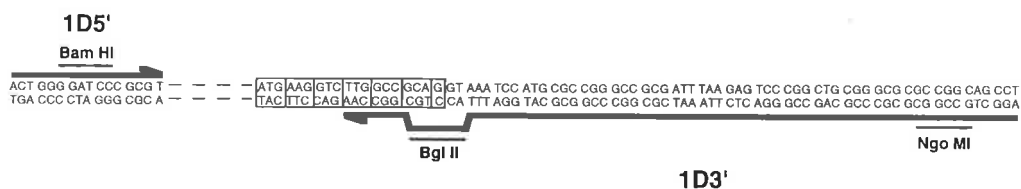
i

Construction of pBLUELIF



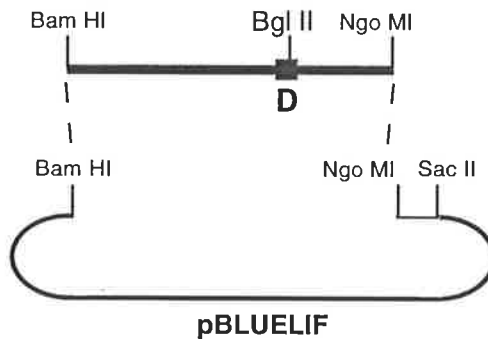
ii

Construction of p1DBLUELIF



iii

Bam HI/Ngo MI cut PCR product 1D



iv

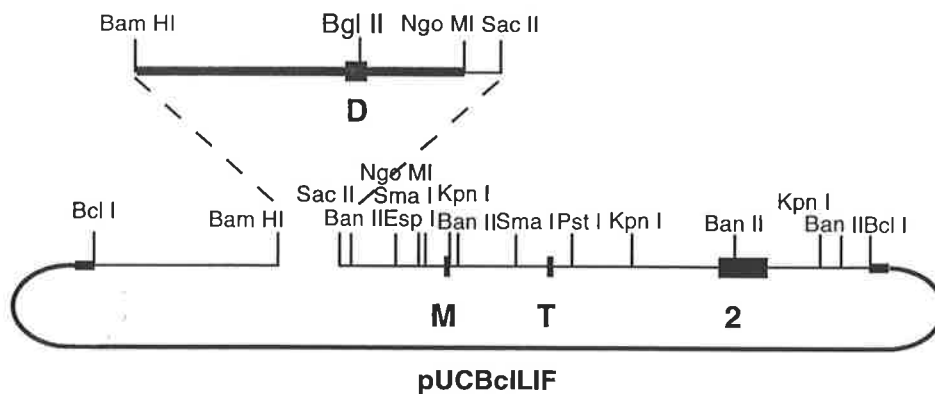
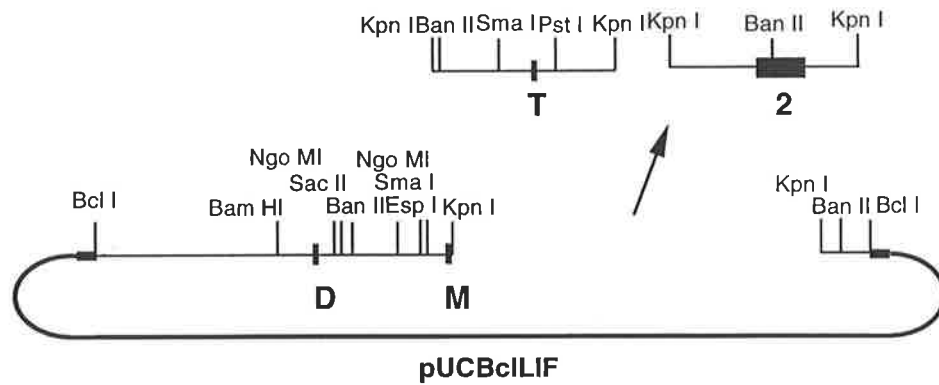


Figure 5.9 Construction of pUCBclLIF-Kpnf and p1MBclLIF

Representation of (i) the removal of the Kpn I fragments from pUCBclLIF forming pUCBclLIF-Kpnf (ii) the PCR product 1M, (iii) cloning of the Esp I/Kpn I cut PCR product 1M into pUCBclLIF-Kpnf and (iv) cloning of the Kpn I partial fragment, containing the PCR product 1M into pUCBclLIF-Kpnf forming p1MBclLIF. The primers 1M5' and 1M3' used for amplification of 1M and the restriction sites for Bam HI, Ban II, Bcl I, Bgl II, Esp I, Kpn I, Ngo MI, Pst I, Sac II and Sma I are shown. Thin line, genomic DNA; thick line, plasmid vector DNA with Bcl I linkers shown; thick stippled line, PCR product 1D; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; boxed codons, translated codons of LIF exon 1M.

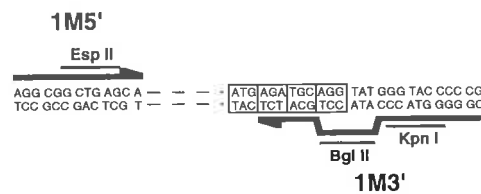
Construction of pUCBclLIF-Kpnf

i



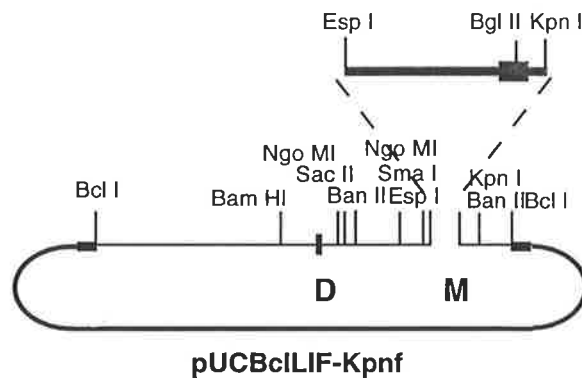
Construction of p1MBclLIF

ii

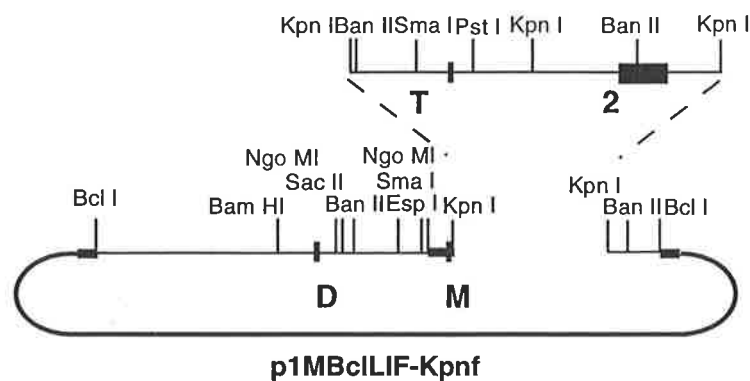


iii

Esp I/Kpn I cut PCR product 1M



iv



1435 kb Kpn I partial fragment of LIF (Fig. 5.10 i; section 5.5.4) was cloned into the Kpn I site of pUC-(H/S/B) forming the vector pKpnLIF-(H/S/B). PCR product 1T (230 bp, amplified with the primers 1T5' and 1T3' indicated in Fig. 5.10 ii), carrying the Bgl II site in exon 1T for insertion of the ATG⁺ reporter construct, was digested with Sma I and Pst I and cloned into pKpnLIF-(H/S/B) digested with Sma I and Pst I (Fig. 5.10 iii) to form p1TKpnLIF. The 1435 Kpn I partial fragment, carrying the introduced Bgl II site, was isolated from p1TKpnLIF and cloned into pUCBclLIF (Fig 5.10 iv) that had been digested with Kpn I to form p1TBclLIF.

5.5.6 *Alteration of LIF Exon 2*

The PCR product E2 (1095 bp, amplified with the primers E25' and E23' indicated in Fig. 5.11 i), carrying the Bgl II site for insertion of the ATG⁺ reporter gene in exon 2, was digested with Ban II and ligated into the Ban II site of pKpnLIF-(H/S/B) (section 5.5.5; Fig. 5.11 ii) to form pE2KpnLIF. The 1435 Kpn I partial fragment of LIF (section 5.5.4), carrying the introduced Bgl II site, was subcloned into pUCBclLIF to form pE2BclLIF (Fig. 5.11 iii).

5.5.7 *Reconstruction of the LIF sequence in pLIFXT7t*

The LIF sequences from the vectors p1DBclLIF, p1MBclLIF, p1TBclLIF, pE2BclLIF, carrying the introduced Bgl II sites, were isolated as Bcl I fragments and cloned into the Bcl I site of pLIFXT7t to form the vectors p1DLIFX, p1MLIFX, p1TLIFX and pE2LIFX (Fig. 5.12). The altered LIF exons were analysed by sequencing to confirm correct localisation of the Bgl II sites.

5.6 The Exchange Step Targeting Vectors

5.6.1 *Construction of the Exchange Step Vectors*

The (ATG⁻) *lacZ* cassette from pnlacF14 was isolated by digestion with Bgl II and cloned into the unique Bgl II site in LIF exon 1D and 1M of p1DLIFX and p1MLIFX respectively to form the exchange step targeting vectors p1DLacLIF and p1MLacLIF (Fig. 5.13 a). The (ATG⁺) *lacZ* cassette from pnlacF12 was likewise cloned into the unique Bgl II sites of LIF

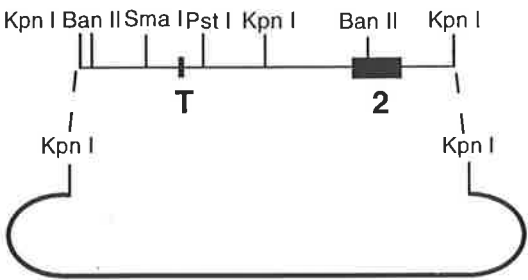
Figure 5.10 Construction of pKpnLIF-(H/S/B) and p1TBclLIF

Representation of (i) the cloning of the Kpn I partial fragment of pUCBclLIF into pUC-(H/S/B) to form pKpnLIF-(H/S/B), (ii) the PCR product 1T, (iv) cloning of the Sma I/Pst I cut PCR product 1T into pKpnLIF-(H/S/B) forming p1TKpnLIF and (v) cloning of the Kpn I partial fragment, containing PCR product 1T, into pUCBclLIF to form p1TBclLIF. The primers 1T5' and 1T3' used for amplification of 1T and the restriction sites for Bam HI, Ban II, Bcl I, Bgl II, Esp I, Kpn I, Ngo MI, Pst I, Sac II and Sma I are shown. Thin line, genomic DNA; thick line, plasmid vector DNA with Bcl I linkers shown; thick stippled line, PCR product 1T; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; boxed codons, translated codons of LIF exon 1T.

Construction of pKpnLIF-(H/S/B)

i

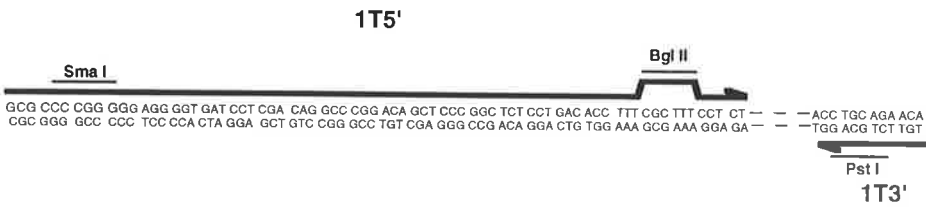
Kpn I partial fragment from pUCBclLIF



pUC-(H/S/B)

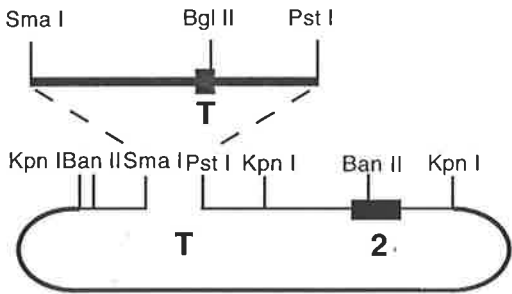
Construction of p1TBclLIF

ii



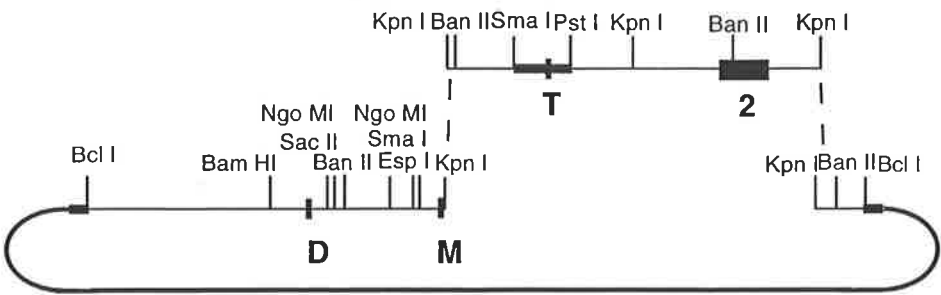
iii

Sma I/Pst I cut PCR product 1T



pKpnLIF-(H/S/B)

iv



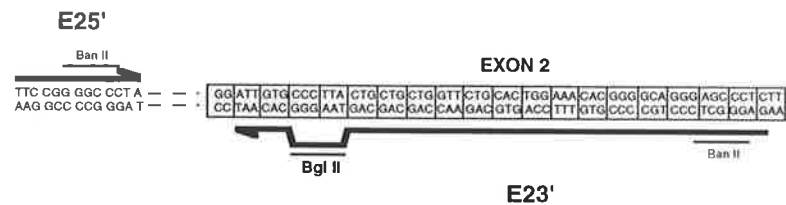
pUCBclLIF

Figure 5.11 Construction of pE2BclLIF

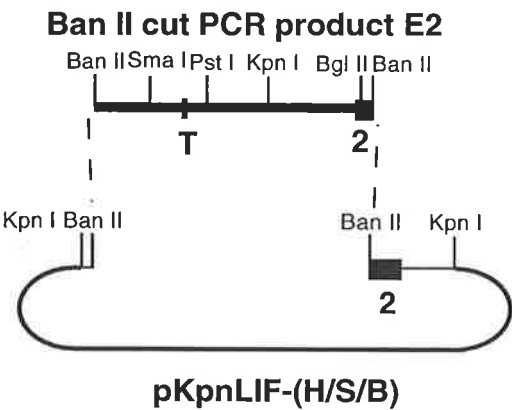
Representation of (i) the PCR product E2, (ii) cloning of the Ban II cut PCR product E2 into pKpnLIF-(H/B/S) and (iii) cloning of the Kpn I partial fragment, containing PCR product E2, into pUCBclLIF to form pE2BclLIF. The primers E25' and E23' used for amplification of E2 and the restriction sites for Bam HI, Ban II, Bcl I, Bgl II, Esp I, Kpn I, Ngo MI, Pst I, Sac II and Sma I are shown. Thin line, genomic DNA; thick line, plasmid vector DNA with Bcl I linkers shown; thick stippled line, PCR product E2; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; boxed codons, translated codons of LIF exon E2.

Construction of pE2BclLIF

i



ii



iii

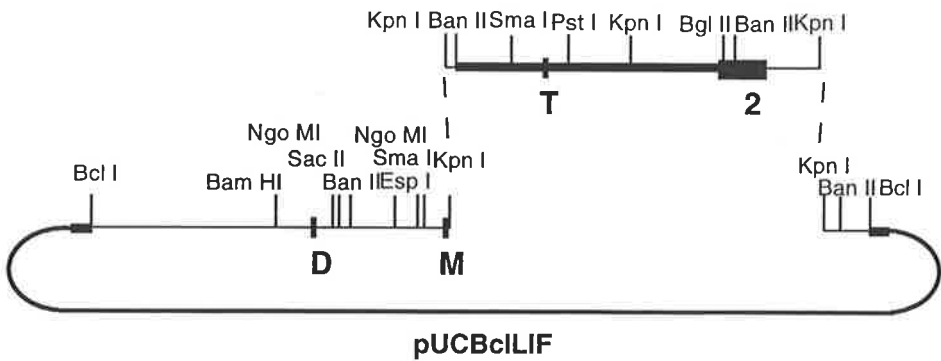


Figure 5.12 Reconstruction of LIF sequence in pLIFXT7t

Representation of the cloning of the altered LIF sequences from vectors p1DBclLIF, p1MBclLIF, p1TBclLIF and pE2BclLIF, isolated as Bcl I fragments into the vector pLIFXT7t linearised by Bcl I digestion. The restriction sites for Bcl I and Bgl II are shown. Thin line, genomic LIF DNA; thick line, plasmid vector DNA; thick stippled lines, PCR products; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR.

Reconstruction of the LIF Sequence in pLIFXT7t

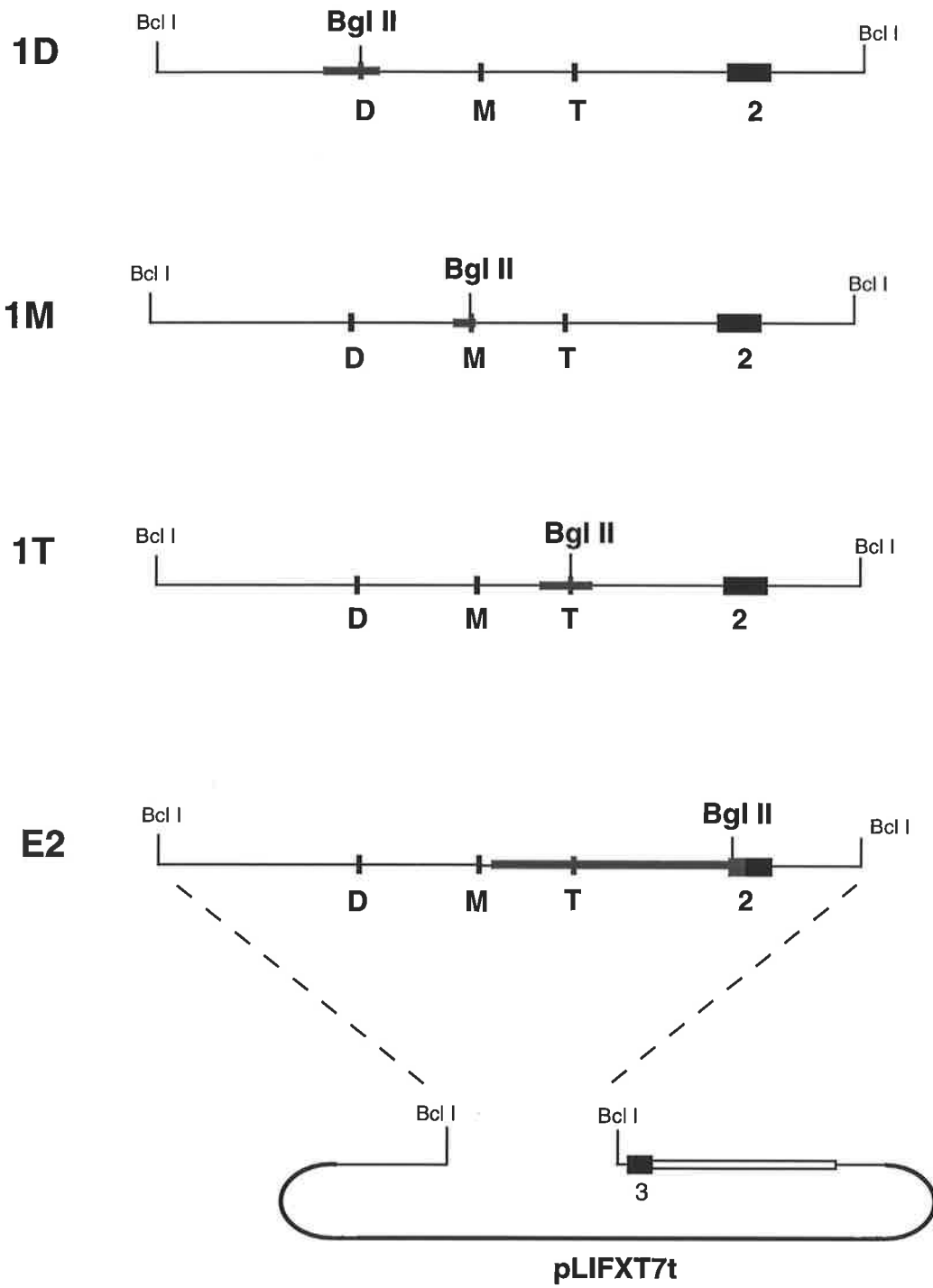
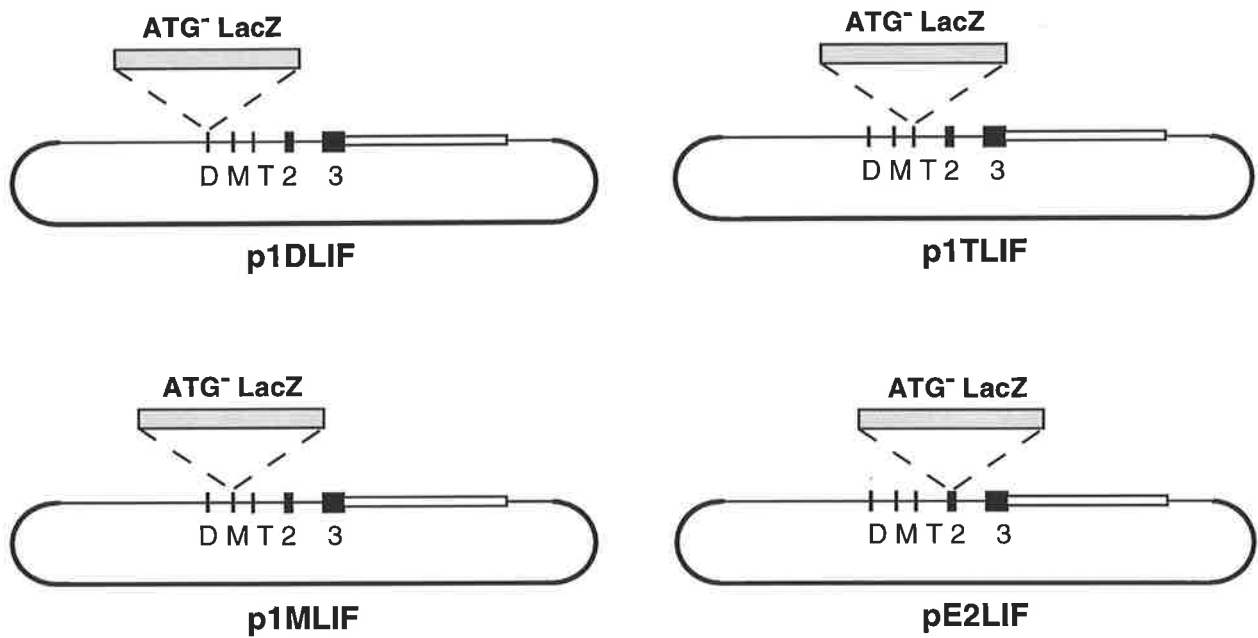


Figure 5.13 Construction and analysis of the exchange step targeting vectors

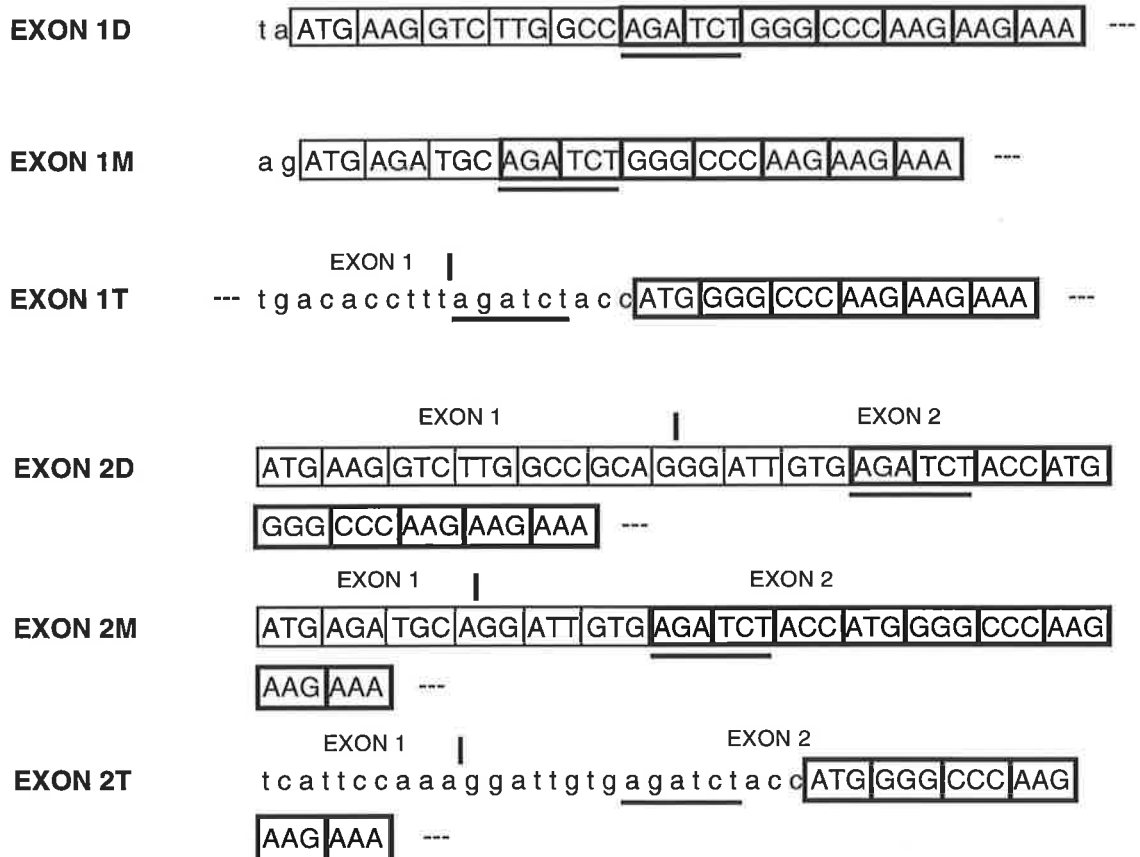
a Representation of the construction of the exchange step targeting vectors showing the cloning of the Bgl II ATG⁻ *lacZ* cassette from pnlacF14 into the Bgl II sites of vectors p1DLIF and p1MLIF (left) and the Bgl II ATG⁺*lacZ* cassette from pnlacF12 into the Bgl II sites of p1TLIF and pE2LIF (right).

b Sequence of the LIF-*lacZ* fusions in lines carrying the *lacZ* reporter gene in exon one of LIF D, M or T or in exon 2.

a Construction of the Exchange Step Targeting Vectors



b



exons 1T and 2, forming the vectors p1TLacLIF and pE2LacLIF respectively (Fig. 5.13 a). Sequence analysis of each clone was carried out to confirm that the *lacZ* cassettes were cloned in frame (Fig. 5.13 b). A restriction map of p1DLacLIF and example of restriction enzyme analysis of each exchange step targeting vector are shown in figure 5.14.

5.7 Investigation of the Conditions Required for 6-TG Selection

Selection for loss of the HPRT tag gene can be achieved with the purine analogue 6-TG, which is toxic to cells expressing HPRT activity. The dose and time required for ES cell selection by 6-TG was investigated.

5.7.1 Concentration of 6-TG Required for Cell Toxicity

To determine the minimum dose of 6-TG required for effective selection against HPRT⁺ cells, the following experiment was carried out. ES cell lines E14TG2a (HPRT⁻, section 3.3.2) or HG60 (tagged line carrying the HPRT cassette, section 4.3.3) were seeded into 6 well plates at a density of 1×10^5 cells per well in complete ES cell medium. Medium containing 6-TG at concentrations of 0, 10, 12, 15, 18, 20, 25, 30, 40 μ M was added twenty four hours later and replaced every forty eight hours. Even the highest concentration of 6-TG did not kill the HPRT deficient ES cell line E14TG2a, while 20 μ M was sufficient to completely kill the HG60 ES cell line after seven days. Subsequent selection with 6-TG was therefore carried out using 6-TG at a final concentration of 20 μ M.

5.8 Exchange Step Gene Targeting

5.8.1 6-TG Selection and Metabolic Co-operation

Metabolic co-operation is the process whereby low molecular weight molecules and ions can move intercellularly via cell-cell junctions known as gap junctions (Subak-Sharpe *et al.*, 1966, 1969). During exchange step gene targeting, direct selection for replacement of the HPRT tag cassette by the exchange step vector is achieved by selection with 6-TG. Transfer of toxic nucleotide analogues, to cells which have lost HPRT expression, from surrounding HPRT⁺ cells with which they are in physical contact (Hooper and Slack, 1977), can lead to death of these cells. This would lead to death of potentially targeted lines in a gene targeting

Figure 5.14 Analysis of exchange step targeting vectors

b Representation of the exchange step targeting vector p1DLacLIF. The restriction sites for Bgl II, Eco RI, Nhe I and Sac II are shown.; Black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIFopen box, 3' UTR, light stippled box, *lacZ* cassette.

c Restriction enzyme analysis of the vectors p1DLacLIF, p1MLacLIF, p1TLacLIF, pE2LacLIF. Plasmid DNA (300 ng) was digested (section 9.3.1) and electrophoresed on 1% TAE agarose gels (section 9.3.5). D, p1DLacLIF; M, p1MLacLIF; T, p1TLacLIF; 2, pE2LacLIF; S, Eco RI digested SPP-1 size markers: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36 kb.

Expected sizes from the Nhe I/Eco RI digests are as follows:

p1DLacLIF: 5.7, 4.0, 3.1, 0.7, 0.6, 0.3 kb

p1MLacLIF: 6.2, 3.5, 3.1, 0.7, 0.6, 0.3 kb

p1TLacLIF: 6.5, 3.2, 3.1, 0.7, 0.6, 0.3 kb

pE2LacLIF: 7.2, 3.1, 2.5, 0.7, 0.6, 0.3 kb.

Expected sizes from the Nhe I/Sac I digests are as follows

p1DLacLIF: 6.4, 4.0, 3.4, 0.7 kb

p1MLacLIF, p1TLacLIF, pE2LacLIF: 7.6, 3.4, 2.8, 0.7 kb.

Expected sizes from the Bgl II/Nhe I digests are as follows

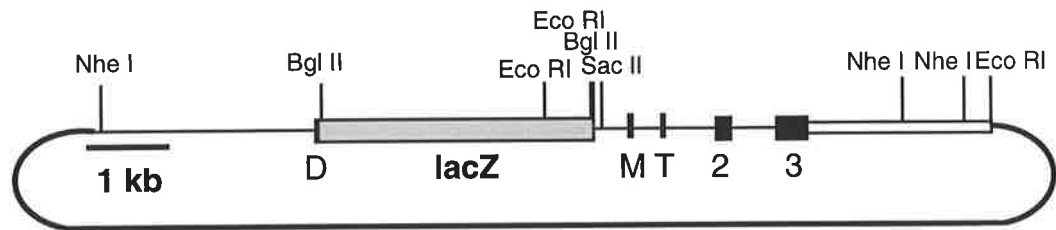
p1DLacLIF: 4.0, 3.6, 3.4, 2.7, 0.7 kb

p1MLacLIF, p1TLacLIF: 3.6, 3.5, 3.4, 3.2, 0.7 kb

pE2LacLIF: 4.4, 3.6, 3.4, 2.5 0.7 kb.

a

p1DLacLIF



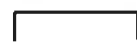
b

Nhe I/Eco RI Nhe I/Sac II

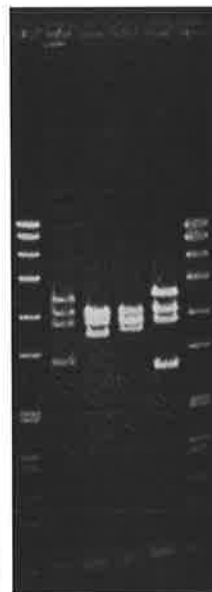


S D M T 2 S S D M T 2 S

Bgl II/Nhe I



S D M T 2 S



experiment. To prevent this, cells must be grown up to and throughout 6-TG selection without physically contacting surrounding colonies. This was achieved by reducing the seeding density used for exchange step gene targeting.

5.8.2 *Transfection of Tag Targeted ES cell line HG60 with the Exchange Step Targeting Vectors*

The germline competent (section 7.3.2), tag step gene targeted ES cell line HG60 was transfected with Nhe I linearised exchange step targeting vectors p1DLacLIF, p1MLacLIF, p1TLacLIF, pE2LacLIF or Bam HI linearised pCH110 as a control. The targeting vectors contained 6.7 kb of homologous DNA and the *lacZ* reporter gene located in the first unique exons of LIF D, M, T or in the common second exon of LIF respectively. Digestion with Nhe I linearises the vectors within the genomic LIF sequence at both the 5' and 3' ends. The electroporation conditions used were the same as for the tag step gene targeting (section 4.3.2).

Cells were seeded into 10 cm diameter plates at a density of 2.5×10^6 cells per plate or 2.5×10^7 cells per 175 cm² flask. Cells in plates were split on day five or six after electroporation and selection in either 6-TG (selection for loss of the tag gene) or 6-TG plus G418 (selection for loss of the tag gene and maintenance of the stability cassette) initiated on day six. Cells seeded into the flasks were split into 10 cm diameter plates on day five and selection in either 6-TG or 6-TG plus G418 initiated on day six. Resistant colonies arising in cultures transfected with the targeting vectors and pCH110 were picked from day 10 of selection. Cells were expanded in the absence of selection and genomic DNA was analysed by Southern blot. Cell lines produced in this experiment were non-independent clones since the transfected cells were passaged before clones were picked.

DNA was digested with Hind III and Eco RV and filters were analysed using a range of probes to detect correctly targeted lines. Figure 5.15 a shows the expected sizes of the wildtype and altered LIF alleles and the probes used for analysis. Four probes were used in this analysis to identify correctly targeted lines.

Probe 1 The 5' external LIF probe was used to detect exchange step targeted lines (also used for the detection of tag step targeted lines, section 4.3.3). The probe hybridises to a fragment of 13.2 kb from the untargeted LIF allele (Fig. 5.15 b i, all lanes), a fragment of 11.8 kb from the tagged LIF allele (Fig. 5.15 b i, lanes 1, 2, 22) and a fragment of 7.6, 8.1 (Fig. 5.15 b i, lanes 3-12), 8.3 or 9.1 kb (Fig. 5.16 b i, lanes 14 - 18) from the exchange step targeting vectors 1DLacLIF, p1MLacLIF, p1TLacLIF or pE2LacLIF respectively.

