

Implementing inducible gene expression in ES cells using the tet-system

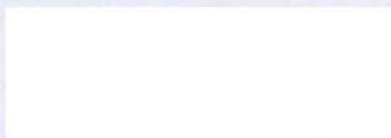
DAWN FISHER

Thesis presented for the degree of Doctor of Philosophy
University of Edinburgh

2005



I declare that the work presented in this thesis is my own,
except where otherwise stated



Dawn Fisher

Acknowledgments

I would firstly like to thank my supervisor Andrew Smith for his guidance throughout my project, and also my second supervisor Andy Docherty for his continued support. I would like to thank the members of the lab, particularly Barry Heavey for his advice and interesting tet-on/tet-off conversation, and Melville Richardson for her expert cloning tips! Thanks to Marios Stavridis for suggestions on my thesis and Austin Smith for his direction in my project. Special thanks to my Mum, Dad, Ray, Lucy, Jacq and Catherine, and my friends for their love and support. I would also like to thank the MRC and Celltech plc for funding this project.

Abstract

This thesis describes the optimisation and implementation of a conditional transgene expression system in mouse embryonic stem (ES) cells for the investigation of signalling pathways controlling *in-vitro* differentiation. This system, the tetracycline-induced conditional gene expression system (tet-system) allows exquisite control over transgene expression *in vitro* and *in vivo*. The system itself is comprised of two components, a transactivator gene, and a transactivator-response element (TRE) which is linked to the transgene to be regulated. The transactivator gene encodes a fusion protein consisting of a modified tetracycline repressor and transcriptional activation domain moieties, and is constitutively expressed. In the presence of doxycycline (dox) the transactivator protein binds to the TRE sequence, which itself consists of an array of operator sequences from the tetracycline operon linked to a minimal promoter region. Transcription of the transgene consequently ensues. This version of the tet-system is referred to as the tet-on system.

In order to establish high and homogenous expression of inducible transgenes in ES cells using the tet-on system, the utilisation of different promoters to direct expression of the transactivator component (rtTA2^s-M2) was investigated, as was the use of different loci at which to target the system components.

It was found that the promoter sequence CAG (which contains a CMV enhancer, a β -actin promoter and a β -globin intron), when linked to the transactivator rtTA2^s-M2 sequence and targeted to the 5' region of the housekeeping gene *hypoxanthine ribosyltransferase (hprt)* gave rise to homogenous expression of the transactivator. However upon targeting the 3' region of *hprt* with the TRE and a linked DsRED2 reporter transgene, addition of dox resulted in heterogenous expression of the transgene.

In an alternative strategy, cells expressing the rtTA2^s-M2 transactivator from the endogenous ROSA26 promoter (R26rtTA2^s-M21F cells), were targeted to integrate the TRE and linked DsRED2 reporter to the 5' region of *hprt*. In cell lines harbouring both components of the tet-system, addition of dox resulted in high and homogenous expression of the DsRED2 transgene. This latter strategy of targeting components to independent loci was therefore superior in terms of inducible transgene expression, and was subsequently used to establish the inducible expression of an activated form of the MAPK pathway component MEK1 (MEK_{act}). Cell lines overexpressing MEK_{act} were thereby analysed for activation of the MAPK pathway in the absence of growth factor and cytokine stimulation, with the view to analysing the involvement of the MAPK pathway in the *in vitro* differentiation of ES cells into neural progenitor cells. However, analysis of ES cell lines overexpressing MEK_{act} revealed that the levels of expression achieved were not sufficient to activate the pathway above basal levels – as determined

by analysis of levels of phosphorylated (and therefore activated) ERK1/2, the only known substrate of MEK1/2.

Abbreviations

6-TG	6-thioguanine
bp	base pair
cDNA	complementary DNA
DNA	deoxyribonucleic acid
Dox	doxycycline
ERK	extracellular-signal regulated kinase
ES cells	Embryonic stem cells
HAT	hypoxanthine, aminopterin and thymidine
<i>hprt</i>	<i>hypoxanthine ribosyltransferase</i>
HSV	herpes simplex virus
ICM	Inner cell mass
IRES	Internal ribosome entry site
kb	kilobase
MAPK	Mitogen-activated protein kinase
MEK	MAPK/ERK kinase
PCR	Polymerase chain reaction
rtTA	reverse tetracycline-controlled transactivator
RNA	ribonucleic acid
SA	splice acceptor
SD	splice donor
Tc	tetracycline
<i>tetO</i>	tetracycline operator sequences
TetR	tetracycline repressor
TRE	transactivator-response element
tTA	tetracycline-controlled transactivator

Table of Contents

Declaration	
Acknowledgments	
Abstract	
Abbreviations	
Table of contents	
List of figures	

Chapter 1	1
INTRODUCTION	1
1.1 Genetic engineering <i>in vitro</i>	1
1.2 The creation of transgenic organisms	4
1.3 The application of conditional technology	6
1.4 The technology used to conditionally control gene expression	7
1.4.1 Recombinase mediated site-specific recombination	8
1.4.2 Inducible gene expression systems	9
1.5 The tetracycline-repressor based conditional gene expression system.....	15
1.5.1 The tet-system – general overview.....	16
1.5.2 Development of the tet-off system.....	16
1.5.3 Development of the tet-on system	19
1.5.4 Utilisation of the tet-on or tet-off system.....	20
1.5.5 The use of VP16 minimal activation domains.....	22
1.5.6 The development of rtTA2 ^S -M2 and rtTA2 ^S -S2	23
1.5.7 The development of transcriptional silencers	24
1.5.8 Application of the tet-system	26
1.6 Previous establishment of the tet system in ES cells	26
1.7 The targeting of both tet system components in ES cells	29
Chapter 2	32
THE TARGETING OF THE TRANSACTIVATOR TO THE 5' <i>HPRT</i> LOCUS IN E14TG2A CELLS	32
2.1 Introduction.....	32
2.2 The cytomegalovirus promoter and enhancer to drive expression of the rtTA2 ^S -M2 transactivator	33
2.3 The <i>hprt</i> targeting vectors.....	33
2.4 Cloning steps involved in the construction of the reverse transactivator containing targeting vectors.....	35
2.5 The targeting of p5' <i>hprt</i> CMV ^{rep} and p5' <i>hprt</i> CMV ^{ins} into E14TG2a cells	36
2.6 Analysis of rtTA2 ^S -M2 expression by RT-PCR	38

2.7	Analysis of rtTA2 ^s -M2 function by transient luciferase assay	38
2.8	Flow cytometry of targeted clones.....	39
2.9	The CAG promoter to drive expression of the reverse transactivator	40
2.10	Flow cytometry of the CAGrep cell lines.....	41
2.11	Discussion.....	42
Chapter 3	48
	INTEGRATION OF THE TRANSACTIVATOR-RESPONSE ELEMENT INTO THE 3' CODING REGION OF <i>HPRT</i>	48
3.1	Introduction.....	48
3.2	Construction of the 3'hpTREDsRED2F and 3'hpTREDsRED2R targeting vectors and electroporation into CAGrepB7 cells	49
3.3	Flow cytometry of ihpTREDsRED2 cell lines	51
3.4	DsRED2 induction in a gradient of concentrations of dox.	52
3.5	Sorting of the DsRED2 positive cells	52
3.6	Induction of DsRED2 after different periods of culture in dox.....	54
3.7	Discussion.....	54
Chapter 4	65
	A BIALLELIC APPROACH TO IMPLEMENTING THE TET-SYSTEM	65
4.1	Introduction.....	65
4.2	Construction of the 5'hpTREDsRED2F and 5'hpTREDsRED2R targeting vectors and electroporation into R26- rtTA2 ^s -M2 1F cells.....	66
4.3	Flow cytometry analysis of iDsRED2 cell lines in the presence and absence of dox	67
4.4	Induction of DsRED2 in a gradient of doxycycline concentrations	68
4.5	Induction of DsRED2 after different periods of culture in dox	69
4.6	Discussion.....	69
Chapter 5	73
	THE ESTABLISHMENT OF INDUCIBLE MEK1 _{ACT} EXPRESSION IN ES CELLS FOR INVESTIGATION INTO THE INVOLVEMENT OF THE MAPK PATHWAY IN NEURAL INDUCTION.....	73
5.1	Introduction.....	73
5.2	Construction of the 5'hpTREMEEK ^{act} and 5'hpTREMEEK ^{wt} targeting vectors	75
5.3	The targeting of the vectors to the 5' region of <i>hpT</i>	77
5.4	Flow cytometry in the presence and absence of dox	78
5.5	Induction in a gradient of dox concentrations and timecourse	78
5.6	Retargeting R26-rtTA2 ^s -M21F cells with iDsRED2 vectors	79
5.7	Immunohistochemistry of iMEKact and iMEKwt cell lines.....	79
5.8	Immunoblotting of iMEKact and iMEKwt cell lines	80
5.9	Discussion.....	82

Concluding notes.....	88
------------------------------	-----------

Materials and Methods.....	92
7.1.1 Molecular Biology Materials	92
7.1.2 Enzymes.....	93
7.1.3 Plasmids.....	93
7.1.4 Probes for Southern blot hybridisation	93
7.1.5 Primers	94
7.1.6 Bacterial strains.....	94
7.1.7 Bacterial growth media.....	94
7.1.8 Antibiotics.....	94
7.1.9 Standard stock solutions and buffers	95
7.2.1 Molecular Biology Methods	95
7.2.2 Preperation of nucleic acids.....	95
7.2.3 General cloning methods	96
7.2.4 Preparation of chemically-competent cells.....	98
7.2.5 Transformation of chemically-competent XL-1 Escherichia coli	98
7.2.6 Transformation of chemically-competent recombinant Stb12 TM Escherichia coli.....	99
7.2.7 Screening of bacterial colonies – Colony lifts	99
7.2.8 Lysis of ES cells and extraction and digestion of genomic DNA.....	100
7.2.9 Electrophoresis of genomic DNA and transfer onto nylon membrane	101
7.2.10 Random-primed ³² P-labelling of oligonucleotides	102
7.2.11 Hybridisation of southern blot membrane	102
7.2.12 Reverse transcription	102
7.2.13 Polymerase Chain Reaction (PCR).....	102
7.2.14 Antibody staining of ES cells	103
7.2.15 Immunoblotting by Western protocol.....	103
7.2.16 Sequencing.....	104
7.3.1 Tissue Culture	104
7.3.2 E14TG2a cell line	105
7.3.3 General tissue culture solutions	105
7.4.1 Growth conditions during ES cell culture.....	106
7.4.2 Passage, expansion and freezing of ES cells	106
7.4.3 ES cell passage, expansion and freezing in 96 well plate format (1:5 passage).....	107
7.4.4 Revival of frozen ES cells.....	108
7.4.5 Revival of frozen ES clones from a 96 well plate	108
7.4.6 Vector preparation of plasmid DNA for targeting experiments	108
7.4.7 Transfection of ES cells by electroporation protocol	109
7.4.8 HAT selection of targeted E14TG2a ES cells	109
7.4.9 6-TG selection of targeted ES cells	110
7.4.10 Picking ES cell colonies.....	110
7.4.11 Transient transfection of ES cells by lipofectamine TM protocol.....	110
7.4.12 Luciferase analysis.....	111
7.4.13 Preperation of cells for immunoblotting.....	111

Appendix I.....	112
Bibliography.....	116

Figures

- Figure 1.1 The targeting of endogenous loci
- Figure 1.2 Examples of Cre-recombinase site-specific recombination
- Figure 1.3 Binary inducible gene expression systems
- Figure 1.4 The ecdysone and GLVP inducible gene expression systems
- Figure 1.5 The lac-repressor based and CreER^{T2} inducible gene expression systems
- Figure 1.6 The tet-on and tet-off systems
- Figure 1.7 A selected panel of transactivators and reverse transactivators
- Figure 1.8 The tet-responsive element
- Figure 1.9 Dose response of transactivators to tetracycline or doxycycline
-
- Figure 2.1 The pDWM101 and pSKB1 targeting vectors
- Figure 2.2 Targeting of pDWM101 and pSKB1 to *hpert*
- Figure 2.3 Cloning of CMV-rtTA2^s-M2-IRES-hrGFP-pA and insertion into targeting vectors
- Figure 2.4 Targeting of p5'*hpert*CMVins to the *hpert* locus
- Figure 2.5 Targeting of p5'*hpert*CMVrep to the *hpert* locus
- Figure 2.6 Analysis of rtTA2^s-M2 expression
- Figure 2.7 Flow cytometry of targeted cell lines
- Figure 2.8 Cloning of CAG-rtTA2^s-M2-IRES-hrGFP-pA and insertion into targeting vector
- Figure 2.9 Targeting of p5'*hpert*CAG to *hpert*
- Figure 2.10 Flow cytometry of targeted cell lines
-
- Figure 3.1 Cloning of TRE and linked reporter sequences into the M42 vector
- Figure 3.2 Electroporation of 3'*hpert*TREdsRED2F and 3'*hpert*TREdsRED2R into CAGrepB7 cells
- Figure 3.3 Flow cytometry of ihprtDsRED2F cell lines
- Figure 3.4 Flow cytometry of ihprtDsRED2R cell lines
- Figure 3.5 DsRED2 expression after culture in a gradient of dox concentrations

- Figure 3.6 Sorting of high DsRED2-expressing cells
- Figure 3.7 Timecourse experiment
-
- Figure 4.1 Targeting of the reverse transactivator to the ROSA26 locus
- Figure 4.2 Cloning of TRE and linked reporter sequences into the pSKB1 vector
- Figure 4.3 Electroporation of 5'hprtDsRED2F and 5'hprtDsRED2R into the R26-rtTA2^S-M21F cell line
- Figure 4.4 Flow cytometry of iDsRED2F cell lines
- Figure 4.5 Flow cytometry of iDsRED2R cell lines
- Figure 4.6 DsRED2 expression after culture in a gradient of dox concentrations
- Figure 4.7 Scatter plots of DsRED fluorescence
- Figure 4.8 Timecourse experiment
-
- Figure 5.1 The MAPK/ERK pathway
- Figure 5.2 Juxtaposition of HA-MEK to the TRE sequence
- Figure 5.3 Creation of the 5'hprtTREMEEKact and 5'hprtTREMEEKwt vectors
- Figure 5.4 Electroporation of 5'hprtTREMEEKact and 5'hprtTREMEEKwt into rtTA2^S-M21F cells
- Figure 5.5 Flow cytometry of iMEKact and iMEKwt cell lines
- Figure 5.6 DsRED2 expression after culture in a gradient of dox concentrations
- Figure 5.7 Timecourse experiment
- Figure 5.8 Scatter plots of DsRED fluorescence
- Figure 5.9 Immunohistochemistry of iMEKact and iMEKwt cell lines
- Figure 5.10 Immunoblotting of targeted cell lines
-
- Figure 6.1 Single-allele approach to implementing the tet-system
- Figure 6.2 Cassette-exchange strategy

Chapter 1

Introduction

The application of reverse genetics in the study of gene function has greatly improved the understanding of many areas of biology, and the genetic engineering technologies that can be used in this approach are growing evermore sophisticated. The purpose of this introduction is to present the reader with an understanding of the strategies available for manipulating gene function in mammals both *in vitro* and *in vivo*, starting with a review of the basic types of genetic modification, and then leading to a description of more recently developed approaches, with particular focus on methods for induced conditional gene expression.