Probe 2 The *lacZ* probe was used to detect lines carrying the exchange step targeting vectors at non-homologous sites within the genome. The probe hybridises to fragments of 9.2, 8.7 (Fig. 5.15 b, ii lanes 3-12), 8.5 or 7.6 kb (Fig. 5.15 b, ii lanes 14 - 18) from correctly targeted lines carrying 1DLacLIF, p1MLacLIF, p1TLacLIF or pE2LacLIF respectively and fragments of other sizes if integrated at non-homologous sites, as the 3' Eco RV site is external to the exchange step vector.

Probe 3 The 3' external *neo* probe was used to detect the *neo* cassette in tag or exchange step targeted lines. This probe hybridises to the same fragments as detected by probe 2 in exchange step targeted lines or a fragment of 7.4 kb in tag step targeted lines (results not shown).

Probe 4 The HPRT probe was used to detect loss of the HPRT tag cassette. This probe hybridises to a fragment of 11.8 kb if the tag gene is present (results not shown).

The results of the targeting experiment are summarised in Table 5.3. Over ninety percent of the cell lines selected in 6-TG plus G418 were targeted correctly. In contrast to this, only five percent of the cell lines selected in 6-TG alone were correctly targeted. This finding demonstrates that dual selection can overcome a high background of HPRT⁻ cells which are G418^S/6-TG^r.

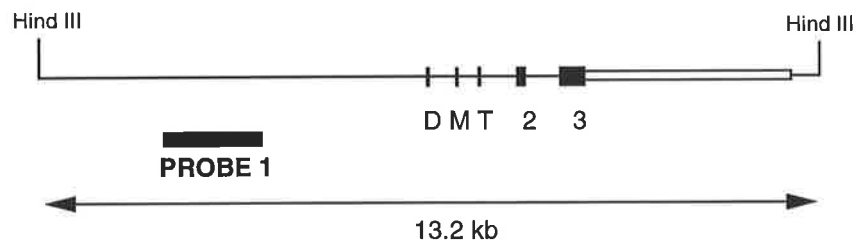
Figure 5.15 a Southern analysis of exchange step gene targeting

Representation of the expected sizes of restriction fragments resulting from (i) the untargeted LIF allele, (ii) the tag step targeted LIF allele and (iii) the exchange step targeted LIF allele after digestion with Eco RV/Hind III showing the probes used for Southern analysis. Probe 1, 5' external probe; probe 2, *lacZ* probe; probe 3, *neo* probe; probe 4, HPRT probe; thin line and thick stippled line, genomic DNA; black boxes, exons: D, exon 1 of LIF D; M, exon 1 of LIF M; T, exon 1 of LIF T; 2, exon 2 of LIF and 3, exon 3 of LIF; open box, 3' UTR, stippled box, HPRT cassette; hatched box, *neo* cassette; light stippled box, *lacZ* gene.

a Southern Analysis of Exchange Step Targeting

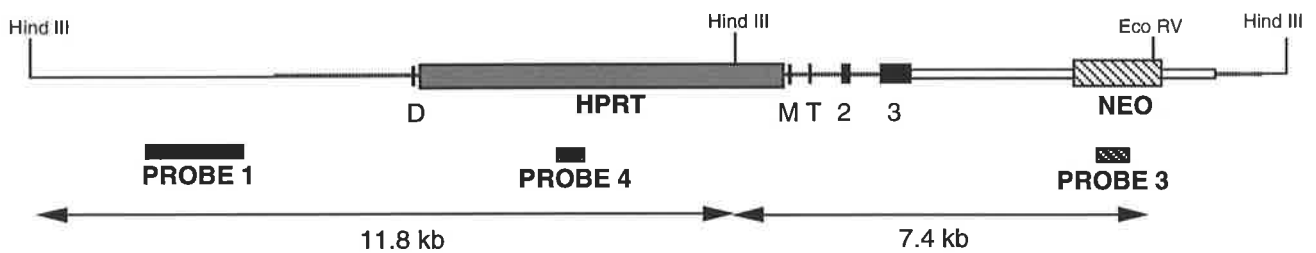
i

Untargeted LIF Allele



ii

Tag Step Targeted LIF Allele



iii

Exchange Step Targeted LIF Allele

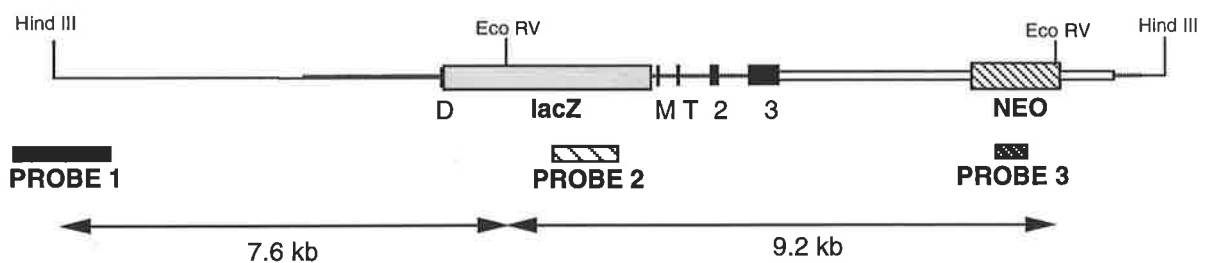


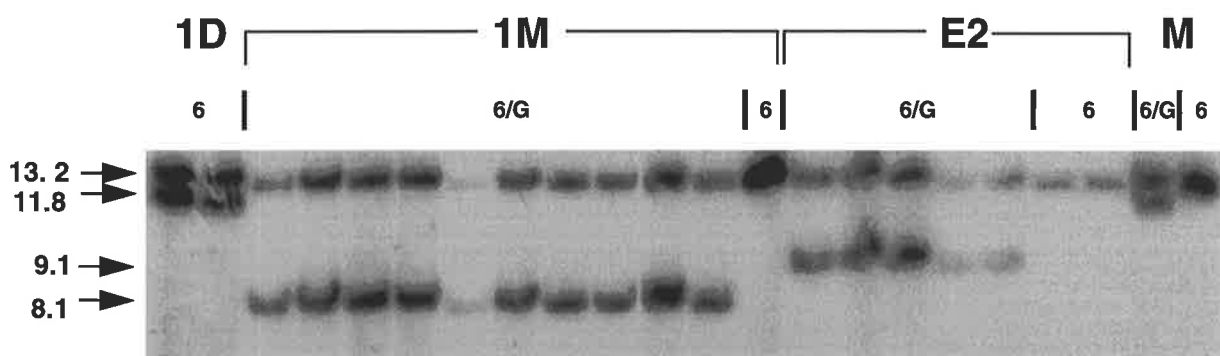
Figure 5.15 b Southern analysis of exchange step gene targeting

Approximately 20 µg of genomic DNA from 6-TG^r and 6TG^r/G418^r clones was digested with Eco RV/Hind III, electrophoresed on a 0.8 % TAE agarose gel and blotted (section 9.3.21). The Southern was probed with (i) the 5' external Bgl II LIF probe of 1.7 kb (probe 1) or (ii) a Bam HI/Sac I *lacZ* probe of 1.1 kb (probe 2) and washed in 0.1 x SSC/0.1 % SDS at 65° for 30 minutes (section 9.3.11; 9.3.22). The sizes of the fragments to which the LIF probes hybridises are 13.2 kb from the untargeted allele, 11.8 kb from the tag step targeted allele and 8.1 and 9.1 kb from the exchange step targeted lines carrying p1MLacLIF and pE2LacLIF respectively. The fragments to which the *lacZ* probes hybridises are 8.7 and 7.6 kb in targeted lines carrying p1MLacLIF and pE2LacLIF, respectively. 1D, lines transfected with p1DLacLIF; 1M, lines transfected with p1MLacLIF; E2, lines transfected with pE2LacLIF; M, lines transfected with pCH110; 6, lines selected in 6-TG; 6/G, lines selected in 6-TG/G418.

b

i

5'External Probe (1)



ii

***lacZ* Probe (2)**

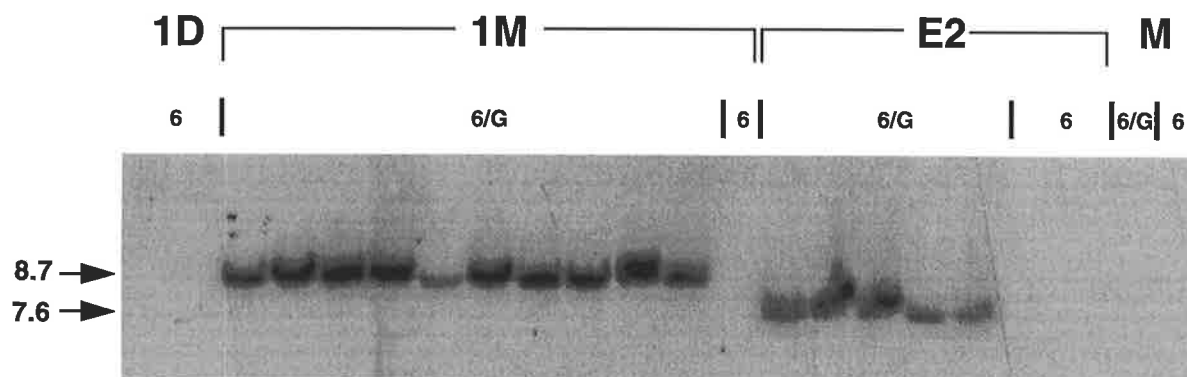


Table 5.3

Exchange step gene targeting results obtained the vectors p1DLacLIF, p1MLacLIF, p1TLacLIF and pE2LacLIF (as described in sections, 5.8.2; 9.4.12). Screening was carried out by Southern blot analysis (9.3.11; 9.3.21-22).

Table 5.3

	Targeted/Screened 6-TG^r	Targeted/Screened 6-TG^r/G418^r	Tag Gene Loss /Screened 6-TG^r	Tag Gene Loss /Screened 6-TG^r/G418^r
IDLacLIF	0/10	1/1	0/10	0/1
IMLacLIF	1/4	19/20	3/4	0/20
ITLacLIF	0/2	1/3	2/2	0/3
E2LacLIF	0/4	10/10	4/4	0/10
TOTAL	1/20 (5%)	31/34 (91%)	9/20 (45%)	0/34 (0%)

Forty five percent of the 6-TG^r lines had become 6-TG resistant due to physical loss of the tag gene by means other than homologous recombination while the remaining lines carried the tag gene but had presumably lost expression of it as detected by Southern analysis. By contrast, none of the 6-TG^r/G418^r lines had physically lost the tag gene by means other than homologous recombination. The background of 6-TG^r/G418^r non-targeted lines of only nine percent resulted from loss of tag gene expression. Thus non-targeted lines which had undergone physical loss of the tag gene by means other than homologous recombination were eliminated by selection for the stability cassette during the exchange step gene targeting.

Resistant lines appearing in cultures transfected with the non-homologous plasmid pCH110 reflected the background of non-targeted lines appearing in transfections with the exchange step targeting vectors. Of the lines transfected with pCH110 that were screened, the 6-TG^r line had physically lost the tag and stability cassettes by means other than homologous recombination while the two 6-TG^r/G418^r lines carried both genes but had presumably lost expression of the tag gene.

Because the cells in this experiment were passaged after transfection, the clones obtained did not represent a single targeting or tag gene loss event. The efficiency of the exchange step gene targeting and the contribution of physical loss of the tag gene by means other than homologous recombination to the background of 6-TG^r lines was therefore investigated in a further round of gene targeting in which cells were grown throughout exchange step targeting without passaging.

5.8.3 *Cell Density Required to Prevent Metabolic Co-operation*

In order to gain accurate information about the efficiency of the stable tag-exchange approach, it was necessary to grow the exchange step targeted cells without passaging up to and throughout the selective process. This would ensure that each resistant colony resulted from a single targeting or tag gene loss event. To determine the highest seeding density that prevented physical contact between colonies grown for seven days without passaging (section 5.8.1), cells were plated at densities of 2×10^6 , 1.5×10^6 , 1×10^6 , 5.7×10^5 , 2.5

$\times 10^5$ and 1.5×10^5 cells/150 cm², after electroporation in the absence of DNA. Cells were grown for 7 days without passaging, fixed, stained and viewed. The highest density at which single colonies predominated was 2.5×10^5 cells/150 cm² (results not shown). This density was used in subsequent targeting experiments.

5.8.4 *Investigation of the Efficiency of Exchange Step Gene Targeting*

The tag step gene targeted ES cell line HG60 (section 4.3.3) was transfected with Nhe I linearised p1DLacLIF or Bam HI linearised pCH110 as a control. One tenth of the amount of cells transfected with pCH110 were plated out compared to cells transfected with p1DLacLIF. Cells transfected with p1DLacLIF were treated in one of two ways. To ensure that each resistant colony resulting after selection in 6-TG or 6-TG/G418 represented a single targeting or tag gene loss event, half of the transfected cells were grown for five days (time required for HPRT activity to decay) without passaging prior to the addition of selection medium. In parallel, half of the transfected cells were seeded into flasks, passaged on day three and a fraction seeded into selective medium on day six. This was carried out to ensure that germline competent targeted lines were obtained.

Resistant colonies arising in cultures transfected with pCH110 or p1DLacLIF were picked from day 10 of selection, expanded and analysed as in section 5.8.2. The results shown in Table 5.4 and discussed below, are those of the first approach, in which each targeted colony represents a single targeting event. An example of the Southern blot analysis is shown in figure 5.17.

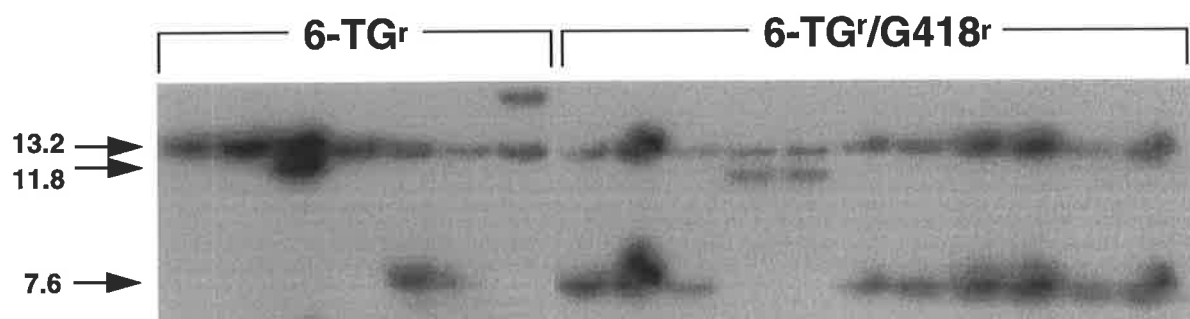
Selection for fourteen days resulted in the appearance of one hundred and three 6-TG^r colonies and twenty three 6-TG^r/G418^r colonies from the transfection with p1DLacLIF (Table 5.4). These values are significantly different (χ^2 test, at a level of significance of .0005). This suggested that dual selection in 6-TG and G418 overcame a background of 6-TG^r lines which had lost tag gene expression by means other than homologous recombination. Genomic DNA from resistant lines was digested with Hind III and Eco RV and probed as in section 5.8.2. Southern blot analysis showed that the majority of 6-TG^r

Figure 5.16 Southern analysis of exchange step gene targeting

Approximately 20 µg of genomic DNA from 6-TG^r and 6TG^r/G418^r clones transfected with p1DLacLIF was digested with Eco RV/Hind III, electrophoresed on a 0.8 % TAE agarose gel and blotted (section 9.3.21). The Southern was probed with (i) the 5' external Bgl II LIF probe of 1.7 kb (probe 1) or (ii) a Bam HI/Sac I *lacZ* probe of 1.1 kb (probe 2) and washed in 0.1 x SSC/0.1 % SDS at 65° for 30 minutes (sections 9.3.11; 9.3.22). The sizes of the fragments to which the LIF probe hybridises are 13.2 kb from the untargeted allele, 11.8 kb from the tag step targeted allele and 7.6 kb from the exchange step targeted lines carrying p1DLacLIF. The fragment to which the *lacZ* probes hybridises is 9.2 kb.

i

5' External LIF Probe (1)



ii

***lacZ* Probe (2)**



Table 5.4

Exchange step gene targeting results carried out with the vector p1DLacLIF (as described in sections 5.8.4; 9.4.12). Screening was carried out by Southern blot analysis (9.3.11; 9.3.21-22).

Table 5.4

	6-TG^r	6-TG^r/G418^r	Total
Cells Plated	2.5 x 10 ⁶	2.5 x 10 ⁶	5 x 10 ⁶
Cells Surviving	3.8 x 10 ⁵	3.8 x 10 ⁵	7.5 x 10 ⁵
Cells Transfected	1.25 x 10 ³	1.25 x 10 ³	2.5 x 10 ³
Resistant Colonies	103	23	126
Screened	28	16	44
Correctly Targeted	6	14	20
Targeting Efficiency	21%	88%	46%
Tag Gene Loss	17/28 (61%)	0/16 (0 %)	17/44 (39 %)

lines had arisen from physical loss of the tag gene in the absence of homologous recombination.

Twenty eight 6-TG^r and sixteen 6-TG^r/G418^r lines from transfection with p1DLacLIF were screened by Southern blot. Of these, six 6-TG^r (21%) and fourteen 6-TG^r/G418^r (88%) lines were correctly targeted (Table 5.4). This demonstrated that dual selection in 6-TG and G418 enriched for targeted lines by eliminating the high background of non-exchange step targeted 6-TG^r lines. Seventeen out of twenty eight 6-TG^r lines (61%) had physically lost the tag and stability genes by means other than homologous recombination, while the remaining lines (18 %) carried both genes but had presumably lost expression of the tag gene and possibly the stability gene. By comparison, none of the 6-TG^r/G418^r lines screened had physically lost the tag gene by means other than homologous recombination and the remainder (12 %) had presumably lost expression of the tag, but not the stability gene (Table 5.4, Fig. 5.16). Thus, selection for expression of the stability cassette used in this targeting approach eliminated one hundred percent of the 6-TG^r lines that resulted from physical loss of the tag gene by means other than homologous recombination.

Four 6-TG^r and no 6-TG^r/G418^r lines were obtained from one tenth the amount of cells transfected with pCH110. If one assumes that of the 6-TG^r lines transfected with p1DLacLIF, eighty represent non-targeted lines (by comparison with the number of 6-TG^r/G418^r lines), the number of resistant lines expected in the pCH110 transfection would be eight. The number appearing in the pCH110 transfection was less than expected but the numbers are too small to be statistically significant. Screening of one of the 6-TG^r lines transfected with pCH110 indicated that it had arisen from physical loss of the tag gene. The 6-TG^r lines transfected with pCH110 therefore reflected the background of non-targeted 6-TG^r lines appearing in the transfection with p1DLacLIF.

Selection for the stability cassette may have an effect on the number of untargeted lines resulting from loss of tag gene expression. If a mechanism such as methylation is responsible for gene inactivation, this may be essentially a “gross” event which inactivates gene

expression in a large region of a chromosome. Thus lines, which retain expression of the stability cassette, and therefore survive 6-TG/G418 selection, would be unlikely to arise from tag gene inactivation by methylation. If this was the case, one would expect a decrease in the number of lines carrying, but not expressing, the tag gene under dual selection compared to the number of such lines obtained under single selection. Southern analysis indicated that four 6-TG^r lines carried the tag gene but had lost expression of it compared to two 6-TG^r/G418^r lines. These numbers are too small to be statistically significant but suggest that dual selection may reduce the background of untargeted lines resulting from loss of tag gene expression.

The possibility that electroporation of ES cells had an effect on physical loss of the tag gene by means other than homologous recombination was investigated in a further experiment in which ES cells transfected with Bam HI linearised pCH110 were compared to untransfected ES cells. Cells were subjected to 6-TG selection as used in the previous targeting experiments. From the equivalent of 4.5×10^6 cells transfected with pCH110 or untransfected, thirty six and eighteen 6-TG^r lines arose, respectively. These values are significantly different (χ^2 test at a level of significance of 0.01) suggesting that the process of transfection may affect tag gene loss by means other than homologous recombination.

The exchange step gene targeting was repeated to produce sufficient targeted ES lines to ensure that germline competent lines were obtained (results not shown). Exchange step targeted lines carrying a marked copy of LIF D, M, T and of all forms of LIF produced in this chapter were returned to the mouse to assess germline competence (section 7.4.1).

5.9 Discussion

This chapter describes the construction of exchange step targeting vectors designed to introduce reporter genes into the endogenous LIF locus such that expression of each and all isoforms of LIF could be visualised at the cellular level in the mouse. The vectors were used to investigate the efficiency of stable tag-exchange gene targeting at the LIF locus using the

tagged, germline competent (section 7.3.2) ES cell line (HG60) described in chapter 4 (section 4.3.2).

A β -galactosidase expression cassette which expresses nuclear localised β -galactosidase activity was used as the reporter gene in the exchange step targeting vectors. Cells expressing reporter gene activity will produce nuclear β -galactosidase activity, allowing these cells to be easily distinguished from cells expressing cytoplasmic endogenous β -galactosidase activity and thus identified in the complex multicellular environment of the mouse. Expression of the *lacZ* cassettes described in this chapter was analysed in transiently transfected COS-1, ES and X cells. Efficient expression of ATG⁺ β -galactosidase cassette was seen in both ES cells and differentiated cell types. β -galactosidase activity was localised to the nucleus and reflected promoter activity. Efficient, controllable expression in pluripotent and differentiated cell types suggests that the reporter constructs modified in this project will be appropriate for use as reporter genes in the mouse.

The HPRT gene was used as the positive/negative tag gene in this project, since both selection for or against this gene can be applied to ES cells. In other published tag-and-exchange systems two selection cassettes, such as *neo* and *tk*, have been used as the positive/negative tag genes. The effectiveness of such systems can be compromised by poor expression levels of one or both of the selection genes. Askew and co-workers (1993) used independent *tk* promoters to direct expression of the *neo* and *tk* genes. Use of two copies of a promoter on a tag step vector may lead to intraplasmid rearrangements or may reduce the expression levels directed by the promoters (Mountford and Smith, 1995). This would have the effect of limiting the number of resistant lines obtained during the tag step gene targeting, thereby reducing the efficiency of the tag step targeting. Wu and co-workers (1994) used the stronger *Pgk-1* promoter to direct expression of both the *neo* and *tk* genes on a bicistronic mRNA containing an internal ribosome entry site (IRES). However, reduced levels of IRES-mediated translation have been reported (Kaufman *et al.*, 1991; Davies and Kaufman, 1992) and this may lead to false positives surviving in the exchange step.

A virtue of the tag-and-exchange approach is the ability to generate null mutations in the tag step gene targeting step for the production of knock-out mice. However using the *tk* gene as the tag cassette precludes this possibility, since expression of HSV-*tk* in postmeiotic germ cells disrupts sperm maturation and leads to male sterility (Braun *et al.*, 1990). It follows that germline competence of cells tagged with the HSV-*tk* gene can not be assessed until after the exchange round of gene targeting.

The fact that positive and negative selection for the HPRT gene can be applied to ES cells, the availability of feeder-independent HPRT⁻ ES cell lines and the ability to generate mice expressing the HPRT tag gene made the HPRT selection system an ideal choice for use in the stable tag-exchange approach. The possibility that a negative selection rather than a positive/negative selection cassette may be used as a tag cassette (section 4.5) requires further investigation, and if true may simplify the stable tag-exchange targeting approach.

The efficiency of the stable tag-exchange approach was assessed in two independent experiments at the LIF locus in which cells were passaged during targeting (non-independent lines produced) or were grown without passaging (independent lines). Using the selection protocol of the unmodified tag-and-exchange system (selection in 6-TG) and that of the stable tag-exchange approach (selection in 6-TG/G418) it was possible to directly compare the efficiencies of these two approaches at the LIF locus. Targeting efficiencies of only five percent (non-independent lines) and twenty one percent (independent lines) were obtained when the selection protocol of the tag-and-exchange approach was used. In contrast to this, when the selection protocol of the stable tag-exchange approach was used, targeting efficiencies of ninety one percent (non-independent lines) and eighty eight percent (independent lines) were obtained. Thus the stable tag-exchange approach greatly improved the efficiency of the tag-and-exchange targeting approach.

In these experiments, forty five (non-independent) and sixty one percent (independent) of 6-TG^r lines resulted from physical loss of the tag gene by means other than homologous recombination. In comparison, no such lines were seen when selection for the stability

cassette was also applied. Thus the selection protocol of the stable tag-exchange approach is able to eliminate the majority of non-exchange step targeted lines arising during exchange step gene targeting by elimination of lines which have undergone tag gene loss by means other than homologous recombination. This feature of the system vastly improves the previously described tag-and-exchange gene targeting system. Results presented in this chapter suggest that the process of transfection itself may have some effect on the process of physical loss of the tag gene by means other than homologous recombination. Preliminary results suggest that physical loss of the tag gene by means other than homologous recombination occurred more frequently in lines that had undergone transfection.

The possibility that dual selection may reduce the amount of non-exchange step targeted lines arising as a result of loss of tag gene expression in lines which carry the tag gene requires further investigation. Larger numbers of 6-TG^r and 6-TG^r/G418^r lines that carry the tag gene but have lost expression of it must be screened. If dual selection is able to reduce the amount of lines that result from promoter shut down it would provide a means to improve the efficiency gene targeting systems which rely on loss of a selection marker.

Examples of tag-and-exchange-based targeting systems have been published since the commencement of this project and are described in chapter 3 (section 3.2.3). Exchange step targeting efficiencies reported elsewhere have been as low as eight (Stacey *et al.*, 1994), ten (Wu *et al.*, 1994), thirteen (Askew *et al.*, 1993) and seventeen percent (Gondo *et al.*, 1994). In these studies, the high background of resistant lines surviving in the exchange step can be attributed to either loss of tag gene expression or physical loss of the tag gene in the absence of homologous recombination. In a system using the *neo* (positive selection cassette) and *tk* (negative selection cassette) selection cassettes as the tag genes, Askew and co-workers (1993) found that less than one percent of gancyclovir resistant (*tk*⁻) lines had undergone exchange step targeting at the $\alpha 2$ subunit of the ATP-dependent sodium potassium pump (Na, K-ATPase) locus. By selecting for lines that were also G418^s (*neo*⁻) a background of ninety one percent could be eliminated. Even with this enrichment only nine percent of the resulting lines had undergone homologous recombination in the exchange step gene targeting.

Up to ninety one percent of the false positives had arisen from loss of tag gene (*tk*) expression while eight percent of these, or ninety one percent of the *ganc^r/G418^s* lines, had resulted from physical loss of the tag gene.

Using a similar selection system, Wu and co-workers (1994) saw a high background of non-homologous recombinants surviving the exchange step. Sixty percent of the *FIAU^r (tk⁻)* lines were also *G418^s*, and of the *FIAU^r/G418^s* lines just under ten percent carried the alteration introduced in exchange step targeting the remainder having physically lost the tag gene by means other than homologous recombination. In a system using the *HPRT* gene as the positive/negative selection cassette Stacey and co-workers (1994) saw an exchange step gene targeting efficiency of only ten percent. Ninety percent of the lines surviving in the exchange step had physically lost the tag gene by means other than homologous recombination.

From these examples it can be seen the major obstacles to efficient exchange step gene targeting are loss of tag gene expression (due to gene inactivation or promoter shut down) or physical loss of the tag gene by means other than homologous recombination. The first of these can not be readily overcome although the choice of positive/negative selection gene(s) may have some effect. These published examples indicate that the background of non-targeted lines surviving in the exchange step is reduced when the *HPRT* gene is used as the tag gene in comparison to the *HSV-tk* gene (Bronson and Smithies, 1994). The chromosomal site into which the tag gene is introduced may also affect this phenomenon. At the *LIF* locus, loss of *HPRT* gene expression accounted for up to fifty percent of the background in non-independent lines when selection in 6-TG was applied. Physical loss of the tag gene resulted in a background of non-homologous recombinants of over sixty percent (of independent lines) at the *LIF* locus and was effectively overcome by selection for expression of the stability cassette during exchange step gene targeting. A targeting efficiency of approximately ninety percent can be repeatably produced when the stable tag-exchange approach is used at the *LIF* locus.

Possible mechanisms for physical loss of the tag gene by means other than homologous recombination include chromosome loss or deletion or gene conversion. The mechanism(s) responsible for this phenomenon were investigated and the results are reported in chapter 6 of this thesis.

Since commencement of this project an alternate two-step targeting approach has been described by Detloff and co-workers (1994), termed the “Plug and Socket” approach. In this system, positive selection for homologous recombination is used during both targeting steps. In the first targeting step, a positive selection cassette and an incomplete second positive selection cassette, termed the socket (HPRT gene), are introduced close to the locus to be altered. In a subsequent round of gene targeting the positive selection cassette is replaced by the alteration of interest, while the missing portion of the socket gene, known as the plug, is also introduced. Generation of a functional plug and socket gene provides direct selection for homologous recombination in the second targeting step. Therefore this approach overcomes the problems associated with selection for loss of gene function in ES cells which can occur by means other than homologous recombination. Detloff and co-workers (1994) found that although one hundred percent of lines surviving in the second round of targeting had undergone homologous recombination leading to production of a active HPRT gene, only twenty-five percent of these lines also carried the alteration introduced in the second round of gene targeting. The incorrectly targeted lines could be eliminated by screening for loss of the positive selection cassette introduced in the first round of targeting by replica plating. Improved vector design may also reduce the background of incorrectly targeted lines produced by this approach (Dettloff *et al.*, 1994). The plug and socket approach therefore provides an alternative approach to the stable tag-exchange system for the efficient introduction of repeated alterations into a single locus.

The major limitation of the plug and socket approach is that the plug and socket gene remains close to the gene under investigation to ensure that a functional plug and socket gene is not introduced into the locus in the absence of the alteration of interest. The close proximity of the plug and socket gene to the gene under investigation may interfere with expression of the

endogenous gene. However, using a gene such as the HPRT gene (for which positive or negative selection can be applied) as the plug and socket gene it is possible to select for lines which have lost this gene. However this requires a further round of selection and loss of the plug and socket gene may result from gene conversion in which the gene under investigation is also restored to wildtype. In addition to this, a partial selection cassette is required on second round vectors used in the plug and socket approach. This increases the difficulty of vector construction. In contrast, in the stable tag-exchange approach it is possible to position the stability cassette well away from the gene under investigation and exchange step targeting vectors do not require exogenous selection cassettes.

The stable tag-exchange targeting approach provides an efficient means by which to repeatedly alter a single site within the mouse genome. The stable tag-exchange approach can provide an enrichment of up to 18 fold during the exchange step over the unmodified tag-and-exchange approach and as a consequence is at least 5 times more efficient than the most efficient published example of the tag-and-exchange gene targeting. The improved efficiency results from the inclusion of the stability cassette in the stable tag-exchange approach. The fact that close to one hundred percent of lines surviving in the exchange step are correctly targeted makes possible the efficient introduction of repeated alterations to any tagged site within the mouse genome using this technique. The features of the system allow the introduction of precise and/or subtle alterations to the mouse genome which will facilitate genetic study of the mouse. This and other similar systems should allow systematic alteration of the mouse genome, permitting direct analysis of predetermined genetic alterations in the context of the ES cell system and the whole organism.

Chapter 6

6 Analysis of ES cell Lines Resulting from Exchange Step Gene Targeting

6.1 Introduction

Selection in 6-TG during exchange step targeting at the LIF locus, in the unmodified tag-and-exchange approach, resulted in a high background of 6-TG resistant lines which did not carry the exchange step vector homologously recombined at the tagged LIF locus (sections 5.8.2, 5.8.4). Southern analysis of these lines indicated that many had resulted from physical loss of the tag gene by a means other than homologous recombination (sections 5.8.2, 5.8.4). Possible mechanisms resulting in physical loss of the tag gene in the absence of homologous recombination are (i) loss of the chromosome containing the tagged LIF allele (ii) chromosome deletion resulting in loss of the tagged LIF allele or (iii) gene conversion resulting in restoration of the tagged LIF allele to wildtype. Determining the mechanism responsible for tag gene loss should lead to an understanding of how the stable tag-exchange approach is able to eliminate these lines and should aid future design of targeting constructs used in the stable tag-exchange approach. Further analysis of non-exchange step targeted lines by karyotypic and Southern analysis was carried out to determine the likely mechanism of tag gene loss.

To determine whether the reporter genes reproduced faithfully the known pattern of LIF expression in vitro, at both the transcript and protein level, reporter gene expression was assessed in exchange step targeted ES cells and a differentiated ES cell derivative. This was carried out to justify use of the cells for the generation of mice.

6.2 Investigation of the Mechanism of Tag Gene Loss

A range of 6-TG resistant lines that survived in exchange step gene targeting due to physical loss of the tag gene by means other than homologous recombination (for examples see Fig. 5.15 b, lines selected in 6-TG) were investigated. Southern analysis (section 5.8.2, 5.8.4) indicated that all of the lines which had physically lost the tag cassette had also physically lost

the stability cassette. To investigate the physical integrity of the LIF locus in these lines karyotypic and Southern blot analysis was carried out.

6.2.1 *Karyotypic Analysis of 6-TG Resistant ES Cell Lines Resulting from Physical Loss of the Tag Gene*

The mouse LIF gene has been localised to chromosome 11, near the centromere at sub-bands A1-A2 (Kola *et al.*, 1990) as illustrated in figure 6.1 a. Karyotypic analysis was employed to determine if physical loss of the tag gene and stability genes had resulted from loss of chromosome 11 or from gross deletion in which the exogenous selection cassettes and surrounding DNA had been removed from chromosome 11. The preparation of chromosome spreads (Robertson, 1987) and chloroquine banding (Caspersson *et al.*, 1970) allowed the 2N karyotype of the lines to be assessed, identification of chromosome 11 and the integrity of chromosome 11 to be investigated (resolution of 20 - 50 x 10⁶ bp). Chromosome spreads were prepared from five ES cell lines which had physically lost the tag and stability cassettes (L6.5, L8.2, L9.1, L13.11, L15.2) and one line which carried the tag and stability cassettes but had an altered banding pattern by Southern analysis suggestive of either a translocation or alteration at the tagged LIF locus (such that an internal Hind III site was lost, Fig. 5.16, lane 7). Chromosome spreads were visualised after chloroquine banding in collaboration with Dr. Graham Webb (Department of Animal Science, Waite Campus, University of Adelaide).

Each of the lines had a 2N karyotype of 40, with two copies of chromosome 11 present. This finding eliminated the possibility of tag gene loss by chromosome loss.

In each case, the two copies of chromosome 11 were grossly normal as assessed by chloroquine banding, indicating that the lines did not carry a large deletion in the region of the LIF gene at sub-bands A1-A2. An example of the chromosome spreads is shown in figure 6.1 b, in which the two copies of chromosome 11 are indicated.

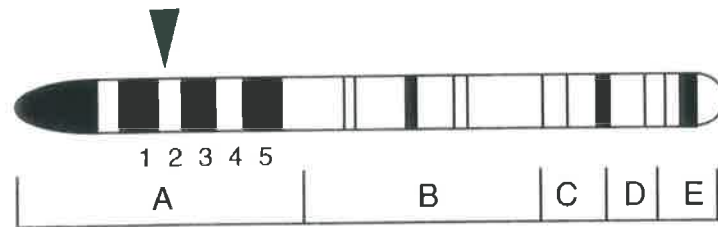
Figure 6.1 Karyotypic analysis of non-targeted lines resulting from physical loss of the tag gene

a Representation of the location of the LIF gene on mouse chromosome 11. The LIF gene (indicated with the arrow) is located near the centromere at sub-bands A1-A2 (Kola *et al.*, 1990).

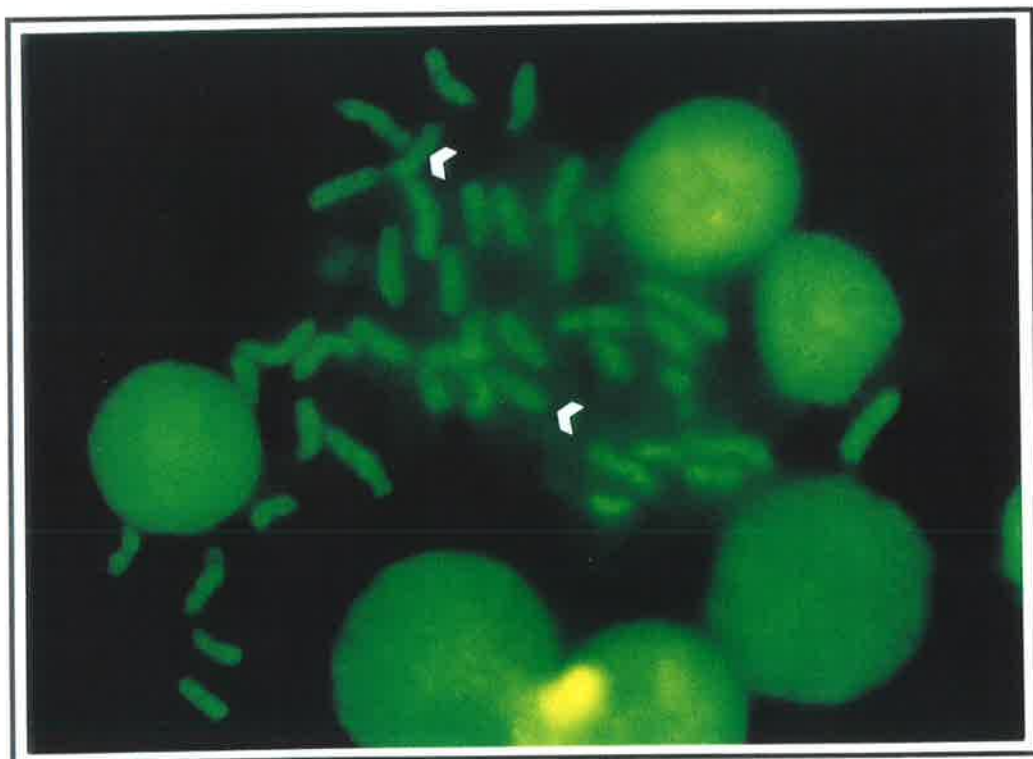
Adapted from Kola *et al.*, (1990)

b Example of karyotypic analysis of lines which have undergone physical loss of the tag gene by means other than homologous recombination. The photograph shows a chromosome spread from the cell line L9.1 in which the 2N karyotype was 40 and both copies of chromosome 11 (indicated with arrows) are present. Chromosome spreads were prepared by the method of Robertson, (1987) and were chloroquine banded by the method of Caspersson and co-workers (1970) and viewed under fluorescence optics. The photograph was taken on a Zeiss Axioplan Universal microscope.

a Chromosomal Location of the Mouse LIF Gene



b



6.2.2 Southern Analysis of 6-TG Resistant ES Cell Lines Resulting from Physical Loss of the Tag Gene

The findings of the previous section suggested that loss of the tag and stability cassettes had occurred either by a deletion of less than 20×10^6 bp, not obvious at the karyotypic level, or by gene conversion. To distinguish between these possibilities, the copy number and integrity of the LIF gene was determined by Southern analysis (able to detect restriction fragment differences in the order of 100-200 bp) in a range of lines which had undergone physical loss of the tag and stability genes by means other than homologous recombination. Southern analysis was carried out using a LIF probe and an internal control probe specific for the autosomal gene *Hesx1* (Thomas *et al.*, 1995), located on mouse chromosome 14 (Webb *et al.*, 1993). The ratio of the *Hesx1* to LIF alleles was determined from the band intensities in control lines which contain two *Hesx1* alleles and one or two wildtype LIF alleles. Comparison of the *Hesx1* to LIF ratio in the test and control lines was used to determine the copy number of the LIF gene in test lines. A wildtype LIF copy number of one would indicate that physical loss of the tag gene had most likely occurred by deletion in which the tag, stability cassettes and surrounding DNA were removed from the tagged allele. A wildtype LIF copy number of two would indicate that physical loss of the tag gene had most likely occurred by gene conversion in which only the tag and stability cassettes were removed from the tagged allele.

Control cell lines carrying (i) two wildtype *Hesx1* and LIF alleles (E14TG2a) or (ii) two wildtype *Hesx1* alleles and one wildtype LIF allele, one tagged LIF allele (HG60) were included in the analysis. Eight lines which had undergone tag gene loss were analysed (L6.5, L7.2, L8.2, L9.1, L10.2, L13.4, L13.11, L15.2). DNA from each of the control and test lines was digested with Bam HI and Southern analysis was carried out using the Eco RI/Pst I *Hesx1* and Sac II/Pst I LIF probes (Fig. 6.2 a). The *Hesx1* probe hybridises to a band of 12 kb, while the LIF probe hybridises to the wildtype LIF allele band of 2.9 kb and a tagged allele band of 9.0 kb between the tag and stability cassettes (Fig. 6.2 a, b). Quantitation of band intensity was carried out by volume integration using the ImageQuant software package.

Figure 6.2 Southern analysis of non-targeted lines resulting from physical loss of the tag gene

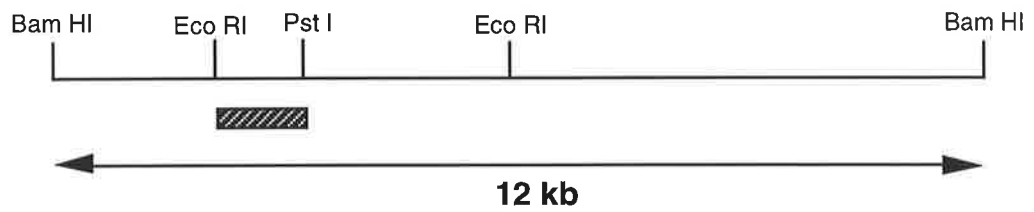
a Representation of the expected sizes of restriction fragments resulting from (i) the *Hesx1* allele, (ii) the tag step targeted LIF allele or (iii) the wildtype LIF allele after digestion with Bam HI. The probes used for Southern analysis and the restriction sites for Bam HI, Eco RI, Pst I, Sac I are shown. Thin line, genomic DNA; black boxes, exons; stippled box, HPRT cassette; hatched box, *Hesx1* probe; light hatched box, LIF probe (exons of the *Hesx1* gene are not shown).

b Southern analysis of ES cell lines which have undergone physical loss of the tag gene by means other than homologous recombination. Approximately 20 µg of genomic DNA from was digested with Bam HI, electrophoresed on a 1.0 % TAE agarose gel and blotted (section 9.3.21). The Southern was probed with the Eco RI/Pst I *Hesx1* probe of 1.1 kb (Fig. 6.2 i) and Sac II/Pst I LIF probe of 0.9 kb (Fig. 6.2 a ii, iii) and washed in 0.1 x SSC/0.1 % SDS at 65° for 30 minutes (9.3.11; 9.3.22). The size of the fragment to which the *Hesx1* probe hybridises is 12.0 kb and fragments to which the LIF probe hybridises are 9.0 and 2.9 kb from the tag step targeted and untargeted LIF alleles, respectively. Lanes are as follows C1, E14TG2a; C2, HG60, 1, L6.5; 2, L7.2; 3, L8.2; 4, L9.1; 5, L10.2; 6, L13.4; 7, L13.11; 8, L15.2.

a Southern Analysis of *Hesx1*/LIF Ratio

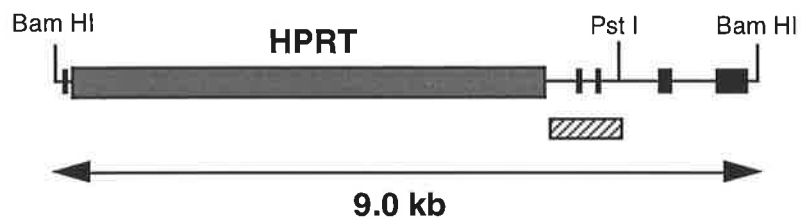
i

Hesx1 allele



ii

Tagged LIF allele

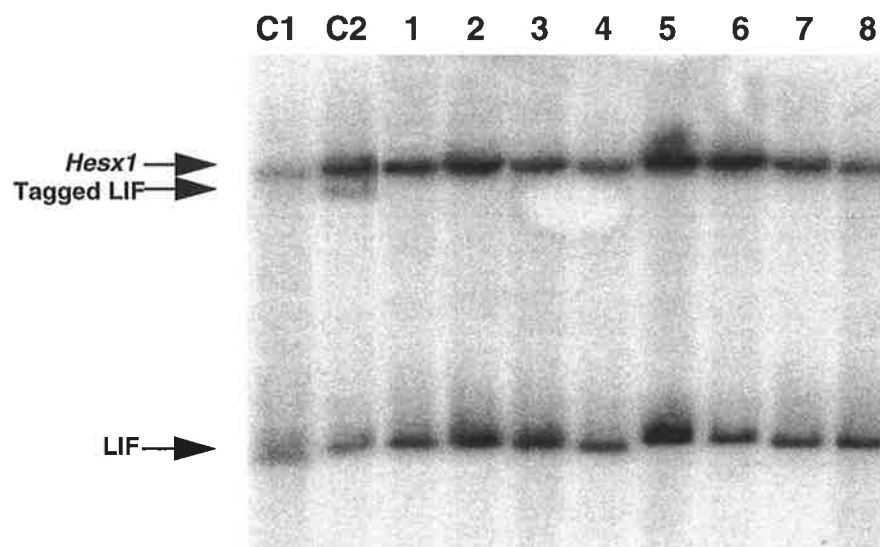


iii

LIF allele



b



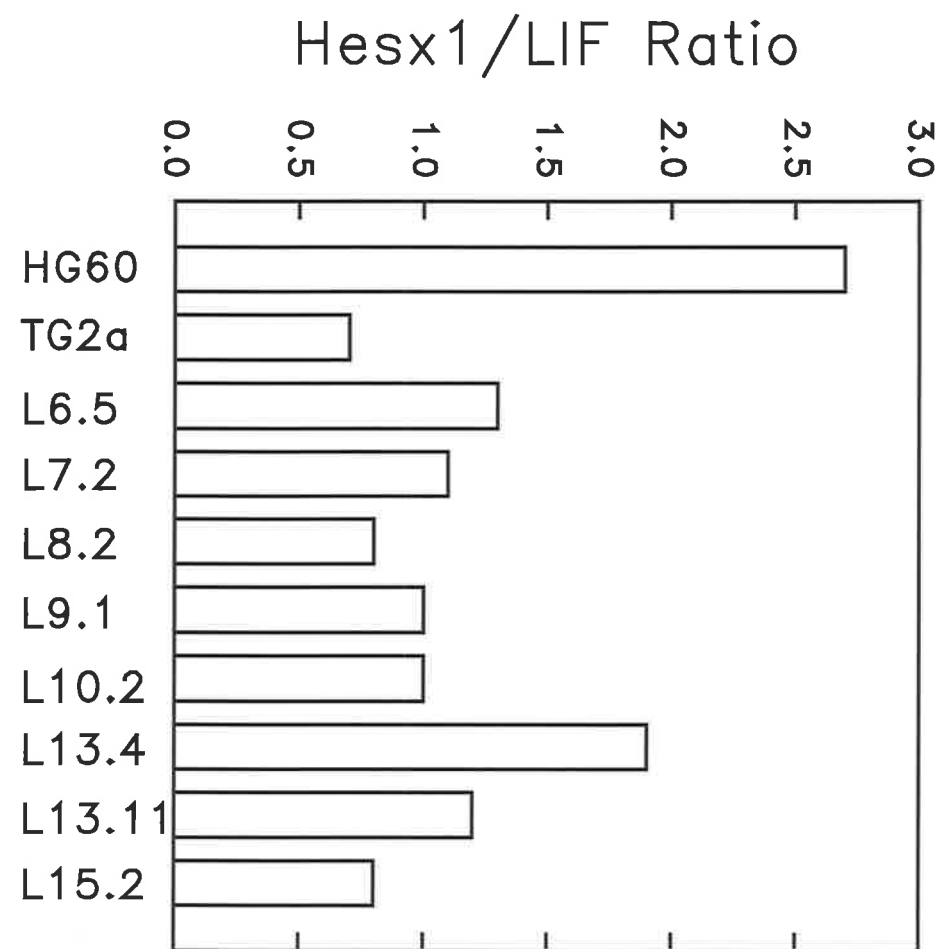
The ratio of *Hesx1* to LIF alleles for each line are graphed in figure 6.3. It can be seen that the ratio in each of the test lines, with the possible exception of L13.4, is closest to that of the control line, E14TG2a which contains two copies of the *Hesx1* gene and two copies of the LIF genes. This indicates that at least seven of the eight lines tested carry two wildtype LIF alleles. Even with the inclusion of line L13.4 the average *Hesx1*/LIF ratio of the test lines was not significantly different from that of the control line E14TG2a (Student's t test, level of significance of 0.001) while it was significantly different from the *Hesx1*/LIF ratio of the control line HG60 (Student's t test, level of significance of 0.001). The most likely explanation for this is that the majority (> 85 %) or all of the affected lines had undergone gene conversion in which the tag and stability cassettes have been removed from the tagged LIF allele without loss of the intervening LIF sequence. This suggests that the mechanism of tag gene loss is part of a repair process which removes exogenous DNA from the mouse genome. Although the results for line L13.4 are ambiguous, the *Hesx1* /LIF ratio for this line is closer to that of the control line HG60 which contains one LIF allele. This line may have lost the intervening LIF sequence between the tag and stability cassettes possibly as a result of inaccurate gene conversion or deletion.

6.3 Analysis of Reporter Gene Expression in Exchange Step Targeted Lines

The exchange step targeted lines carrying the LIF reporter constructs (section 5.8) were analysed to determine whether the pattern of reporter gene expression in these cells reflected the known LIF expression pattern in ES cells and their differentiated progeny to justify use of the cells for the production of mice. RNase protection analysis has shown that ES cells grown in culture express very low levels of the LIF D and M transcripts, while treatment with the chemical inducer MBA (section 1.6.5), leads to an upregulation of both these transcripts (Rathjen *et al.*, 1990b). Expression of the *lacZ* reporter gene was therefore assessed in ES cells and MBA differentiated derivatives at both the transcript and protein level.

Figure 6.3 Graph of the ratio of Hesx1 to LIF alleles

Graph of the ratio of *Hesx1* to untargeted LIF alleles in ES cell lines which have undergone physical loss of the tag gene by means other than homologous recombination. Quantitation of band intensities was carried out by volume integration using the ImageQuant software package. Values were normalised with respect to the *Hesx1* allele and the ratio of *Hesx1*/untargeted LIF band intensity is plotted for each cell line. The control line HG60 carries one wildtype LIF allele and two *Hesx1* alleles and the control line E14TG2a contains two wildtype LIF and *Hesx1* alleles (as described in section 6.2.2).



6.3.1 *RT-PCR Detection of LIF D Transcription in Exchange Targeted Lines Carrying the Reporter Gene in the Second Exon of LIF*

An RT-PCR approach was used to detect the expression of LIF D in two independent exchange step targeted lines carrying the *lacZ* reporter gene in the second exon of LIF (section 5.8). RNA was isolated from ES cells or from ES cells which had been differentiated with MBA. ES cells or cells to be treated with MBA were seeded at a density of 1×10^4 cells/cm². ES cells were grown for five days in complete ES cell medium prior to the isolation of RNA. ES cells treated with MBA were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF prior to isolation of RNA.

In order to distinguish between RT-PCR products derived from residual genomic DNA and the fusion transcript, lines carrying the reporter gene in exon two of LIF (E2 34.5, SE2 2.4) were used since products amplified from genomic DNA or RNA in these lines differ in size by over 1500 bp (Fig. 6.4 a). A primer specific for the first exon of LIF D (1Dlac5') and the *lacZ* reporter gene (E2lac3') were used to amplify the LIF D-*lacZ* fusion transcript (Fig. 6.4 a). A spliced product (384 bp) was amplified from both the ES cell and MBA differentiated ES cells while the genomic product (1946 bp) was not seen (Fig. 6.4 b). Southern analysis using a *Sac* II/*Ban* I *lacZ* probe (approx. 270 bp; Fig. 6.4, hatched box) was used to confirm that the 384 bp product was specific for the fusion transcript (Fig. 6.4 b ii). The *lacZ* probe also detects the *lacZ* gene contained in the pUC19 markers (lanes H). This result confirmed expression of LIF D by ES cells and that the LIF D transcript was correctly spliced. The relative level of expression in ES and MBA differentiated ES cells was not assessed using this non-quantitative PCR approach.

6.3.2 *In Vitro Expression of β -galactosidase Activity in Exchange Step Targeted Lines*

A functional assay for the LIF-LacZ or LacZ protein produced by the exchange step targeted lines was carried out. A histochemical assay for β -galactosidase activity using 5-bromo-4-chloro-3-indol- β -D-galactopyranoside (X-gal) as the substrate was performed on ES cells and ES cells differentiated with MBA as above (section 6.3.1). Lines carrying the *lacZ*

Figure 6.4 Analysis of LIF D transcription in exchange step targeted lines

a Representation of the expected products from RT-PCR analysis of LIF D transcription in lines carrying the *lacZ* reporter gene in exon 2 of the LIF gene. The expected sizes of the products, the primers used for PCR amplification (1Dlac5' binds to LIF exon 1D and E2lac3' binds to a region within the *lacZ* coding sequence) and the SacII/Ban I *lacZ* probe from pnlacF used for Southern analysis are shown. The expected product amplified from the fusion transcript is 384 bp long (upper) and that amplified from the genomic sequence is 1946 bp (lower). Arrows, primers used for PCR amplification, black boxes, LIF exons, discontinuous line, LIF intron; stippled box, *lacZ* gene, hatched boxes, *lacZ* probe.

b RT-PCR analysis (section 9.3.13) of LIF D transcription in exchange step targeted lines. RT-PCR reactions were carried out using 2 µg RNA and approximately half of each sample was electrophoresed on 2.8 % TAE agarose gels (section 9.3.5). The product amplified from the fusion transcript is indicated with the arrow. The origin of the upper band is not clear since it is not the product expected from amplification of the genomic sequence. (i) Lanes are as follows

H: Hpa digested pUC19 size markers; sizes: 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34 and 26 bp.

1: negative control reaction using water

2: reaction using ES cell RNA from line SE2 2.4

3: reaction using MBA differentiated ES cell RNA from line SE2 2.4

H: Hpa digested pUC19 size markers

S: Eco RI digested SPP-1 size markers: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48 and 0.36 kb.

(ii) Southern analysis of LIF D expression showing gel (left) and Southern analysis results (right). Gel was blotted probed with a Sac I/Ban I *lacZ* probe of 270 bp (Fig. 6.4 a, hatched box; section 9.3.21; 9.3.11) and washed in 0.1 x SSC/0.1 % SDS at 65° for 30 minutes (section 9.3.22). Lanes are as follows:

H: Hpa digested pUC19 size markers

1: reaction using ES cell RNA from line SE2 2.4

2: reaction using MBA differentiated ES cell RNA from line SE2 2.4

3: negative control reaction using water

H: Hpa digested pUC19 size markers

S: Eco RI digested SPP-1 size markers

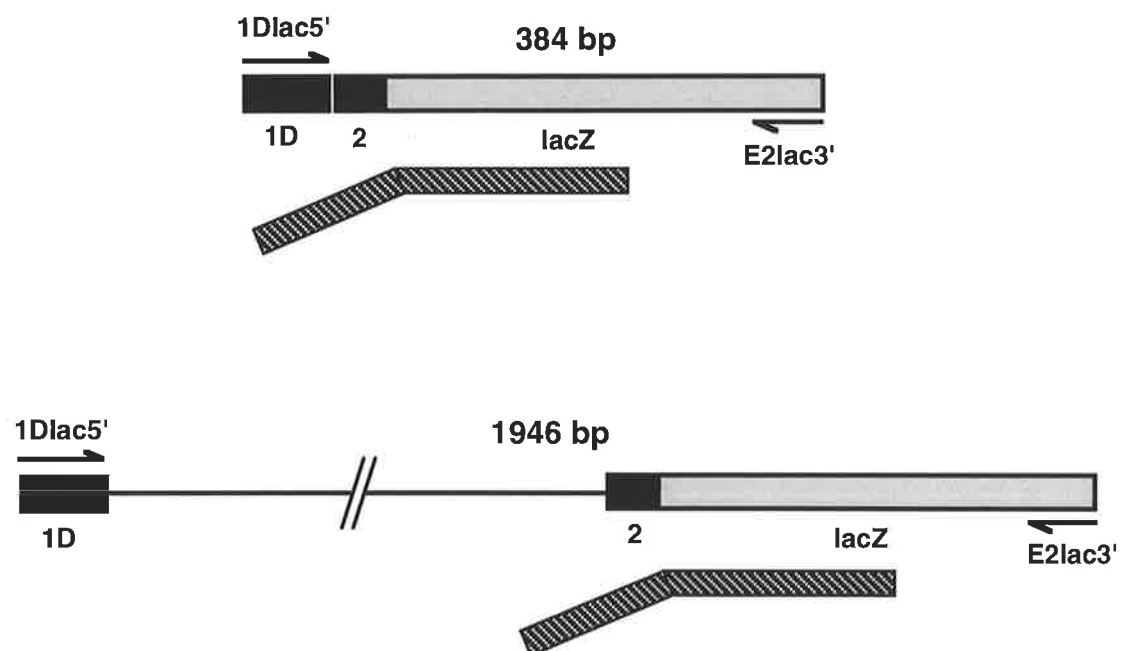
4: reaction using ES cell RNA from line E2 34.5

5: reaction using MBA differentiated ES cell RNA from line E2 34.5

6: negative control reaction using water

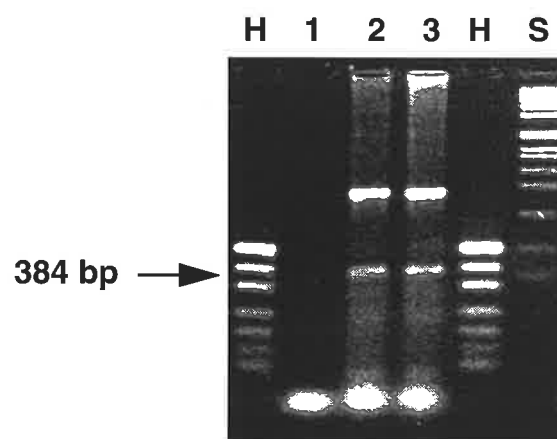
H: Hpa cut pUC19 size markers

a RT-PCR Analysis of LIF D Expression

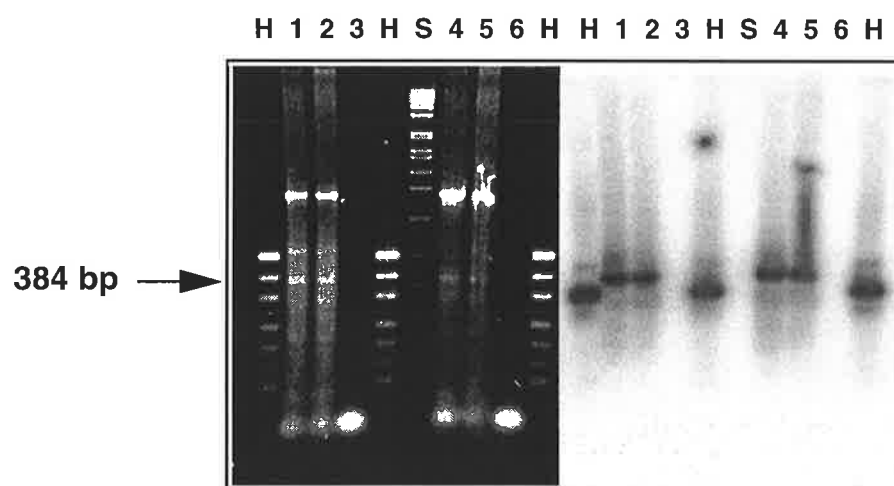


b

i



ii



reporter gene in the first exon of LIF D, M, T or in the common second exon of LIF, lines 1D 16.1, 1M 14.4, 1T 9.1 and SE2 2.4 respectively, were assessed by direct visualisation. A tag step targeted ES cell line lacking a *lacZ* reporter gene (HG60) was included in this analysis as a negative control. β -galactosidase activity was not detected in any of the undifferentiated ES cell lines and this presumably reflects the sensitivity of this assay since LIF transcription in ES cells has been shown previously (Rathjen *et al.*, 1990b) and confirmed in this work (section 6.3.1). Mountford and co-workers (1994) used gene targeting in ES cells to introduce a single *lacZ* reporter gene into the LIF gene to allow visualisation of LIF expression in vitro and in vivo. Using X-gal as a substrate for β -galactosidase activity LIF expression in undifferentiated ES cells was also undetectable in that study (Mountford, *et al.*, 1994).

β -galactosidase activity was observed in each of the ES cell lines carrying a marked form of LIF D, M, T or of all LIF isoforms after MBA differentiation but was not seen in parallel cultures of the control line HG60. Cells displaying high or low levels of β -galactosidase activity were seen. Relatively high level β -galactosidase activity was seen in differentiated cells, which were generally located in isolation. These cells, displaying large nuclei, and a large flattened morphology, were seen in cultures of 1T 9.1 and SE2 2.4. β -galactosidase activity was clearly localised to the nucleus of these cells (Fig. 6.5 a, b).

Cells displaying high levels of β -galactosidase activity, located within stem cell-like colonies, were identified in each of the MBA treated cultures of 1D 16.1, 1M 14.4, 1T 9.1 and SE2 2.4 (Fig. 6.6 a - d) but not in cultures of HG60 (Fig. 6.6 e). It was not obvious from initial inspection whether β -galactosidase activity in these colonies had arisen from a single cell or in a group of closely associated cells. Concurrent visualisation of the nucleus (using Hoechst stain) and β -galactosidase activity in line SE2 2.4 indicated that a single cell was likely to be responsible for this staining pattern (Fig. 6.7 a, b). From this result it can be seen that β -galactosidase positive cells located within stem cell-like colonies were similar in appearance to the isolated cells, displaying an enlarged, flattened nucleus. It is therefore possible that the cell types located within stem cell-like colonies and in isolation are the same, although this

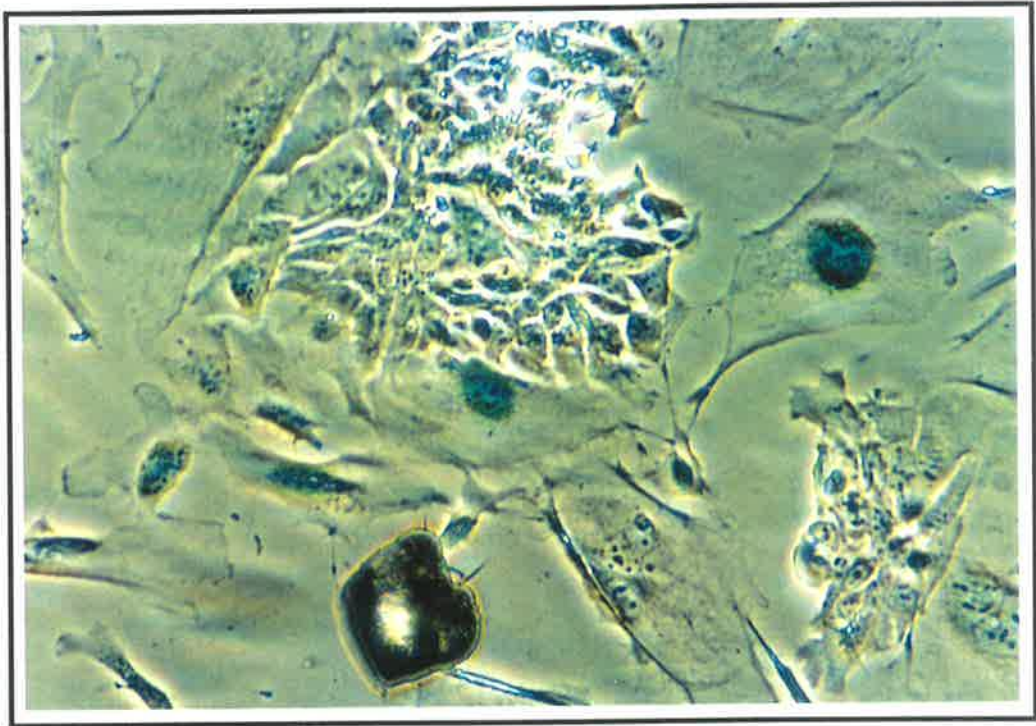
Figure 6.5 Nuclear localised β -galactosidase activity in exchange step targeted line SE2 2.4

a Photograph of β -galactosidase activity in the exchange step targeted line SE2 2.4 after differentiation with MBA showing nuclear localised β -galactosidase activity (photographed under phase contrast optics at 50 x magnification).

b Photograph of β -galactosidase activity in the exchange step targeted line SE2 2.4 after differentiation with MBA showing nuclear localised β -galactosidase activity in a single cell (photographed under phase contrast optics 100 x magnification).

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15) prior to fixation and histochemical staining for β -galactosidase activity, using X-gal (1 mg/ml) as the substrate (section 9.4.13). Photographs were taken on the Nikon Diaphot microscope.

a



b

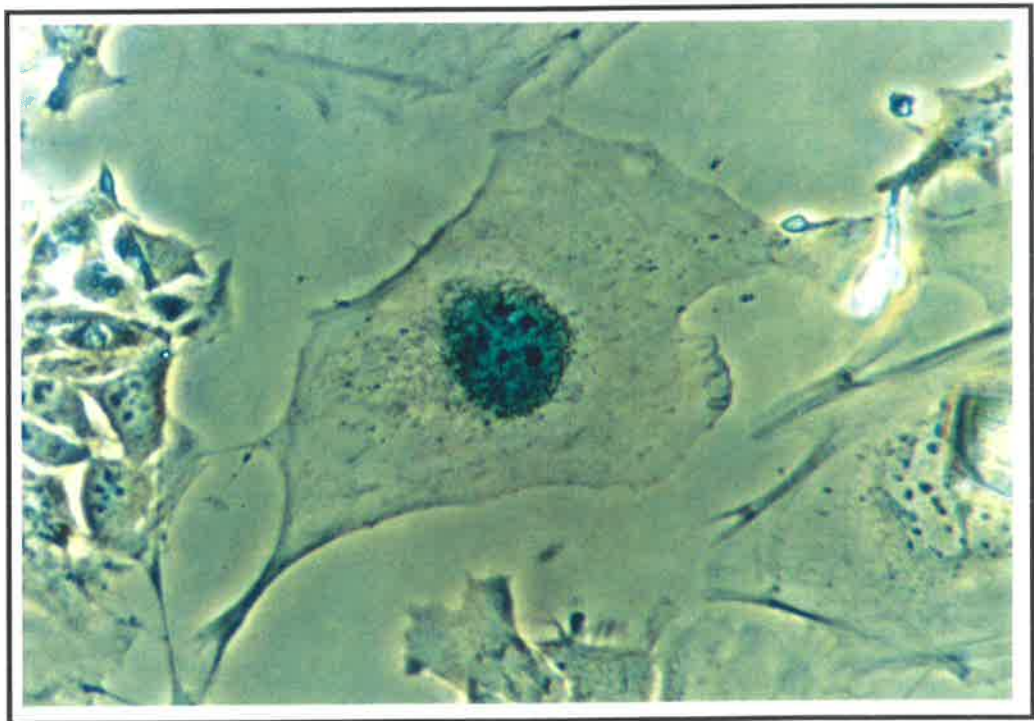
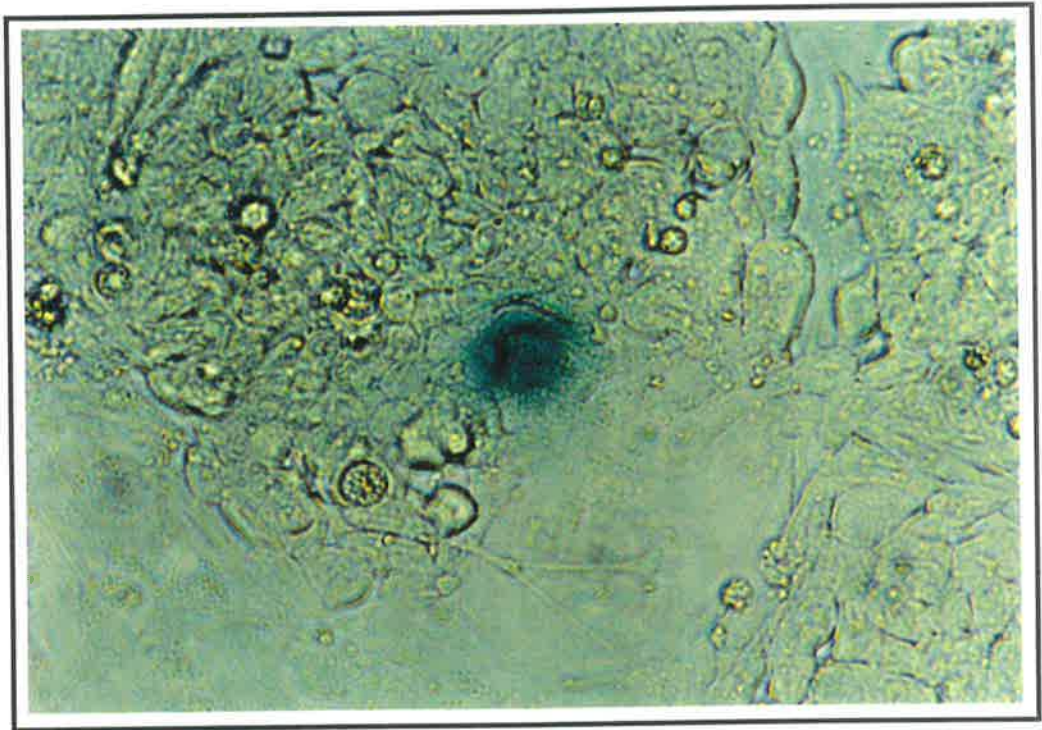


Figure 6.6 High level β -galactosidase activity in exchange step targeted lines

a-b Photograph of β -galactosidase activity in exchange step targeted lines after differentiation with MBA **a** 1D 16.1 (photographed under bright field optics at 50 x magnification) and **b** 1M 14.4 (photographed under bright field optics at 25 x magnification).