1.1 Genetic engineering *in vitro*

The cellular genome can be genetically manipulated in a wide variety of ways, ranging from the relatively simple scenario of constitutively overexpressing a transgene, to the sophisticated implementation of a conditional gene expression system (the description of which will be described later on). Such manipulations usually involve the transfection of

foreign sequences into cells followed by selection for cells harbouring stable integration of the exogenous sequences into the genome. This selection is usually made possible either by the incorporation of an antibiotic resistance selection marker into the sequences to be inserted or by co-transfection of a separate antibiotic resistance gene encoding plasmid. Those cells resistant to the applied antibiotic will have incorporated the exogenous sequences and these can be further expanded in appropriate selective media. An exception to this is the use of retroviral vectors in the delivery of foreign sequences into genomic DNA. This form of sequence delivery can be very efficient and therefore in some instances there is no need to select for integration.

The constitutive overexpression of a complementary DNA (cDNA) sequence is often achieved by the transfection of DNA encoding a promoter and downstream transgene, into cells (along with a selectable marker as mentioned previously), followed by suitable selection. This will result in the isolation of a number of clones all harbouring insertions at different locations within the genome and therefore this process is referred to as 'random' integration. As there is no control over the site of sequence insertion, random integration carries with it the risk of insertional mutagenesis, due to the possibility of insertion into, or in close proximity to endogenous genes, leading to abrogation of, or misregulation of gene expression. In addition, independent clones harbouring the same exogenous sequences are likely to express different levels of the transgene as the foreign sequences will be subject to different position effects due to their random integration throughout the genome. This calls for the screening of large numbers of clones in the search for those that express the cDNA appropriately. Multiple copy number at the site of insertion also frequently occurs.

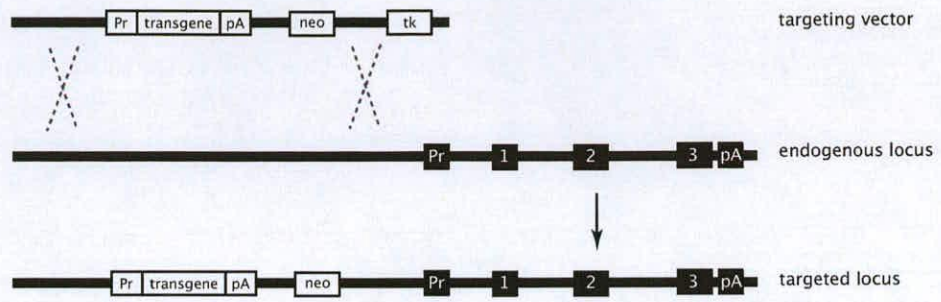
These potential problems can be overcome by the 'targeted' insertion of transgenic sequences to a predetermined locus which is known to be optimal for transgene expression. This form of insertion usually results in the integration of just one copy of exogenous sequence into the genome. For targeted insertion to a predetermined locus, foreign sequences are inserted into a vector that encodes homology to that locus. Upon transfection of the targeting vector, in a small proportion of cells, crossing-over events

occur between regions of homology encoded within the endogenous locus and within the vector sequences (Figure 1.1a). Those cells harbouring integrated foreign sequences can subsequently be isolated in appropriate selection media and clones in which correct targeting of the locus has occurred can be identified by analysis of their genomic DNA. Homologous recombination can be used to target a cDNA to a locus known to be permissive for transgene expression (Figure 1.1a). The resulting clones should all express similar levels of transgene as they are genetically identical, and this therefore obviates the need to screen large numbers of targeted clones for appropriate expression of the transgene.

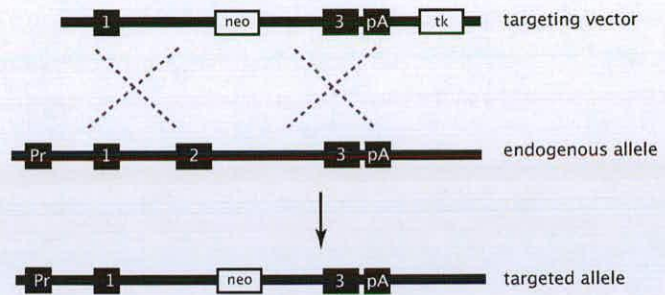
In some cases the disruption of an endogenous gene is desired, for example in the creation of a null allele. Gene targeting technology makes this possible as targeting vectors can be designed to 'knock-out' endogenous exons (Figure 1.1b). A second round of targeting to the remaining wild-type allele results in the establishment of cell-lines homozygous for the mutation in the gene of interest. This enables investigation into the function of this gene by the direct comparison of these modified cells to cells possessing wild type alleles. Homologous recombination can also be used to add sequences to endogenous genes, resulting in the 'tagging' of endogenous protein with antigenic epitopes or fluorescent moieties which may aid in the detection of that protein (Figure 1.1c). The targeting vector has to be carefully designed to insert the exogenous sequence in frame with the endogenous gene at the desired locus (the exogenous sequence is usually inserted at the beginning or the end of the ORF). Transcription and subsequent translation of this unit therefore results in the production of a fusion protein, with one moiety comprising the cellular protein and the other, the foreign protein. In an alternative scenario gene targeting can be used to insert transgenes downstream of endogenous regulatory sequences so that the expression of the foreign sequences mimics that of the endogenous gene (Figure 1.1.d). This strategy is often used to spatially restrict transgene expression *in vivo* (see below).

Figure 1.1 The targeting of endogenous loci. a) The hypothetical targeting of a transgene to an endogenous locus that is known to be permissive for transgene expression. The targeting vector encodes the transgene sequence with linked promoter and polyA sequence and also a neomycin resistance gene (depicted as neo) and a herpesvirus thymidine kinase gene (depicted as tk). Promoter and polyA sequences are not shown for either of the two selection genes. The endogenous locus incorporates the exogenous sequences after two cross-over events occur between homologous regions of the locus and targeting vector. This event can be selected for using positive G418 selection. Illegitimate recombination can be selected against by using negative ganciclovir selection (in a correct targeting event the tk gene is not incorporated). b) Creation of a knock-out or null allele by the removal of a coding region after homologous recombination between the targeting vector and endogenous locus. Correctly targeted clones can be selected for using positive G418 selection and negative ganciclovir selection. c) 'Tagging' of an endogenous gene by insertion of an epitope tag downstream of the coding region. For simplicity antibiotic resistance transgenes are not shown. d) 'Knock-in' of a reporter sequence to an endogenous locus. Correct targeting of the reporter transgene results in its regulation by the endogenous promoter. The endogenous allele is no longer expressed due to the replacement of vital coding sequences with the transgene ORF. Antibiotic resistance transgenes are not shown.

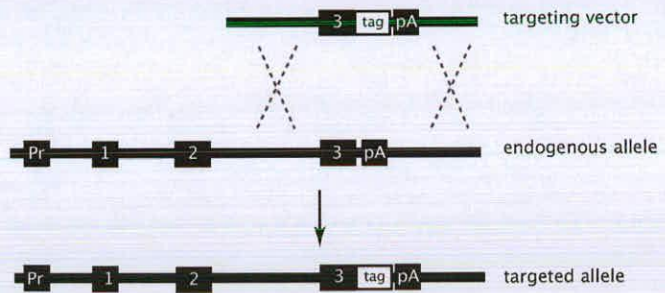
a) Targeted constitutive gene expression



b) Creation of a 'knock-out' allele



c) Creation of a 'tagged' endogenous allele



d) The 'knock-in' of a reporter transgene to an endogenous locus



1.2 The creation of transgenic organisms

Currently transgenic mice can be created by three different routes, pronuclear injection, an ES cell-mediated route, and nuclear transfer. Pronuclear injection is not only restricted to mice, but can be used to create transgenic mammals of any species and is carried out by the microinjection of foreign DNA (usually a transgene and linked promoter), into either pronuclei of a one cell embryo. Animals that have incorporated the foreign DNA into their genome and carry the manipulation through the germline (the tissue that produces gametes), are referred to as founder animals, and these can be used to establish transgenic lines heterozygous and homozygous (if this is not lethal) for the manipulation. All founder animals are genetically dissimilar as the foreign DNA integrates at random positions within the genome. As a result transgene expression is likely to vary between founder animals as the exogenous sequences are subject to different position effects (in a similar manner to the random integration of exogenous sequences *in vitro*). In addition concatemerization of the foreign sequences before integration into the genome often takes place, resulting in the incorporation of multiple copies of the transgene (Brinster *et al.*, 1981). Large numbers of repeats have previously been reported to lead to transgene silencing (Mehtali *et al.*, 1990).

The creation of transgenic organisms via pronuclear injection is a well established and generally applicable technology. However the type of genetic modification is limited to the random integration of foreign sequences. The establishment of murine embryonic stem (ES) cell lines in 1981 (Martin., 1981; Evans and Kaufman., 1981), paved the way for the creation of transgenic mice containing directed genetic alterations of endogenous genes. ES cells are derived from the inner cell mass (ICM) of the early embryo, a transient population of cells that eventually gives rise to all the differentiated cells types found in the adult mouse. ES cell lines are established by plating out early blastocysts *in vitro* and subsequently disaggregating and culturing undifferentiated outgrowths arising from these blastocysts. Upon replating, these outgrowths often give rise to undifferentiated colonies that can be expanded further to give rise to ES cell lines which can be continuously passaged in the presence of a feeder layer of fibroblasts, or

leukaemia inhibitory factor (LIF) cytokine, and serum. These ES cell lines retain the pluripotent nature of their ICM precursors. When injected into early blastocysts they are able to contribute to all cell lineages including the germ layer. Unlike the embryonic ICM which is a transient population, ES cells can be cultured indefinitely, and thus provide an invaluable tool for the creation of transgenic mice. The genetic manipulations created in ES cells, for example by the gene targeting methods described above, can be transferred to the mouse germline by injecting these cells into early blastocysts. These embryos will give rise to chimeric mice which contain two genetically distinct populations of cells. In a successful chimera cells derived from the manipulated ES cells will colonise the germline, and the chimera can be used to create by subsequent breeding heterozygous and homozygous transgenic mouse lines. ES cells can therefore be used as a vehicle to transfer complex genetic manipulations to the murine germline, and as a result the advent of this technology revolutionised the field of reverse genetics.

The third route of creating transgenic animals is the process of nuclear transfer (NT), which involves the transfer of a nucleus from a genetically manipulated cell (somatic or embryonic) into an enucleated egg. Although several species of transgenic animal have been created by this method, including mice (Rideout *et al.*, 2000), goats (Keefer *et al.*, 2001), cows (Krisher *et al.*, 1995) and pigs (Uhm *et al.*, 2000), the efficiency of this process is very low, with success rates of around 2-4 live-born animals per 100 NT embryos (Wolf *et al.*, 2001). In addition postnatal abnormalities often arise, such as increased body weight and premature ageing. Whilst NT enables the manipulation of endogenous loci by homologous recombination in those animals which have not successfully generated ES cell lines (Denning *et al.*, 2001, Dai *et al.*, 2002), the low success rate of obtaining live-born animals and prevalence of postnatal problems at present prevent this technology from becoming a practical alternative to ES cell-based transgenesis in mice.

1.3 The application of conditional technology

The approaches described above have all described *constitutive* genetic manipulations, that is the permanent overexpression, or mutation of a gene upon stable engineering of a cell line, or in an *in vivo* setting, the modification or inactivation of a gene in all tissues of the body, for the entire lifespan of the organism. For reasons that will be elaborated on further, constitutive genetic manipulations are not always desirable when attempting to elucidate gene function or establishing models for human disease. Recently methods have been devised to spatially and temporally restrict either the inactivation of an endogenous gene, or the expression of an exogenous transgene. The technology used to bring about such conditional manipulations will be discussed in detail below. However the advantages of using conditional strategies over constitutive approaches will first be discussed.

In the study of gene function *in vivo*, the spatially-restricted inactivation of a widely expressed endogenous gene enables the study of the function of that gene in the tissue of interest. This approach is superior to the use of constitutive null mutants as interpretation of gene function can be hindered by the indirect effects of inactivating the gene in surrounding tissues. In addition the timing of gene inactivation can be very important as in extreme cases constitutive inactivation of a gene will result in embryonic lethality which will preclude analysis into its requirement at later stages of development. Both these points are exemplified in a study investigating FGF8 signalling in the developing limb bud in which *fgf8* gene inactivation was conditionally restricted to early limb ectoderm (Lewandoski *et al.*, 2000). This approach not only circumvented the requirement for FGF8 during gastrulation but it also restricted *fgf8* ablation to the tissue of interest without compromising its other functions in the developing organism.

Studies of gene function by the constitutive inactivation of an endogenous gene can sometimes be hindered by the compensatory effects of closely related genes. Knock-out mice, for example sometimes exhibit less severe phenotypes than their conditional counterparts, owing to adaptive mechanisms that occur during development. Such

compensatory mechanisms were recently reported in a study by Sage *et al.*, which highlighted the problem of using transgenic animals containing constitutive inactivating mutations as models for cancer. It was found that mouse embryonic fibroblasts (MEFs) that underwent an acute (conditional) inactivating mutation of the retinoblastoma (Rb) gene, responded differently to various imposed cellular states (such as senescence and quiescence), than MEFs harbouring a constitutive mutation. This was found to be due partly to compensation by the upregulation of the Rb-related p107 gene in the constitutively mutant Rb line. Although such compensation also occurred in MEFs exhibiting acute Rb mutation after time in culture, the analysis of this cell line immediately after mutation more accurately recapitulated the situation in cancer, which is often caused by the sporadic genetic mutation of Rb and other oncogenes involved in the regulation of the cell cycle.

Conditional manipulations are not solely limited to the inactivation of genes. Indeed the spatiotemporal restriction of transgene expression can lead to far more sophisticated experimental approaches than those possible by simple constitutive overexpression. Ultimately the experimental approach will depend upon the study in question. The availability of a vast repertoire of conditional technology however provides greater experimental flexibility.

1.4 The technology used to conditionally control gene expression

Genetic manipulations can be rendered conditional, principally by the utilisation of two types of technology, one which is largely used to permanently inactivate alleles or reactivate a transgene in a conditional manner (using recombinase-mediated site-specific recombination), and the other which is used to manipulate the expression of a transgene in a reversible manner, the timing of which is up to the discretion of the experimenter (this involves the use of conditional inducible gene expression systems). Here both types of technology shall be discussed with a focus on inducible gene expression systems. It should be noted that recently the separate identity of these two technologies has become blurred with the development of inducible recombinase-based technology (Imai *et al.*,

2001, el Marjou *et al.*, 2004) and recombinase activated inducible systems (Belteki *et al.*, 2005, Guo *et al.*, 2005).

1.4.1 Recombinase mediated site-specific recombination

Recombinase mediated site-specific recombination involves the use of either Cre-recombinase (found in bacteriophage P1) or Flp recombinase (found in *Saccharomyces cerevisiae*), which act on pairs of corresponding *loxP* and *FRT* sites respectively. Both systems are based on the same principle, therefore only the Cre/*loxP* system shall be expanded on here.

Cre-recombinase is a site-specific protein that binds to 34 bp *loxP* recognition sites (two Cre proteins bind per site), and is able to mediate conservative recombination between a pair of *loxP* sites. The recognition sequences comprise 13 bp inverted repeat sequences, which flank an asymmetric 8 bp core sequence conferring directionality to the site. The positioning of *loxP* sites either side of coding regions (or non-coding regions that are essential to the expression of that gene), orientated in the same direction, will lead to inactivation of the gene due to excision of the intervening DNA sequence between them upon Cre-mediated recombination. This technique is referred to as '*floxed*' a gene and is widely used to achieve conditional gene inactivation by expressing Cre in a temporally and/or spatially restricted manner.

Cre-mediated inactivation of an endogenous gene can be achieved *in vitro* by introducing Cre-encoding sequence into a cell line harbouring a '*floxed*' allele, via transfection of a Cre-expressing plasmid, or introduction of the sequence using a viral vector (Sage *et al.*, 2003). Recently a technique has been developed for efficiently introducing proteins into cells via fusion to the protein transduction domain (PTD) of the HIV tat protein, (Schwarze *et al.*, 1999). The fusion of Cre to this PTD domain therefore offers an alternative mode of delivery of the recombinase.