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15) prior to fixation and histochemical staining for β -galactosidase activity using X-gal (1 mg/ml) as the substrate (section 9.4.13).

a



b

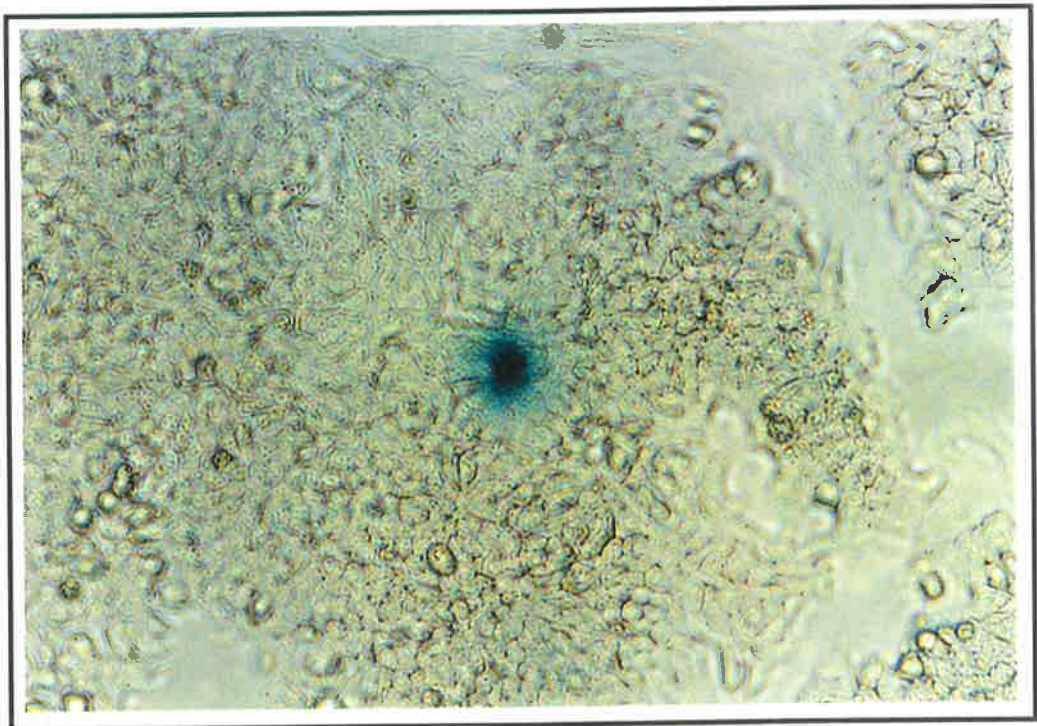
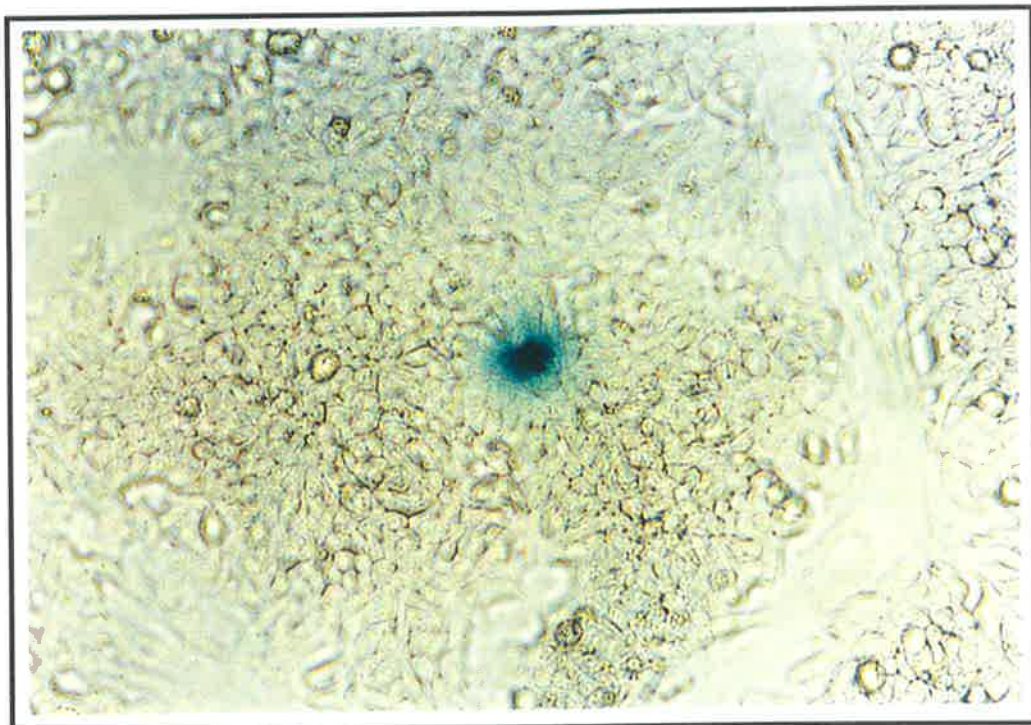


Figure 6.6 High level β -galactosidase activity in exchange step targeted lines

c-d Photograph of β -galactosidase activity in exchange step targeted lines after differentiation with MBA **c** 1T 9.1 (photographed under bright field optics at 25 x magnification) and **d** SE2 2.4 (photographed under bright field optics at 25 x magnification).

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15) prior to fixation and histochemical staining for β -galactosidase activity using X-gal (1 mg/ml) as the substrate (section 9.4.13).

c



d

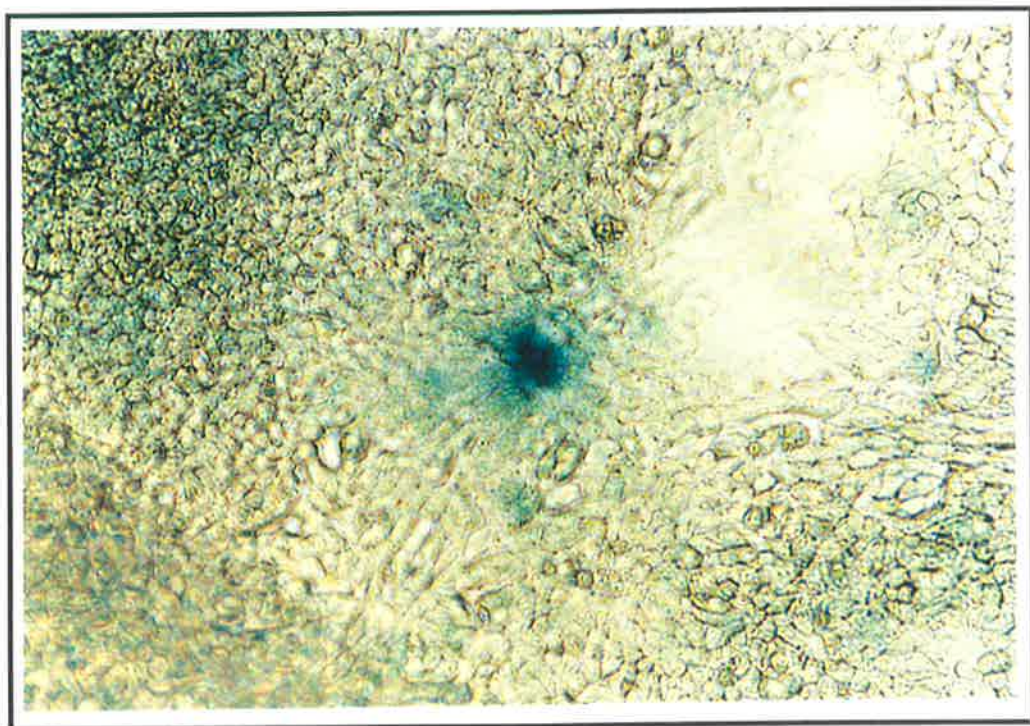


Figure 6.6 High level β -galactosidase activity in exchange step targeted lines

e Photograph of β -galactosidase activity in the control line HG60 after differentiation with MBA (photographed under bright field optics at 25 x magnification).

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15) prior to fixation and histochemical staining for β -galactosidase activity, using X-gal (1 mg/ml) as the substrate (section 9.4.13).

e

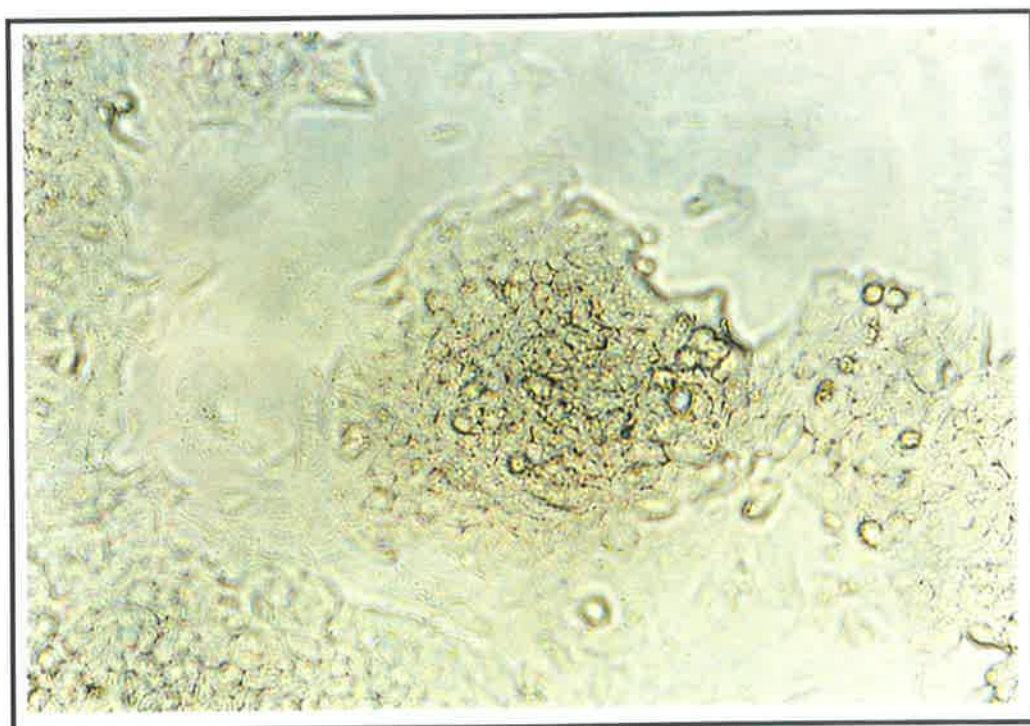


Figure 6.7 Concurrent visualisation of β -galactosidase activity and nuclei

a Photograph of exchange step targeted line SE2 2.4 after differentiation with MBA showing β -galactosidase activity contained within a stem cell like colony (photographed under bright field optics at 100 x magnification).

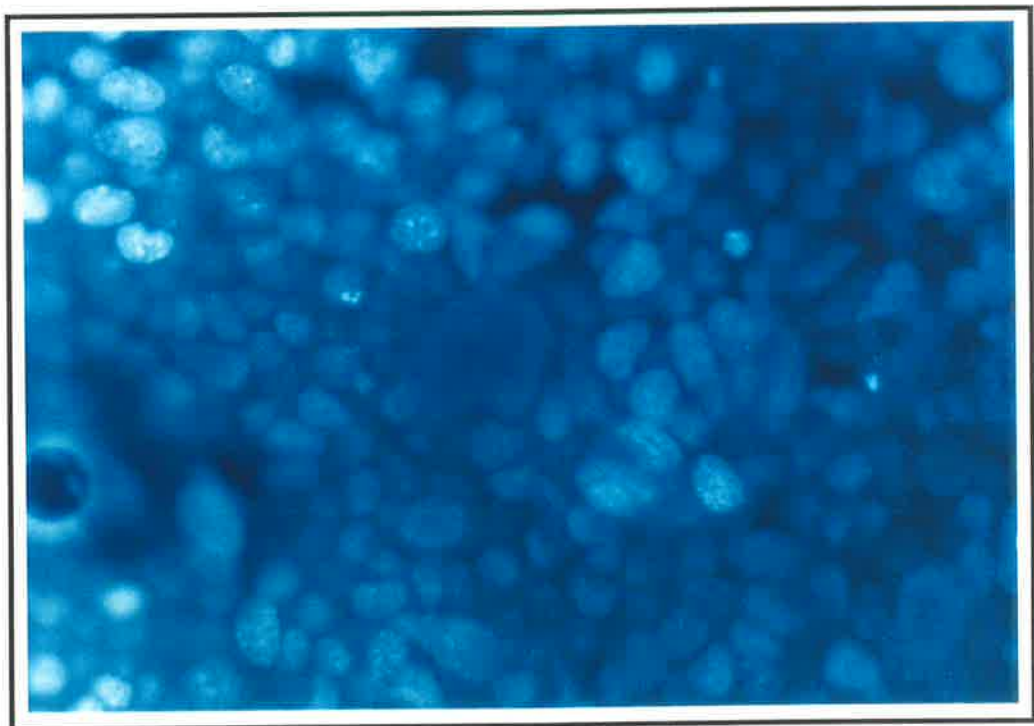
b Photograph of exchange step targeted line SE2 2.4 after differentiation with MBA (as above) showing Hoechst stained nuclei (photographed under fluorescence optics at 100 x magnification).

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15 prior to fixation and histochemical staining for β -galactosidase activity using X-gal (1 mg/ml) as the substrate (section 9.4.13) and staining for the nuclei with Hoechst dye (section 9.4.16). Photographs were taken on the Zeiss Axioplan Universal microscope.

a



b



possibility has not been investigated further. If this is not the case it would suggest that LIF T is expressed in a different, more migratory cell type from that in which LIF D and M are expressed.

Cells displaying low levels of β -galactosidase activity were also seen and were particularly apparent in cultures of lines 1M 14.4, 1T 9.1 (Fig. 6.8 a, b) and SE2 2.4 (Fig. 6. 6 d, 6.7 a, 6.8 c) but were not seen in cultures of the line HG60. In each of these lines low level staining was seen in cells clustered in groups within stem cell-like colonies. Upregulation of LIF in a subpopulation of cells within stem cell-like colonies indicates that expression of each LIF transcript is specifically regulated. The fact that this sort of staining was not seen in undifferentiated ES cell cultures indicates that the staining is unlikely to represent LIF expression in undifferentiated ES cells. However, the appearance and location of these cells suggests that this pattern of expression reflects LIF expression in a closely related differentiated derivative or intermediate of ES cells. In the example shown of low level staining in the culture of 1M 14.4 (Fig 6.8 a) the β -galactosidase positive cells are located at the periphery of the stem cell-like colony. This is suggestive of a 'stem cell niche' in which LIF expression in surrounding differentiated ES cells maintains a stem cell population as has been proposed previously (Rathjen *et al.*, 1990b; Mountford *et al.*, 1994).

The number of β -galactosidase positive cells varied within and between experiments due to variability in the cell types produced and MBA induced toxicity. This made it difficult to compare the absolute numbers of β -galactosidase positive cells between the cell lines, although general trends were seen. The number of cells displaying high and low level β -galactosidase expression was repeatedly significantly greater in cultures of SE2 2.4 (Fig. 6.5 a, 6.6 d, 6.7 a, 6.8 c, 6.9), than in cultures of 1D 16.1, 1M 14.4 and 1T 9.1 in which a single isoform of LIF could be visualised. In the example shown in figure 6.9 many differentiated cells displaying high levels of β -galactosidase activity are clustered together. This may reflect differentiation of a particular cell type in a localised area. The increase in the number of β -galactosidase positive cells in line SE2 2.4 was greater than that expected from concurrent visualisation of each LIF isoform. This finding is consistent with the possibility

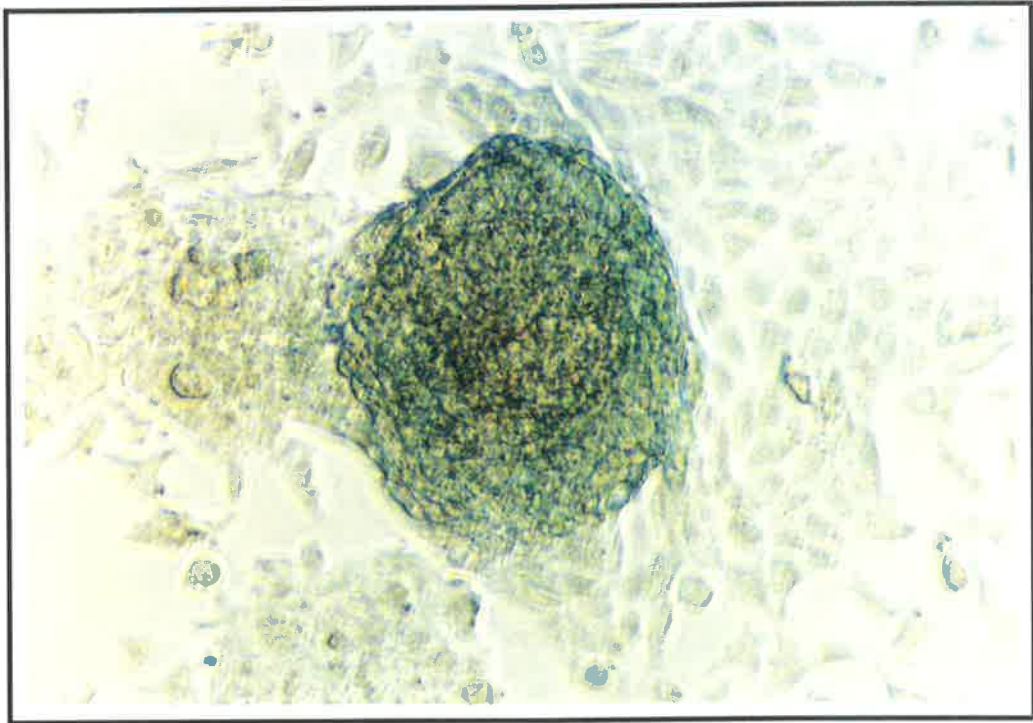
Figure 6.8 Low level β -galactosidase activity in exchange step targeted lines

a Photograph of β -galactosidase activity in exchange step targeted line 1M 14.4 after differentiation with MBA showing staining at the periphery of the stem cell-like colony (photographed under bright field optics at 50 x magnification).

b Photograph of β -galactosidase activity in exchange step targeted line 1T 9.1 after differentiation with MBA showing staining within a stem cell-like colony (photographed under bright field optics at 50 x magnification).

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15) prior to fixation and histochemical staining for β -galactosidase activity, using X-gal (1 mg/ml) as the substrate (section 9.4.13). Photographs were taken on the Nikon Diaphot microscope.

a



b

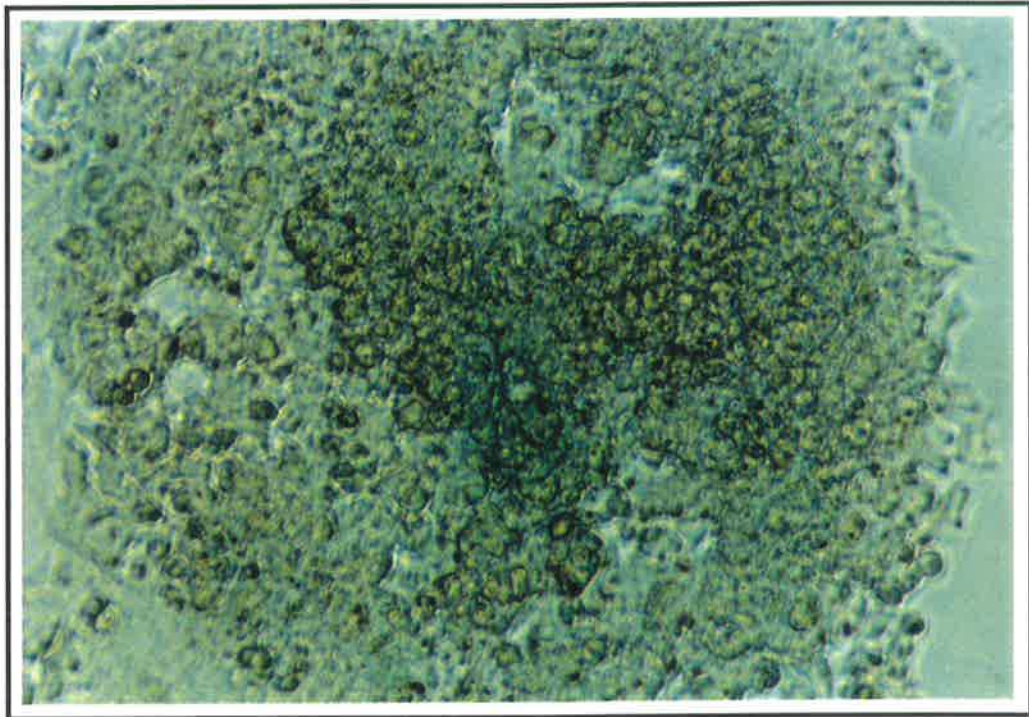


Figure 6.8 Low level β -galactosidase activity in exchange step targeted lines

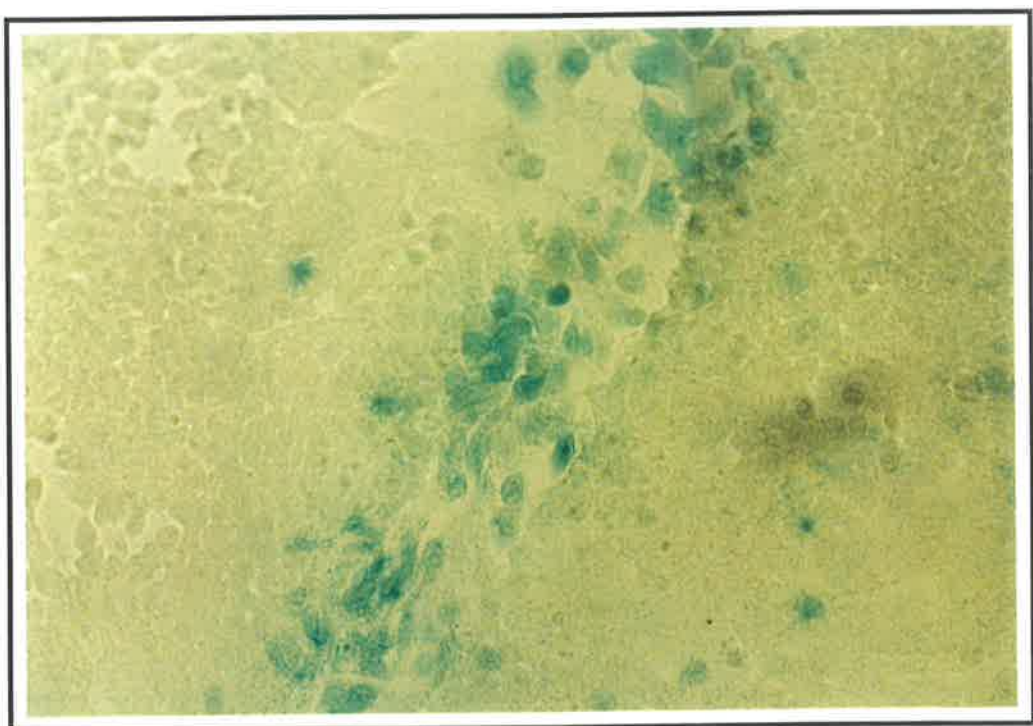
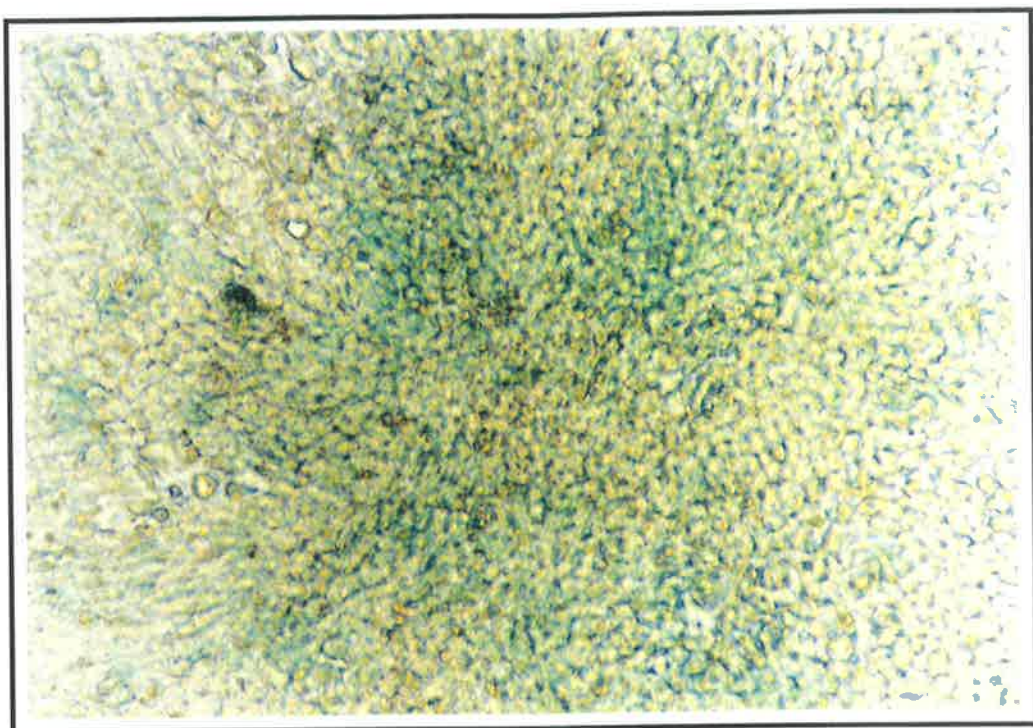
c Photograph of β -galactosidase activity in exchange step targeted line SE2 2.4 after differentiation with MBA showing staining within a stem cell-like colony (photographed under bright field optics at 50 x magnification). Photograph was taken on the Nikon Diaphot microscope.

Figure 6.9 β -galactosidase activity in exchange step targeted line SE2 2.4

Photograph of β -galactosidase activity in exchange step targeted line SE2 2.4 after differentiation with MBA showing many high level staining cells (photographed under bright field optics at 50 x magnification). Photograph was taken on the Zeiss Axioplan Universal microscope.

Cells were grown for eighteen hours in reduced LIF medium (100 U/ml), for two days in reduced LIF medium containing MBA (5 mM), and for two days in medium lacking LIF (section 9.4.15) prior to fixation and histochemical staining for β -galactosidase activity, using X-gal (1 mg/ml) as the substrate.

c



that some cells carrying a single marked isoform of LIF produce β -galactosidase expression at levels below detection by the histochemical β -galactosidase assay. In line SE2 2.4, in which expression of all known LIF isoforms can be visualised, concurrent low level expression of two or more isoforms of LIF in a single cell results in levels of β -galactosidase activity which are detectable by this assay. This results in an increase in β -galactosidase positive cells seen in this culture compared to that expected from the sum of β -galactosidase positive cells seen in the cultures in which a single isoform can be visualised.

The results presented in this chapter indicate that the LIF reporter system reproduces faithfully LIF expression under in vitro growth conditions and suggests that exchange step targeted ES cells will be suitable for LIF expression mapping in the mouse.

6.5 Discussion

In published examples of the unmodified tag-and-exchange gene targeting approach, a high background of 6-TG^r cells which are not exchange step targeted (Askew *et al.*, 1993; Gondo *et al.*, 1994; Stacey *et al.*, 1994; Wu *et al.*, 1994) has rendered the system unsuitable for the repeated introduction of alterations into a single site within the mouse genome. This background results from loss of tag gene expression presumably caused by promoter shutdown or gene inactivation, or alternatively physical loss of the tag gene by means other than homologous recombination. In this project, up to seventy seven percent of the non-targeted 6-TG^r lines had arisen due to physical loss of the tag gene by means other than homologous recombination from the LIF locus. Of the lines examined further, over eighty percent had physically lost the tag gene and stability cassettes by a process which restored the tagged LIF allele to wildtype. This presumably occurred by gene conversion in which the exogenous DNA was removed from the tagged LIF allele. This is consistent with the findings of Wu and co-workers (1994). In a tag-and-exchange approach, physical loss of the tag gene at the *Colla-1* locus contributed to the background of non-homologous recombinants which survived selection in the exchange step. Southern blot analysis showed that greater than ninety seven percent of these lines had physically lost the tag gene by gene conversion. In

other reports in which physical loss of the tag gene has been reported the mechanisms responsible for this phenomenon were not investigated.

The stability cassette was included in the stable tag-exchange approach to prevent physical loss of the tag gene by chromosome loss or deletion. However use of the stability cassette was also able to overcome loss of the tag gene by gene conversion. In each of the cell lines investigated both the tag and stability cassettes had been removed by this process. This suggests that gene conversion is a 'large scale process' in which the tag and stability cassettes were always removed in parallel. As a result of this, dual selection for loss of the tag gene and maintenance of the stability cassette, eliminates lines which have lost the tag gene by gene conversion and enriches for lines which have undergone homologous recombination.

Reporter gene expression was assessed in correctly targeted cell lines produced in exchange step gene targeting to justify use of these cells for production of mice (section 7.4.1). Expression of the LIF D-*lacZ* fusion transcript in two targeted ES cell lines carrying the reporter gene in exon two of the LIF gene was detected in ES cells and MBA differentiated derivatives. This indicated that reporter gene transcription reflected that of the known LIF D expression pattern in these cells in vitro (Rathjen *et al.*, 1990b).

Analysis of β -galactosidase activity in ES cells and their MBA differentiated derivatives indicated that expression of the reporter protein reflected the known LIF expression pattern in vitro (Rathjen *et al.*, 1990b; Mountford *et al.*, 1994). The β -galactosidase-fusion proteins retained enzymic activity in differentiated ES cell derivatives and the NLS was functional in MBA differentiated derivatives of ES cells. The *lacZ* reporter system should therefore be suitable for use in the differentiated cell types of the mouse and enable the identification of individual β -galactosidase positive cells within complex tissues of the mouse. Use of this reporter system to map the expression of LIF in the mouse has advantages over techniques used previously for LIF expression mapping (Conquet and Brûlet, 1990; Murray *et al.*, 1990; Bhatt *et al.*, 1991; Shen and Leder, 1992; Robertson *et al.*, 1993) since it distinguishes between each LIF isoform and provides cellular resolution.



The inability to detect β -galactosidase activity in undifferentiated cultures of ES cells indicates that the low level of LIF expression in these cells (Rathjen *et al.*, 1990b) is below the limits of detection by the histochemical assay used in this work. This is consistent with the findings of Mountford and co-workers (1994) and may be overcome with use of more sensitive systems for the detection of β -galactosidase (section 3.7.2).

Several conclusions about LIF biology can be drawn from the findings presented in this chapter.

- i. Cells expressing relatively high or low levels of LIF expression were seen in MBA differentiated ES cell cultures, with lower levels of expression seen within or surrounding stem cell-like colonies. This indicates that level of LIF expression is differentially regulated in differentiated cell types *in vitro*.
- ii. Expression of LIF D, M and T in a subpopulation of cells produced by treatment with MBA indicates that the expression of each LIF isoform is regulated specifically *in vitro* implying that each LIF isoform has biological roles.
- iii. Many more β -galactosidase positive cells were seen in cultures in which expression of all LIF isoforms could be visualised in comparison to cultures in which expression of a single LIF isoform could be visualised. The number of β -galactosidase positive cells in these cultures was greater than the sum of the β -galactosidase positive cells in each of the lines in which a single LIF isoform could be visualised. This suggests that more than one isoform of LIF is expressed concurrently in a single cell type *in vitro*. This implies that expression in a given tissue of the mouse may be non-uniform.
- iv. Translation of the LIF-LacZ fusion protein in lines in which LIF D or M can be visualised indicates that the initiation ATG in the first exons of these forms of LIF are utilised. This is of particular importance in the case of the first exon of LIF M which is

conserved in all species analysed but only contains an initiation ATG in the case of the mouse.

Direct in vivo analysis of some classes of developmental control genes can be hampered by the extremely low levels of gene expression. In addition to this, the production of more than one biologically active form of gene products such as LIF (Rathjen *et al.*, 1990a) and Int-2 (Acland *et al.*, 1990), may further complicate this analysis. The gene targeting and reporter systems used in this project may therefore provide a paradigm for the analysis of such gene products. The targeting approach provides a system by which multiple alterations can be made to the gene of interest with ease, while the reporter system provides a means by which to identify the cellular sites of expression of multiple forms of the gene product, both in vitro and in vivo. Application of this approach to the analysis of similar developmental control genes should lead to a detailed understanding of the function of developmental control gene products in the context of the whole organism.

Chapter 7

7 Assessment of Germline Competence of the E14TG2a ES Cell Line and Targeted Derivatives

7.1 Introduction

In the first step of the stable tag-exchange targeting approach, ES cell lines can be generated for the production of null mutant mice. Using tag step targeted ES cells, precise and/or subtle alterations can be made repeatedly to the tagged allele efficiently in exchange step gene targeting. The phenotypic consequences of these alterations can be assessed in the resultant ES cells cultured *in vitro* or *in vivo* after return to the mouse. For the production of mice genetically altered ES cells produced by the stable tag-exchange approach must retain germline competence throughout two rounds of gene targeting.