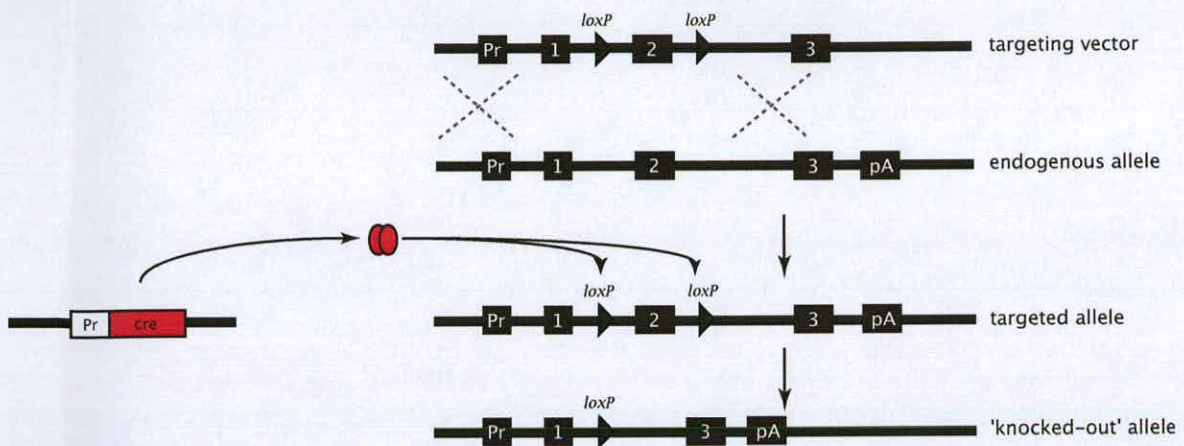
Site-specific recombination has been widely applied *in vivo* to inactivate endogenous genes in a spatiotemporal manner (Figure 1.2a). This is usually achieved by creating two transgenic lines, one containing the ‘*floxed*’ allele, and the other expressing Cre-recombinase in a specific tissue (this is done by utilising a tissue-specific promoter to drive Cre expression) and in addition harbouring a mutated ‘null’ allele of the gene of interest. The breeding of these two lines will result in a proportion of the progeny expressing Cre-recombinase and containing one endogenous ‘*floxed*’ allele and one null allele. This will result in inactivation of the conditional allele in every Cre expressing tissue (Gu *et al.*, 1994, Trumpp *et al.*, 1999).

This recombinase-mediated technology can also be applied to reactivate transgenes or endogenous genes that have been engineered so that expression is blocked by the insertion of a foreign ‘stop’ sequence, which is usually a selection cassette which contains a transcriptional or translational stop (Figure 1.2b). For the activation of a transgene, the transgenic line is usually engineered for near ubiquitous expression of the gene, and once established this is crossed to an appropriate tissue-specific Cre-expressing line (Grieshammer *et al.*, 1998, Heidt *et al.*, 2005). Alternatively endogenous loci can be manipulated in a similar manner. Endogenous gene activation has been used to create mouse models of human skin-blistering disorders by ubiquitously mutating endogenous genes involved in these disorders and then expressing the mutant alleles in a tissue-specific manner (Cao *et al.*, 2001).

1.4.2 Inducible gene expression systems

The first inducible gene expression systems developed were single transgenes that were linked to a promoter responsive to an external stimulus or inducer, such as heat shock, heavy metals or hormones. However, the use of such activating stimuli and inducing agents resulted in pleiotropic effects in the target cells due to interference with endogenous cellular pathways. Furthermore, promoters driving the transgenic sequences were also prone to activation by endogenous transcription factors resulting in high basal transgene expression in the absence of inducer and therefore the use of these systems in

a) The establishment of a floxed endogenous allele and subsequent conditional inactivation



b) The inactivation of an endogenous allele and subsequent conditional activation

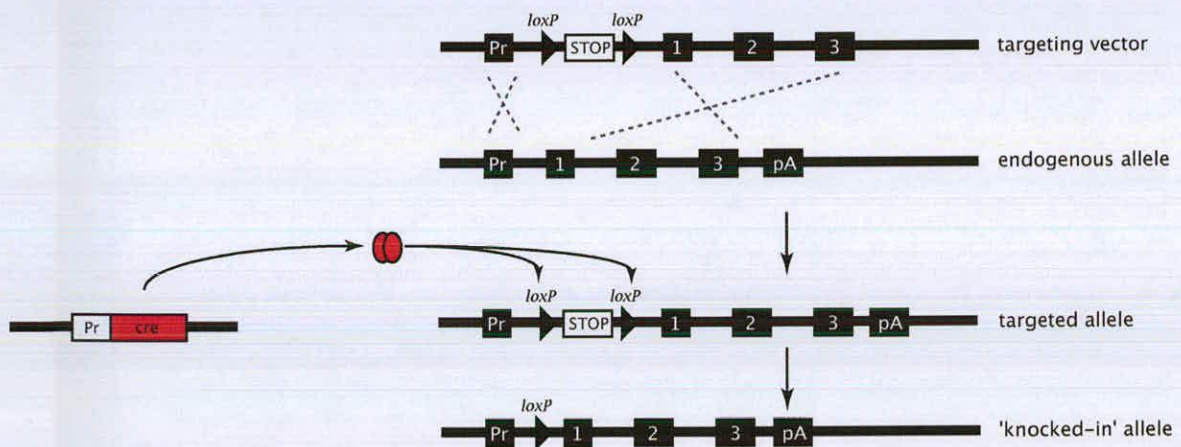


Figure 1.2 Examples of Cre-recombinase mediated site-specific recombination. a) Creation of a hypothetical conditional knock-out allele. A targeting vector encoding sequence homologous to the endogenous locus and two homospesific *loxP* sites flanking exon 2 is transfected into cells. Legitimate recombination results in the 'floxed' of exon 2. The presence of cre-recombinase results in the excision of exon 2 by site-specific recombination between the *loxP* sites. b) Hypothetical reactivation of an endogenous gene. A targeting vector encoding a transcriptional 'stop' cassette flanked by two *loxP* sites is transfected into cells. Legitimate recombination results in the integration of the 'floxed' stop cassette between the endogenous promoter and exon 1, creating a 'null' allele due to early termination of transcription. Cre-recombinase action results in the removal of the stop cassette and expression of the gene. For simplicity selection cassettes are not depicted for either of the strategies.

the conditional expression of transgenes was limited (these systems are reviewed in Yarranton., 1992).

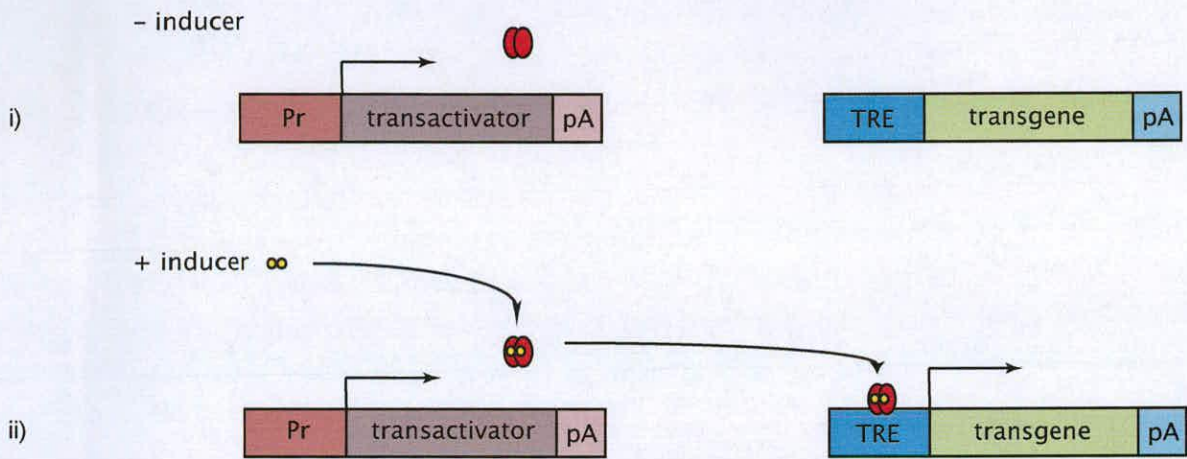
More advanced forms of conditional gene expression were subsequently developed, based on the interaction of two components. Most of these systems utilise a transactivator that responds to the addition or withdrawal of inducer by binding specifically to a transactivator-response element (TRE) (Figure 1.3). cDNA encoding a transgene is linked to the TRE and upon sequence specific binding by the transactivator, expression of the transgene is induced, (a repressor-based system has also been developed, which will be described at a later point).

The early versions of these binary systems utilised viral proteins as transactivators and were prone to high levels of basal transgene expression and tumorigenesis (Khillan *et al.*, 1988). Today most inducible gene expression systems rely on chimeric transactivator proteins comprised of a combination of eukaryotic/bacterial and viral sequences, which are engineered so as to not interfere with endogenous cellular metabolism. In addition careful selection and modification of transactivator response elements aim to ensure that these sequences are only regulated by the system transactivator and not by endogenous host factors.

Such binary inducible gene expression systems can be established *in vitro* by the integration of both the transactivator (TA), and transactivator-response element (TRE) into the genome. Transcription of the TRE-linked cDNA is initiated upon addition, or withdrawal (depending on the system in question) of inducer to the culture. As expression of the transgene is wholly dependent on the presence or absence of inducer, the experimenter has exquisite control over the transcriptional state of the cDNA, and can therefore turn the transgene on or off as required.

Binary systems can be utilised *in vivo* by generating two transgenic lines, one that expresses the transactivator, and the other that harbours the TRE and linked transgene. The progeny of this cross will result in animals containing both elements of the inducible

a) Addition of inducer activates transgene expression



b) Addition of inducer terminates transgene expression

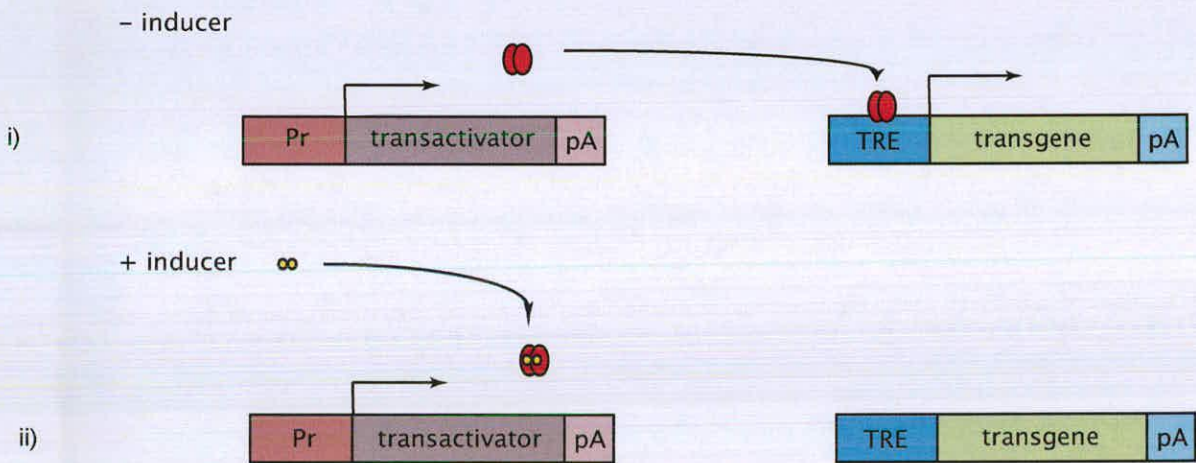


Figure 1.3 Binary inducible gene expression systems. The systems are comprised of two elements, one encodes a transactivator and the other contains the transactivator-response element (TRE) linked to the cDNA to be regulated. The transactivator is expressed irrespective of the presence or absence of inducer. a) A hypothetical system that requires inducer for transgene expression. i) In the absence of inducer the transactivator is not able to bind the TRE, ii) upon addition of inducer an inducer/transactivator complex is formed that binds the TRE and activates transgene expression. b) i) A hypothetical system that overexpresses the transgene in the absence of inducer. ii) The inducer binds the transactivator rendering it incapable of binding to the TRE, thus transgene expression is terminated.

system, and transgene expression can be regulated by the addition of inducer to the animals food or water, intraperitoneal or intravenous injection, or by topical application (the route of administration depends upon the system used and also the experimental design). Tissue-specific expression of the transgene can be achieved in a similar manner to that of the Cre-recombinase system mentioned above, by using a tissue-restricted promoter to drive expression of the transactivator. The crossing of a line expressing tissue-restricted transactivator, to a line with potential ubiquitous TRE-linked cDNA expression results in tissue-specific transgene activation upon application or withdrawal of inducer, (Saam *et al.*, 1999, Shigehara *et al.*, 2003, Nguyen *et al.*, 2005).

The ability to induce the expression of a transgene and switch it off at a later timepoints has opened up new experimental opportunities in many areas of biology. In the field of cancer research for example, the induction and further silencing of oncogenes allows investigation into the requirement of an oncogene not only for initial tumorigenesis, but also for tumour maintenance (Felsher *et al.*, 1999, Chin *et al.*, 1999). Similarly in the creation of transgenic models of disease, the subsequent repression of an overexpressed pathogenic transgene gives insight into the permanency of the disease phenotype and in cases where symptoms gradually diminish, provide models for remission. This was exemplified in a study in which the conditional overexpression of a Gi-coupled human heptahelical receptor in mice was found to model human congestive heart failure (Redfern *et al.*, 1999). It was observed that when the transgene was subsequently switched off there was a dramatic reversal of the illness, which provided a useful model in which to investigate the processes underlying disease regression. Similarly the symptoms of Huntington disease were alleviated in a mouse model expressing an inducible *huntingtin* gene after expression of this transgene was terminated, highlighting the possibility of developing therapeutic treatments to specifically target this gene product, and in addition providing a useful model in which to test any future therapy (Yamamoto *et al.*, 2000).

Inducible transgene technology also offers an exciting tool for the investigation of signalling pathways and regulatory genes involved in early development, and particularly

in the study of cell fate decisions that occur during embryogenesis or during the differentiation of ES cells. This will be discussed further in section 1.7.

1.4.2.1 Types of inducible gene expression systems

Several conditional inducible gene expression systems have been developed and are in use today, however the most successful and widely cited system is the tetracycline inducible gene expression system (hereon referred to as the tet-system) developed by Hermann Bujard and Manfred Gossen. The next section will briefly describe a selection of the inducible binary gene expression systems available with the aim of introducing the reader to the variety of ways of achieving inducible gene expression. This will be followed by a detailed description of the function and development of the tet-system.

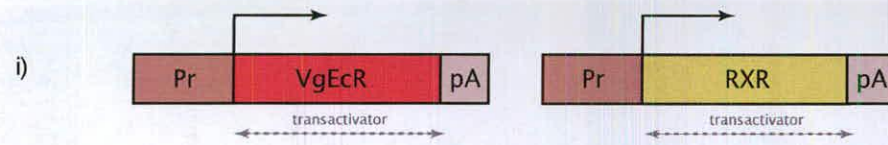
Ecdysone system

The ecdysone-regulated gene expression system is a binary system that was originally operated by the addition of the insect molting hormone ecdysone. The transactivator is comprised of two receptor components, the modified insect ecdysone receptor VgEcR (containing an N-terminal truncation fused to a VP16 activation domain) and the mammalian retinoid X receptor RXR. These two receptors heterodimerise and bind to the E/GRE composite response element upon the addition of ecdysone steroid, an interaction which initiates the transcription of the E/GRE linked transgene (Figure 1.4a).

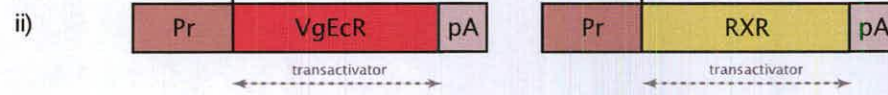
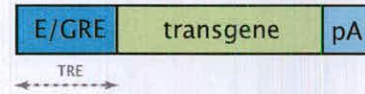
The ecdysone inducible system can activate homogenous expression of β -gal *in vitro* (No *et al.*, 1996), and throughout the mammary epithelium *in vivo* (Albanese *et al.*, 2000). The steroid inducers are able to mediate very rapid inductions of transgenes owing to their ability to efficiently transpass cellular membranes as a result of their lipophilic nature. They are also rapidly eliminated both *in vitro* and *in vivo* to ensure speedy termination of transgene expression (Saez *et al.*, 2000; Hoppe *et al.*, 2000). Although the insect hormone ecdysone and the alternative phytoecdysteroids inducers (ecdysteroids released by plants that 'mimic' ecdysone), have no effect on mammalian physiology, an

Figure 1.4 a) The ecdysone conditional gene expression system. i) The two transactivating elements VgEcR and RXR are expressed constitutively but are unable to bind the transactivator-response element E/GRE due to the absence of inducer. ii) Upon addition of ponasterone A inducer, VgEcR and RXR heterodimerise to form a functional transactivator and bind to E/GRE, activating transcription of the regulatable transgene. b) The GLVP inducible gene expression system. GLVP, a chimeric fusion protein between VP16, GAL4 and a mutant progesterone receptor is expressed constitutively, but unable to bind the UASPr response element due to the sequestration of the fusion protein by heat-shock proteins in the cytoplasm. ii) Addition of RU486 inducer renders GLVP free to translocate into the nucleus, where it activates transgene expression by binding the UASPr transactivator-response element. Abbreviations: Pr, promoter; VgEcR, ecdysone receptor; RXR, retinoid X receptor, E/GRE ecdysone composite response element; pA, polyA; mPr, mutated progesterone receptor; HSP, heat-shock protein; UASPr, GLVP transactivator-response element containing four UAS sequences; TRE, transactivator-response element.

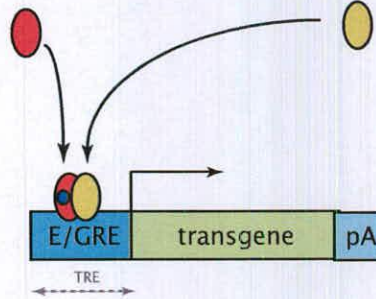
a) Ecdysone system



- ponasterone A

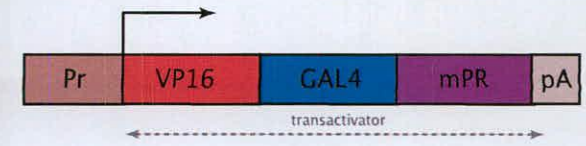


+ ponasterone A ●

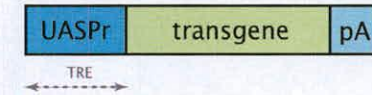


b) GLVP system

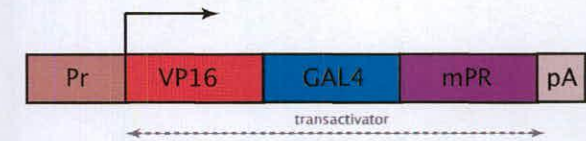
i)



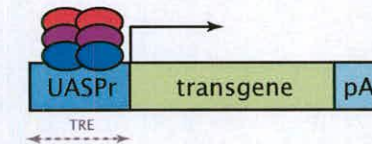
- RU486



ii)



+ RU486



initial concern was the availability of the most potent steroid inducer – muristerone A. The phytoecdysteroid ponasterone A was however found to have the same inductive capabilities as muristerone A and its abundance in the leaves of many species of plant ensures its availability (Saez *et al.*, 2000).

Despite reported inductions of over four orders of magnitude in stably derived clones *in vitro* (No *et al.*, 1996), such levels of induction have not yet been obtained *in vivo*, (Saez *et al.*, 2000; Hoppe *et al.*, 1999). Indeed there is a scarcity of papers describing any physiological effect from overexpressing transgenes using the ecdysone system in cell culture or in transgenic organisms. Also, whilst the penetrance of steroidal inducers are advantageous in certain situations, the downside of this is that sustained levels of transgene expression is difficult to achieve *in vivo* as the inducer is rapidly eliminated from the bloodstream and host tissues. The most promising delivery of inducer to date for continuous transgene expression is via the subcutaneous implantation of pellets (Albanese *et al.*, 2000).

GLVP system

A promising system that has been developed in recent years and has seen widespread application is the GLVP conditional gene expression system. GLVP is the chimeric transactivator component of the system, and is comprised of the GAL4 DNA binding domain, a mutated progesterone receptor and the viral protein VP16 activation domain (Figure 1.4b). GLVP is constitutively expressed but unable to activate transgene expression due to its sequestration by heatshock proteins in the cytoplasm (an interaction that is mediated by the binding of the mutated progesterone receptor moiety). Addition of a progesterone antagonist (such as the synthetic steroid RU486), results in the release of GLVP and enables it to bind to the GLVP response element in the nucleus (that consists of four GAL4 binding sites juxtaposed to a minimal thymidine kinase promoter). This results in activation of transcription of the downstream cDNA.

Numerous examples of utilisation of the GLVP gene expression system have been documented *in vivo* (Wang *et al.*, 1999; Zhao *et al.*, 2001; Pierson *et al.*, 2000), with no,

or low basal levels of transgene expression detected in the absence of inducer. A major disadvantage of this system however, is its restriction to use in adult mice due to the abortigenic properties of synthetic steroids (doses of above 100 µg/kg bodyweight induce abortion in mice – Chua *et al.*, 1998, thus precluding its use in the study of early development). It is not yet clear if this system leads to uniform transgene expression, as it has largely been used to study the overexpression of secreted proteins, however inductions of up to 1500-fold have been obtained *in vivo* (Chua *et al.*, 1998), and levels of expression have been reported to be dose dependent (Zhao *et al.*, 2001).

Lac Repressor-based system

Whilst the ideology of utilising transactivators to mediate transgene expression has been the most popular way of achieving inducible expression in the binary systems developed to date, it is worth commenting on alternative ways that have been explored.

One such strategy is to use a repressor to block expression of an otherwise constitutively expressed transgene by tight binding to part of the transcription unit (Cronin *et al.*, 2001; Dancz *et al.*, 2002). Although this strategy is not as popular as the previous systems described, Cronin *et al* have utilised a system based on the lac repressor (lacI) and lac operator (lacO) interaction to achieve inducible expression of the murine *tyrosinase* gene in mice (Figure 1.5a). Tyrosinase catalyses the first step in melanin biosynthesis, and the authors showed that derepression of the tyrosinase transgene after IPTG inducer ingestion resulted in a coat colour change in adult mice (IPTG binds to lacI and renders it incapable of binding the lacO sequences).

This experiment yielded a very obvious and impressive phenotype, however a limitation of this repressor-based system is the necessary modification to the promoter element driving transgene expression. In order for transgene expression to be repressed, lac operator sequences need to be integrated into the promoter element in a way that will lead to transgene repression upon lacI binding, yet still enable appropriate transgene expression on the application IPTG inducer to the system. It is therefore important that the incorporation of lacO sequences into the promoter element does not impair promoter

function or promoter specificity. In addition the relatively low binding affinity of IPTG inducer for the lac repressor ($\kappa_a 10^6 \sim M^{-1}$), in comparison to other systems, results in only moderate induction (Hu *et al.*, 1987).

Mutant estrogen-receptor based system

Inducible systems based on the mutant estrogen-receptor have largely been used to create ligand-inducible forms of cre-recombinase by developing fusion proteins that comprise of the mutant estrogen-receptor and Cre-recombinase moieties (Cre-ER), as shown in figure 1.5b. Mutation of the estrogen receptor renders it insensitive to activation by endogenous β -estradiol agonist, but not 4-OH-tamoxifen - a synthetic antagonist, (this system is therefore not subject to activation by endogenous pathways). When cDNA sequence encoding this fusion protein is linked to a suitable promoter, expression of the fusion protein is achieved, and the protein is sequestered in the cytoplasm due to the interaction of the estrogen binding domain with the Hsp90 complex (Schwenk *et al.*, 1998). However, upon application of 4-OH-tamoxifen, the fusion protein is released from the complex and able to translocate into the nucleus where the cre-recombinase mediates site specific recombination between identical *loxP* sequences.

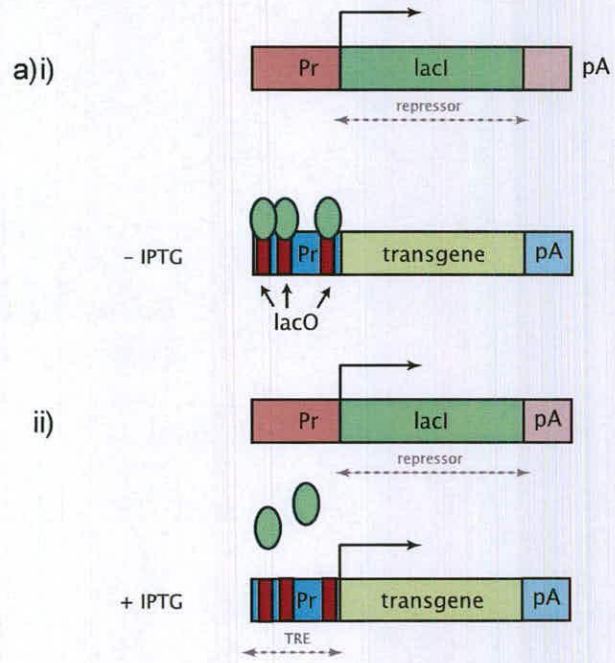
This inducible system has been widely used *in vivo* to temporally regulate somatic mutagenesis by injecting transgenic animals containing 'floxed' alleles and expressing Cre-ER, with tamoxifen inducer (Vooijs *et al.*, 2001, Weber *et al.*, 2001 to name two examples). An improved inducible Cre-recombinase (CreER^{T2}), has recently been developed that is 10-fold more sensitive to tamoxifen induction (Vallier *et al.*, 2001; Leone *et al.*, 2003).

1.5 The tetracycline-repressor based conditional gene expression system

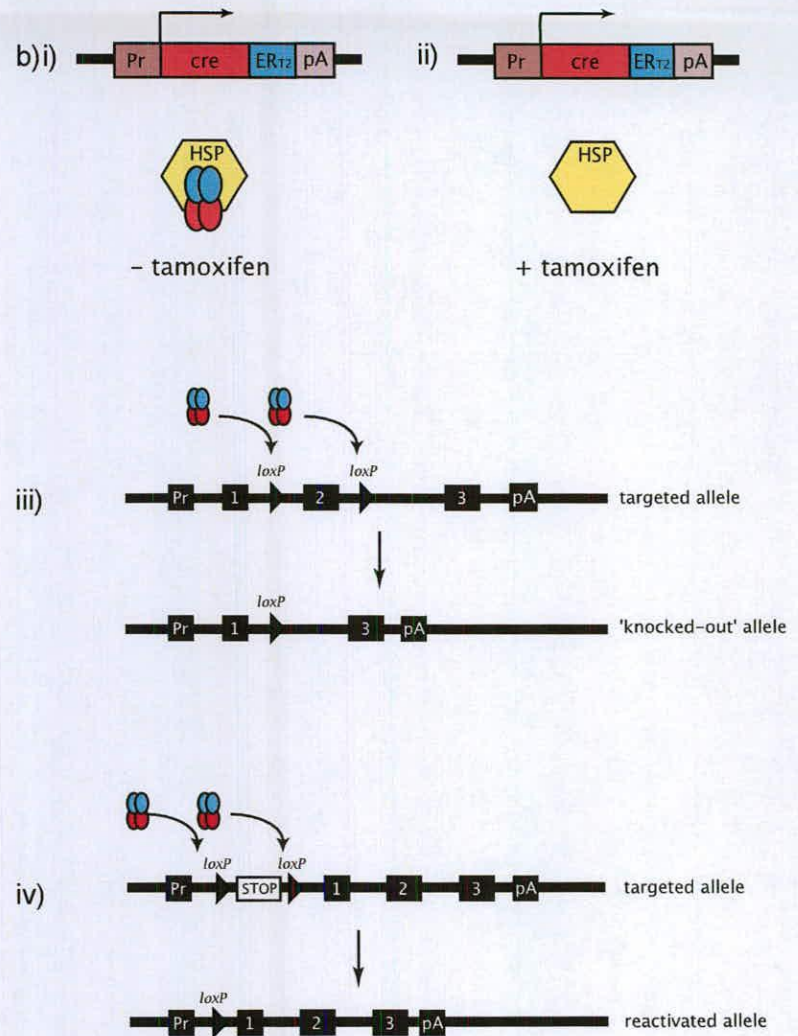
The tetracycline-dependent conditional gene expression systems owe their popularity to their versatility, effectiveness, and ease of manipulation over the other conditional expression systems available. Since the first papers emerged detailing the use of such

Figure 1.5 a) The lac repressor-based inducible gene expression system. i) In the absence of IPTG inducer the lac-repressor (*lacI*) binds to the *lac* operator sequences embedded in the promoter and thus represses transcription. ii) When inducer is added, it binds to the lac repressor and renders it unable to bind the responsive promoter, thus enabling transcription of the transgene to occur. b) The inducible creER^{T2} conditional expression system. i) A fusion protein between cre-recombinase and the mutated estrogen receptor is expressed constitutively but unable to act in the nucleus due to sequestration in the cytoplasm by heat-shock proteins. ii) Addition of tamoxifen inducer renders the chimeric protein free to translocate into the nucleus and mediate site-specific recombination between homospecific *loxP* sites to inactivate or 'knock-out' an allele (shown in biii), or reactivate an endogenous allele or transgene (shown in biv). Abbreviations: Pr, promoter; *lacI*, lac repressor; *lacOPr*, *lacI* responsive promoter containing lac operator sequences; pA, polyA; HSP, heat-shock promoter; ER^{T2}, mutated estrogen receptor; STOP, transcriptional stop cassette (promoter and polyA sequence not shown).

Lac-repressor based system



Inducible CreERT2 system



systems over 10 years ago (Gossen *et al.*, 1992), a lot of work has been carried out to improve the expression system's integral components, firstly in order to achieve tighter control of transgene expression, and secondly to make the tet-systems more attractive for use in a wider number of applications. Here, I will briefly discuss their basic operation, and detail the most important advances that have occurred in the development of these systems to date.