In contrast to teratocarcinoma derived EC cells, which are almost exclusively germline incompetent, (Robertson, 1986), a large range of ES cell lines retain germline competence after genetic manipulation (Hogan *et al.*, 1994). The reason for this difference has been attributed primarily to the non-euploid karyotype of EC cells and the normal euploid karyotype of ES cells (Robertson., 1986). A normal euploid karyotype is required by ES cell lines if they are to form functional gametes at high efficiency (Hogan *et al.*, 1994). Therefore, it is essential to check the karyotype of ES cell lines or targeted derivatives before chimaeric mice are produced with the cells.

The culture conditions used for ES cell maintenance can affect their karyotype. Suboptimal growth conditions can lead to selection for variants which have undergone chromosome rearrangements and/or mutations which are not readily detectable at the karyotypic level (Hogan *et al.*, 1994). Thus, it is also essential to check the germline competence of ES cells before targeting is commenced. Although culture in medium supplemented with LIF is sufficient to support the growth of undifferentiated ES cells, many laboratories co-culture ES cells with feeder cells (Hogan *et al.*, 1994), particularly when germline competent cells are required. Co-culture with feeder cells has several disadvantages. It increases the time and cost associated with ES cell maintenance, increases the risk of contamination, reduces the

repeatability of experiments and can hamper or preclude the use of certain selection protocols. In this project, ES cells were cultured in the absence of feeder cells and with an unpurified source of LIF (Smith., 1991). This was the first time such conditions had been used for ES cell culture in Adelaide. It was therefore of particular importance to establish that the unmanipulated E14TG2a and targeted derivatives used or produced in this project retained germline competence.

To assess pluripotency, ES cells are returned to the mouse blastocyst by microinjection or by blastocyst co-culture (Hogan *et al.*, 1994). The ability of ES cells to contribute to tissues of the mouse can be assessed by the identification of ES cell genetic markers in the resultant chimaeric mice. Chimaerism can be assessed by the identification of coat colour markers and by the identification of enzyme markers, such as isoforms of glucosephosphate isomerase (GPI) (Robertson., 1986) in various tissues. Germline competence can be assessed by the identification of coat colour markers transmitted to the offspring of chimaeric mice or by the identification of enzyme markers in sperm of the chimaerae (Robertson., 1986).

7.2 Assessment of Germline Competence of Unmanipulated E14TG2a ES Cell Line

To determine if the culture conditions used in this project were sufficient to maintain pluripotent ES cells, E14TG2a ES cells described in chapter 3 (section 3.3.2), were returned to the mouse blastocyst by microinjection, chimaerae were identified by coat colour assessment and the degree of chimaerism determined. Matings of chimaeric male mice were carried out to assess germline competence. Microinjection of ES cells was carried out in collaboration with Anita Peura and Steve McIlpatrick (Bresatec Ltd., Adelaide).

7.2.1 Return of E14TG2a ES Cells to the Mouse

E14TG2a ES cells maintained in the absence of feeders and in unpurified LIF, were assessed by karyotypic analysis and were found to have a normal 2N karyotype of forty (a minimum of 15 spreads were counted; normal karyotype: > 65 % of spreads 2N = 40). Cells were microinjected into host CBA-F2 (CBA/C57BL) blastocysts and transferred into

pseudopregnant females. An example of microinjection of ES cells into the mouse blastocyst can be seen in figure 7.1 in which a microinjected blastocyst (Fig 7.1 a) and a microinjected blastocyst before transfer into a pseudopregnant female (Fig. 7.1 b) can be seen. Four female and twelve male offspring resulted from these transfers. The distortion in the sex ratio resulting from use of male (XY) ES cells seen in this work has been reported previously (Bradley *et al.*, 1984; Robertson *et al.*, 1986).

Chimaerae were identified by coat colour assessment. Detailed descriptions of coat colour genetics can be found in Silvers (1979) and Lyon and Serle (1994) and genotypes of the mouse in Festing (1992). Briefly, the 129/Ola derived E14TG2a ES cells are homozygous for the chinchilla allele (c^{ch}/c^{ch}) (Hooper *et al.*, 1987), while the CBA F2 hosts are homozygous for the wildtype allele at this locus (C/C ; Lyon and Serle 1994). Coat derived from the host blastocyst is either agouti or black in colour and that derived from the ES cells is either chinchilla or agouti depending on surrounding tissue. All but one of the male offspring showed overt coat colour chimaerism. An example of the chimaeric mice is shown in figure 7.2 a and the percentage coat colour chimaerism of each mouse is represented in figure 7.2 b.

7.2.2 Investigation of Germline Competence of E14TG2a ES Cells

The 11 chimaeric males produced with the E14TG2a ES cells (section 7.2.1) were mated to BALB/c females to assess germline transmission. The transmission of ES cell coat colour markers to offspring results in grey mice from this cross whereas offspring resulting from host blastocyst are agouti or black.

The two males showing the highest level of chimaerism by coat colour (60 %, mice 11 and 13) transmitted the ES cell genotype through the germline. The percentage germline transmission is illustrated in figure 7.2 c. One of the germline transmitting male chimaerae, the BALB/c female with which it was mated and an example of offspring carrying the ES cell genotype can be seen in figure 7.3 a. An example of the offspring carrying the ES cell genotype is shown in comparison to a normal litter mate in figure 7.3 b.

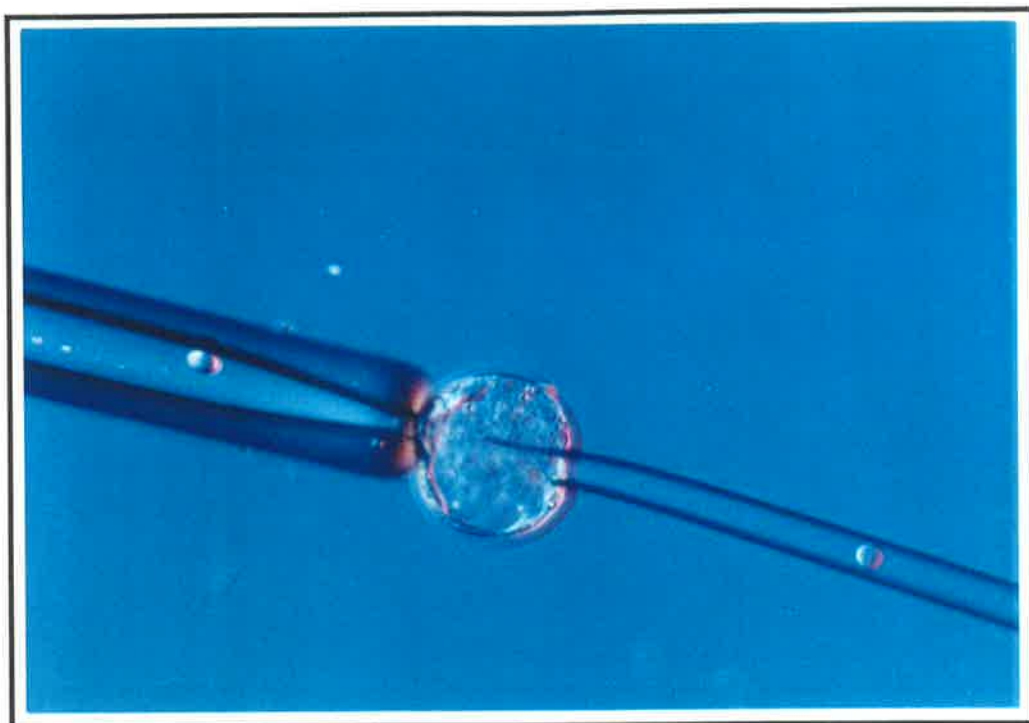
Figure 7.1 Return of ES cells to the mouse blastocyst

a Photograph of microinjection of ES cells into the preimplantation mouse blastocyst showing the injection of cells into the blastocoelic cavity using an injection pipette (photographed under differential interference contrast optics at 50 x magnification).

b Photograph of a microinjected blastocyst prior to transfer into a pseudopregnant female (photographed under differential interference contrast optics at 50 x magnification).

Photographs taken on the Nikon Diaphot microscope were kindly provided by Joe Wrin.

a



b

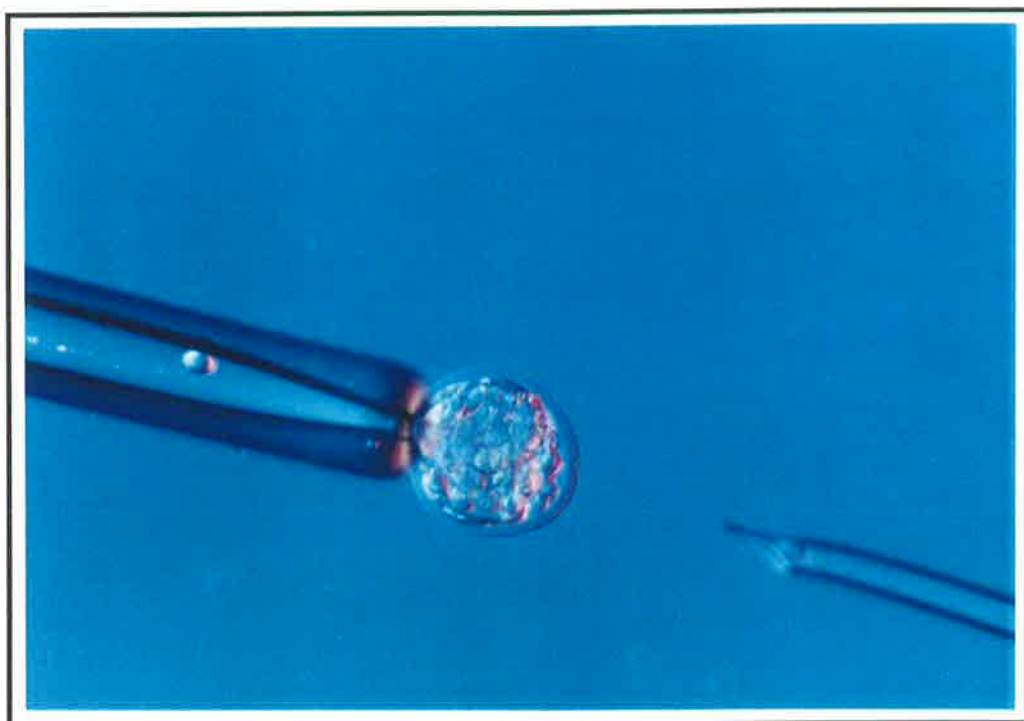


Figure 7.2 Analysis of chimaeric mice produced with unmanipulated E14TG2a cells

a Photograph of a chimaeric mouse produced with unmanipulated E14TG2a ES cells. Chimaeric mice were produced as described in section 9.5.1.

b Graph showing the percentage contribution of the chinchilla, agouti and black coat colours to chimaeric female and male mice produced with unmanipulated E14TG2a ES cells. The numbers below each column represent individual mice.

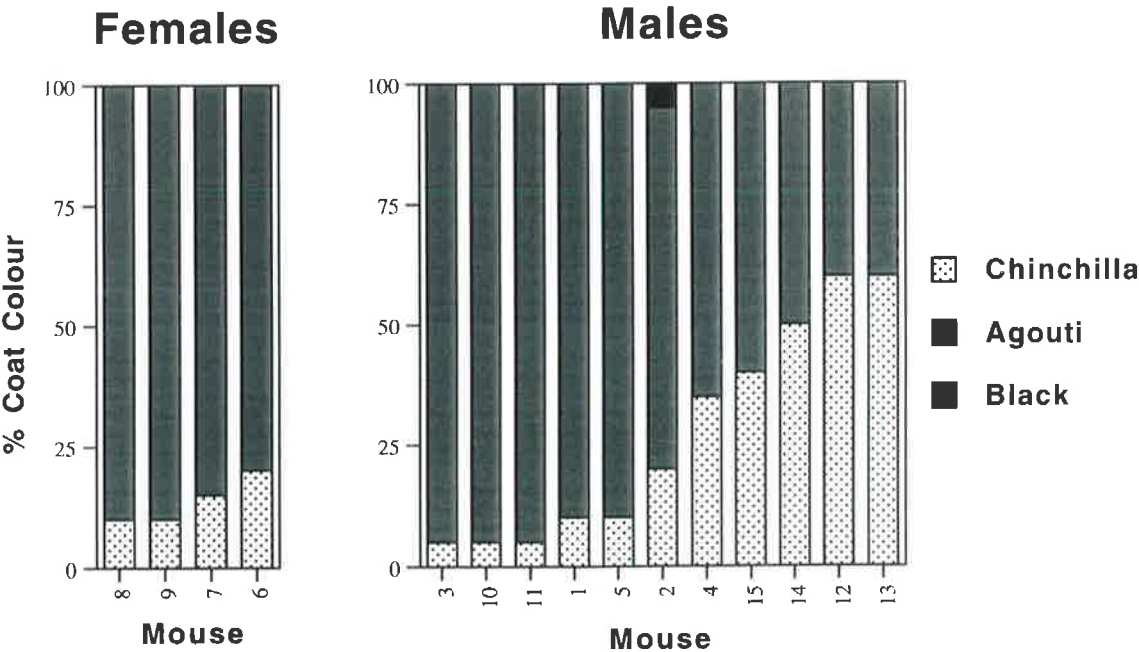
c Graph representing the percentage germline transmission for chimaeric male mice produced with unmanipulated E14TG2a ES cells as assessed by breeding. The numbers below each column represent individual mice (as above).

Chimaeric Analysis: E14TG2a

a



b



c

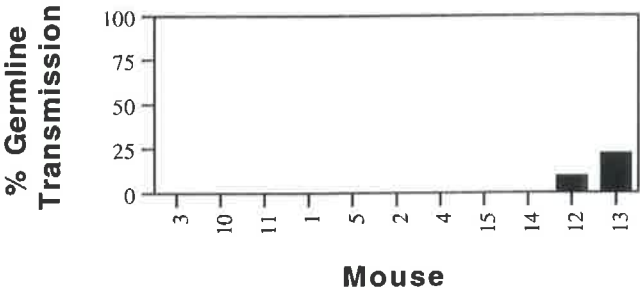


Figure 7.3 Offspring carrying the ES cell markers transmitted through the germline

a Photograph of chimaeric male mouse #12 (chinchilla 60 %/ agouti 40 %, left), the BALB/c female (albino, right) with which it was mated and an example of the offspring produced by the transmission of ES cell markers through the germline (grey, middle).

a Photograph showing an example of the offspring produced by the transmission of ES cell markers through the germline (affectionately known as Lucky, behind) in comparison to an agouti sibling (in front).

Chimaeric mice were produced as described in section 9.5.1.

a



b



Four of the chimaeric mice produced with the E14TG2a cells had eye abnormalities which resulted in blindness in one or both eyes. The two chimaerae that transmitted the ES cell genotype through the germline were blind in one or both eyes, but this abnormality was not transmitted to heterozygous offspring. Similar eye defects were seen in the chimaerae produced with each targeted derivatives of the E14TG2a cell line produced in this work.

7.3 Assessment of Germline Competence of Tag Step Targeted ES Cells

Tag step targeted ES cell lines produced in this work (section 4.3.2) were returned to the mouse blastocyst by microinjection. Chimaerae were identified by coat colour assessment, while germline transmission was assessed by the glucosephosphate isomerase (GPI) assay (DeLorenzo and Ruddle, 1969). The GPI assay was used since this a quick way to determine germline competence and this result was required before exchange step targeting could be carried out.

7.3.1 *Return of Tag Step Targeted ES Cells to the Mouse*

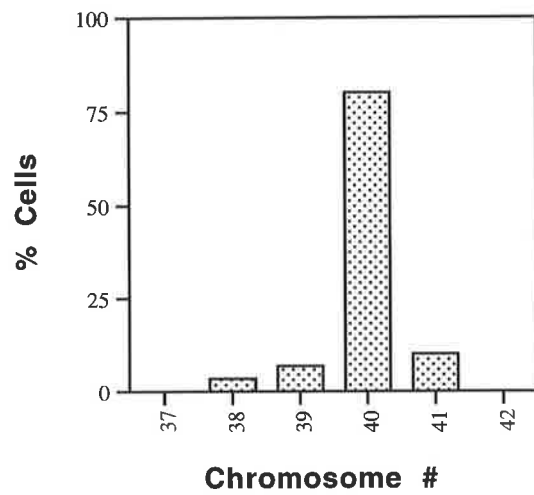
Karyotypic analysis of targeted lines derived from tag step gene targeting was carried out. Two correctly targeted lines selected with HAT (lines H9 and H20; Fig. 4.8 b-d, lanes 9 and 20) and two selected with HAT plus G418 medium (lines HG26 and HG60; Fig. 4.8 b-d, lanes 26 and 60), with normal karyotypes (Fig 7.4 a - d) were microinjected into CBA F2 or C57BL F1 (C57BL/6 x C57BL/10) blastocysts and transferred into pseudopregnant females. Six offspring (female:male, 3:3) resulted from CBA F2 blastocysts injected with the line H20 all of which were chimaeric as assessed by coat colour. Chimaerism ranged from five to twenty percent (results not shown). Eight offspring (female:male, 3:5) resulted from CBA F2 blastocysts injected with the line HG26. At least two were chimaeric by coat colour assessment and chimaerism ranged from five to fifteen percent (results not shown). Seventeen offspring (female:male, 5:12) resulted from CBA F2 or C57BL F1 blastocysts injected with the line HG60 and all were chimaeric as assessed by coat colour (Fig. 7.5 a). No offspring resulted from CBA F2 or C57BL F1 blastocysts injected with the line H9.

Figure 7.4 Karyotypic analysis of tag step targeted ES cells

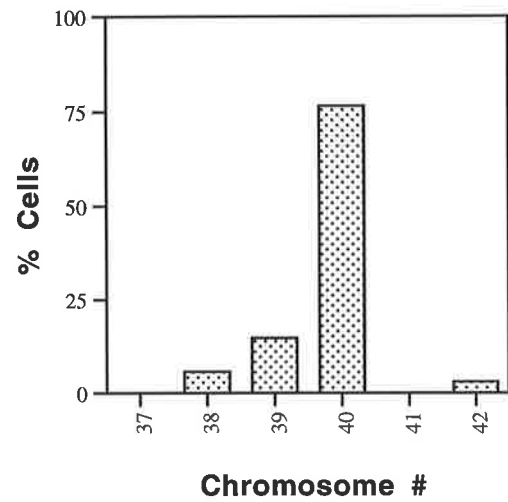
a-d Graphs showing the number of 2N chromosomes counted in examples of the tag step targeted cell lines **a** H9, **b** H20, **c** HG26 and **d** HG60, represented as percentages.

Chromosome spreads were prepared and stained by the method of Robertson, (1987) (section 9.4.16) and viewed under bright field optics on the Zeiss Axioplan Universal microscope.

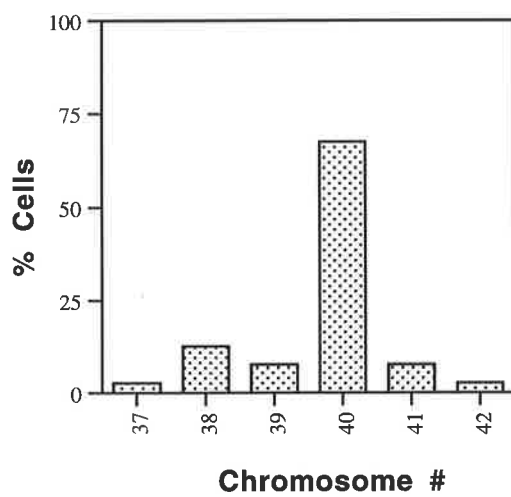
a Karyotype of H9



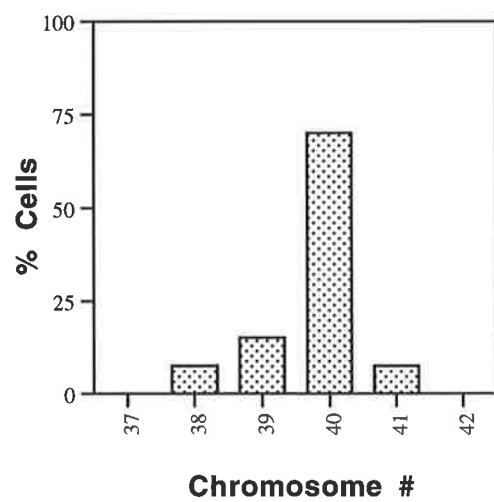
b Karyotype of H20



c Karyotype of HG26



d Karyotype of HG60



7.3.2 Investigation of Germline Competence of Tag Step Targeted ES Cell Lines

The *Gpi-1s* gene is expressed widely in tissues of the mouse (Lyon and Serle, 1994). Two alleles of this enzyme, *Gpi-1s^a* and *Gpi-1s^b*, differ in their electrophoretic mobility and can be distinguished by starch-gel electrophoresis followed by specific staining (DeLorenzo and Ruddle, 1969). *Gpi-1s^a* migrates more slowly than *Gpi-1s^b*. The distribution of these alleles differs in various strains of the mouse and in inbred strains the alleles are homozygous. The *Gpi-1s^a* allele occurs in strains including BALB/c (DeLorenzo and Ruddle, 1969) and 129 (Hooper *et al.*, 1987) while the *Gpi-1s^b* allele occurs in strains including C57BL and CBA (DeLorenzo and Ruddle, 1969). Thus, the contribution of E14TG2a derived ES cells (129 strain) to various tissues of chimaeric mice produced by introduction of ES cells to C57BL and CBA blastocysts can be determined by GPI analysis.

Chimaeric males produced with the tag step targeted lines H20, HG26 and HG60 were sacrificed, sperm isolated in collaboration with Jenny Washington and *Gpi-1s* analysis carried out. The results are shown in figure 7.5 b. Controls containing sperm samples carrying the *Gpi-1s^a* enzyme (Fig. 7.5 b, Lanes +SP) or *Gpi-1s^b* enzyme (Fig. 7.5 b, lanes -SP), or a mixture of blood samples carrying the *Gpi-1s^a* or *Gpi-1s^b* enzyme (sample carrying *Gpi-1s^a* 90 % and *Gpi-1s^b* 10 %; Fig. 7.5 b, lanes +BL; sample carrying *Gpi-1s^b* 90 % and *Gpi-1s^a* 10 %; Fig. 7.5 b, lanes -BL) were included. Samples from the chimaeric males produced with H20 (lanes 2, 3, 4), HG26 (Lane 8) and HG60 (Lanes 5 - 12, 14-17) were assessed. Sperm samples from four of the chimaeric males produced with the line HG60 contained the *Gpi-1s^a* enzyme variant (upper arrow indicated with 129). This indicated that sperm derived from the introduced ES cells was produced in mice 7, 10, 14 and 16 (lanes 7, 10, 14, 16). The percentage transmission for each of these mice was estimated by determining the relative intensities of the *Gpi-1s^a* and *Gpi-1s^b* bands by volume integration using the ImageQuant II software. Levels of transmission of approximately twenty, twenty five, thirty and ninety percent were seen for mice seven, ten, sixteen and fourteen respectively (Fig. 7.5 c). The germline competent line HG60 was used for exchange step gene targeting.

Figure 7.5 Analysis of chimaeric mice produced with tag step targeted ES cell lines

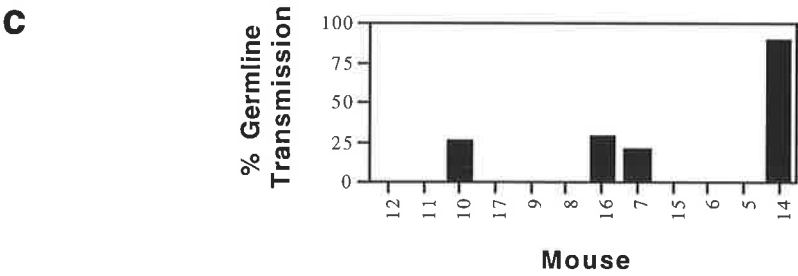
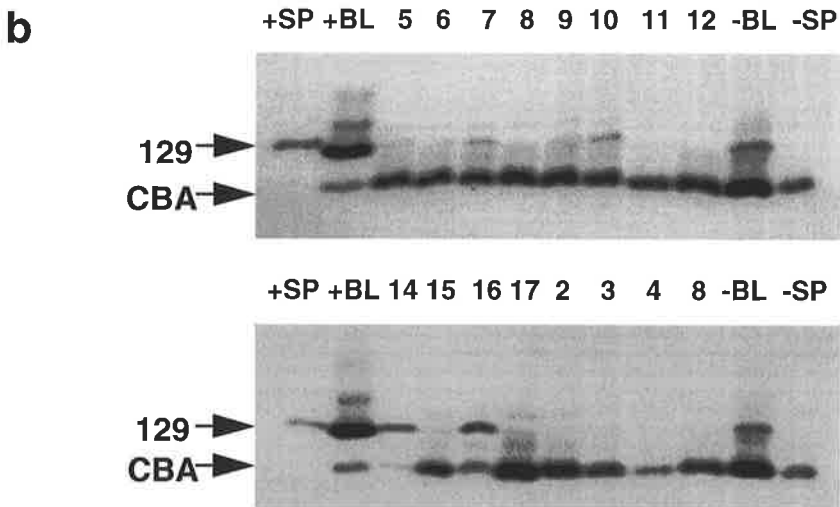
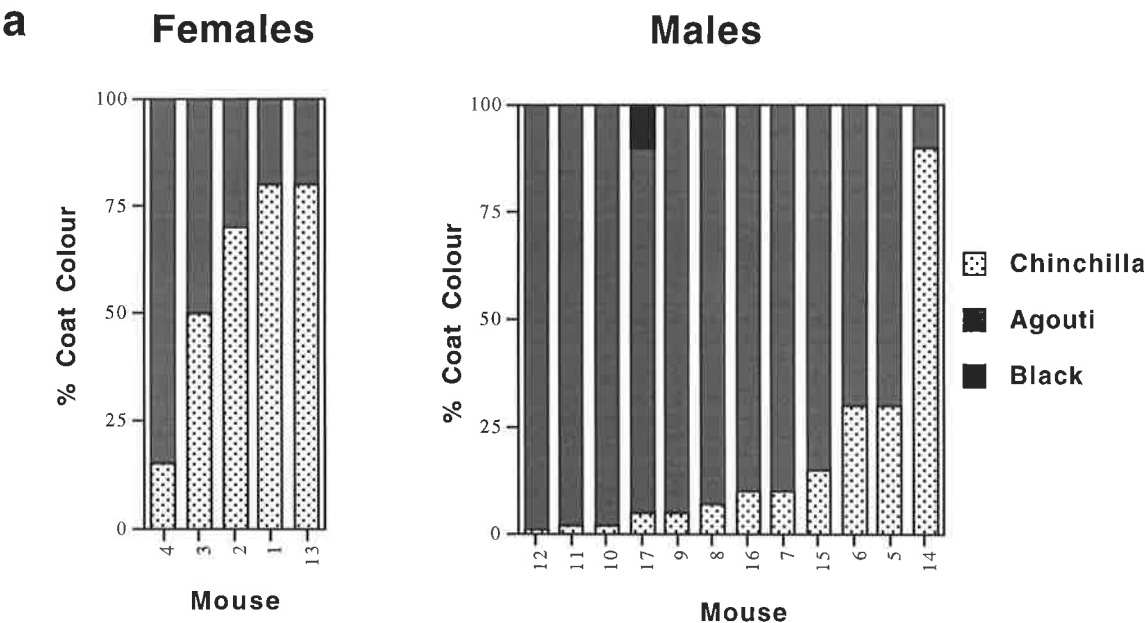
a Graph showing the percentage contribution of the chinchilla, agouti and black coat colours to chimaeric female and male mice produced with tag step targeted ES cell line HG60. The numbers below each column represent individual mice.

b Results of GPI analysis of sperm from chimaeric male mice produced with tag step targeted lines (section 9.5.2). The ES cell derived allele from the 129 strain *Gpi-1^a* is indicated with the arrow marked 129 and the CBA allele *Gpi-1^b* is indicated with the arrow marked CBA. The lanes are marked with the corresponding mouse numbers or with control labels. Controls containing sperm samples from 129 (lanes +SP) or CBA strains (lanes -SP), or a mixture of blood samples from the 129 and CBA strains (129 strain 90 % and CBA strain 10 %; lanes +BL; CBA strain 90 % and 129 strain 10 %; lanes -BL) were included. Samples from tag step targeted line H20 (lanes 2-4), HG26 (lane 8) and HG60 (lanes 5-12, 14-17) are shown.

c Graph representing the percentage germline transmission for chimaeric male mice produced with ES cell line HG60 as assessed by GPI analysis. The numbers below each column represent individual mice (as above).

Chimaeric Analysis: Tagged Lines

Chimaeric Analysis: HG60



7.4 Assessment of Germline Competence of Exchange Step Targeted ES cells

7.4.1 Return of Exchange Step Targeted ES Cell Lines Carrying pE2LacLIF to the Mouse

Karyotypic analysis of ES cell lines carrying the *lacZ* reporter gene in exon two of the LIF gene, produced in exchange step gene targeting (section 5.8.2) was carried out. Three correctly targeted lines selected in 6-TG plus G418 medium with normal karyotypes (SE2 2.2, Fig. 7.6 a; SE2 2.4, Fig. 7.6 b and E2 34.5, Fig 7.6 c) were microinjected into CBA F2 host blastocysts. Ten offspring (female:male, 2:8) resulted from microinjection of blastocysts with the line SE2 2.2 and six offspring resulted (female:male, 3:3) from blastocysts microinjected with the line SE2 2.4 (Fig. 7.7 a). All offspring were chimaeric as assessed by coat colour and the percentage chimaerism ranged from five to ninety percent. This is represented diagrammatically for line SE2 2.4 (Fig. 7.7 b). Only one offspring resulted from blastocysts injected with the cell line E234.5 but did not survive.

7.4.2 Investigation of Germline Competence of Exchange Step Targeted ES Cell Lines Carrying pE2LacLIF

The 3 chimaeric males produced with the cell line SE2 2.4 were mated to BALB/c females to assess germline competence of this cell line. One of these mice, which was only thirty percent chimaeric as assessed by coat colour (Fig. 7.7 c, mouse 4) transmitted the ES cell genotype through the germline. Of eleven offspring resulting from this cross, all carried the ES cell derived coat colour markers (Fig. 7.7 d, e) indicating that germline transmission had occurred one hundred percent of the time. Fifty percent of the progeny of this male should be heterozygous for the LIF allele carrying the *lacZ* gene in the second exon. This is likely to cause a null mutation in the LIF gene. All mice derived from this male developed normally as has been seen previously for LIF⁺/LIF⁻ mice (Stewart *et al.*, 1992; Escary, *et al.*, 1993; Fig. 7.7 c, d). The line SE2 2.2 has also transmitted the ES cell genotype through the germline.

7.5 Discussion

It is now routinely possible to generate germline competent gene targeted ES cells (Hogan *et al.*, 1994; Brandon *et al.*, 1995a-c). When using a two step gene targeting approach, such as

Figure 7.7 Analysis of chimaeric mice produced with exchange step targeted ES cell line SE2 2.4

a Photograph of chimaeric mice produced with the ES cell line SE2 2.4. Mice were produced as described in section 9.5.1.

b Graph showing the percentage contribution of the chinchilla, agouti and black coat colours to chimaeric female and male mice produced with the ES cell line SE2 2.4. The numbers below each column represent individual mice.

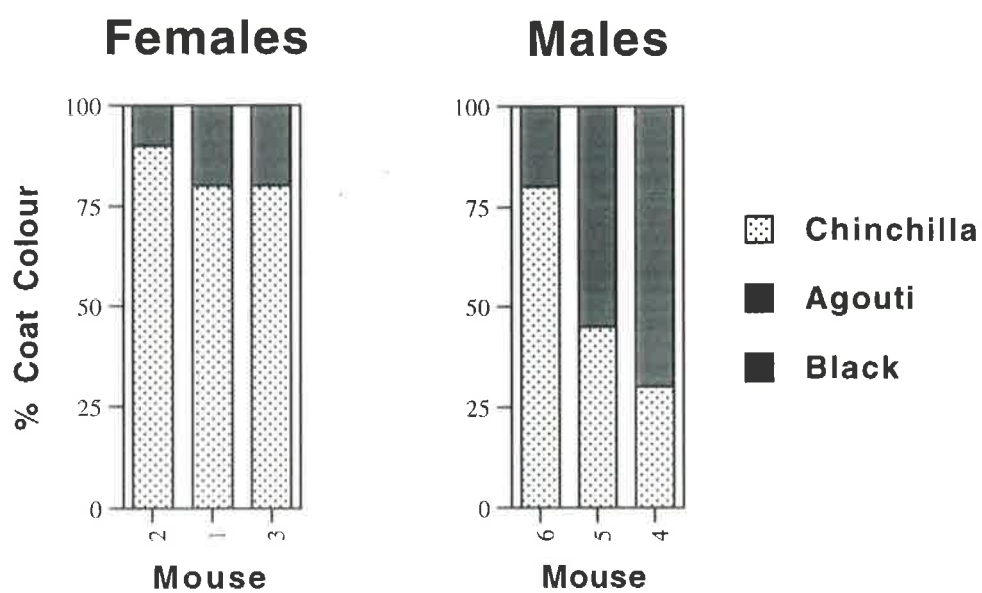
c Graph representing the percentage germline transmission for chimaeric male mice produced with ES cell line SE2 2.4 as assessed by breeding. The numbers below each column represent individual mice (as above).

a

Chimaeric Analysis: SE2 2.4



b



c

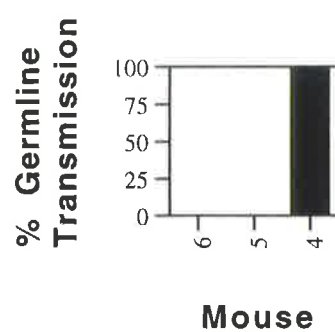


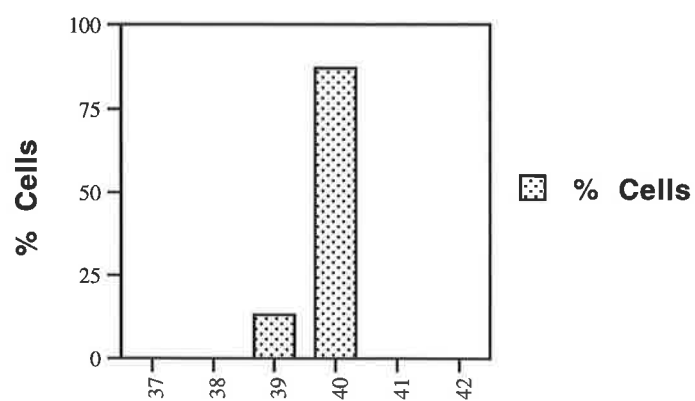
Figure 7.6 Karyotypic analysis of exchange step targeted ES cells

a-c Graphs showing the number of 2N chromosomes counted in examples of the tag step targeted cell lines **a** SE2 2.2, **b** SE2 2.4, **c** E2 34.5, expressed as percentages.