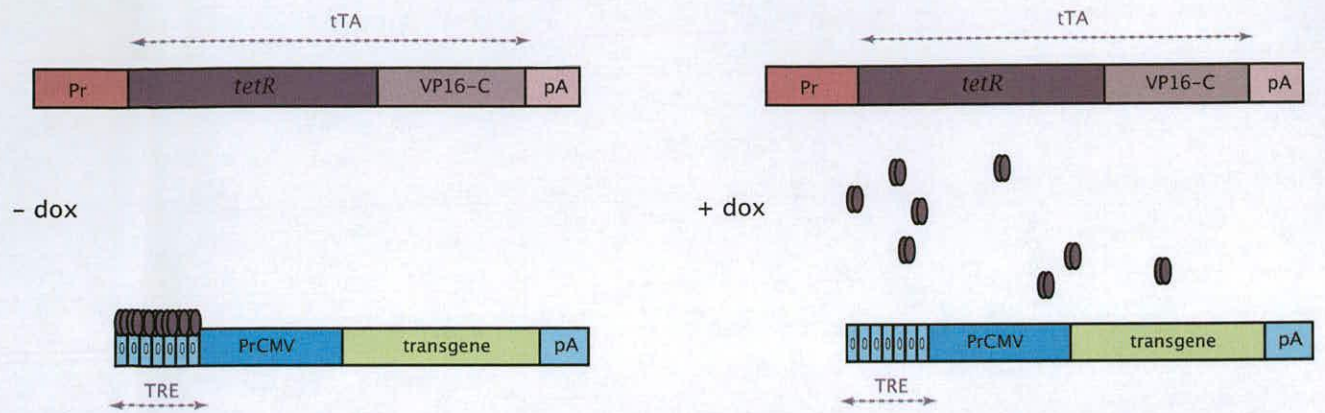
1.5.1 The tet-system – general overview

The tet system is a transactivatory gene expression system which is comprised of a chimeric transactivator and a transactivator-response element (TRE). In accordance with the binary systems previously described, the cDNA to be overexpressed is linked to the TRE element and its expression is induced upon binding of the transactivator to the response element. There are two versions of the tet-system which are commonly used, the tet-off (Figure 1.6a) and the tet-on (Figure 1.6b) systems. The transactivators of both systems are fusion proteins comprising of tet repressor (found in *E.coli* see below), and viral VP16 moieties. The transactivator used in the tet-on system however has reverse DNA binding properties to that of its counterpart utilised in the tet-off system and is therefore referred to as a reverse transactivator. The kinetics of transgene activation and inactivation between the two systems differ, and therefore choice of tet-system depends upon the experimental setup (this will be expanded on below).

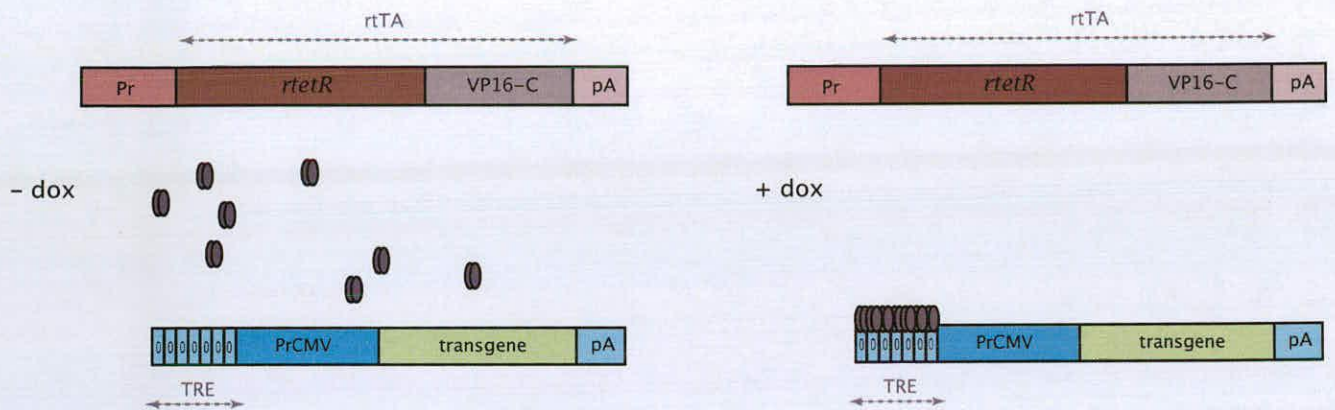
1.5.2 Development of the tet-off system

The original paper detailing the development of the tetracycline inducible conditional gene expression system was published in 1992 by Manfred Gossen and Hermann Bujard. Having previously focussed on the Lac repressor-based inducible system, dissatisfaction with the efficacy of IPTG-mediated transgene induction prompted the authors to develop a novel system, this time utilising regulatory elements of the Tn10-specified tetracycline-resistance operon of *E.coli*. The integral components of this system are the tetracycline repressor (*tetR*) (Figure 1.6c), and the tetracycline operator (*tetO*) sequences. The tetracycline repressor functions in *E.coli* to inhibit the expression of the tetracycline

a) Tet-off system



b) Tet-on system



c) Stereo plot of TetR structure

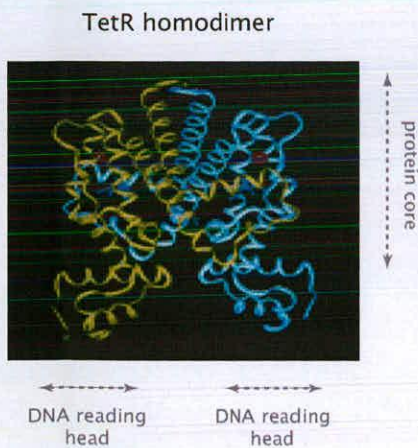


Figure 1.6 The tet-on and tet-off systems a) Schematic of the tet-off system. In the absence of inducer, the TetR moiety of the tTA transactivator binds to the *tetO* sequences encoded within the transactivator-response element (TRE). The VP16 moiety of tTA recruits transcriptase complex and transcription of the transgene ensues. Upon addition of inducer it binds the tTA inducing a conformational change that renders it unable to bind the TRE, and thus expression of the transgene is abolished. b) Schematic of the tet-on system. In the absence of dox the reverse transactivator (rtTA) is unable to bind the TRE and activate transcription of the transgene. Upon the addition of dox rtTA binds the *tetO* sequences within the TRE resulting in transgene expression. c) Stereo plot of TetR homodimer with the two TetR monomers shown in different colours. The two DNA reading heads bind to the 19 bp *tetO* sequence (tcctatcagtgatagaga). Figure adapted from Gossen *et al.*, 1995.

resistance gene (*tetA*) in the absence of tetracycline (Tc) antibiotic. TetR binds as a homodimer to a palindromic 19 base pair *tetO* sequence upstream of the resistance gene thereby repressing its transcription when no antibiotic is present. Upon influx of tetracycline (Tc) into the cell, it binds TetR with high affinity, inducing a conformational change in the dimer which increases the distance between the two tetR DNA binding heads from 34Å to 39Å. This renders TetR unable to bind to the adjacent major grooves within the *tetO* sequence and this interaction is subsequently extinguished.

The high affinity (K_a 10^{12} - 10^{13} M^{-1} , Hillen *et al.*, 1983), and high specificity of this TetR/*tetO* interaction was harnessed by Gossen and Bujard to create an inducible binary transactivatory system (Gossen & Bujard 1992). To create the tetracycline-controlled transactivator (tTA), sequence encoding full length *tetR* was fused to sequence encoding the C-terminal domain of HSV-VP16 (which when translated is comprised of 130 amino acids. Figure 1.7). To produce the transactivator response element to which tTA binds to activate transgene transcription, an array of 7 *tetO* sequences were fused to a minimal CMV promoter fragment (which extends 53 bp upstream and 75 bp downstream of the transcription initiation site and lacks the enhancer element. Figure 1.8). The tTA was placed under the control of a CMV promoter and enhancer sequence, and the TRE was linked to a firefly luciferase reporter gene after which both elements were stably introduced into HeLa cells by random integration. Upon the addition of Tc inducer some clones exhibited negligible basal levels of luciferase expression yet induced the reporter to over five orders of magnitude upon withdrawal of Tc (Gossen & Bujard., 1992). This provided evidence that the system was able to activate high levels of transgene expression without exhibiting 'leaky' transgene expression in the absence of tTA binding (the TRE) due to activation by endogenous factors.

Some clones were shown to exhibit high levels of basal transcription in the presence of Tc. Due to the random integration of both system components this was most likely due to the insertion of the TRE into a region that was under the influence of an enhancer element or in close proximity to an endogenous promoter. As clones were also derived that exhibited undetectable levels of basal transgene expression this suggested it was not

Figure 1.7 A selected panel of transactivators and reverse transactivators. Not all transactivators that have been developed are depicted. There is a series of tTAs and rtTAs that contain different numbers of VP16 minimal activation domains and differ in activation potential (Baron *et al.*, 1997). In addition a novel reverse transactivator has been developed that has an increased range of dox sensitivity (Akagi *et al.*, 2001). Abbreviations: aa, amino acid; F, HSV minimal activation domain termed F domain (amino acid sequence is shown for tTA2).

		Amino acid substitutions in original TetR peptide sequence	Further modifications	Reference
tTA		original sequence	N/A	Gossen & Bujard., 1992
rtTA		Glu71-Lys71 Asp95-Asn95 Leu101-Ser101 Gly102-Asp102	N/A	Gossen <i>et al.</i> , 1995
tTA2		original sequence	VP16 C-terminal domain substituted with 3 repeats of VP16 minimal activation domains (domain sequence is shown)	Baron <i>et al.</i> , 1997
rtTA2		Glu71-Lys71 Asp95-Asn95 Leu101-Ser101 Gly102-Asp102	VP16 C-terminal domain substituted with 3 repeats of VP16 minimal activation domains	Baron <i>et al.</i> , 1997
rtTA2s-M2		Ser12-Gly12 Glu19-Gly19 Ala56-Pro56 Asp148-Glu148 His179-Arg179	VP16 C-terminal domain substituted with 3 repeats of VP16 minimal activation domains. Coding sequence humanised, and optimised for increased stability	Urilinger <i>et al.</i> , 2000
rtTA2s-S2		Glu19-Gly19 Ala56-Pro56 Asp148-Glu148 His179-Arg179	VP16 C-terminal domain substituted with 3 repeats of VP16 minimal activation domains. Coding sequence humanised, and optimised for increased stability	Urilinger <i>et al.</i> , 2000

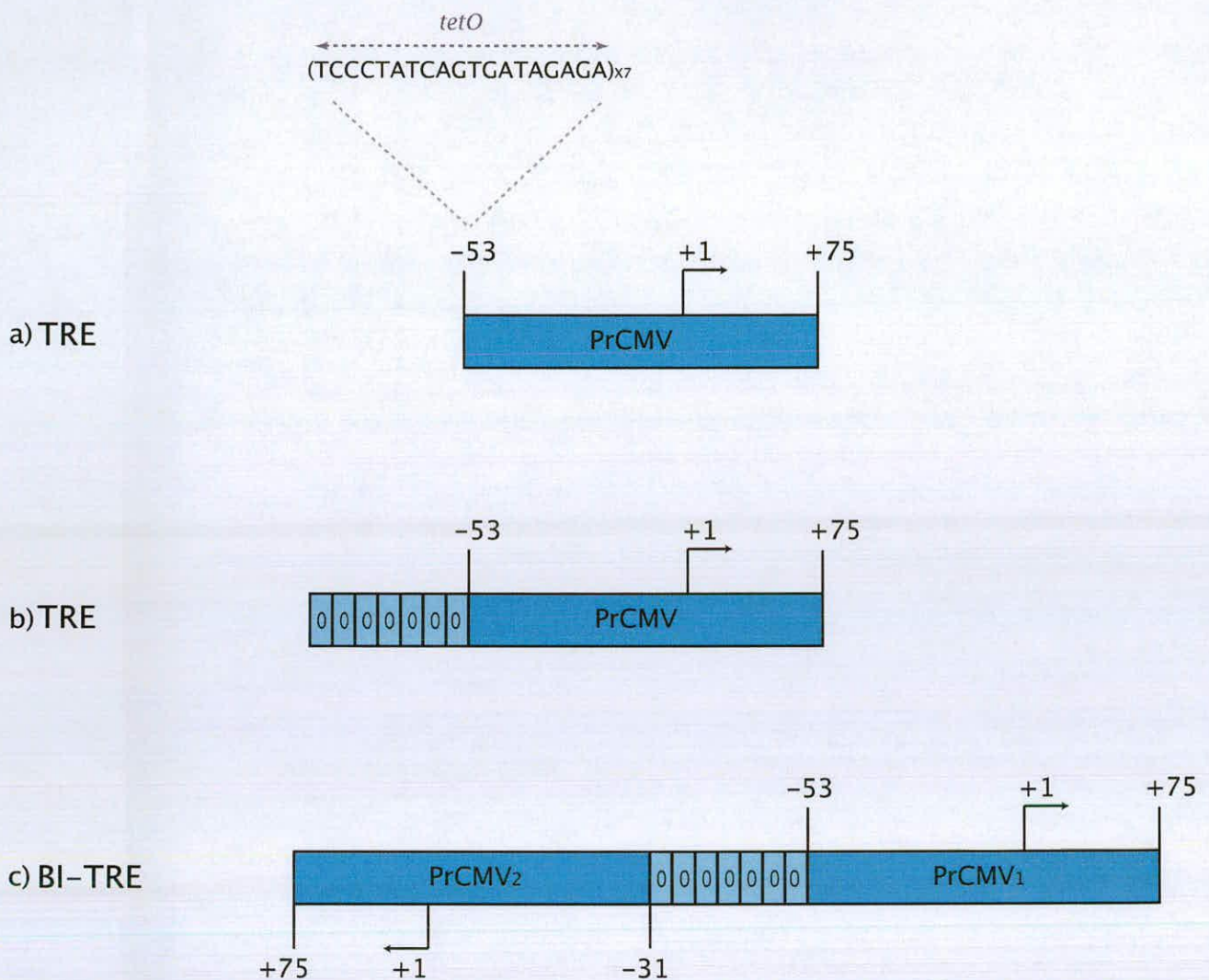


Figure 1.8 Representation of tet-responsive elements (TREs). a) The original TRE is comprised of an array of 7 *tetO* sequences fused to a CMV minimal promoter fragment spanning 53 bp upstream of the transcription initiation site and 75 bp downstream of the initiation site. This minimal promoter element lacks the CMV enhancer and therefore transcription is only initiated upon binding of the *tetO* sequences by transactivator. The same TRE is shown schematically in b. c) The bi-directional TRE contains two CMV promoter fragments, which lead to the transcription of linked transgenes in convergent directions after the central *tetO* sequences are bound by transactivator.

a problem intrinsic to the system itself. Similarly clones derived exhibiting low transgene induction levels may have harboured either or both elements integrated into regions of chromatin that were not permissive for high levels of transcription (in the case of tTA), or inaccessible to transcription factors (in the case of the TRE).

Transgene inactivation by addition of Tc to the culture media was found to be very rapid, with luciferase levels reaching 2% of their original values 12 hours after addition of the antibiotic (at a concentration of 1 µg/ml). Activation of the luciferase transgene was not as efficient as this required the efflux of Tc from the intracellular environment, however over 20% of fully induced levels were reached within 12 hours of withdrawal, and inductions of over 70% after a further 12 hours. It was found that a range of luciferase expression levels could be obtained between the concentrations of 0 and 1 µg/ml Tc, a phenomenon which the authors ascribe to the partial activation of the TRE at intermediate Tc concentrations (Figure 1.9). However, as the luciferase assay is based on lysis of the whole cell population this conclusion was premature as it is plausible that the detection of intermediate luciferase levels was the result of only a subpopulation of cells expressing the transgene with the remaining cells negative for luciferase expression, rather than the detection of homogenous population-wide levels of intermediate luciferase transcription.

The initial publication of the tet-off inducible system was promptly followed by a report of the feasibility of its use *in vivo*, (Furth *et al.*, 1994). Several founder lines containing either CMV-driven tTA, TRE-linked luciferase, or TRE-linked β-galactosidase were created by pronuclear injection. Transactivator-expressing lines were then crossed with lines harbouring TRE-linked transgenes and transgene expression in the presence and absence of Tc (administered via slow-release subcutaneous pellets) was assayed. Whilst many of the double-transgenic lines arising from the tTA/TRE-luciferase cross exhibited luciferase expression in most tissues in the absence of Tc, and low expression in the presence of Tc, transgene inductions were not of the same magnitude as those achieved *in vitro* (the highest reported being around 10³ fold). Inductions were also shown to vary greatly between different tissues, a phenomenon which is probably due to the positioning

of the tet-system elements within chromatin regions that possess varying permissivity for gene transcription depending upon cell type. This would be expected to lead to variation in tTA transcription and in the case of the TRE, variation in its recruitment of tTA between different tissues of the same organism.

The progeny of the tTA/TRE β -galactosidase crosses exhibited mosaicism of β -galactosidase expression in all tissues analysed, suggesting that both transgenes were only expressed in a subset of cells. This was thought to be due to mosaic properties of the CMV promoter used to drive tTA expression rather than a problem intrinsic to the tet-system components. Indeed a later publication reported homogenous and hepatic-specific expression of β -galactosidase when tTA was placed under the regulation of a LAP promoter (Kistner *et al.*, 1996). Inductions of over five orders of magnitude were also reported when these tTA lines were crossed with TRE-luciferase reporter lines, indicating that *in vivo* transgene expression can be induced to levels comparable to those obtained *in vitro*.

1.5.3 Development of the tet-on system

The system described above requires the continuous administration of tetracycline to keep transgene expression switched off (tet-off system). Whilst the concentration of Tc required to keep the transgene silent (0.1 $\mu\text{g/ml}$ in HeLa cell culture) has no obvious deleterious effect on HeLa cells over prolonged periods of time, such a situation is suboptimal in certain experimental setups, for example sustained transgene inactivation when using the system *in vivo*, or when it is imperative that a transgene is activated rapidly.

The development of a transactivator which possessed reversed DNA binding properties to that of the original tTA gave rise to the tet-on version of the tet-system (Gossen *et al.*, 1995). This reverse transactivator (rtTA) *requires* Tc to initiate transgene expression (Figure 1.6b). It was created by randomly mutagenising *tetR* sequence and then using a genetic screen established in *E.coli* to select for the repression of a reporter in the

presence of Tc (Gossen *et al.*, 1995). The most promising mutant isolated contained four amino acid substitutions, three of which were predicted to be involved in the conformational change of tetR during induction (Figure 1.7). This mutated *tetR* sequence was then fused to the C-terminal VP16 sequence (which was identical to the sequence present in tTA) to produce the reverse Tc-dependent transactivator.

In contrast to the tet-off system (which utilises tTA), the tet-on system is favoured in situations in which a transgene has to be activated rapidly. HeLa clones stably harbouring both rtTA (driven by a CMV promoter and enhancer), and the TRE-linked luciferase element, exhibited maximum levels of reporter induction in less than 24 hours after addition of inducer (Gossen *et al.*, 1995). However, in comparison to the levels achieved with tTA, the use of rtTA gave rise to inductions of only a 1000 fold (Figure 1.9). Studies *in vivo* however show that rtTA is able to activate the expression of a TRE-linked luciferase reporter to over five orders of magnitude above basal levels upon the administration of inducer (Kistner *et al.*, 1996).

The lower transgene inductions observed *in vitro* when utilising the reverse transactivator prompted investigation into the search for other more potent inducer compounds. Anhydrotetracycline (Atc) was found to be a very effective inducer (Lederer *et al.*, 1996) of the tet-on system, as was Doxycycline (Dox). Currently the standard inducer of choice for both versions of the tet-system is Dox, owing to both its high activation capabilities and its commercial availability.

1.5.4 Utilisation of the tet-on or tet-off system

As previously described, the tet-off system activates transgene expression in the absence of dox. The use of this system is therefore favourable in experimental situations where it is necessary that the tet-inducible transgene is active for long periods of time (as this does not require the presence of dox). This is especially relevant *in vivo*, where the continued administration of inducer (via whichever route) is not a trivial matter, as the dosage and timing of drug administration needs to be carefully calculated to produce the required

internal bioavailability of dox. In those situations in which it is necessary for the transgene to remain silent for prolonged periods, the utilisation of the tet-on system is therefore desirable.

Another factor in determining what version of the tet-system is most suitable for the experimental context is whether rapid activation, or inactivation of the transgene is required. This is dependent upon the kinetics of cellular saturation, or elimination of dox inducer. *In vitro*, antibiotic can be washed out of culture efficiently and therefore when using the tet-off system, transgene repression (which requires dox uptake) is only slightly more rapid than transgene induction (which requires the elimination of dox from cells). *In vivo* however transgene induction using the tet-off system is substantially slower than transgene repression as it requires the depletion of dox from the tissue under observation. This can take between 3 and 10 days depending on the organ in question, and is most likely due to the time required to eliminate dox depots that have accumulated in various tissues (Baron & Bujard 2000). In situations that require efficient transgene induction *in vivo*, the tet-on version is therefore the system of choice.

A study into the requirement of endothelin receptor B (EDNRB) during development utilised both the tet-off and tet-on systems and highlighted their different functional capabilities (Shin *et al.*, 1999). Mice homozygous for a mutation in the *Ednrb* gene die as juveniles as a result of the inability of neural crest-derived enteric neurons to colonize the distal colon leading to the condition megacolon. In an effort to establish the precise window during development at which EDNRB signalling is required, *Ednrb* was conditionally expressed using either the tTA or rtTA transactivators, so that transcription of *Ednrb* could be manipulated by the administration or withdrawal of dox to the pregnant mother (*Ednrb* expression was entirely dox-dependent with both endogenous alleles rendered null). The tet-on system was used to investigate the earliest requirement for EDNRB function in development by inducing *Ednrb* expression at different times during development (with *Ednrb* transcription off previous to this). The rapid induction of the endogenous gene (between 9 and 13 hours after antibiotic administration as determined by northern blotting) during mid-gestation enabled the

earliest requirement to be estimated at between E10 and E11. Subsequently the latest requirement for *Ednrb* function was identified using the tet-off system. *Ednrb* was expressed throughout early development and then rapidly inactivated by administering dox to pregnant females (complete inactivation occurred 6 hours after dox administration). Using this technique it was found that *Ednrb* expression was required before E12.5 but not after this time (Shin et al., 1999).

1.5.5 The use of VP16 minimal activation domains

When transactivator-encoding elements are randomly integrated into chromatin, the subsequent screening process results in the selection of clones which express the transactivator at sufficiently high enough concentrations to activate high levels of the TRE-linked transgene, yet low enough to prevent deleterious effects such as squelching (which is a toxic effect sometimes exhibited when overexpressing exogenous components, due to their interaction with endogenous factors and subsequent sequestration of these factors from their intracellular pools). Despite this screening process, squelching has been reported to be a problem in some cell lines overexpressing the tetracycline-transactivator (Markusic *et al.*, 2005), and is likely to result from the aberrant interaction of the VP16 moiety with the intracellular regulatory factors. Indeed it has been shown previously that the overexpression of VP16 as a fusion protein with GAL4 is not tolerated by cells (Berger *et al.*, 1992).

Additionally it was envisaged that the targeting of the transactivator to predetermined loci in an effort to place it under the control of a tissue-restricted promoter would result in intracellular transactivator concentrations that were dependent on the activity of the regulatory promoter. If the promoter in question was highly active it may result in the inability to derive stably targeted clones if such high levels of transactivator are toxic to the cell. With this in mind a series of transactivators were developed which were more readily tolerated by the intracellular environment. This was achieved by creating a panel of transactivators in which the C-terminal moiety of VP16 was replaced with varying numbers of 12-amino acid VP16 minimal activation domains (Baron *et al.*, 1997).

Analysis of protein extracts, prepared from pools of cells stably expressing the original and the novel transactivators showed that in comparison to tTA, the newly developed transactivators were present at higher intracellular concentrations, and therefore are presumed to be more readily tolerated. The most promising transactivator was tTA2 which contained three minimal activation domains (Figure 1.7), and exhibited activational properties similar in magnitude to the original tTA.

1.5.6 The development of rtTA2^s-M2 and rtTA2^s-S2

Whilst tTA is able to induce transgene expression to over five orders of magnitude, the original rtTA was less capable of activating transgene expression to the same extent. This was mainly due to residual affinity of rtTA for the TRE in the absence of inducer, which often resulted in high basal activity of the responder element. As a result of this there was therefore a need to develop an improved reverse transactivator which exhibited no affinity for the TRE in the absence of inducer yet was still able to activate high levels of transgene expression.

The *de novo* creation of a reverse transactivator possessing these properties was carried out using a novel screening assay set up in *S.cerevisiae* (Urlinger *et al.*, 2000). The *tetR* gene was first amplified by PCR under mutagenic conditions and then integrated into a plasmid upstream of the C-terminal VP16 sequence. The plasmid also contained a TRE and juxtaposed GFP coding region. After transformation into yeast, TetR proteins that possessed a reverse phenotype were selected by screening for GFP positive colonies under UV light. It was envisaged that such a eukaryotic based screen would overcome a major problem of the previous *E.coli* based screen, that is the loss or reduction in phenotype of the TetR after fusion to the VP16 moiety and establishment in a eukaryotic environment.

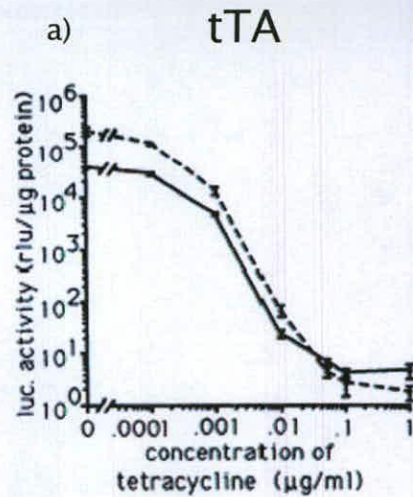
One of the most promising clones identified in this first round of screening was rtTA-S2, which contained four amino acid exchanges within the TetR moiety and possessed an activation potential similar in magnitude to that of the original rtTA, yet exhibited lower

background activation in the absence of dox. In an effort to examine the functional contribution of the different amino acids, stretches of sequence encoding either the two most N-terminal or two most C-terminal amino acid substitutions were exchanged with the corresponding sequence from the original tTA and these chimeric transactivators were analysed for reverse phenotype and activation potential. It was observed that just two of the amino acid substitutions were sufficient to confer a reverse phenotype and this coding sequence was therefore subject to a second round of mutagenesis and screening, yielding the clone rtTA-M1. The reintroduction of the two amino acid substitutions that were removed before the second round of mutagenesis led to production of rtTA-M2 possessing elevated induction levels whilst still retaining the reverse DNA binding properties.

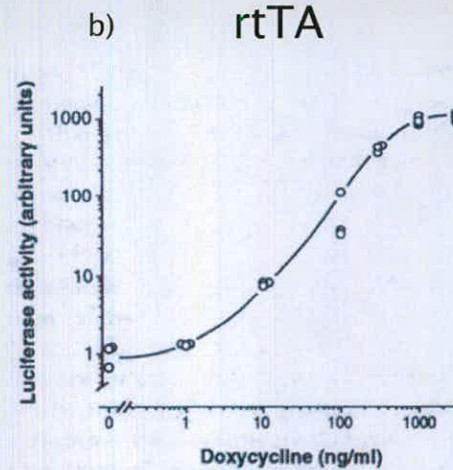
In order to improve the stability of the novel rtTA-S2 and rtTA-M2 reverse transactivators the VP16 C-terminal sequence was replaced by three minimal activation domains and codon frequencies were altered to resemble those found in mammalian cells to produce rtTA^S-S2 and rtTA^S-M2 respectively (Figure 1.7). They were then stably introduced into a HeLa cell line that contained an already integrated TRE-linked luciferase. It was found that both reverse transactivators exhibited many improved characteristics over the original rtTA, such as a comparable activation potential with no background activity of luciferase in the absence of dox (Figure 1.9). rtTA-M2 exhibited increased sensitivity to dox and was able to fully activate transgene expression at concentrations of just over 100 ng/ml dox in comparison to the original rtTA which requires 1 µg/ml (Figure 1.10). A recent study in MCF-7 cells reported that continual expression of this novel transactivator can be tolerated by cells for months in culture indicating the reduced toxicity effects of these new generation reverse transactivators over their original counterparts (Qu *et al.*, 2004).

1.5.7 The development of transcriptional silencers

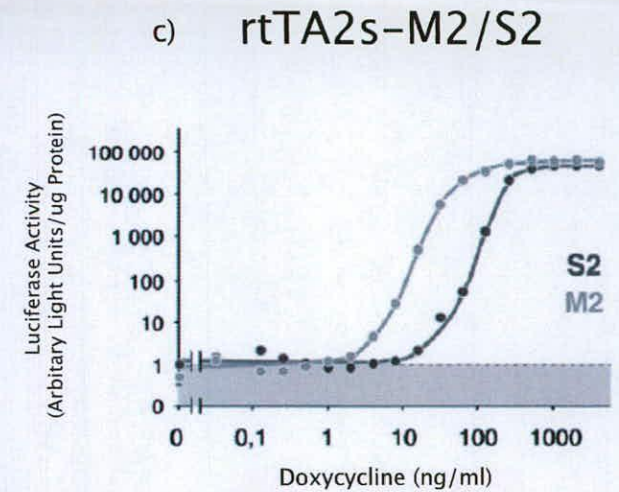
The tet-responsive transcriptional silencers tTS^{kid} (Freundleib *et al.*, 1999) and TetR-KRAB (Deuschle *et al.*, 1995), bind and silence the TRE in the absence of inducer.



Gossen & Bujard., 1992



Gossen *et al.*, 1995



Urlinger *et al.*, 2000

Figure 1.9 Dose response at varying concentrations of Tc or dox inducer. All results are based upon the stable integration of both transactivator and luciferase linked tet-response elements into HeLa cells. a) Luciferase reduction with increasing concentration of Tc inducer using tTA. The dose response profile is depicted for two clones. b) Dose response profile for rtTA expressing cells. Note maximal luciferase induction is lower than that achieved when using tTA. c) Profiles for the two reverse transactivators rtTA2s-M2 and rtTA2s-S2. These novel transactivators exhibit a larger range of transgene induction when compared to rtTA and can also activate maximal luciferase expression at lower concentrations of dox, with the rtTA2s-M2 transactivator being the more sensitive of the two.

These silencers were originally developed to address the problem of leaky transgene expression sometimes seen when using the earlier version of the tet-on system. With the advent of rtTA2^S-M2 the use of tet-responsive silencers is now largely irrelevant. However, there are certain applications in which they are still useful (see below), and therefore the development of tTS^{kid} (the most widely used of the silencers) will now be described.

The silencer (tTS^{kid}), is a fusion between the Tet repressor and transcriptional silencing domains derived from the human kidney protein Kid-1 (Freundlieb *et al.*, 1999). During the development of tTS^{kid} its dimerisation sequence was exchanged for the dimerisation sequence encoded by the E class tet repressor, (the TetR moiety of the transactivators belongs to the B class), in order to prevent heterodimerisation of the silencer and transactivator monomers and formation of non-functional dimers. In the absence of Dox, tTS^{kid} binds to the TRE and inhibits the formation of transcriptase complex around the minimal promoter. It is able to bind and repress TRE activity at dox concentrations of up to 10 ng/ml when used in conjunction with the tet-on system, without negatively impacting on maximum transgene levels induced by rtTA.

The tTS^{kid} silencer has been shown to act to reduce background levels of transgene expression both *in vitro* and *in vivo* using the original rtTA-tet-on system (Rossi *et al.*, 1998, Zhu *et al.*, 2001). Nowadays basal transgene levels are generally not found to be a problem when utilising rtTA2^S-M2. However, there is one field in which tTS^{kid} can still play an important role and that is in the area of gene therapy. Despite the improved properties of rtTA2^S-M2 over the original rtTA, upon transient transfection of this novel transactivator and TRE-reporter element into cells, basal levels of the reporter in the absence of dox can sometimes be detected (Lamartina *et al.*, 2003). One application in which this could pose a problem is in the extra-chromosomal regulation of therapeutic genes, where it is necessary that the gene product is tightly regulated. A study investigating the expression of tet-regulated Epo after intramuscular injection of the tet-system components found that levels were undetectable when tTS^{kid} encoding sequence was also injected (Lamartina *et al.*, 2003).