Chromosome spreads were prepared and stained by the method of Robertson, (1987) (section 9.4.16) and viewed under bright field optics on the Zeiss Axioplan Universal microscope.

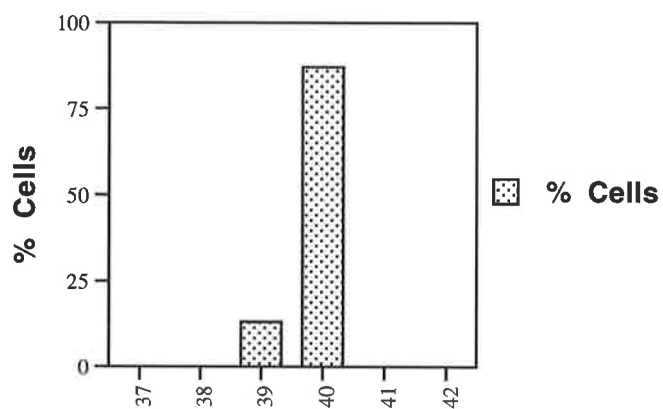
a

Karyotype of SE2 2.2



b

Karyotype of SE2 2.4



c

Karyotype of E2 34.5

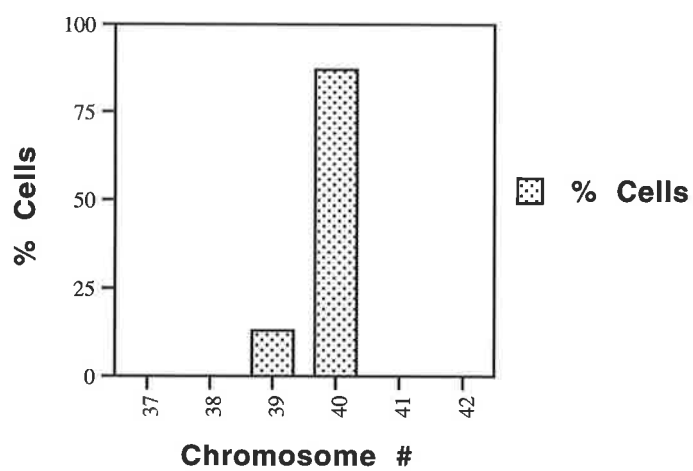


Figure 7.7 Analysis of chimaeric mice produced with exchange step targeted ES cell line SE2 2.4

d Photograph of neonates produced from the chimaeric male mouse # 4 (chinchilla 30 %/agouti 70 %) which carry the ES cell coat markers. Fifty percent of these mice should carry the *lacZ* marker gene in exon two of the LIF gene.

e Photograph of mice produced from the chimaeric male mouse # 4 (chinchilla 30 %/ agouti 70 %) which carry the ES cell coat markers. Fifty percent of these mice should carry the *lacZ* marker gene in exon two of the LIF gene (as above).

Mice were produced as described in section 9.5.1.

d



e



the stable tag-exchange approach, it is particularly important to ensure that the ES cells retain germline competence throughout each round of gene targeting. This prevents inadvertent use of ES cell lines which are unsuitable for the production of mice. Using the HPRT tag gene it is possible to assess germline competence of both the tag and exchange step targeted lines. This is not a possibility using tag gene selection systems based on the HSV-*tk* gene (section 5.9) since HSV-*tk* expression in the male gonads leads to sterility (Braun *et al.*, 1990).

Assessing the germline competence of the ES cells used throughout this project was particularly important as the germline transmission of the E14TG2a ES cells grown under the conditions used throughout this work had not been assessed in the Department of Biochemistry at Adelaide University, and suboptimal growth conditions may lead to karyotypic abnormalities which preclude germline transmission (Hogan *et al.*, 1994). The results of work presented in this chapter indicates that the growth conditions used in this work are able to support the growth of germline competent ES cells as has been shown by Mountford and co-workers (1994).

A normal euploid karyotype is essential for efficient colonisation of the germline by ES cells (Hogan *et al.*, 1994). The growth conditions and two rounds of gene targeting did not adversely affect the karyotype of the E14TG2a cell line. Chimaerae of up to sixty percent with E14TG2a, ninety percent with the tag step targeted line HG60 and ninety percent with exchange step targeted line SE2 2.4 were produced in this work, as assessed by coat colour. These data indicate that the E14TG2a cell line retained pluripotency after two rounds of gene targeting. Germline transmission of the unmanipulated E14TG2a, tag step targeted and exchange step targeted ES cells was assessed by breeding or GPI analysis. This indicated that the E14TG2a cell line retained germline competence throughout two rounds of gene targeting and under stringent growth conditions. Germline transmission of ES cells which have undergone two rounds of gene targeting has also been reported for J1 (Wu *et al.*, 1994), HM-1 (Stacey *et al.*, 1994) and CCE (Gondo *et al.*, 1994) lines, suggesting that two step targeting systems for the production of genetically altered mice will be applicable to a wide range of ES cell lines.

Eye abnormalities, sometimes resulting in blindness in one or both eyes, were seen in some of the chimaeric mice produced with untargeted E14TG2a, tag step and exchange step targeted ES cell lines. Such defects were not seen in non-chimaeric agouti siblings, indicating that the likely source of the abnormalities was the E14TG2a cell line. Similar defects have not been noted by workers in the laboratory of Dr. Austin Smith who culture the E14TG2a cell line under similar conditions (Smith, 1991; A. Smith, personal communication). This suggests that a mutation(s) in the E14TG2a ES cell line has arisen but probably not as a result of the culture conditions used in this work. The fact that the eye abnormalities were not transmitted to the heterozygous offspring generated with these cells, indicates that the gene or genes responsible for this phenomenon are likely to be recessive. Because the analysis of LIF expression is to be carried out in heterozygous mice produced with these cells, this defect was not considered a problem.

Germline transmission of an exchange step targeted ES cell line suggests that each of the exchange step targeted lines produced in this work will be suitable for the production of mice. The heterozygous mice produced will be used for expression mapping of each and all isoforms of LIF (section 3.7.2) which will develop normally due to the presence of one normal LIF allele (Stewart *et al.*, 1992; Escary *et al.*, 1993). Normal development of these mice has been seen (Fig. 7.7 d, e). The mice in which LIF expression is to be assessed will be normal with respect to the HPRT allele since heterozygous mice will be produced from matings between chimaeric males and wildtype females. The female offspring of this mating will be heterozygous for the mutant HPRT allele and hence normal (Hooper *et al.*, 1987) while male offspring will carry a maternal wildtype X chromosome and so will also be normal with respect to the HPRT allele.

The results presented in this chapter indicate that the stable tag-exchange targeting approach should be generally suitable for the production of germline competent ES cell lines carrying precise and/or subtle alterations. Coupled with the efficiency of the targeting system itself,

this indicates that the stable tag-exchange targeting approach will be a valuable tool for genetic analysis of the mouse.

Chapter 8

8 General Discussion

8.1

The underlying theme throughout early mammalian development is successive differentiation of pluripotent cells of the embryo which leads to a restriction of cell fate. The fate of a given cell in the mammalian organism is primarily determined by the extracellular signalling molecules with which the cell comes into contact (Heath and Smith, 1988). These signals are transduced to the nucleus of the responding cell where they lead to alterations in the particular array of genes expressed and consequently determine cell fate (sections 1.4.2-1.4.3). A precise understanding of the genes which affect developmental events, known as developmental control genes, is required if we are to reach a detailed understanding of mammalian embryogenesis. The establishment of a culture system for pluripotential stem cells of the mouse (section 1.6.6) and development of the technique of gene targeting in these cells (sections 1.6.8) has made the mouse an attractive model for the study of mammalian development. The *in vitro* culture of ES cells provides a system in which developmental control genes, which act on pluripotent stem cells, can be identified and analysed. In addition, gene targeting in ES cells provides a means to introduce predetermined genetic alterations into the mouse genome. This permits the analysis of developmental control genes in the context of the mammalian organism. This analysis is reliant on systems suitable for use in the ES cell system. The aim of this thesis was to produce systems suitable for the analysis of developmental control genes in ES cells both *in vitro* and *in vivo*.

Analysis of gain of function mutations in developmental control genes can provide useful information about the potential roles of such gene products. The ES cell system provides an experimentally tractable, model system for early mouse embryogenesis in which to assess pluripotent cell behaviour. Using this system, effects on ES cell proliferation, migration, differentiation and death can be assessed after exogenous application of factors which act at the extracellular level. Investigation of genes for which a source of the gene product is unavailable, or factors which act at the intracellular level, requires expression systems suitable for use in ES cells. Developmental control gene products are likely to exert effects at

very low thresholds and to cause different biological effects at various concentrations (section 1.7.2). As a consequence of this, the expression of such genes in ES cells requires specialised systems.

In this project, a system which permits precisely controllable heterologous gene expression in ES cells was characterised in stably transfected ES cell lines. The inductive signal, IFN, did not have any measurable effect on the growth rate of ES cells and did not compromise their pluripotency (section 2.2). The 6-16-based expression vectors, were able to direct IFN-inducible heterologous gene expression in transiently and stably transfected ES cells (section 2.2). Expression in the absence of induction was undetectable, the precise level of expression could be controlled by the amount of IFN used for induction and high levels of expression could be maintained for periods of up to two days using these vectors. These features make the 6-16-based expression system suitable for the over-/inappropriate expression of developmental control genes in ES cells *in vitro*.

Clonal cultures of ES cells stably transfected with a 6-16-based vector can be produced and the effects of gene expression on ES cells cultured *in vitro* could be compared in parallel induced or uninduced cultures. Effects on ES cell proliferation, differentiation, migration and death can then be assessed *in vitro*. Precise control over the level of expression directed by the 6-16-based vectors will facilitate analysis of factors which exert alternative effects in a concentration dependent manner.

Analysis of loss of function mutants is also widely used to gain an understanding of gene function in the mouse, and is preferable to the analysis of gain of function mutations since it provides information about the normal roles of the gene under investigation. Gene targeting in ES cells (section 1.6.8) provides the means to introduce loss of function mutations into the mouse genome, the consequences of which can then be assessed in ES cells or in mice generated from the modified ES cells. In the case of developmental control genes, *in vivo* loss of function mutations are often embryonic lethals, complicating or preventing analysis of the mutant phenotype. More sophisticated gene targeting approaches are therefore required

for the analysis of these genes. The introduction of subtle modifications into developmental control genes will be required to dissect the normal roles of such genes at the molecular level. The factors which control expression of the gene, the structural features of the gene product which enable it to mediate its function and the targets on which it acts could all be investigated by the repeated introduction of alterations into developmental control genes. The ability to create precise, subtle alterations in the mammalian genome will also be required for the generation of accurate mouse models for human disease, alteration of commercial products for pharmaceutical and therapeutic purposes and the generation of livestock with improved features.

A system based on the tag-and-exchange approach first described by Reid and co-workers (1990), which permits the systematic introduction of precise and/or subtle modifications into the mouse genome was described and characterised in work presented in this thesis. Recently published examples of the tag-and-exchange approach, have shown it to be inefficient and therefore unsuitable for the repeated introduction of alterations to a single site within the mouse genome. The modified approach developed in this project, termed the stable tag-exchange approach, overcomes one of the major inefficiencies of this system. In the stable tag-exchange approach, the site within the genome to be altered is tagged with a tag and stability cassette in an initial round of gene targeting. The tag gene provides direct selection for subsequent homologous recombination at this site, and is replaced by the alteration of interest in a second round of gene targeting. The stability cassette remains on the altered allele and prevents physical loss of the tag gene in the absence of homologous recombination during exchange step gene targeting (sections 3.3.1-2). The stable tag-exchange targeting approach was assessed at the LIF locus in ES cells.

In characterisation of the stable tag-exchange approach, a number of findings relating to gene targeting and the mechanism of homologous recombination were made.

- i. In the tag step gene targeting it was found that two selection cassettes, separated by almost five kilobases of homologous DNA, could be introduced into the mouse genome on a

single replacement vector when selection for only one of the cassettes was applied. The efficiency of tag step gene targeting at the LIF locus was 16.7 percent (section 4.3.3).

This finding led to the proposal of a mechanistic model for homologous recombination of replacement vectors into the ES cell genome (section 4.4). Novel features of the model include the involvement of the targeting vector ends in initiation of the recombination process and the role of extended regions of heterology near the termini of the vector in preventing passage of the recombination machinery (section 4.4). The model proposes that recombination is initiated by the vector ends and regions of non-homology near the termini of the vector prevent crossover events occurring in the region of homologous DNA between them by blocking the passage of the recombination machinery. If the proposed model is shown to be valid, greater flexibility may be applied to replacement vector design making it possible to introduce multiple regions of non-homology into the mouse genome on a single replacement vector.

For example, it may be possible to generate gain of function mutations by the introduction of multiple copies of a gene of interest into the endogenous locus. As opposed to non-targeting approaches to the production of gain of function mutations, direct modification of the endogenous locus should lead to overexpression of the gene under investigation in a spatially and temporally correct manner. This would be particularly useful for the study of developmental control genes which are likely to exert distinct roles at various stages of development and in various locations, and would also be applicable to the study of genetic diseases which are caused by overexpression of disease associated genes. It may also be possible to introduce multiple genes into a single site of the genome in an analogous manner. This would be particularly useful to the study of complex human diseases, the symptoms of which are caused by range of genes (Bronson and Smithies, 1994). It may be possible to produce mouse models for such diseases by the introduction of all the genes responsible for the symptoms of the disease into a particular site within the genome. Similarly it may be possible to introduce genes required in a particular synthetic pathway or developmental hierarchy into a particular site within the genome. Products of pharmaceutical or therapeutic

value may be produced and analysed, or developmental systems analysed in the context of the whole organism.

The efficiency of the exchange step of tag-and-exchange systems is compromised by a high background of lines which result from physical loss of the tag gene in the absence of homologous recombination (Askew *et al.*, 1993; Stacey *et al.*, 1994; Wu *et al.*, 1994; Gondo *et al.*, 1994). The stability cassette, used in the stable tag-exchange system, was included on the tag step vector in an effort to minimise this background.

ii. Comparison, at the LIF locus, of the unmodified tag-and-exchange targeting approach with that of the stable tag-exchange approach, achieved by using the selection protocol of either approach, indicated that the major cause of untargeted lines surviving in the unmodified approach was physical loss of the tag gene (sections 5.8.2, 5.8.4). Selection for the stability cassette during the exchange step of the stable tag-exchange approach, eliminated these lines (sections 5.8.2, 5.8.4). This increases the efficiency of the targeting approach, making it suitable for the repeated introduction of alterations into the mouse genome.

iii. The mechanism by which loss of the tag gene occurred during exchange step targeting was gene conversion, in which the tagged LIF locus was restored to wildtype by removal of the exogenous DNA (section 6.2.2). In no case were lines seen which had lost either the tag or stability cassette by gene conversion, implying that gene conversion is a 'large scale' process in which both the exogenous selection cassettes were removed from the genome in parallel.

In work presented in this thesis, the cell line HG60 was produced in tag step gene targeting and will be used for the efficient introduction of subtle and/or precise alterations into the LIF locus. As a result of the improved efficiency of the stable tag-exchange approach, it was possible to repeatedly introduce alterations into the LIF locus with ease (section 5.9) and as a result, lines were produced which will be used to map expression of each and all isoforms of LIF. The targeted ES cell lines retained germline competence (sections 7.3.2, 7.4.2) and can

therefore be used for the generation of mice. Line HG60 will be used in subsequent gene targeting for the production of mice carrying predetermined genetic alterations at the LIF locus while the exchange step targeted lines will be used to generate mice in which to map the cellular sites of each and all forms of LIF expression.

Using the exchange step targeted ES cell lines carrying a marked form of each or all LIF isoforms it has been possible to analyse the expression of the LIF isoforms separately in vitro in MBA differentiated derivatives of ES cells. A number of findings relating to the biology of LIF have been made. The expression of each LIF isoform is specifically regulated in vitro, implying that each has biologically significant roles. Various levels of LIF expression in a population of differentiated cells and concurrent expression in a single type suggests that LIF expression may not be uniform throughout the cells of a given tissue in vivo.

The systems described and characterised in this work fulfil the aims of this thesis. The 6-16-based expression system provides a means to direct precisely controllable heterologous gene expression in ES cells while the stable tag-exchange targeting approach provides an efficient means by which to introduce repeated non-selectable alterations to any site within the mouse genome. These systems will increase the flexibility and sophistication of ES cells as a system for the analysis of developmental biology.

8.2 Future Work

8.2.1 6-16-Based Expression Vectors

The 6-16-based expression vectors characterised in this work are able to direct IFN-inducible expression in ES cells and in a differentiated derivative known as X cells. An expression system with the characteristics of the 6-16 system but with a widespread host cell range would be of great value. The possibility that the 6-16-based expression vectors may be able to function in differentiated cell types therefore needs to be assessed. The ubiquitous nature of the IFN response in vivo (reviewed in Williams 1991) coupled with the known activity of the 6-16 promoter in a range of cell types suggests that 6-16-based expression vectors will be

suitable for use in many differentiated cell types. To test this, ES cells transfected with the expression vectors carrying reporter gene such as the *lacZ* gene (p1lac12X, section 5.3), could be differentiated by LIF withdrawal, chemical induction or growth in suspension cultures (section 1.6.5), prior to the assessment of β -galactosidase activity in parallel induced or uninduced cultures. Using this readily visualisable reporter gene, expression vector activity in differentiated cell types could be distinguished easily from that in residual stem cells which remain in these culture systems.

The ability of the 6-16-based expression vectors to direct precisely controlled gene expression in differentiated cells could be coupled with the ability to induce lineage specific differentiation of ES cells in vitro. With application of the appropriate factors or conditioned media to ES cells, it may be possible to generate model systems for developmental events such as haematopoiesis, neurogenesis, osteogenesis and organogenesis by lineage specific differentiation (section 1.6.5). Analysis of the genes controlling later developmental events in such model systems could then be achieved with use of the 6-16-based expression vectors.

It would also be of value to characterise further the expression directed by the 6-16-based expression vectors. The full length and truncated 6-16 promoters, 1 and 3 respectively (section 2.2), contained in the 6-16-based expression vectors carry a splice donor site present at the 3' end of the first untranslated exon of the 6-16 gene. The promoters also contain portions of the first intron of the 6-16 gene in which copies of the sequence ATG exist. It is not known if these factors affect expression directed by the vectors. Translation directed by the 6-16 vectors could be assessed further with the use of reporter genes, such as the β -galactosidase or luciferase gene, which like the CAT gene, permit expression levels to be quantified accurately. Comparison of expression at the RNA and protein level by RNase protection and assaying for enzyme activity could be carried out in parallel induced and uninduced cultures to determine the efficiency of translation directed by the 6-16-based expression vectors.

The attributes of the 6-16-based expression vectors make them suitable for use in the ES cell system. The ability of the vectors to direct (i) precisely controllable levels of heterologous gene expression in (ii) stably transfected ES cells (iii) in the absence of detectable basal expression make them ideal for the analysis of developmental control genes in the ES cell system. The vectors may also be suitable for use in a wide range of mammalian cell types making them generally useful for genetic analysis in mammalian cells in vitro.

8.2.2 *Stable Tag-Exchange Gene Targeting*

The model proposed to describe homologous integration of replacement vectors (section 4.4), based on the results of targeting with the tag step vector pHNLX, is speculative and should be further tested. The hypothesis that two selection cassettes separated by a large amount of homologous DNA, are both integrated into the ES cell genome with high frequency when selection for only one is applied could be tested by screening additional lines transfected with tag step targeting vectors. The G418^r lines, transfected with pHNLX in work presented in this thesis, could be screened by Southern analysis to determine if correctly targeted G418^r lines carry both the HPRT and *neo* genes. In addition, tag step targeting at an alternative site within the genome with a vector similar in design to pHNLX should be carried out to ensure that the findings relating to the integration of replacement vectors are not locus specific.

If the mechanisms proposed in the model are correct they can be exploited for the future design of targeting vectors. To achieve this, the parameters governing the integration of tag step vectors must be further defined. (i) The amount of heterologous DNA that is required to “block” passage of the mammalian recombination machinery, (ii) the maximum amount of DNA positioned between two regions of non-homology that can be introduced into the genome and (iii) the number of separate regions of non-homology that can be introduced into the genome should all be investigated. Targeting experiments could be carried out at a single site within the genome, using a series of vectors designed to test each of these parameters.

A tag step targeted, germline competent ES cell line has been produced in work presented in this thesis which will be suitable for the efficient introduction of repeated alterations to the

LIF locus in subsequent rounds of gene targeting. Alterations to the control elements of the LIF gene or to the coding sequence such as to regions encoding structural motifs, functional domains and even single amino acids can be envisaged. Phenotypic analysis of such alterations could then be carried out in homozygous mice produced with these ES cells to determine the effect of such alterations on the normal functions of LIF.

Exchange step targeted lines produced in this work, which carry a reporter gene in a single allele of each or all forms of the LIF gene, will be used for the *in vitro* assessment of LIF expression and for the production of mice in which the cellular sites of LIF expression can be visualised. *In vitro*, these cell lines will be used to determine the cellular sites of LIF expression in ES cells differentiated into embryoid bodies (Robertson, 1987). Embryoid body formation provides a model system in which processes such as neurogenesis and haematopoiesis can be studied (Evans and Kaufman, 1981; Martin, 1981; Doetschman *et al.*, 1985; Robertson, 1987; Shen and Leder, 1992). Since LIF has been implicated in both these processes (reviewed in Hilton, 1992; Escary *et al.*, 1993), it will therefore be of interest to determine the expression pattern of the various forms of LIF during these developmental processes.

In vivo expression of each LIF isoform will be assessed throughout development from the earliest stages of embryonic development to post-natal stages in mice which carry a single marked allele of LIF. This should give insight into the normal roles of each form of LIF during embryogenesis, which has not been possible from the analysis of LIF null mutant mice (Stewart *et al.*, 1992; Escary *et al.*, 1993). Expression of LIF will also be analysed in the adult mouse. Many tissues of the adult are known to express LIF D and/or LIF M (Robertson *et al.*, 1994). The highest levels of expression have been detected in self-renewing tissues of the mouse, leading to speculation that LIF expression in a subpopulation of differentiated cells maintains a "stem cell niche" in these tissues (Robertson *et al.*, 1993). Identification of the cellular sites of expression of each and all forms of LIF in tissues of the adult should indicate whether this hypothesis is correct and lead to a detailed understanding of the normal *in vivo* roles of each isoform of this protein.

Interruption of the coding region of LIF D and M with the *lacZ* cassette in the exchange step targeted lines produced in this work should result in null alleles of these isoforms. By breeding mice carrying a null allele of LIF D or M to homozygosity, mice lacking LIF D or M expression can be produced. It may also be possible to produce mice lacking expression of LIF T in the same way, although translation of this form of LIF commences downstream of the *lacZ* cassette introduced in exchange step gene targeting and it is therefore possible that active LIF T protein would still be produced from the altered LIF T allele. The phenotypic consequences of null mutations of the LIF isoforms could then be assessed in isolation throughout development and in the adult. This would be particularly interesting in the case of LIF T for which definitive roles have not yet been assigned. However, such analysis may be inconclusive due to compensation by the other LIF isoforms or other gp130 ligands.

Teratocarcinomas are thought to arise from pluripotent stem cells or primordial germ cells which have escaped normal regulatory signals. The ability of LIF to maintain pluripotent stem cells of the mouse (Smith *et al.*, 1988; Williams *et al.*, 1988) and to act as a mitogen for primordial germ cells (Matsui, 1992) in vitro implicates LIF in teratocarcinogenesis. Deregulated LIF expression has been proposed by two independent groups to underlie teratocarcinogenesis (Kola *et al.*, 1990; Rathjen *et al.*, 1990b). Teratocarcinogenesis is tractable to experimental analysis since teratocarcinoma formation can be induced experimentally by the introduction of ES cells to ectopic sites in the mouse (Hogan *et al.*, 1994). The involvement of LIF in teratocarcinogenesis could be assessed by comparison of teratocarcinoma formation induced by unmanipulated ES cells, ES cells which lack expression of all isoforms of LIF (A. Smith, unpublished) or ES cells which lack a single isoform of LIF. A reduced propensity to form teratocarcinomas by ES cells lacking one or all isoforms of LIF would provide direct evidence that LIF can be involved in the process of teratocarcinogenesis.

The features of the stable tag-exchange gene targeting approach, demonstrated at the LIF locus, indicate that the system will be generally useful for the precise modification of the

mouse genome. The efficiency of the system and the ability to use it to introduce precise non-selectable alterations into the mouse genome make the system ideal for the repeated introduction of non-selectable subtle changes into the mouse genome. This will facilitate rigorous genetic analysis of the mouse and will be particularly useful to the study of mammalian embryogenesis.

Chapter 9

9 Materials and Methods

9.1 Abbreviations

A	amps
Ac	acetate
ATP	adenosine triphosphate
BES	N,N-bis(2-hydroxyethyl)-2-aminoethane-sulphonic acid
bisacrylamide	N'N'-methylene-bisacrylamide
bp	base pair
CAT	chloramphenicol acetyl transferase
CIP	calf intestinal phosphatase
c.p.m.	counts per minute
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
DMEM	Dulbecco's modified Eagle medium
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
d.p.c.	days post coitum
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylene diamine tetra acetic acid
EGTA	ethylene glycol-bis(b-aminoethyl ether)-N,N,N',N'-tetra acetic acid
FCS	foetal calf serum
GLB	gel loading buffer
G418	G418 sulphate
G6PDH	glucose-6-phosphate dehydrogenase

HAT	hypoxanthine amethopterin thymidine
HEPES	N-2-hydroxyethyl piperazine-N-ethane sulphonic acid
IFN	interferon
IPTG	isopropyl- β -D-thiogalactopyranoside
I.U.	international units
kb	kilobase pairs
LIF.....	leukaemia inhibitory factor
LB.....	luria broth
MBA	3-methoxybenzamide
min.	minute
MOPS	3-([N-morpholino]propane sulphonic acid)
MQ	Milli-Q filtered
MTT.....	tetrazolium
NADP.....	nicotinamide adenine dinucleotide phosphate
O.D.....	optical density
O/N.....	overnight
ONPG.....	o-nitrophenyl- β -D-galactosidase
PBS.....	phosphate buffered saline
PCR.....	polymerase chain reaction
PEG	polyethylene glycol
PMS.....	phenazine methosulphate
RNA.....	ribonucleic acid
RNAse	ribonuclease
RNAsin.....	ribonuclease inhibitor
r.p.m.	revolutions per minute
RT	room temperature
SDS	sodium dodecyl sulphate
sec.....	second
TEMED	N, N, N', N'-tetramethyl-ethenediamine
TLC	thin layer chromatography

6-TG.....	6-thioguanine
u.....	unit
UV.....	ultra violet
v.....	volume
V.....	volts
w.....	weight
X-gal	5-bromo-4-chloro-3-indole- β -D-galactopyranoside

9.2 Materials

9.2.1 Chemicals and Reagents

All chemicals and reagents were of analytical grade and were obtained from BDH Chemicals Ltd. with the exception of the following which were obtained from the following suppliers:

acetyl Co-A.....	Sigma Chemical Co.
agarose.....	Sigma Chemical Co.
amethopterin	Sigma Chemical Co.
ATP	Boehringer Mannheim Biochemicals
bacto-agar	Difco Labs Ltd.
bacto-tryptone.....	Difco Labs Ltd.
BES.....	Calbiochem
bromophenol blue	Sigma Chemical Co.
C ¹⁴ -threo-chloramphenicol	DuPont
DMEM	Gibco BRL
DTT	Boehringer Mannheim Biochemicals
EGTA.....	Sigma Chemical Co.
ESGRO.....	Amrad
ethidium bromide	Sigma Chemical Co.
FCS	Commonwealth Serum Laboratories
fructose-6-phosphate	Sigma Chemical Co.
G418	Gibco BRL

gelatin.....	Sigma Chemical Co.
gentamycin	Schering Corporation
G6PDH.....	Sigma Chemical Co.
hypoxanthine.....	Sigma Chemical Co.
L-glutamine	Sigma Chemical Co.
Hams F12.....	Gibco BRL
IPTG	Boehringer Mannheim Biochemicals
magnesium acetate.....	Sigma Chemical Co.
β -mercaptoethanol	Sigma Chemical Co.
mouse α & β fibroblast IFN	Sigma Chemical Co.
MTT.....	Sigma Chemical Co.
NADP.....	Sigma Chemical Co.
PMS.....	Sigma Chemical Co.
retinoic acid.....	Sigma Chemical Co.
sephadex G50	Pharmacia LKB Biotechnology Inc.
sepharose CL-6B	Pharmacia LKB Biotechnology Inc.
6-TG.....	Sigma Chemical Co.
thymidine	Sigma Chemical Co.
X-gal	Scimar
yeast extract	Difco Labs. Ltd.

9.2.2 *Enzymes*

Enzymes were obtained from the following suppliers :

AMV reverse transcriptase.....	Molecular Genetic Resources
CIP	Boehringer Mannheim Biochemicals
CAT.....	P-L Biochemicals
<i>E. coli</i> DNA polymerase KF-1.....	Bresatec Ltd.
lysozyme.....	Sigma Chemical Co.
proteinase K.....	E. Merck

restriction endonucleases	Pharmacia LKB Biotechnology Inc. and New England Biolabs
RNAse A.....	Sigma Chemical Co.
RNAsin.....	Bresatec Ltd.
Taq polymerase	Bresatec Ltd.
T4 DNA ligase	Bresatec Ltd.
T7 DNA polymerase	Pharmacia LKB Biotechnology Inc.

9.2.3 *Nucleic Acid Molecular Weight Standards*

Hpa I digested pUC19 size markers were obtained from Bresatec Ltd.. Band sizes: 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34, 26 bp.

Eco RI digested SPP-1 size markers were obtained from Bresatec Ltd.. Band sizes: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36 kb.

Eco RI/Hind III lambda DNA size markers were prepared by digestion of lamda phage DNA (New England Biolabs). Band sizes: 21.23, 5.15, 4.97, 4.27, 3.53, 2.03, 1.90, 1.58, 1.38, 0.97, 0.83, 0.56, 0.13 kb.

9.2.4 *Radionucleotides*

[alpha-³²P] dATP (3000 Ci/mmol), [alpha-³⁵S] dATP (1000-1500 Ci/mmol) were supplied by Bresatec Ltd.

9.2.5 *Kits*

The GeneClean II™	Bresatec Ltd.
Gigaprime kit.....	Bresatec Ltd.
T ⁷ Sequenase kit.....	Pharmacia LKB Biotechnology Inc.
Super Z Applicator kit	Helena Laboratories

9.2.6 *Miscellaneous Materials*

bottle top filter (0.22 µM)	Corning

3MM chromatography paper.....	Whatman Inc.
cellulosenitrate membrane (0.45 μ M).....	Schleicher and Schuell
chambered slides.....	Nunc Inc.
freezing vials.....	Nunc Inc.
Hybond-N ⁺ nylon membrane (0.45 μ M).....	Amersham Corporation
medical grade X-ray film.....	Eastman Kodak
Nytran nylon membrane (0.45 μ M)	Schleicher and Schuell
Phosphorimager screens.....	Molecular Dynamics
Silica gel 60 F254 TLC plates.....	E Merck
tissue culture grade plates.....	Falcon
tissue culture grade flasks.....	Falcon
tissue culture grade pipettes.....	Falcon
tissue culture grade 4-, 6-, 24-well plates	Falcon

9.2.7 Plasmid Vectors

Cloning and expression vectors were obtained from the following:

pBluescript.....	Pharmacia LKB Biotechnology Inc.
pCH110	Pharmacia LKB Biotechnology Inc.
pT7T319U.....	Stratagene Inc.
pUC18.....	New England Biolabs

Recombinant DNA vectors were obtained from the following:

p1OT (Whyatt <i>et al.</i> , 1993).....	Dr. Peter Rathjen
p3OT (Whyatt <i>et al.</i> , 1993).....	Dr. Peter Rathjen
pBLCAT3 (Luckow and Shutz, 1987)	Dr. Peter Rathjen
pnI2(I1s) (Reid <i>et al.</i> , 1990)	Prof. John. Heath
pnlacF (Jacques Peschon).....	Dr. Peter Rathjen
pOT (Whyatt <i>et al.</i> , 1993).....	Dr. Peter Rathjen
pPGKneob (Dylan Edwards)	Dr. Peter Rathjen
pgkNEOpA (Bruce Mann).....	Dr.. Varaporn Thonglairoam

pXMT2 (Whyatt *et al.*, 1993)..... Dr. Peter Rathjen

Dr. Varaporn Thonglairoam was formerly, and Dr. Peter Rathjen is currently at the Department of Biochemistry, University of Adelaide, Adelaide, Australia and Prof. John Heath is at the Department of Biochemistry, University of Birmingham, Birmingham United Kingdom.

9.2.8 Oligonucleotides

DNA oligonucleotides were synthesized by Bresatec using an Applied Biosystems DNA Synthesizer (Model 380B). All primer sequences are 5'-3'.

Sequencing Primers:

T7: TAATACGACTCACTATAGGGAGA

T3: ATTAACCCTCACTAAAGGGA

USP (M13 -20): GTAAAACGACGGCCAGT

USP (M13 -40): GTTTTCCCAGTCACGAC

RSP (M13 -24): AACAGCTATGACCATG

RSP (M13 -48): AGCGGATAACAATTCACACAGGA

E2SP: TTGAGCTGTGCCAGTT

Primers for alteration of pnlacF:

lacATG5': AGATCTACCATG

lac5': AGATCTGGGCCCAA

lac3': GACGGGATCCCCCACGCTC

Primers for alteration of the LIF exons:

Bcl linker: CTGATCAG

1D5': ACTGGGGATCCCGCGT

1D3': AGGCTGCCGGCGCGCCCGCAGCCGGGACTCTTAAATCGCGGCCCGGC

GCATGGATTAGATCTGGCCAA

1M5': AGGCGGCTGAGCA

1M3': CGGGGGTACCCAGATCTGCATCT

1T5': GCGCCCCGGGGGAGGGGTGATCCTCGACAGGCCCGGACAGCTCCCGG

CTCTCCTGACACCTTTAGATCTCCTCT

1T3': TGTTCCTGCAGGT

E25': TTCCGGGGCCCTA

E23': AAGAGGGCTCCCTGCCCCGTGTTTCCAGTGCAGAACCAGCAGCAGAGATCT

CACAAT

Primers for RT-PCR Detection of LIF-D Expression:

1Dlac5': ATGAAGGTCTTGCCCGCAG

E2lac3': TTGACCGTAATGGGATAG

9.2.9 Water and Sterilisation

Water was filtered using the Milli-Q[®] Reagent Water System and for use in tissue culture solutions, the final filtration was carried out using the Millipak 40 filter (0.22 μ M pore size). Water, solutions and buffers were autoclaved (20 psi at 140°C, 25 min.) prior to use. Tissue culture solutions (section 9.4) were filtered in vacuo using Corning disposable bottle top filters (0.22 μ M pore size).