1.5.8 Application of the tet-system

The efficacy of the tet-system is exemplified by its widespread application in a variety of different organisms including plants, bacteria, worms and mammals. Its ability to function well both in mammalian cell lines and transgenic organisms has led to its varied application. Besides the temporal-spatial overexpression of exogenous genes, it has been used to overexpress endogenous genes (Bond *et al.*, 2000), ablate subpopulations of olfactory neurons *in vivo* (Gogos *et al.*, 2000), mark populations of epithelial stem cells in the postnatal mouse (Blanpain *et al.*, 2004), and create temporal knockouts (Saam *et al.*, 1999), to name but a few examples. There are also a number of ways of establishing the system, including the random integration of elements, the targeting of elements to predetermined loci, and the integration of elements utilising viral vectors (Hofmann *et al.*, 1996, Markusic *et al.*, 2005).

Whilst tet-systems enjoy varied application, the focus of this project is on its implementation in ES cells, therefore its use in this pluripotent cell type will now be examined.

1.6 Previous establishment of the tet system in ES cells

Embryonic stem cells, owing to their pluripotent nature, offer a model of differentiation for potentially every lineage. Indeed, *in vitro* differentiation of ES cells has been shown to generate many differentiated cell types including pancreatic, hepatic, cardiac and neural. These differentiation procedures not only provide an opportunity to gain insight into the pathways involved in the specification and acquisition of cell fate but also allow the information gained to be harnessed in order to direct differentiation either by altering the external environment or genetically manipulating the cells with the aim of producing unlimited sources of cells useful for therapeutic purposes.

The differentiation of a precursor cell to that of a more restricted fate involves the expression of endogenous cell fate determining factors at discrete windows in time. In analysing the pathways involved, it is therefore often desirable to overexpress genes at different stages in differentiation. In such experimental situations inducible gene expression technology is the perfect tool, as it allows the experimenter exquisite control over the expression of a transgene. In a recent study by Kyba *et al*, investigation into the potential contribution of primitive embryonic blood progenitor cells to definitive hematopoiesis was investigated by constitutively overexpressing the HoxB4 transcription factor in cells isolated from precirculation murine yolk sac. After subsequent culture on an OP9 stromal cell layer the cells were engrafted into lethally irradiated recipients, whereby they were found to reconstitute the hematopoietic system (Kyba *et al.*, 2002). Engraftment of donor-derived bone marrow from primary into secondary recipients however, showed that constitutive expression of HoxB4 was inhibitory to the differentiation of the lymphoid lineage. The authors thereby established conditional expression of the transcription factor in ES cells using the tet-on system. These cells were differentiated via an embryoid body stage with overexpression of HoxB4 induced only between days 4 to 6 of differentiation. The cells were cultured on a stromal layer before engraftment into lethally irradiated recipients. This conditional overexpression of HoxB4 enabled the cells to give rise to multilineage engraftment in both primary and secondary recipients. Therefore, it was concluded that induced overexpression of HoxB4 was only required for a transient period in order to confer definitive multilineage engraftment potential on ES cells.

Another study in which the tet-system provided the means to temporally overexpress a transcription factor involved in differentiation was conducted by Sonntag *et al*, who investigated the link between the acquisition of a midbrain dopaminergic cell fate and differentiation into a dopaminergic cell type. It was found that overexpression of the transcription factor Nurr1 (which is necessary for the specification of the mesencephalic dopaminergic phenotype during embryogenesis) *in vitro*, in non-neuronal cells, led to a dopaminergic biochemical identity. This was achieved by establishing tet-regulatable Nurr1 expression in ES cells and overexpressing the transcription factor at different

stages during an *in vitro* differentiation protocol. Overexpression of Nurr1 at later stages of differentiation was shown to give rise to non-neuronal dopaminergic cell types suggesting that dopaminergic phenotype can be acquired independently of a neuronal context (Sonntag *et al.*, 2004).

The directed differentiation of ES cells down defined lineages has important implications for cell replacement therapies, which aim to replace diseased or damaged cells *in vivo* with viable counterparts. One way to direct differentiation in this manner is to genetically manipulate ES cells to overexpress key regulators of differentiation. Miyazaki *et al.*, reported that tet-regulated overexpression of pdx-1 (a transcription factor which is an essential regulator of pancreatic development), at discrete timeframes during a previously established embryoid-body based differentiation protocol, promoted the differentiation of insulin-producing cells. It is worth noting that conditional expression of pdx-1 was necessary as previous attempts to establish stable ES cell lines overexpressing pdx-1 had failed, probably as a result of toxicity effects of pdx-1 on ES cells.

Inducible systems have also been used to study the self-renewal properties of ES cells themselves. It is important to gain an understanding of the mechanisms behind ES cell self-renewal and their pluripotent nature in order to determine how this unique cell type can ultimately give rise to every cell type in the adult organism. The Oct4 transcription factor was shown to be a key regulator in the maintenance of the pluripotent state of ES cells in a study where conditional expression of Oct4 was achieved via the tet-system (Niwa *et al.*, 2000). It was reported that the level of Oct4 expression was crucial in maintaining pluripotency as increased expression of Oct4 (to over 150% of original levels) resulted in differentiation of the cells into primitive ectoderm and mesoderm, whereas a decrease in Oct4 levels resulted in differentiation into trophectoderm.

The search for gene products which play an important role in the maintenance of the ES cell state may be greatly facilitated by the recent development of a strategy used to create an ES cell library which contains genome-wide bi-allelic mutations (Yusa *et al.*, 2004).

The loss of function of Bloom's syndrome gene (*Blm*) results in increased rates of loss of heterozygosity owing to a high rate of recombination between non-sister chromatids of homologous chromosomes in G2 of the cell cycle. In a recent report both *Blm* alleles in ES cells were manipulated so that their transcription was dependent upon the withdrawal of doxycycline (with the utilisation of the tet-off system). Subsequent transient loss of *Blm* expression coupled with N-ethyl-N-nitrosourea mutagenesis led to the production of an ES cell library containing an estimated 4.2 independent clones per locus bearing bi-allelic mutations. The usefulness of the library in phenotype-based genetic screening was validated by screening for mutants of glycosylphosphatidylinositol-anchor biosynthesis.

1.7 The targeting of both tet system components in ES cells

High basal levels of transgene expression in the absence of inducer is a frequently reported problem upon random integration of tet-system elements, (Izumi and Gilbert., 1999, Qu *et al.*, 2004.), and as a result of this it is necessary to screen large numbers of clones. This background transcription can be attributed to the action of endogenous enhancer sequences upon the TRE component (and in those cases where the original rtTA is used, residual binding of the transactivator to the TRE is also an issue as discussed in 1.5.5). In order to overcome the problem of background expression, it was decided in this project to target both elements of the tet-system to pre-characterised loci in an attempt to find suitable genomic sites that would allow for tightly regulated expression of cDNAs. This strategy has the added advantage that once such a system has been established it can be used to directly compare the overexpression of independent cDNAs. Such an approach was undertaken by Wutz *et al.*, to investigate the domains of *Xist* RNA that are responsible for the silencing and localisation functions of *Xist* during X-inactivation. A series of *Xist* cDNA deletion mutants were overexpressed with the tet-system in ES cells. As both components of the system were targeted into the genome the array of clones expressing different mutant forms of *Xist* could be directly compared (Wutz *et al.*, 2002).

The X-linked hypoxanthine phosphoribosyltransferase (*hprt*) gene was initially chosen to incorporate both elements of the tet-system (as described in chapters 2 and 3 of this thesis). HPRT is an enzyme which catalyses one of the first steps in the salvage pathway of nucleoside biosynthesis in mammalian cells. The gene itself spans 33 kilobases and is comprised of 9 exons. The *hprt* locus is well characterised, and there are a number of vectors (Thompson *et al.*, 1989; Reid *et al.*, 1991), and cells lines (Hooper *et al.*, 1987; Magin *et al.*, 1992) available to facilitate the targeting of exogenous sequences to the locus. Indeed there are targeting vectors available to target both the 5' region and the 3' region of the *hprt* locus. The targeting of the reverse transactivator to the 5' region and the TRE to the 3' region of *hprt* results in an intervening distance of 30 kb between both elements.

At such a distance it was considered that the promoter driving the expression of the transactivator was unlikely to activate expression of the TRE-linked transgene in the absence of inducer, a problem which has been reported when both tet-components are integrated in a single construct, (Hofmann *et al.*, 1996, Markusic *et al.*, 2005).

Several transgenes have previously been targeted to the *hprt* locus. Kirby *et al* used the pMPN8 replacement vector to target a truncated receptor for the growth factor erythropoietin upstream of the *hprt* coding region. When treated with exogenous erythropoietin these mice exhibited a increase in multipotent, clonogenic haematopoietic cells (Kirby *et al.*, 1996). In an independent study Bronson *et al.*, 1996, targeted a single copy of the murine *bcl-2* transgene under the control of either a human or chicken β -actin promoter to the 5' region of *hprt* using the pMP8SKB correction vector. Widespread expression of the transgene was detected in most tissues analysed in the resulting transgenic mice. The expression of transgenic sequences from the *hprt* locus in ES cells have also been described (Kyba *et al.*, 2002; Wutz *et al.*, 2002), supporting the notion of *hprt* being an ideal candidate for the expression of exogenous sequences.

The utilisation of the ROSA26 locus is also described in this project. This locus was originally discovered during a gene trap screen carried out in ES cells (the gene trap

encoded a β -gal reporter). The derivation of transgenic mice using ES cells harbouring the 'gene trapped' ROSA26 locus led to the discovery that β -gal expression was present throughout all the tissues of the embryo (Zambrowicz *et al.*, 1997). Subsequent analysis of the locus allowed the construction of a vector which yields efficient targeting of ROSA26. This vector contains the original splice acceptor site that was contained within the gene trap and therefore transgene expression can be driven from the endogenous ROSA26 promoter (this is discussed in greater detail in chapter 4). The ubiquitous expression of transgenes targeted with this vector was exemplified by the creation of a transgenic mouse line that expressed Cre-recombinase from the endogenous ROSA26 promoter (Soriano *et al.*, 1999). When this line was crossed to a reporter line, which contained a 'floxed' β -gal sequence (also targeted to ROSA26), the resulting mice were uniformly blue upon staining with X-gal. The endogenous ROSA26 promoter has previously been used to drive the expression of the rtTA in ES cells (Wutz *et al.*, 2002, Kyba *et al.*, 2002), which exemplifies its suitability in conjunction with the tet-system.

It has previously been described that the application of the tet-system in ES cells can greatly facilitate the study of lineage differentiation, and even pluripotency itself. The aim of this project was initially to establish the tet-on system in ES cells, using a fluorescent reporter transgene to validate the system (this will be described in chapters 2 to 4 of this thesis). Pre-characterised loci were targeted with the components of the tet-system to obtain tight and uniform expression of the inducible reporter. Transgene expression levels over a range of concentrations of inducer were also investigated in order to determine whether or not intermediate levels of transgene expression could be obtained. Upon the establishment of a suitable strategy cDNAs were overexpressed to investigate the involvement of the mitogen activated protein kinase (MAPK) pathway in neural induction using an established *in vitro* differentiation cell monolayer protocol (Ying *et al.*, 2003), the results of which are presented in chapter 5 of this thesis.

Chapter 2

The targeting of the transactivator to the 5' *hprt* locus in E14TG2a cells

2.1 Introduction

The initial strategy for integrating both the driver and responder elements of the tet-system into ES cells was to target both sequences to the hypoxanthine ribosyltransferase (*hprt*) locus, the reverse transactivator (rtTA^{2s}-M2) into a 5' proximal position, and the transactivator-response element (TRE) into a 3' position. In this chapter, the targeting of the rtTA²-M2 cDNA, driven by the different promoters CMV and CAG is described. The subsequent analysis shows that the composite CAG promoter is far superior in terms of driving homogenous levels of rtTA^{2s}-M2 throughout an ES cell population.

2.2 The cytomegalovirus promoter and enhancer to drive expression of the rtTA2^s-M2 transactivator

In the aim of establishing a conditional gene expression system in ES cells, which allows complete control over the induction and levels of transgene expression, it is optimal that the transgene in question is expressed at homogenous levels, ubiquitously throughout the cell population. It is therefore important to choose a promoter that in turn, drives constitutive and homogenous levels of the reverse transactivator. Throughout the early development of the tet-system, the cytomegalovirus CMV promoter and enhancer element (hereon just referred to as the CMV promoter) was used to drive the expression of tTA and rtTA. After the initial development of tTA, it was reported that stable integration of both the transactivator and luciferase-linked transactivator response element into HeLa cells, gave rise to luciferase inductions of over 10⁵ fold in magnitude, upon the withdrawal of tetracycline inducer (Gossen & Bujard, 1992). In more recent studies, the CMV promoter element has been shown to give rise to homogenous expression of TRE-linked transgenes when used to drive tTA expression in CHO (Chinese hamster ovary) cell lines (Izumi *et al.*, 1999), and rtTA2^s-M2 expression in ML20 (a derivative of MCF-7) cell lines (Qu *et al.*, 2004). The CMV promoter was therefore the first candidate of choice for regulating rtTA2^s-M2 expression in ES cells.

2.3 The *hprt* targeting vectors

E14TG2a cells were used to incorporate both elements of the tet system. This ES cell line is karyotypically male, and lacks endogenous HPRT activity. The E14TG2a cell line was derived from the parental E14 line, by selection for an HPRT negative phenotype in medium containing 6-thioguanine (which is converted to a toxic metabolite by HPRT). This resulted in the isolation of clones which contained spontaneous deletions at the *hprt* locus (Hooper *et al.*, 1987). E14TG2a was one of those clones isolated. These cells harbour a 55 kb deletion in the 5' region *hprt*, which encompasses the promoter region, and exons 1 and 2 (Tsuda *et al.*, 1997).

Two correctional vectors were available to independently target the reverse transactivator-encoding sequences to this locus. Both vectors contain the genomic *hpvt* sequences deleted upon the derivation of the E14TG2a cell line, and the targeting of these vectors to *hpvt* results in the reestablishment of the ORF and restoration of HPRT activity. Clones exhibiting HPRT activity can be selected for in HAT medium (which contains hypoxanthine, aminopterin and thymidine). Aminopterin blocks the *de novo* nucleoside synthesis pathway and only those clones possessing functional salvage pathway metabolism (of which HPRT is an essential component) are able to generate nucleotide triphosphates.

The two vectors utilised possessed different modes of integration. The insertion vector pDWM101 (Thompson *et al.*, 1989) contains genomic sequence from the wild type *hpvt* allele which extends from 650 base pairs upstream of the *hpvt* promoter, to an *EcoRI* site located within intron 3 (Figure 2.1a). During the construction of this vector, intron 1 of *hpvt* was reduced from 10.8 kb (the size found in the wild type *hpvt* allele) to 4.1 kb. A region encompassing exon 3 and the immediate flanking area is homologous to the corresponding region in E14TG2a cells (the size of this homology is between 2.3 and 4.2 kb depending on the exact endpoint of the *hpvt* deletion). The linearisation of this vector at a *XhoI* site within the region of homology, and subsequent targeting, results in the formation of a functional *hpvt* allele and in addition the duplication of exon 3 (Figure 2.2a). A correct targeting event results in the positioning of transgenic sequences encoded within the vector, 2.5 kilobases upstream of the newly integrated *hpvt* promoter (Figure 2.2a). These transgenic and promoter sequences are separated by plasmid backbone sequence that also integrates into the locus.