9.2.10 Bacterial Strains

The *E. coli* bacterial strains DH5alpha and GM2929, obtained from Dr. Justin Dibbens, were stored as glycerol stocks. *E. coli* bacterial strain DH5alpha was used as the host for all recombinant plasmids, with the exception of plasmids which were to be digested with dam methylation sensitive restriction enzymes, in which case the strain GM2929 was used.

DH5alpha:*supE44 delta lac U169 (phi80 lacZdelta M15) hsdR17 recA1 endA1 gyrrA96 thi-1 relA1*

GM2929:*dam-13::Tn9 dcm-6 hsdR2 recF143 McrA⁻ McrB⁻ galK2 galT22 ara-14 lacY1 xyl-5 thi-1 tonA31 rpsL136 hisG4 tsx-78 mtl-1 glnV44 leuB6 rfbD1*

Dr. Justin Dibbens is currently at the Hanson Centre for Cancer Research, Institute of Medical and Veterinary Science, Adelaide, Australia.

9.2.11 Bacterial Growth Media

FTB:

Bacto yeast extract..... 2.4% (w/v)

Bacto-tryptone..... 1.2% (w/v)

Glycerol..... 0.4% (v/v)

KH₂PO₄.....17 mM

K₂HPO₄.....72 mM

FTB broth was autoclaved prior to the addition of KH₂PO₄ and K₂HPO₄.

LB:

NaCl 1% (w/v)

bacto-tryptone 1% (w/v)

yeast extract..... 0.5% (w/v)

The pH was adjusted to 7.0 with NaOH.

Psi broth:

Bacto yeast extract..... 0.5% (w/v)

Bacto tryptone..... 2% (w/v)

MgSO₄..... 0.5% (w/v)

The pH was adjusted to 7.6 with KOH.

L plates:

NaCl 0.5% (w/v)

bacto-tryptone 1% (w/v)

yeast extract	0.5% (w/v)
bacto-agar	1.2% (w/v)

Ampicillin and chloramphenicol resistant bacterial selection was achieved using 100 µg/ml ampicillin and 30 µg/ml chloramphenicol, respectively.

9.3 Recombinant DNA Techniques

9.3.1 Restriction Endonuclease Digestions

Plasmid DNA suspended in water was digested with 2 units of restriction enzyme per µg of DNA in a 10µl reaction (if DNA fragments were to be isolated for cloning DNA was overdigested with 10 u enzyme/µg DNA in a 20µl reaction) at 37°C for one hour in the presence of 0.1 v of the appropriate New England Biolabs buffer.

9.3.2 Endfill Reaction

Up to 5 µg of DNA with 5' overhangs, to be used in blunt end ligations, was endfilled in the following reaction allowed to proceed at 37°C for 1 hour.

dATP, dCTP, dGTP, dTTP	0.5	mM
10 x Endfill buffer	0.1	v
DTT	10	mM
DNA polymerase (KF-1).	10	u

10 x Endfill Buffer:

Tris-HCl (pH 7.5)	100	mM
DNA	500	ng
MgCl ₂	100	mM
NaCl	100	mM

9.3.3 *Endchew Reaction*

Up to 5 µg of DNA with 3' overhangs to be used in blunt ligations, had 3' overhangs removed by exonuclease digestion in the following reaction, allowed to proceed at 37°C for 15 minutes.

10 x T4 Polymerase Buffer	0.1	v
dATP, dCTP, dGTP, dTTP	0.5	mM
T4 DNA polymerase	5	u

10 x T4 DNA polymerase buffer:

Tris-HCl (pH 7.4)	700	mM
MgCl ₂	100	mM
DTT	50	mM

T4 DNA polymerase was inactivated by incubation at 70°C for 10 minutes prior to use of the DNA in ligation reactions.

9.3.4 *CIP Reaction*

To restriction endonuclease digestions CIP (1 u), 10 x CIP buffer (0.5 µl, Boehringer Mannheim Biochemicals) and MQ water to a final volume of 25 µl were added and incubated at 37°C for 30 minutes.

9.3.5 *Agarose Gel Electrophoresis*

Horizontal agarose gels (0.8, 1.0, 2.0, 2.5, 2.8 % TBE or TAE agarose) submerged in TBE or TAE were used. DNA samples, containing 1 x GLB, were electrophoresed at 100 mA. Nucleic acid was visualised by staining gels with ethidium bromide (5 µg/ml) and photographed under short wavelength UV light, unless it was a preparative gel, in which case long wavelength UV light was used. Appropriate bands were isolated from preparative gels using sterile scalpel blades.

10 x GLB:

glycerol.....50% (v/v)
SDS.....0.1% (w/v)
bromophenol blue.....0.05% (w/v)
xylene cyanol.....0.05% (w/v)

TAE:

Tris-acetate 40 mM
Sodium acetate 20 mM
Na₂EDTA 1 mM

The pH of TAE was adjusted to 8.2 with acetic acid.

TBE:

Tris-borate.....100 mM
Na₂EDTA 2.5 mM

The pH of TBE was 8.3.

9.3.6 *Isolation of DNA from Agarose Gels*

Appropriate DNA bands were excised from agarose gels using sterile scalpel blades and the DNA eluted with the GeneClean II™ kit using the protocol of the supplier.

9.3.7 *Ligation Reactions*

Vector (50 ng), insert DNA (3 fold molar excess), 1 x ligation buffer, rATP (1 mM), DTT (1 mM) and T4 DNA ligase (1 u, cohesive end ligation) or T4 DNA ligase (10 u, blunt end ligation) were mixed in a volume of 20 µl. Reactions were allowed to proceed at 20°C for 2 to 4 hours.

10 x Ligation Buffer:

Tris-HCl (pH 7.5).....500 mM
MgCl₂.....100 mM

9.3.8 Preparation of Competent Bacterial Cells

E. coli cells were streaked out on L plates and grown at 37°C overnight. A single colony was used to inoculate Psi broth (5 ml, grown with shaking, 37°C, O/N). A 1:30 subculture was made in Psi broth (15 ml, grown for 90 min., with shaking, 37°C or until O.D.₆₀₀ 0.6 was obtained). A 1:20 subculture was made in Psi broth (100 ml, grown for 90 min., with shaking, 37°C or until O.D. 0.5-0.6 was obtained). The bacterial cells were chilled on ice for five minutes and pelleted by centrifugation (6000 r.p.m., 4°C, 5 min., SS34 rotor in Sorvall RC-5 centrifuge). After resuspension in Tfb 1 (0.4 v) the cells were chilled on ice for a further 5 minutes. The cells were pelleted by centrifugation (as above) and resuspended in Tfb 2 (0.04 v). After incubation on ice for 15 minutes aliquots of cell suspension (100 µl) were transferred to Eppendorf tubes and snap frozen in a dry ice/ethanol bath. The cells were stored at -80°C until required.

Tfb 1:

KAc	30	mM
RbCl.....	100	mM
CaCl ₂	10	mM
Glycerol	15%	(v/v)

The pH was adjusted to 5.8 with acetic acid.

Tfb 2:

MOPS.....	10	mM
CaCl ₂	75	mM
RbCl.....	10	mM
glycerol.....	15%	(v/v)

The pH was adjusted to 6.5 with KOH.

9.3.9 Transformation of Competent Bacterial Cells

Approximately 10 ng of circular, plasmid DNA or one half volume of ligation reactions was mixed with competent cells (100 µl) and left on ice for 30 minutes. The cells were heat

shocked at 42°C for 2 minutes and incubated in LB (1 ml, with shaking, 37°C 60 min.). The cells were pelleted by centrifugation (14000 r.p.m., 30 sec., in an Eppendorf 5415 microfuge), resuspended in MQ water (50 µl) and spread on an L plate supplemented with ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml) where required. Transformant colonies were grown at 37°C overnight. Where blue/white colour selection was required plates were spread with IPTG (50 µl, 50 µg/ml) and X-gal (20 µl, 50 µg/ml) before transformant cells were plated.

9.3.10 Bacterial Colony Lift Procedure

A circular nitrocellulose filter (Cellulosenitrate, Schleicher and Schuell) was placed briefly onto the surface of the bacterial plate and was marked in order to orient the filter. The filter was then placed on 3MM filter paper (Whatman Inc.) soaked in the following solutions.

10% SDS	3	mins.
NaCl 1.5 M, NaOH 0.5 M	5	mins.
NaCl 1.5 M, Tris-HCl (pH 7.5)...	5	mins.
2 x SSC.....	5	mins.

The nucleic acid was immobilized on the filter by UV crosslinking (120 J, Stratagene Inc., UV Stratalinker™ 1800) and the bacterial debris washed off in 2 x SSC.

SSC:

NaCl	150	mM
sodium citrate.....	15	mM

9.3.11 *alpha*-³²P]-labelled DNA Probes

Single stranded DNA probes were prepared using the Gigaprime labelling kit and the protocol of the supplier with 10-60 µCi [*alpha*-³²P] dATP. The reaction was made up to 100 µl with MQ water and unincorporated label removed from the probe by centrifugation on a Sephadex G-50 column in a 1 ml syringe cylinder (3200 r.p.m., 4 min.).

9.3.12 Small Scale Isolation of Plasmid DNA

Small scale isolation of plasmid DNA was carried out by the following method based on the rapid "miniprep" procedure (Holmes and Quigley, 1981). LB (2ml, 100 µg/ml ampicillin), was inoculated with a single transformant colony and incubated at 37°C with shaking overnight. The cells were pelleted by centrifugation (14000 r.p.m., 15 seconds in an Eppendorf 5415 centrifuge), and resuspended in STET buffer (400 µl). To the suspension lysozyme (200 µg) was added, and the cells lysed by boiling for 45 seconds. The cell debris was pelleted by centrifugation (14000 r.p.m., 1 min., as above). The supernatant was retained and phenol/chloroform extracted. The nucleic acid was precipitated with an equal volume of iso-propanol (-20°C, 10 mins.), pelleted by centrifugation (14000 r.p.m., 15 mins., as above), and air-dried before resuspension in MQ water (30 µl). A 1/6 volume of the nucleic acid was analysed by agarose gel electrophoresis.

STET :

Sucrose.....	8%	(w/v)
Triton X-100.....	5%	(v/v)
EDTA.....	50	mM
Tris-HCl (pH 8).....	50	mM

9.3.13 Polymerase Chain Reaction Technique

Polymerase chain reactions were set up in capillary tubes (in a reaction volume of 20 µl) as follows. Circular plasmid template DNA (10 µl of 1ng/µl) or cDNA (5 µl of a 2/5 dilution of cDNA, section 9.3.25) was mixed in the capillary tube with the appropriate primers (1 µl each of 10 ng/µl), PCR buffer, each dNTP (0.2 mM) and Taq Polymerase (0.5 u) . Oligonucleotide primers over 50 bp in size were used after purification on 15% denaturing acrylamide gels.

10 x PCR buffer:

Tris-HCl (pH 8.3).....	100	mM
KCl.....	500	mM

MgCl₂.....25 mM
 gelatin 0.01% (w/v)

Reactions were carried out in an FTS-1 Thermal Cyclor (Corbett Research). A typical thermal profile for PCR was:

(1).....94°C, 2 min.; 50°C, 5 sec.; 72°C, 1 min.
 (2-39) 94°C, 5 sec.; 50°C, 15 sec.; 72°C, 45 sec.
 (40).....72°C, 2 min.; 25°C.

PCR products were purified on non-denaturing 15% polyacrylamide gels or were visualised on 2.8 % TAE gels.

9.3.14 *Non-denaturing Polyacrylamide Gel Electrophoresis*

15% polyacrylamide gels were prepared from a 40% acrylamide stock (38:2 acrylamide:bis-acrylamide). To 20 x TBE (2.5 ml), acrylamide stock (18.75 ml) and MQ water (28.25 ml) were added. 10% ammonium persulphate (500 µl) and TEMED (60 µl) was added just prior to pouring between glass plates of the appropriate size separated by 1 mm spacers. Gels were allowed to set for 30-60 minutes and pre-electrophoresed at 25 mA in 1 x TBE for 15 minutes. DNA samples containing 1x GLB (section 9.3.5) were electrophoresed at 25 mA.

9.3.15 *Denaturing Polyacrylamide Gel Electrophoresis*

6% and 12% polyacrylamide denaturing gels were prepared from a 40% acrylamide stock solution. Urea (21 g) was dissolved in MQ water and 20 x TBE (2.5 ml) and an appropriate volume of acrylamide stock was added in a final volume of 50 ml. 10% ammonium persulphate (500 µl) and TEMED (60 µl) was added just prior to pouring between glass plates (20 X 40 cm) separated by 0.4 mm spacers. The gels were allowed to set for 30 minutes and pre-electrophoresed for 30 minutes at 1400 V. DNA in 1 x GLB (section 9.3.5) was electrophoresed at 1400 V.

9.3.16 Isolation of DNA Fragments from Polyacrylamide Gels

After electrophoresis, the glass plates were separated and gels stained with ethidium bromide (5 µg/ml), destained for 10 min and the DNA bands excised under longwave UV light using sterile scalpel blades. DNA was eluted in gel elution buffer (400 µl, 37°C, O/N), and ethanol precipitated in the presence glycogen (40 µg).

Gel Elution Buffer:

SDS 0.1% (v/v)

NH₄Ac 0.5 M

MgAc 10 mM

9.3.17 Denaturation of Double Stranded Plasmid DNA for Sequencing

Double stranded plasmid DNA (4 µg in 19 µl) was treated with RNase A (1 µl, 10 mg/ml, 37°C, 15 min). The DNA was denatured by incubation in the presence of denaturing solution (5 µl, 1 M NaOH, 1mM EDTA, 37°C, 15 min). Denatured plasmid DNA was purified by centrifugation (1800 r.p.m., 3 mins.) on a sepharose CL-6B column (in a 0.5 ml Eppendorf tube).

9.3.18 Sequencing of Double Stranded Plasmid DNA

Dideoxy sequencing of denatured DNA was carried out using the Pharmacia LKB Biotechnology Inc. T₇ polymerase sequencing kit using the protocol of the supplier. The reaction products were separated on 6% denaturing polyacrylamide gels (section 9.3.15). The gels were transferred to dry 3MM filter paper (Whatman Inc.) before drying at 65°C in vacuo. Radioactivity was detected by exposure to X-ray film at -80°C with intensifying screens, or by exposure to phosphorimager screens.

9.3.19 Large Scale Isolation of Pure Plasmid DNA

Ultra pure plasmid DNA was prepared by the following method for DNA manipulation and transfection of eukaryotic cells. FTB (500 ml, 100 µg/ml ampicillin) was inoculated with a

single transformant colony and incubated overnight with shaking at 37°C. The cells were pelleted by centrifugation (7000 r.p.m., 4°C, 5 mins., GSA rotor in a Beckman J-21 centrifuge), and the pellet resuspended in GTE (13 ml). To each half of the resulting suspension, lysis solution (13 ml, 0.2 M NaOH/1% SDS), was added and swirled on ice for 5 minutes. The lysis suspension was neutralized by the addition of acetate solution (6.5 ml, 3M KAc/2M acetic acid pH 5.8). After a further ten minutes on ice, the cell debris was pelleted by centrifugation (12000 r.p.m., 4°C, 10 mins., SS34 rotor in a Sorvall RC5 centrifuge). The supernatant was retained and nucleic acid precipitated by the addition of cold iso-propanol (0.6 v). After mixing, the nucleic acid was pelleted by centrifugation (8000 r.p.m., 4°C, 5 mins., SS34 rotor as above). The pellet was drained well and resuspended in TE (7 ml), CsCl (7 g) and ethidium bromide (700 µl, 10 mg/ml). Debris was removed by centrifugation (3000 r.p.m., 10 mins. in a Heraeus Christ labofuge). The supernatant was transferred to an unused Oakridge tube (polycarbonate, 10 ml) and the tubes balanced with paraffin. A caesium chloride density gradient was formed by centrifugation (45000 r.p.m., 24 hours, 20°C, Ti50 rotor in a Beckman L5-50 ultracentrifuge). Plasmid DNA (1 ml) was harvested, under long wavelength UV light into an equal volume of iso-propanol (equilibrated against NaCl saturated TE). The DNA was extracted with equal volumes of iso-propanol until the ethidium bromide was removed. The DNA was diluted 1 in 4 with MQ water and ethanol precipitated (0.1 v NaAc, 2.5 v C₂H₅OH). The DNA was finally resuspended in MQ water (100 µl). The concentration of plasmid DNA was determined by spectrophotometric analysis (O.D.₂₆₀ 1=50 µg/ml).

GTE:

Glucose.....	50	mM
Tris-HCl (pH 7.6).....	25	mM
EDTA.....	10	mM

9.3.20 *Small Scale Preparation of High Molecular Weight DNA from Mammalian Cells*

To ES cells grown in 24-well plates, lysis solution (300 µl) was added and incubated at 37°C overnight. The cell lysate was transferred to an Eppendorftube and the nucleic acid drawn out

of solution with an equal volume of isopropanol by agitation (the preparation was stored at -20°C until required). The nucleic acid was transferred into 70% ethanol (500 μl) and precipitated by centrifugation (14000 r.p.m., 10 mins in an Eppendorf 5415 centrifuge). The pellet was air dried and resuspended in 1 x TE (50 μl , O/N, 37°C). The DNA (0.4 μg) was analysed by Southern Blot.

Lysis Solution:

Tris-HCl (pH 8.5).....	100	mM
EDTA.....	50	mM
SDS.....	0.2%	(v/v)
NaCl.....	200	mM
Proteinase K.....	0.1	mg/ml

9.3.21 Southern Blot Analysis

Approximately 20 μg of high molecular weight DNA was digested using 2-4 u of enzyme/ μg DNA enzyme (in a reaction volume of 25 μl , 37°C , 18 hours). The DNA in 1 x GLB and RNase A (2 mg/ml) was run on horizontal 0.8 or 1% TAE agarose gels prepared as described (section 9.3.5) at 40 mA, overnight in an Owl Scientific horizontal tank (A2 model). Gels were photographed under longwave UV light, depurinated (0.25 M HCl, 15 min.), rinsed in MQ water and transferred to Hybond- N^{+} nylon membranes (Amersham) using the protocol of the supplier. The DNA was immobilized by UV crosslinking (120 J, Stratagene Inc., UV StratalinkerTM 1800).

9.3.22 Hybridisation of Radioactive Probes to Hybond- N^{+} Membranes

Hybridisation reactions at 65°C were carried out in a Hybaid midi hybridisation oven. Filters were prehybridised overnight at 65°C in prehybridisation solution (10 ml/filter). Probes were boiled for 10 minutes with sonicated salmon sperm DNA (1 mg), snap-cooled on ice before addition to hybridisation cylinders. Filters were probed overnight at 65°C before washing successively for 30 minutes in 2 x SSC/1% SDS, 0.5 x SSC/1% SDS, and 0.1 x SSC/1% SDS at 65°C .

Prehybridisation Solution:

NaPO ₄ (pH 7.2)	250	mM
SDS	7%	(v/v)
PEG 6000	10%	(w/v)
EDTA.....	1	mM
sonicated DNA	0.1	mg

Radioactivity was detected as described in section 9.6.

9.3.23 *Hybridisation of Radioactive Probes to Nitrocellulose Membranes*

Filters (section 9.3.10) were prehybridized in sealed petri dishes at 42°C for 2 to 4 hours in prehybridisation solution. Probes were boiled (10 mins.) with sonicated salmon sperm DNA (1 mg), snap-cooled on ice before addition to petri dishes. Filters were probed overnight at 42°C before washing as above (section 9.3.22)

Prehybridisation Solution:

deionized formamide	40%	(w/v)
SDS	1%	(v/v)
PEG 6000	10%	(w/v)
Tris-HCl (pH 7.4).....	50	mM
50x Denhardt's solution.....	5%	(w/v)
NaCl	5.85	g

9.3.24 *Cytoplasmic RNA Preparation*

Cytoplasmic RNA was isolated by the method of Edwards and co-workers (1985) as follows. Cells were harvested by trypsinization, washed once in PBS and resuspended in TNM (2 ml, on ice). Cells were lysed by vigorous pipetting and were left on ice (5 mins). Nuclei were pelleted by centrifugation (3000 rpm, 4°C, 5 min.) and supernatant decanted onto TUNES (2 ml) before mixing. The mixture was extracted twice with phenol/chloroform

equilibrated with Tris-HCl (0.1 M, pH 8.0, 3000 r.p.m., 5 min.). The aqueous phase was precipitated (0.1 v NaAc, pH 5.5 and 2.0 v C₂H₅OH 100 %, -80°C, 30 min.). The nucleic acid was resuspended in MQ water (0.4 ml) and ethanol precipitated before final resuspension in MQ water.

TNM:

Tris-HCl (pH 7.6)	30	mM
NaCl	150	mM
MgCl ₂	15	mM
NP40	0.4%	(v/v)

TUNES:

Tris-HCl (pH 8.0)	10	mM
Urea	7	M
NaCl	0.35	mM
EDTA	1	mM
SDS	2%	(v/v)

9.3.25 *cDNA Synthesis*

RNA (2 µg) was heated (65°C, 5 min.) and snap cooled on ice (2 min.). The following were added to the RNA and the reaction allowed to proceed at 41°C for one hour.

10 x RT buffer	0.1	v
dATP, dCTP, dGTP, dTTP	0.2	mM each
RNAsin	40	u
oligo d(T) primer	0.5	µg
AMV RTase	8	u

10 x RT Buffer:

Tris-HCl	0.5	M
MgCl ₂	60	mM

KCl..... 0.4 M
 DTT.....10 mM

9.4 Tissue Culture Techniques

9.4.1 Cell Lines

The cell lines used throughout the course of this work were obtained from:

MBL-5 (Dr. Lindsay Williams, Ludwig Institute, Melbourne, Australia), E14TG2a (Dr. Austin Smith, CGR Edinburgh, UK) and COS-1 (ATCC).

The isolation of E14TG2a and MBL-5 ES cells is described in Hooper *et al.*, (1987) and Pease *et al.*, (1990), respectively.

9.4.2 Solutions

1 x PBS:

NaCl 80 g
 KCl..... 0.2 g
 KH₂PO₄..... 0.2 g
 Na₂HPO₄..... 1.15 g

PBS was made up to 1 litre with MQ water and sterilised by autoclaving (section 9.2.9).

Trypsin EDTA:

trypsin (1:250) 1 g
 10 x Versene 100 ml

Trypsin EDTA was filter sterilised (section 9.2.9).

10 x Versene:

EDTA..... 2 g
 NaCl 80 g
 KCl..... 2 g
 KH₂PO₄..... 2 g

Na₂HPO₄.....11.5 g

Versene was made up to 1 litre with MQ water, filter sterilised (section 9.2.9) and stored at 4°C.

β-mercaptoethanol/PBS:

β-mercaptoethanol..... 100 mM

PBS 14.1 ml

β-mercaptoethanol/PBS was stored at 4°C for up to two weeks.

LIF conditioned medium:

COS cell medium conditioned with LIF was prepared as described by Smith (1991) with the exception that COS-1 cells were used and transfections were carried out by electroporation.

9.4.3 Media

Incomplete ES cell medium:

Incomplete ES cell medium was made with DMEM powder (high glucose, with L-glutamine, without sodium bicarbonate and sodium pyruvate).

DMEM.....67.4 g

NaHCO₃.....18.5 g

gentamycin..... 6.25 ml

The pH of incomplete ES cell medium was adjusted to 7.4 with HCl. Incomplete ES cell medium made up to 5 litres with MQ water, filter sterilised (section 9.2.9) and stored at 4°C.

Complete ES cell medium:

incomplete ES cell medium 85% (v/v)

FCS.....15% (v/v)

LIF conditoned medium..... 100 u/ml

β-mercaptoethanol..... 0.1 mM

L-glutamine..... 1 mM

Incomplete COS-1 cell medium:

incomplete COS-1 cell medium was made with DMEM powder (high glucose, with L-glutamine, without sodium bicarbonate and sodium pyruvate).

DMEM.....67.4 g
 HEPES.....23.8 g
 gentamycin..... 6.25 ml

The pH of incomplete COS-1 cell medium was adjusted to 7.4 with HCl. Incomplete COS-1 cell medium was made up to 5 litres with MQ water, filter sterilised (section 9.2.9) and stored at 4°C.

Complete COS-1 cell medium:

incomplete COS-1 cell medium..90 % (v/v)
 FCS.....10 % (v/v)

Hams F12 medium :

Hams F12 was made with Hams F12 powder (with L-glutamine and without sodium bicarbonate).

Hams F12 10.6 g
 NaHCO₃..... 1.18 g
 gentamycin..... 1.25 ml

The pH of Hams F12 was adjusted to pH 7.2 with HCl. Hams F12 was made up to 1 litre with MQ water, filter sterilised and stored at 4°C.

9.4.4 *ES and COS-1 Cell Maintenance*

ES cells were maintained routinely as described in (Smith, 1991) with the exception that ES cells were grown in complete ES cell medium containing 15% FCS and under 10 % CO₂. COS-1 cells were maintained in COS-1 cell medium, grown at 37°C in 5% CO₂ and passaged at confluence.

9.4.5 *Transient Transfection of ES cells by Electroporation*

Exponentially growing ES cells, at the appropriate cell density for passaging, were harvested by trypsinization, pelleted by centrifugation (1200 r.p.m., 5 mins in a Hettich Universal benchtop centrifuge). Cells were pooled and respun (1200 r.p.m., as above) before

resuspension in cold ES cell electroporation buffer (10 ml) and a cell count performed. The cells were respun and resuspended at a cell density of 1×10^7 cells/ml. To cell suspension (0.5 ml) in an electroporation cuvette (Bio-Rad, 4mm disposable electroporation cuvette), circular DNA (20 μ g) was added. After mixing the cells were left at 4°C for 10 minutes. The cells were electroporated (Bio-Rad Genepulser, 210 mV, 960 μ F), incubated at room temperature for 10 minutes, seeded at a density of 1.5×10^6 cells/well of a 6-well plate into complete ES cell medium or ES cell medium/MedII medium (1:1) and grown for 48 hours under 10% CO₂.

Electroporation Buffer :

HEPES	20	mM
NaCl	137	mM
KCl.....	5	mM
Na ₂ HPO ₄	0.7	mM
glucose	6	mM
β -mercaptoethanol.....	0.1	mM

The pH of the electroporation buffer was adjusted to 7.0 with NaOH. Electroporation buffer was filter sterilised (9.2.9) and stored at 4°C.

9.4.6 Transient Transfection of COS-1 Cells by Electroporation

COS-1 cells were grown to confluency in COS-1 cell medium (150 cm² flasks). After three rinses in PBS the cells were incubated in prewarmed trypsin (2 ml, 5 min, RT). The trypsin was inactivated with COS-1 cell medium (8 mls, warmed to 37°C). The cells were detached with agitation, pelleted by centrifugation (1000 r.p.m., 5 min.) and resuspended in cold electroporation buffer (20 ml). After a cell count was performed, the cells were pelleted by centrifugation (as above) and resuspended in cold electroporation buffer (10 ml). The cells were pelleted by centrifugation and resuspended at a concentration of 1×10^7 cells/ml. To electroporation cuvettes (Bio-Rad, 4mm disposable electroporation cuvette), containing FCS (50 μ l) and sonicated salmon sperm DNA (50 μ l, 10 mg/ml), and circular plasmid DNA (10 μ l), cell suspension (0.5 ml) was added. After mixing and incubation (10 min, 4°C) the cells

were electroporated (Bio-Rad Genepulser, 270 mV, 250 μ F). After 10 minutes at room temperature, the cells (1/30 v) were plated into COS-1 cell medium (3 mls, warmed to 37°C) in a well of a six well plate at a density of 1.5×10^6 cells/well in Complete COS-1 cell medium. β -galactosidase assays were carried out 2 days later.

9.4.7 *Stable Transfection of ES Cells by Calcium Phosphate Co-precipitation*

Transfections of the CAT expression plasmids were carried out by a modification of the method described by Chen and Okayama (1987). Exponentially growing ES cells were seeded at a density of 2×10^6 cells/100 mm diameter plate, and grown for 24 hours at 37°C under 10% CO₂. Precipitates were formed with plasmid DNA (20 μ g, pPGKneob was co-transfected with the expression vector DNA at a ratio of 1:5), CaCl₂ (500 μ l, 0.25 M) and 2 x BBS (500 μ l, pH 6.99). Cells were transfected for 16 hours under 3% CO₂ in a 50:50 mix of DMEM:Hams F12 containing 10% FCS. The precipitate was washed from the cells and they were incubated for a further 30 hours, under standard culture conditions. At this time selective medium containing 200 μ g/ml G418 was added and was replaced every 48 hours. Resistant colonies were picked when no viable mock transfected cells remained and stable ES cell lines were expanded under conditions of selection.

2 x BBS:

BES (pH 6.99)	50	mM
NaCl	280	mM
Na ₂ HPO ₄	1.5	mM

9.4.8 *Screening of Stable ES Cell lines for CAT Activity*

Stable ES cell lines were screened for CAT expression by the following method. Exponentially growing, stably transfected ES cells, at the appropriate cell density for passaging were seeded into 6 well plates at a cell density of 5×10^5 cells per well and grown at 37°C overnight in complete ES cell medium, under 10% CO₂. Two hours after refeeding with complete ES cell medium, cell lines requiring IFN induction were treated with the

appropriate volume of IFN (270 I.U./ml, mouse α and β IFNs, Sigma Chemical Co.). Cells were grown for 48 hours and harvested (section 9.4.10).

9.4.9 Time Course of IFN Induced CAT Expression

Exponentially growing p1CX ES cells, at the appropriate cell density for passaging, were seeded into 35 mm diameter plates at a density of 5×10^5 cells/well and were grown overnight. Two hours after refeeding with 3 ml of complete ES cell medium, IFN (1000 I.U./ml) was added to the cells. Cells were harvested (section 9.4.10), after 2, 6, 12, 24 and 48 hours of growth in the continuous presence of IFN or after 8 or 10 hours in the presence of IFN.

9.4.10 Preparation of Crude Cell Extract for the CAT Assay

The transfected ES cells were washed three times with PBS. The cells were incubated in TEN buffer (1 ml) at room temperature for 5 minutes, scraped off the dish with a disposable cell scraper and pelleted by centrifugation (14000 r.p.m., 15 seconds in an Eppendorf 5415 microfuge). The cell pellet was resuspended in Tris-HCl (50 μ l, 0.25 M, pH 7.6) and the cells were lysed by three cycles of freeze thawing (alternate immersion in a dry ice bath and a 37°C heating block for 5 mins. followed by vortexing). Cell debris was pelleted by centrifugation (14000 r.p.m., 4°C, 10 mins. as above). The supernatant was retained and stored at -20°C. The protein concentration was determined by the method of Bradford (1976).

TEN Buffer:

Tris-HCl (pH 7.4).....	40	mM
EDTA.....	1	mM
NaCl	150	mM

9.4.11 The Chloramphenicol Acetyl-transferase Assay

CAT assays were carried out by the method of Gorman *et al.*, (1982), as follows. The crude cell extract (section 9.4.10), was heated (65°C, 7 mins.), to destroy the endogenous deacetyl

activity. To the appropriate volume of protein extract, Tris-HCl (70 μ l, 1M, pH 7.6), acetyl-CoA (20 μ l, 4 mM) and ^{14} C-chloramphenicol (4 μ l) was added, and the volume made up to 150 μ l with TD buffer. The chloramphenicol acetyl-transferase reaction was allowed to proceed at 37°C for 1 hour. The reaction was stopped by the addition of ice cold ethyl acetate (1 ml). The mixture was vortexed briefly and then centrifuged (14000 r.p.m., 30 sec. in an Eppendorf 5415 microfuge). The organic upper phase, containing the chloramphenicol, was dried in vacuo. The chloramphenicol was resuspended in ethyl acetate (30 μ l) and vortexed briefly. The samples were spotted onto a silica gel thin layer chromatography (TLC) aluminium sheet. The TLC was developed in ascending chloroform:methanol (95 : 5), for 30 minutes. The TLC was dried at room temperature and was exposed to X-ray film with enhancer screens at -80°C for 24 or 48 hours. The CAT assays were quantitated by liquid scintillation counting, (HiSafe, 1214 Rackbeta liquid scintillation counter LKB WALLAC).

TD Buffer :

Tris-HCl (pH 7.8)	250	mM
DTT	5	mM

9.4.12 Stable Transfection of ES Cells with Targeting Vector DNA by Electroporation

The vector DNA (150 μ g) was linearized by restriction enzyme digestion (37°C, 18 hours) and ethanol precipitated. After three washes in 70 % ethanol the DNA was resuspended in sterile PBS (150 μ l). ES cells were harvested by trypsinisation in ES cell medium (150 cm² flasks, 1200 r.p.m., 5 min), rinsed in complete ES cell medium (20 ml) and resuspended in PBS (20 ml). After a cell count was performed the cells were pelleted by centrifugation (as above) and resuspended in PBS at a concentration of 1.54×10^8 cells/ml. To the cell suspension (650 μ l) in a cuvette (Bio-Rad, 4 mm disposable electroporation cuvette) the resuspended DNA (150 μ l) was added and immediately transfected (Bio-Rad Genepulser, 800 mV, 3 μ F).

ES cells transfected with pHNLX or no DNA (section 4.3.2) were plated into complete ES cell medium (20 ml, 37°C) at 6×10^4 /cm² in 100 mm diameter plates and grown at 37°C in

10% CO₂ as described in table 4.3. Selection with HAT (100 µM hypoxanthine, 0.4 µM amethopterin, 16 µM thymidine), G418 (200 µg/ml) or HAT/G418 was applied 24 hours after transfection. Transfected cells were also plated into complete ES cell medium (10 ml, 37°C) at $6 \times 10^3/\text{cm}^2$ in a 100 mm diameter plate to determine the number of cells surviving electroporation.

ES cells transfected with p1DLacLIF, p1MLacLIF, p1TLacLIF, pE2LacLIF or pCH110 DNA (sections 5.8.2, 5.8.4) were plated into complete ES cell medium as described in sections 5.8.2 and 5.8.4 and table 5.4. Selection with 6-TG (20 mM), or 6-TG/G418 (20 mM/175 µg/ml) was applied after a period of five days to allow for decay of HPRT activity. Transfected cells were also plated into complete ES cell medium (10 ml, 37°C) at $6 \times 10^3/\text{cm}^2$ in a 100 mm diameter plate to determine the number of cells surviving electroporation and to determine the transfection efficiency by determining the number of β -galactosidase positive cells. The cells transfected with pCH110 were fixed and histochemical staining for β -galactosidase activity (section 9.4.13) carried out to determine the transfection efficiency (Table 5.4) two days after transfection.

4.4.13 Histochemical Staining for β -Galactosidase Activity

β -Galactosidase activity expressed by transfected ES or COS-1 cells was detected by the following procedure. Cells were rinsed three times in PBS and fixed with glutaraldehyde (0.2 % in PBS) for 5 minutes. After a further three rinses in PBS, the cells were incubated in β -galactosidase stain solution containing X-gal (400 µg/ml or 1 mg/ml) at 30°C for 16 hours. Cells were viewed under phase contrast and bright field optics using the Nikon Diaphot or Zeiss Axioplan Universal microscopes.

β -galactosidase stain solution:

K ₃ Fe(CN) ₆	0.45	mM
K ₄ Fe(CN) ₆	0.45	mM
MgCl ₂	1	mM

9.4.14 *Differentiation of ES cells to X Cells*

Transiently transfected ES cells (section 9.4.7) were seeded into complete ES cell medium:MedII medium (1:1) (MedII medium was isolated from MedII cells cultured in incomplete ES cell medium for 5 days; J. Rathjen, unpublished) and grown for two to three days to permit X cell formation. After this time, cells requiring induction with IFN were treated with IFN (1000 I.U./ml) and histochemical staining for β -galactosidase activity was carried out 24 hours later.

9.4.15 *MBA Differentiation of ES Cells*

ES cells were differentiated with MBA by the method of Smith (1991). ES cells were seeded at a density of 1×10^4 cells/cm² in complete ES cell medium (reduced LIF 100 u/ml ESGRO) and grown at 37°C overnight. Medium was replaced with complete ES cell medium (reduced LIF 100 u/ml ESGRO) containing 5 mM MBA and cells incubated for a further 48 hours. Medium was replaced complete ES cell medium (lacking LIF) and cells were grown for a further 48 hours. Cells were fixed (section 9.4.13) and stained using X-gal (1 mg/ml). Cells were viewed under phase contrast and bright field optics using the Nikon Diaphot or Zeiss Axioplan Universal microscopes.

9.4.16 *Hoechst Staining of ES cells*

After staining for β -galactosidase activity (section 9.4.13) cells were rinsed and Hoechst dye (diluted 1/1000) was added to the cells for 30 seconds before they were rinsed in PBS. Cells were viewed under fluorescence optics using the Zeiss Axioplan Universal microscope.

9.4.17 *Karyotypic Analysis of ES Cell Lines*

All ES cell lines reintroduced into the mouse blastocyst and cells that had undergone tag gene loss by means other than homologous recombination, were karyotyped by the method of Robertson (1987). Briefly, rapidly growing ES cells were arrested in mitosis by treatment with colcemid (Sigma Chemical Co.), chromosome spreads were prepared and solid stained (Robertson, 1987) or Q banded (Caspersson *et al.*, 1970). Solid stained chromosome

spreads were viewed under phase contrast optics and Q banded spreads were viewed under fluorescence optics on the Zeiss Axioplan Universal microscope.

9.5 Mouse Manipulations and Analysis

9.5.1 *Chimaerae Production*

Chimaeric mice were produced by blastocyst microinjection by the methods described in Hogan and co-workers (1994). E14TG2a ES cells and targeted derivatives were introduced into CBA x C57BL or C57BL/6 x C57BL/10 blastocysts. Chimaerae were identified by coat colour assessment. E14TG2a ES cells are homozygous for the chinchilla allele (c^{ch}/c^{ch}) and chimaerism was detected by the presence of sandy coat color pigmentation.

9.5.2 *GPI Analysis of Sperm*

GPI analysis of sperm obtained from the epididymis of male chimaerae was carried out using the Super Z applicator kit using the protocol of the supplier.

9.5.3 *Breeding*

Male chimaerae produced with E14TG2a ES cells and targeted derivatives were mated with Balb/c female mice and transmission of the ES cell genotype through the germline was revealed by the production of grey offspring.

9.6 Phosphorimager Analysis and Autoradiograph Scanning

Gels and filters were exposed to Storage Phosphor Screens (Molecular Dynamics) and processed using a Molecular Dynamics PhosphorImager running the ImageQuant software package. Autoradiographs were scanned using a Hewlett Packard *ScanJet IICx* scanner running the DeskScan II 2.0 software package. PhosphorImager and DeskScan files were manipulated using the AdobePhotoshop™ and MacDraw Pro programs, and printed using a Hewlett Packard LaserJet 5 MP printer.

9.7 Containment Facilities

9.7.1 *Recombinant DNA Manipulations*

All manipulations involving viable organisms which contained recombinant DNA were carried out in accordance with the regulations and approval of the Australian Academy of Science Committee on Recombinant DNA and the University Council of the University of Adelaide.

9.7.2 *Animal Manipulations*

All procedures involving animals were carried out with the approval of the University of Adelaide Animal Ethics Committee.

References

References

- Acland, P., Dixon, M., Peters, G. and Dickson, C., (1990). *Nature (London)*, **343** pp. 662-665
- Amaya, E., Musci, T. J. and Kirschner, M. W., (1991). *Cell*, **66** pp. 257-270
- Askew, G. R., Doetschman, T. and Lingrel, J. B., (1993). *Mol. Cell. Biol.*, **13** (7), pp. 4115-4124
- ATCC, (1994). *American Type Culture Collection Cell Lines and Hybridomas* Rockville, MD
- Berger, J., Hauber, J., Hauber, R., Geiger, R. and Cullen, B. R., (1988). *Gene*, **66** pp. 1-10
- Bernstine, E. G., Hooper, M. L., Grandchamp, S. and Ephrussi, B., (1973). *Proc. Natl. Acad. Sci. (USA)*, **70** pp. 3899-3903
- Bettess, M., (1993). Honours Thesis, University of Adelaide, Adelaide.
- Bezzubova, O. Y. and Buerstedde, J.-M., (1994). *Experimentia*, **50** pp. 270-275
- Bhatt, H., Brunet, L. J. and Stewart, C. L., (1991). *Proc. Natl. Acad. Sci. USA*, **88** pp. 11408-11412
- Bradley, I., Evans, M., Kaufman, M. H. and Robertson, E. J., (1984). *Nature (London)*, **309** pp. 255-256
- Brandon, E. P., Idzera, R. L. and McKnight, G. S., (1995). *Curr. Biol.*, **5** (6), pp. 625-634
- Brandon, E. P., Idzera, R. L. and McKnight, G. S., (1995). *Curr. Biol.*, **5** (7), pp. 758-765
- Brandon, E. P., Idzera, R. L. and McKnight, G. S., (1995). *Curr. Biol.*, **5** (8), pp. 873-881
- Braun, R. E., Lo, D., Pinkert, C. A., Widera, R. A., Flavell, R. A., Palmiter, R. D. and Brinster, R. L., (1990). *Biol. Reprod.*, **43** pp. 684-693
- Brinster, R. L., (1974). *J. Exp. Med.*, **140** pp. 1049-1056
- Bronson, S. K. and Smithies, O., (1994). *J. Biol. Chem.*, **269** (44), pp. 27155-27158
- Brown, A., (1994). Honours Thesis, University of Adelaide, Adelaide.
- Capecchi, M. R., (1989b). *Trends in Genet.*, **5** (3), pp. 70-76
- Caspersson, T., Zech, L. and Johansson, C., (1970). *Exp. Cell Res.*, **60** pp. 315-319
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C., (1994). *Science*, **263** pp. 802-805
- Chen, C. and Okayama, H., (1987). *Mol. Cell. Biol.*, **7** (8), pp. 2745-52
- Conover, J. C., Ip, N. Y., Poueymirou, W. T., Bates, B., Goldfarb, M. P., DeChiara, T. M. and Yancopoulos, G. D., (1993). *Development*, **110** pp. 559-565
- Conquet, F. and Brulet, P., (1990). *Mol. & Cell. Biol.*, **10** (7), pp. 3801-3805
- Conquet, F., Peyriebras, N., Tiret, L. and Brulet, P., (1992). *Proc. Natl. Acad. Sci., USA*, **89** pp. 8195-8199
- Cosman, D., (1993). *Cytokine*, **5** pp. 95-106

- Cui, C., Maqsood, A. W., Wight, D., Kopchick, J. and Stambrook, P. J., (1994). *Trans. Res.*, **3** pp. 182-194
- Damjanov, I., Damjanov, A. and Solter, D., (1987) in *Teratocarcinomas and Embryonic Stem Cells: a practical approach* E. J. Robertson, Eds. IRL Press Ltd., Oxford, pp. 1-23
- Davies, M. and Kaufman, R., (1992). *Journal of Virol.*, **66** pp. 1924-1932
- Davis, S., Aldrich, T. H., Stahl, N., Pan, L., Taga, T., Kishimoto, T., Ip, N. Y. and Yancopoulos, G. D., (1993). *Science*, **260** pp. 1805-1810
- DeLorenzo, R. J. and Ruddle, F. H., (1969). *Biochem. Genet.*, **3** pp. 151-162
- Deng, C. and Capecchi, M., (1992). *Mol. Cell Biol.*, **12** pp. 3365-3371
- Detloff, P. J., Lewis, J., John, S. W. M., Shehee, W. R., Langenbach, R., Maeda, N. and Smithies, O., (1994). *Mol. Cell. Biol.*, **14** (10), pp. 6936-6943
- Doetschman, T. C., Eisetter, M., Katz, M., Schmidt, W. and Kemler, R. J., (1985). *J. Embryol. Exp. Morphol.*, **87** pp. 27-45
- Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S. and Smithies, O., (1987). *Nature*, **330** pp. 576-578
- Edwards, D. R., Parfett, C. L. J. and Denhardt, D., (1985). *Mol. Cell. Biol.*, **5** pp. 3280-3288
- Ernst, M., Gearing, D. P. and Dunn, A. R., (1994). *EMBO J.*, **13** (7), pp. 1574-1584
- Escary, J. L., Pereau, J., Dumenil, D., Ezine, S. and Brulet, P., (1993). *Nature*, **363** pp. 361-364
- Evans, M. J., (1972). *J. Embryol. Exp. Morphol.*, **28** (163-173), pp.
- Evans, M. J. and Kaufman, M. H., (1981). *Nature (London)*, **292** pp. 154-156
- Feldman, B., Poueymirou, W., Papaioannou, V. E., DeChiara, T. M. and Goldfarb, M., (1995). *Science*, **267** pp. 246-249
- Festing, M. F. W., (1992). *Mouse Genome*, **90** (3), pp. 231-315
- Flach, G., Johnson, M. H., Braude, P. R., Taylor, R. A. S. and Bolton, V. N., (1982). *EMBO Journal*, **1** pp. 681-686
- Folger, K. R., Wong, E. A., Wahl, G. and Capecchi, M. R., (1982). *Mol. Cell. Biol.*, **2** pp. 1372-1387
- Friedman, R. L., Manly, S. P., McMahon, M., Kerr, I. M. and Stark, G. R., (1984). *Cell*, **38** pp. 745-755
- Friedrich, G. and Soriano, P., (1991). *Genes and Development*, **5** pp. 1513-1523
- Fuji, J. T. and Martin, G. R., (1983). *J. Embryol. Exp. Morphol.*, **74** pp. 79-86
- Gardener, R. L. and Rossant, J., (1979). *J. Embryol. Exp. Morphol.*, **52** pp. 141-152
- Gardener, R. L., (1982). *J. Embryol. Exp. Morphol.*, **68** pp. 175-198
- Gardener, R. L., (1983). *Int. Rev. Exp. Pathol.*, **24** pp. 63-133
- Gassmann, M., Donoho, G. and Berg, P., (1995). *Proc. Natl. Acad. Sci. USA*, **92** pp. 1292-1296

- Gearing, D. P., Gough, N. M., King, J. A., Hilton, D. J., Nicola, N. A., Simpson, R. J., Nice, E. C., Kelso, A. and Metcalf, D., (1987). *EMBO J.*, **6** (13), pp. 3995-4002
- Gearing, D. P., Comeau, M. R., Friend, D. J., Gimpel, S. D., Thut, C. J., McGourty, J., Brasher, K. K., King, J. A., Gillis, S., Mosley, B., Ziegler, S. F. and Cosman, D., (1992). *Science*, **255** pp. 1434-1437
- Gearing, D. P. and Bruce, A. C., (1992). *New. Biol.*, **4** pp. 61-64
- Gondo, Y., Nakamura, K., Nakao, K., Sasoka, T., Ito, K., Kimura, M. and Katsuki, M., (1994). *Biochem. Biophys. Res. Com.*, **202** (2), pp. 830-837
- Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbossa, J. A. and Ruddie, F. H., (1980). *Proc. Natl. Acad. Sci. USA*, **77** (7380-7384), pp.
- Gorman, C. M., Moffat, L. F. and Howard, B. H., (1982). *Mol. Cell. Biol.*, **2** pp. 1944-1951
- Green, J. B., Howes, G., Symes, K., Cooke, J. and Smith, J. C., (1990). *Development*, **108** pp. 173-183
- Grez, M., Akgun, E., Hillberg, F. and Ostertag, W., (1990). *Proc. Natl. Acad. Sci. (USA)*, **87** pp. 9202-9206
- Gu, H., Zou, Y. and Rajewsky, K., (1993). *Cell*, **73** pp. 1155-1164
- Hannan, G., Lehnert, S. A., MacAvoy, E. S., Jennings, P. A. and Molloy, P. L., (1993). *Gene*, **130** pp. 233-239
- Harrison, S., (1993). Honours Thesis, University of Adelaide, Adelaide.
- Hasty, P., Ramirez-Solis, R., Krumlauf, R. and Bradley, A., (1991a). *Nature (London)*, **350** pp. 243-246
- Hasty, P., Rivera-Perez, J. and Bradley, A., (1991b). *Mol. Cell Biol.*, **11** (11), pp. 5586-5591
- Hasty, P., Rivera-Perez, J., Chang, C. and Bradley, A., (1991c). *Mol. Cell. Biol.*, **11** (9), pp. 4509-4517
- Hasty, P., Rivera-Perez, J. and Bradley, A., (1992). *Mol. Cell. Biol.*, **12** (6), pp. 2464-2474
- Heath, J. K. and Smith, A. G., (1988). *J. Cell Sci. (Suppl)*, **10** pp. 257-266
- Heyer, W. D., (1994). *Experimentia*, **50** pp. 223-233
- Hilton, D. J., Nicola, N. A. and Metcalf, D., (1988). *Anal. Biochem.*, **172** pp. 359-367
- Hilton, D. J., (1992). *TIBS*, **17** pp. 72-76
- Hilton, D. J., Hilton, A. A., Raicevic, A., Rakar, S., Harrison-Smith, M., Gough, N. M., Begley, C. G., Metcalf, D., Nicola, N. A. and Wilson, T. A., (1994). *EMBO J.*, **13** pp. 4765-4775
- Hogan, B., Beddington, R., Costantini, F. and Lacy, E., (1994), *Manipulating the mouse embryo, a laboratory manual* Cold Spring Harbour Laboratory,
- Holliday, R., (1964). *Genet. Res.*, **5** pp. 5586-5591
- Holmes, D. S. and Quigley, M., (1981). *Anal. Biochem.*, **114** pp. 193-197
- Hooper, M. L. and Slack, C., (1977). *Dev. Biol.*, **55** pp. 271-284

- Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M., (1987). *Nature (London)*, **326** pp. 292-295
- Hsu, L. and Heath, J. K., (1994). *Biochem. J.*, **302** pp. 102-111
- Hughes, S. M. and Blau, H. M., (1990). *Nature (London)*, **345** pp. 350-353
- Ip, N. Y., Nye, S. H., Boulton, T. G., Davis, S., Taga, T., Li, Y., Birren, S. J., Yasukawa, K., Kishimoto, T., Anderson, D. J., Stahl, N. and Yancopoulos, G. D., (1992). *Cell*, **69** pp. 1121-1132
- Israel, D. I. and Kaufman, R. J., (1989). *Nucl. Acids Res.*, **17** (12), pp. 4589-4604
- Jackson, I. J., (1989) in *Genes and Embryos* pp. 165-221
- Johnson, M. H. and Maro, B., (1988) in *Experimental approaches to mammalian development* Cambridge University Press, Cambridge, pp. 35-65
- Johnson, R. S., Sheng, M., Greenberg, M. E., Kolodner, R. D., Papaioannou, V. E. and Spielgelman, B. M., (1989). *Science*, **245** pp. 1234-1236
- Karin, M., (1992). *FASEB Journal*, **6** pp. 2581-2589
- Kaspar, P., Dvorak, M. and Bartunek, P., (1993). *FEBS Lett.*, **319** pp. 159-162
- Kaufman, R. J., Davies, M. W., Wasley, L. C. and Michnick, D., (1991). *Nucl. Acids Res.*, **19** pp. 4485-4490
- Kaufman, M. H., (1992), *The atlas of mouse development* Academic Press, Harcourt Brace Jovanovich, London
- Kelly, J. M., Gilbert, C. S., Stark, G. R. and Kerr, I. M., (1985). *Eur. J. Biochem.*, **153** pp. 367-371
- Kelly, J. M., Porter, A. J. C., Chernajovsky, Y., Gilbert, C. S., Stark, G. R. and Kerr, I. M., (1986). *EMBO Journal*, **5** (7), pp. 1601-1606
- Ko, N. S. H., Takahashi, N., Sugiyama, N. and Takano, T., (1989). *Gene*, **84** pp. 383-389
- Kola, I., Davey, A. and Gough, N. M., (1990). *Growth Factors*, **2** pp. 235-240
- Kucherlapati, R. S., (1984). *Proc. Natl. Acad. Sci. USA*, **81** pp. 3153-3157
- Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D. and Costantini, F. D., (1983). *Cell*, **34** (343-58), pp.
- Lake, J., (1996). Ph.D. Thesis, University of Adelaide, Adelaide.
- Lanford, R. E., White, R. G., Dunham, R. G. and Kanda, P., (1988). *Mol. Cell. Biol.*, **8** (7), pp. 2722-2729
- Lavranos, T. C., Rathjen, P. D. and Seamark, R. F., (1995). *J. Reprod. Fertility*, **Nov. issue** pp.
- Le Mouellic, H., Lallemand, Y. and Brulet, P., (1992). *Cell*, **69** pp. 251-264
- Lewis, N. E. and Rossant, J., (1982). *J. Embryol. Exp. Morphol.*, **72** pp. 169-181
- Li, M., Sendtner, M. and Smith, A., (1995). *Nature*, **378** pp. 724-727
- Lin, F., Sperle, K. and Sternberg, N., (1984). *Mol. Cell. Biol.*, **4** pp. 1020-1034

- Lin, W. C. and Culp, L. A., (1991). *BioTechniques*, **11** pp. 344-351
- Lyon, M. F. and Serle, A. G., (1989), *Genetic Variants and Strains of the Laboratory Mouse* Oxford University Press, Oxford
- Mader, S. and White, J. H., (1993). *Proc. Natl. Acad. Sci. USA*, **90** pp. 5603-5607
- Mansour, S. L., Thomas, K. R. and Capecchi, M. R., (1988). *Nature (London)*, **336** pp. 348-352
- Mansour, S. L., Thomas, K. R., Deng, C. and Capecchi, M., (1990). *Proc. Natl. Acad. Sci. USA*, **87** pp. 7688-7692
- Martin, G. R., (1981). *Proc. Natl. Acad. Sci. USA*, **78** pp. 7634-7638
- Matsui, Y., Zsebo, K. and Hogan, B. L. M., (1992). *Cell*, **70** pp. 841-847
- McMahon, A. P. and Bradley, A., (1990). *Cell*, **62** pp. 1073-1085
- Melton, D. W., (1994). *BioEssays*, **16** (9), pp. 633-638
- Metcalf, D., Nicola, N. A. and Gearing, D. P., (1990). *Blood*, **76** (1), pp. 50-56
- Metcalf, D., Hilton, D. and Nicola, N. A., (1991). *Blood*, **77** (2150-2153), pp.
- Metcalf, D., (1992). *Growth Factors*, **7** pp. 169-173
- Michalska, A. E. and Choo, K. H. A., (1993). *Proc. Natl. Acad. Sci. USA*, **90** pp. 8088-8092
- Mintz, B. and Illmensee, K., (1975). *Proc. Natl. Acad. Sci. (USA)*, **72** pp. 3585-3589
- Morita, T., Yoshimura, Y., Yamamoto, A., Murata, K., Mori, M., Yamamoto, H. and Matsushiro, A., (1993). *Proc. Natl. Acad. Sci. USA*, **90** pp. 6577-6580
- Mortensen, R. M., Conner, D. A., Chao, S., Geisterfer-Lowrance, A. A. T. and Seidman, J. G., (1988). *Mol. Cell. Biol.*, **12** pp. 2391-2395
- Mortensen, R. M., Zubiar, M., Neer, E. J. and Seidman, J. G., (1991). *Proc. Natl. Acad. Sci. (USA)*, **88** pp. 7036-7040
- Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M. and A., S., (1994). *Proc. Natl. Acad. Sci. USA*, **91** pp. 4303-4307
- Mountford, P. S. and Smith, A. G., (1995). *Trends In Genet.*, **11** (5), pp. 179-184
- Murakami, M., Hibi, M., Nakagawa, N., Nakagawa, T., Yasukawa, K., Yamanishi, K., Taga, T. and Kishimoto, T., (1993). *Science*, **260** pp. 1808-1810
- Murray, R., Lee, F. and Chiu, C.-P., (1990). *Mol. Cell. Biol.*, **10** pp. 4953-4956
- Niswander, L. and Martin, G. R., (1992). *Develop.*, **114** pp. 775-768
- Nusslein-Volhard, C. and Wieschaus, E., (1980). *Nature (London)*, **287** pp. 795-801
- Nusslein-Vollhard, C., Fronhofer, H. G. and Lehman, R., (1987). *Science*, **238** pp. 1675-1681
- Ogawa, T., Yu, X., Shinohara, A. and Egelman, E. H., (1993). *Science*, **259** pp. 1896-1899
- Ow, D., Wood, K. V., DeLuca, M., de Wet, J., Helsinki, D. R. and Howell, S. H., (1986). *Science*, **234** pp. 856-859

- Papaioannou, V. E., McBurney, M. W., Gardener, R. L. and Evans, M. J., (1975). *Nature (London)*, **258** pp. 70-73
- Pease, S., Braghetta, P., Gearing, D., Grail, D. and Williams, R. L., (1990). *Dev. Biol.*, **141** pp. 344-352
- Pederson, R. A., (1988) in *Experimental approaches to mammalian development* J. Rossant, R. A. Pedersen, Eds. Cambridge University Press, Cambridge, pp. 3-33
- Pennica, D., Shaw, K. J., Swanson, T. A., Moore, M. W., Shelton, D. L., Zioncheck, K. A., Rosenthal, A., Taga, T., Paoni, N. F. and I. W. W., (1995). *J. Biol. Chem.*, **270** (18), pp. 10915-10922
- Porter, A. C. J., Chernajovsky, Y., Dale, T. C., Gilbert, C., Stark, G. R. and Kerr, I. M., (1988). *EMBO Journal*, **7** (1), pp. 85-92
- Pratt, H. P. M., Bolton, V., N. and Gudgeon, K. A., (1983). *CIBA Found. Symp.*, **98** pp. 197-227
- Rathjen, P. D., Toth, S., Willis, A., Heath, J. K. and Smith, A. G., (1990a). *Cell*, **62** pp. 1105-1114
- Rathjen, P. D., Nichols, J., Toth, S., Edwards, D. R., Heath, J. K. and Smith, A. G., (1990b). *Genes & Dev.*, **4** pp. 2308-2318
- Reid, L. H., Gregg, R. G., Smithies, O. and Koller, B. H., (1990). *Proc. Natl. Acad. Sci. USA*, **87** pp. 4299-4303
- Riddle, R. D., Johnson, R. L., Laufer, E. and Tabin, C., (1993). *Cell*, **75** pp. 1401-1416
- Roberts, C. W., Shutter, J. R. and Korsmeyer, S. J., (1994). *Nature*, **368** pp. 747-749
- Robertson, E. J., (1986). *Trends Genet.*, **2** pp. 9
- Robertson, E., Bradley, A., Kuehn, M. and Evans, M., (1986). *Nature (London)*, **323** pp. 445-448
- Robertson, E. J., (1987) in *Teratocarcinomas and embryonic stem cells: a practical approach* E. J. Robertson, Eds. IRL press, Oxford, Washington DC, pp. 71-112.
- Robertson, M., Chambers, I., Rathjen, P., Nichols, J. and Smith, A., (1993). *Dev. Genet.*, **14** pp. 165-173
- Robinson, R. C., Grey, L. M., Staunton, D., Vankelecom, H., Vernallis, A. B., Moreau, J. F., Stuart, D. I., Heath, J. K. and Jones, E. Y., (1994). *Cell*, **77** (July 1), pp. 1101-1116
- Rose, T. M., Weiford, D. M., Gunderson, N. L. and Bruce, A. G., (1994). *Cytokine*, **6** pp. 48-54
- Rosner, M. H., Vigano, M. A., Rigby, P. W. J., Arnheiter, H. and Staudt, L. M., (1991). *Science*, **253** pp. 144-145
- Rossant, J., (1988) in *Experimental approaches to mammalian embryonic development* J. Rossant, R. A. Pedersen, Eds. Cambridge University Press, Cambridge, pp. 97-120
- Rossant, J. and Joyner, A. L., (1989). *Trends in Genet.*, **5** (8), pp. 277-283
- Rubinstein, M., Japon, M. A. and Low, M. J., (1993). *Nucl. Acids Res.*, **21** (11), pp. 2613-2617
- Rudnicki, M. A., Braun, T., Hinuma, S. and Jaenisch, R., (1992). *Cell*, **71** pp. 383-390

- Schwartzberg, P. I., Goff, S. P. and Robertson, E. J., (1989). *Science*, **246** pp. 799-804
- Schweinfest, C. W., Jorcyk, C. L., Fujiwara, S. and Papas, T. S., (1988). *Gene*, **71** pp. 207-210
- Scott, M. p., Tamkun, J. W. and Hartzell, G. W., (1989). *Biochem. Biophys. Acta*, **989** pp. 25-48
- Shackleford, G. M. and Varmus, H. E., (1987). *Cell*, **50** pp. 89-95
- Shen, M. M. and Leder, P., (1992). *Proc. Natl. Acad. Sci., USA*, **89** pp. 8240-8244
- Shen, M. M., Skoda, R. C., Cardiff, R. D., Campos-Torres, J., Leder, P. and Ornitz, D. M., (1994). *EMBO J.*, **13** (6), pp. 1375-1385
- Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. and Ogawa, T., (1993). *Nature Genet.*, **4** pp. 239-243
- Silvers, W. K., (1979), *The coat colours of mice* Springer-Verlag, New York
- Slack, J., Darlington, B., Heath, J. and Godsave, S., (1987). *Nature (London)*, **326** pp. 197-200
- Smith, A. G. and Hooper, M. L., (1987). *Dev. Biol.*, **121** pp. 1-9
- Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D., (1988). *Nature (London)*, **336** pp. 688-690
- Smith, A. G., (1991). *J. Tiss. Cult. Meth.*, **13** pp. 89-94
- Smith, A. G. and Rathjen, P. D., (1991). *Sem. Dev. Biol.*, **2** pp. 317-328
- Smith, A. G., (1992). *Seminars in cell biology*, **3** pp. 385-399
- Smith, A. G., Nichols, J., Robertson, M. and Rathjen, P. D., (1992). *Dev. Biol.*, **151** pp. 339-351
- Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. and Kucherlapati, R. S., (1985). *Nature, London*, **317** pp. 230-234
- Smithies, O. and Kim, H. S., (1994). *Proc. Natl. Acad. Sci. USA*, **91** pp. 3612-3615
- Snow, M., H. L. and Tam, P. P. L., (1979). *Nature (London)*, **279** pp. 555-557
- Sokol, S., Wong, G. G. and Melton, D. A., (1990). *Science*, **249** pp. 561-564
- Solter, D. and Knowles, B. B., (1978). *Proc. Natl. Acad. Sci. USA*, **75** pp. 5565-5569
- Soriano, P., Friedrich, G. and Lawinger, P., (1991). *J. Virol.*, **65** (5), pp. 5
- Soriano, P., (1995). *Ann. Rev. Neurosci.*, **18** pp. 1-18
- Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A. and Melton, D. W., (1994). *Mol. Cell. Biol.*, **14** (2), pp. 1009-1016
- Stahl, J., Gearing, D. P., Wilson, T. A., Brown, M. A., King, J. A. and Gough, N. M., (1990). *J. Biol. Chem.*, **265** (15), pp. 8833-8841
- Stewart, T. A. and Mintz, B., (1981). *Proc. Natl. Acad. Sci. (USA)*, **78** pp. 6314-6318
- Stewart, C. L., (1982). *J. Embryol. Exp. Morphol.*, **67** pp. 167-179

- Stewart, C. L., Kaspar, P., Brunet, L. J., Bhatt, H., Gadi, I., Kontgen, F. and Abbondanzo, S. J., (1992). *Nature*, **359** pp. 76-79
- Stockli, K. A., Lottspeich, F., Sendtner, M., Masiakowski, P., Carroll, R., Gotz, R., Lindholm, D. and Thoenen, H., (1989). *Nature (London)*, **342** pp. 920-923
- Strittmatter, S. M., Fankhauser, C., Huang, P., Mashimo, H. and Fishman, M., (1995). *Cell*, **80** pp. 445-452
- Struhl, G., (1989). *Nature (London)*, **338** pp. 741-744
- Subak-Sharpe, J. H., Burk, R. R. and Pitts, J. D., (1966). *Heredity*, **21** pp. 342-343
- Subak-Sharpe, J. H., Burk, R. R. and Pitts, J. D., (1969). *J. Cell Sci.*, **4** pp. 353-367
- Summerbell, D., (1983). *J. Embryol. Exp. Morph.*, **78** pp. 269-289
- Szostak, J. W., Orr-Weaver, T. L. and Rothstein, R. J., (1983). *Cell*, **33** pp. 25-35
- Taga, T., Narazaki, M., Yasukawa, K., Saito, T., Miki, D., Hamaguchi, M., Davis, S., Shoyab, M., Yancopoulos, G. D. and Kishimoto, T., (1992). *Proc. Natl. Acad. Sci. USA*, **89** pp. 10998-11001
- Tarkowski, A. K., (1959). *Nature (London)*, **184** pp. 1286-1287
- te Riele, H., Maandag, E. R. and Berns, A., (1992). *Proc. Natl. Acad. Sci. USA*, **89** pp. 5128-5132
- Thomas, K. R. and Capecchi, M. R., (1986). *Nature, London*, **324** pp. 34-38
- Thomas, K. R., Folger, K. R. and Capecchi, M. R., (1986). *Cell*, **44** pp. 419-428
- Thomas, K. R. and Capecchi, M. R., (1987). *Cell*, **51** pp. 503-512
- Thomas, K. R., Deng, C. and Capecchi, M. R., (1992). *Mol. Cell. Biol.*, **12** (7), pp. 2919-2923
- Thomas, P. Q., Johnson, B. V., Rathjen, J. and Rathjen, P. D., (1995). *J. Biol. Chem.*, **270** (8), pp. 3869-3875
- Thomas, P. Q., (1995). Ph.D. Thesis, University of Adelaide, Adelaide.
- Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. and Melton, D. W., (1989). *Cell*, **56** pp. 313-321
- Thonglairoam, V., (1994). Ph.D. Thesis, University of Adelaide, Adelaide, South Australia.
- Tomida, M., Yamamoto-Yamaguchi, Y. and Hozumi, M., (1984). *J. Biol. Chem.*, **259** pp. 10978-10982
- Valancius, V. and Smithies, O., (1991). *Mol. Cell. Biol.*, **11** pp. 1402-1408
- van der Putten, H., Botteri, F. M., Miller, A. D., Rosenfeld, M. G., H., F., Evans, R. M. and Verma, I. M., (1985). *Proc. Natl. Acad. Sci. USA*, **82** pp. 6148-6152
- van Deursen, J. and Wieringa, B., (1992). *Nucl. Acids Res.*, **20** (15), pp. 3815-3820
- Vetter, D., Andrews, B. J., Roberts-Beatty, L. and Sadowski, P. D., (1983). *Proc. Natl. Acad. Sci. USA*, **80** pp. 7284-7288

Ware, C. B., Horowitz, M. C., Renshaw, B. R., Hunt, J. S., Liggitt, D., Koblar, S. A., Gliniak, B. C., McKenna, H. J., Papayannopolou, T., Thoma, B., Cheng, L., Donovan, P. J., Peschon, J., Bartlett, P. F., Willis, C. R., Wright, B. D., Carpenter, M. K., Davidson, B. L. and Gearing, D. P., (1995). *Development*, **121** pp. 1283-1299

Webb, G. C., Thomas, P. Q., Ford, J. H. and Rathjen, P. D., (1993). *Genomics*, **18** pp. 464-466

Weis, J., Fine, S. M., David, C., Savarirayan, S. and Sanes, J. R., (1991). *J. Cell. Biol.*, **113** pp. 1385-1397

Whyatt, L. M., (1991). Honours Thesis, University of Adelaide, Adelaide, South Australia.

Whyatt, L. M., Duwel, A., Smith, A. G. and Rathjen, P. D., (1993). *Mol. Cell. Biol.*, **13** (12), pp. 7971-7976

Wilkinson, D. G., Bailes, J. A. and McMahon, A. P., (1987). *Cell*, **50** pp. 79-88

Williams, R. L., Hilton, D. J., Pease, S., Williams, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A. and Gough, N. M., (1988). *Nature*, **336** pp. 684-687

Williams, B. R. G., (1991). *Eur. J. Biochem.*, **200** pp. 1-11

Wu, H., Liu, X. and Jaenisch, R., (1994). *Proc. Natl. Acad. Sci. USA*, **91** pp. 2819-2823