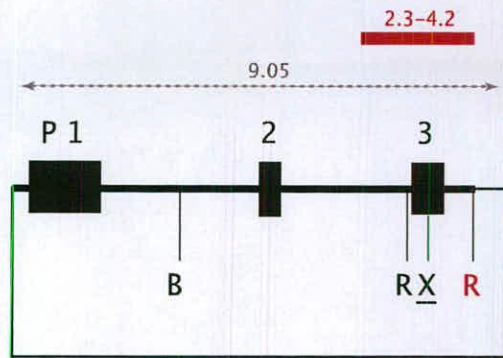
Manipulation of the pDWM101 vector was required to facilitate the incorporation of transgenic sequences into this vector. Plasmid pDWM101 was linearised by partial digestion with *EcoRI*, after which a oligonucleotide sequence encoding *KpnI*, *NotI*, *BamHI* and *SmaI* restriction sites and with *EcoRI* compatible ends was inserted into the cut vector (Figure 2.1b). This resulted in the recreation of the *EcoRI* site at one end of

Figure 2.1 Schematic representation of pDWM101 and pSKB1 targeting vectors.

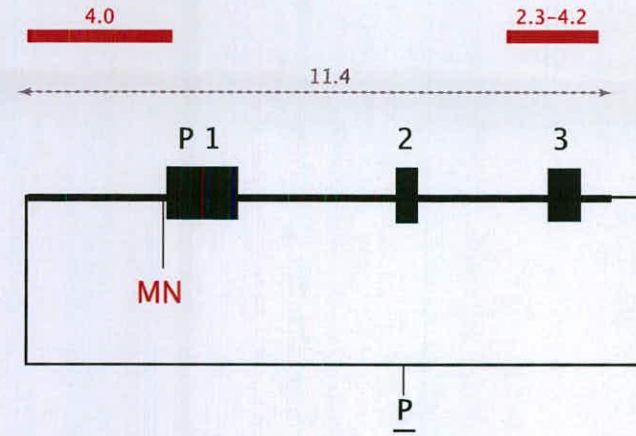
a) Insertion vector pDWM101 with the size of genomic encoding region and the region of homology to the endogenous *hprt* locus in E14TG2a cells indicated. b) Modification of the pDWM101 vector by insertion of a polylinker into the *Eco*R1 site downstream of exon 3 of the genomic insert. c) The pSKB1 replacement vector with the size of the genomic encoding region and the region of homology indicated. Potential restriction enzyme sites which can be used to incorporate exogenous sequence are shown in red. Underlined restriction sites represent sites of linearisation.

pDWM101 insertion vector

a)

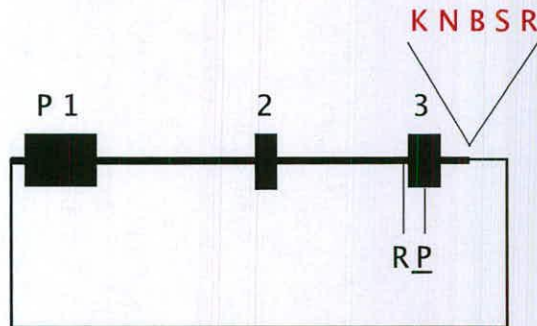


c) pSKB1 replacement vector



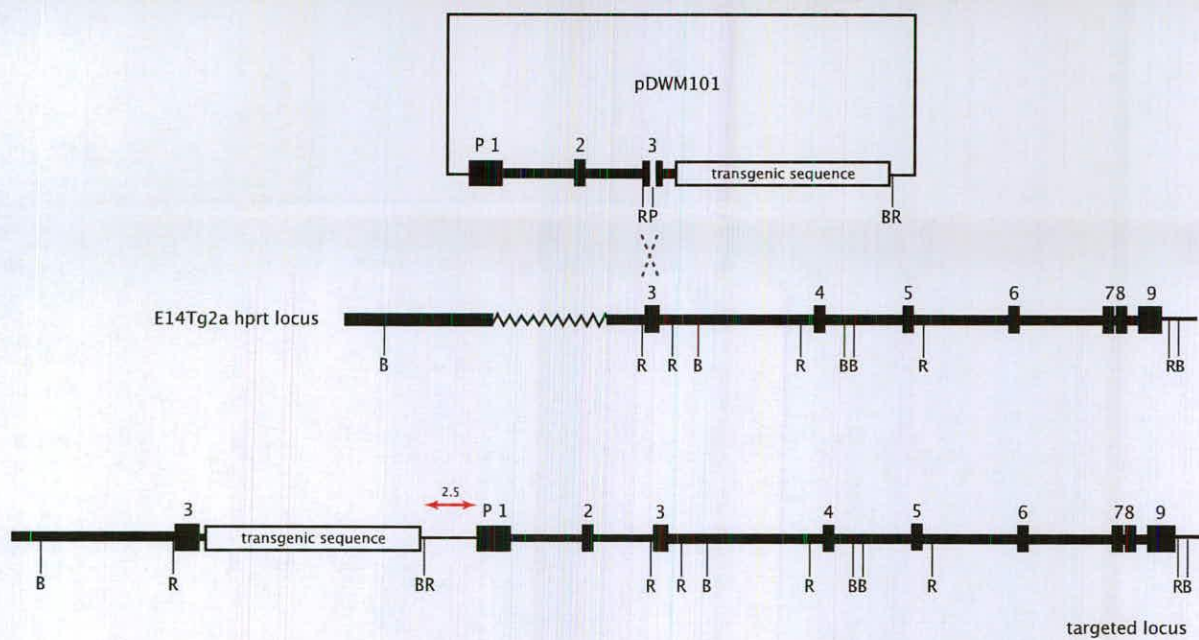
pDWM101 - addition of polylinker

b)



- | | | | |
|------------|--|--|----------------|
| P 1 | | Promoter and exon1 | |
| | | Genomic sequence | |
| | | plasmid backbone | |
| | | size of genomic insert (kb) | |
| | | region of homology to <i>hprt</i> locus in E14TG2a cells | |
| | | | K <i>KpnI</i> |
| | | | N <i>NotI</i> |
| | | | B <i>BamHI</i> |
| | | | R <i>EcoRI</i> |
| | | | P <i>PmeI</i> |
| | | | M <i>MluI</i> |
| | | | S <i>SmaI</i> |

a)



2
█ exon sequence

P
█ promoter sequence

— plasmid sequence

~~~~~ deleted region

X crossover event

B *Bam*HI  
R *Eco*RI  
P *Pme*I

b)

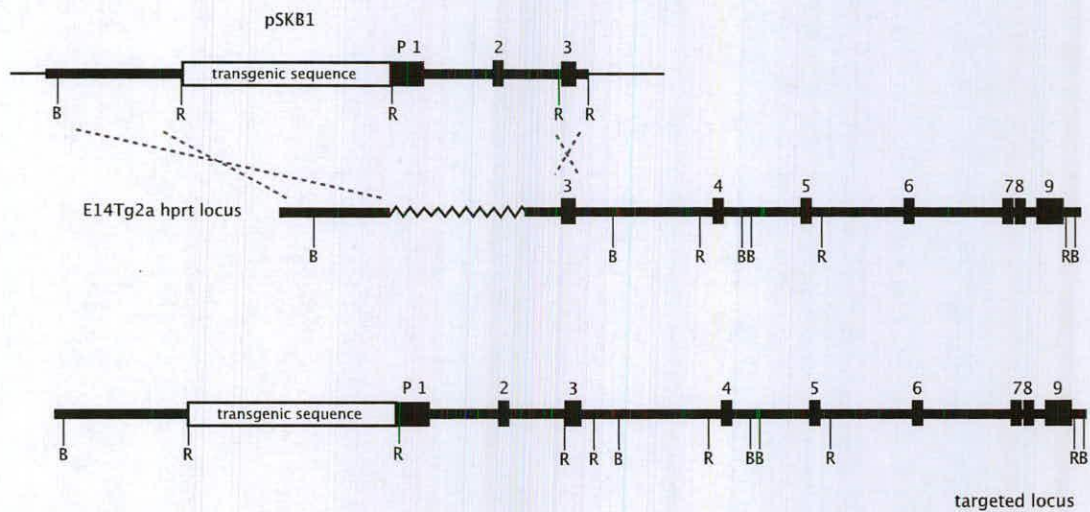


Figure 2.2 Targeting of the pDWM101 and pSKB1 vectors to *hprt*. a) Targeting of the integration vector pDWM101 after linearisation with *Pme*I results in the establishment of a functional *hprt* gene with duplication of exon 3. The pDWM101 plasmid backbone is also integrated. Transgenic sequences lie 2.5 kb from the newly inserted *hprt* promoter (indicated in red). b) Targeting of the replacement vector pSKB1 to *hprt* after linearisation with *Pme*I also results in the formation of a functional *hprt* gene. As two crossover events occur during targeting, plasmid backbone sequence is not incorporated into the *hprt* locus. Transgenic sequences are positioned immediately upstream of the *hprt* promoter.

the inserted oligonucleotide sequence. The orientation of the oligonucleotide within pDWM101 was determined by sequencing and is shown in Figure 2.1b.

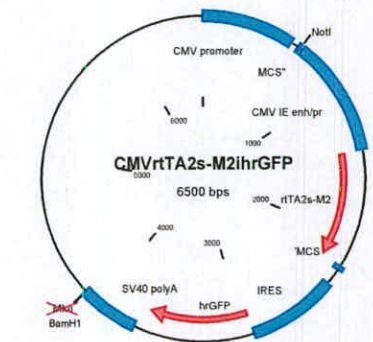
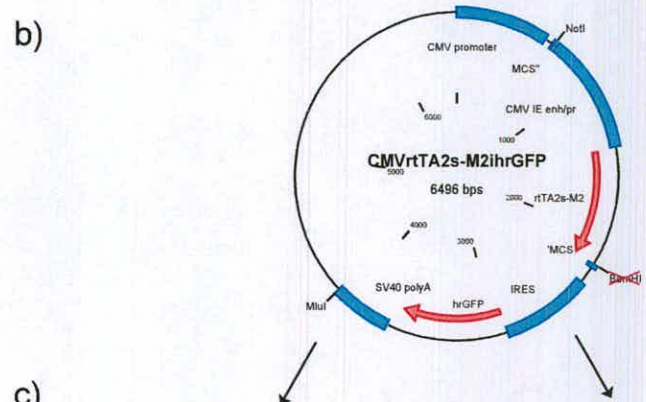
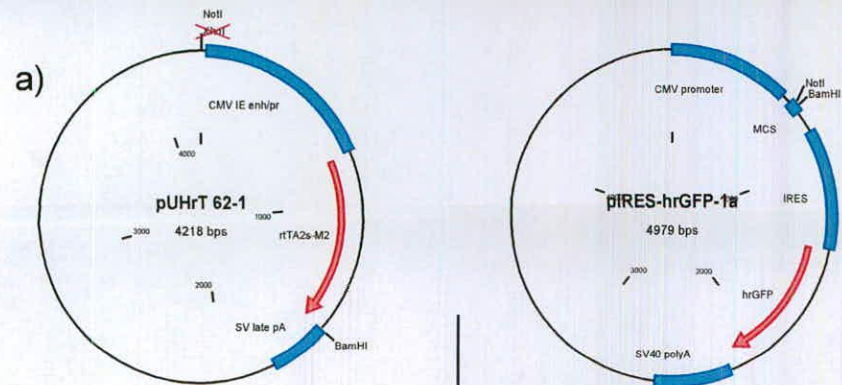
The *Xho*I site for linearisation is located within exon 3 of the *hprt* genomic sequence. It was deemed necessary to change this site to a less common restriction site so as to reduce the chance of the linearisation restriction site also being present within any additional sequences that were to be cloned into the targeting vector. The vector was therefore digested with *Xho*I, and after filling in the recessed 3' ends with klenow polymerase, *Pme*I linkers were inserted to create an eight base pair site of linearisation.

The alternative targeting vector available, pSKB1 (Bronson *et al.*, 1996), not only encodes homology to exon 3 and the surrounding region but also to the region 5' of the *hprt* deletion in E14TG2a cells, and thus can integrate as a replacement vector (Figure 2.2b). Transgenic sequences can be cloned into either the *Not*I site or *Mlu*I site of pSKB1, which are located immediately 5' to the genomic *hprt* promoter (Figure 2.1c). The linearisation of this vector at a *Pme*I site within the plasmid backbone, and subsequent electroporation results in the formation of a functional *hprt* gene. As two crossover events occur during the correct targeting of this vector, no duplication events result from integration and the plasmid backbone sequence is not incorporated into the targeted locus. The pSKB1 vector required no additional manipulation before inserting the transgenic sequences.

#### **2.4 Cloning steps involved in the construction of the reverse transactivator containing targeting vectors**

For the purpose of monitoring the expression of the rtTA2<sup>S</sup>-M2 transcription unit after the targeting of the exogenous sequences, the reverse transactivator coding sequence was linked to an Internal Ribosome Entry Site (IRES) sequence and a fluorescent reporter sequence. The reporter used was humanised *renilla* GFP (hrGFP), a fluorescent protein found in *Renilla reniformis* that has similar emission and excitation spectra to that of

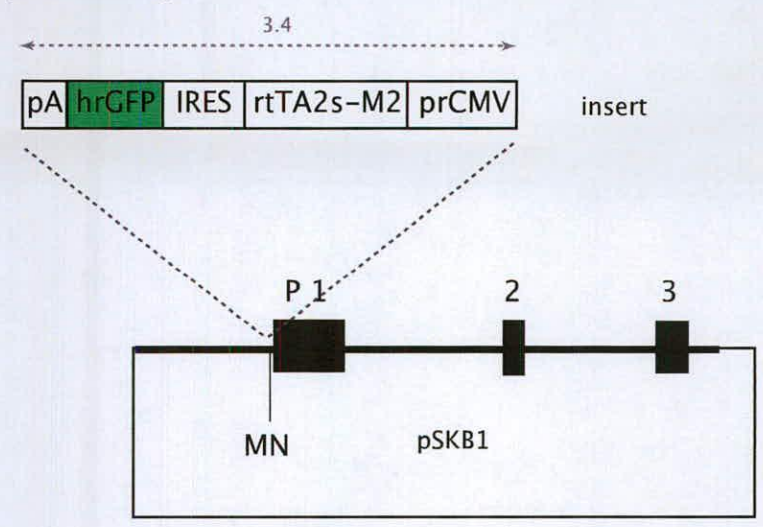
Figure 2.3 Cloning strategy for creating the CMV-rtTA2<sup>s</sup>-M2-IRES-hrGFP-pA fragment and subsequent insertion into the targeting vectors. a) The *Xho*I site upstream of the CMV promoter in pUHrT62-1 was abolished (as indicated by X), and a *Not*I site inserted in its place. Restriction digest with *Not*I plus *Bam*HI yielded the CMV-rtTA2<sup>s</sup>-M2 sequence, and this was subsequently inserted into *Not*I/*Bam*HI digested pIRES-hrGFP-1a plasmid. b) The *Bam*HI site of this resulting plasmid was removed and the CMV-rtTA2<sup>s</sup>-M2-IRES-hrGFP-pA fragment was excised by *Not*I/*Mlu*I digest and inserted into *Not*I/*Mlu*I digested pSKB1 to create the p5'hprrtCMV<sup>rep</sup> targeting vector as shown in d). c) The unique *Mlu*I site found in pCMVrtTA2<sup>s</sup>-M2ihrGFP was replaced with a *Bam*HI site and the CMV-rtTA-IRES-hrGFP-pA sequence was subsequently extracted as a *Not*I/*Bam*HI fragment. This was then ligated into *Not*I/*Bam*HI digested modified pDWM101 vector to create the p5'hprrtCMV<sup>ins</sup> targeting vector depicted in e).



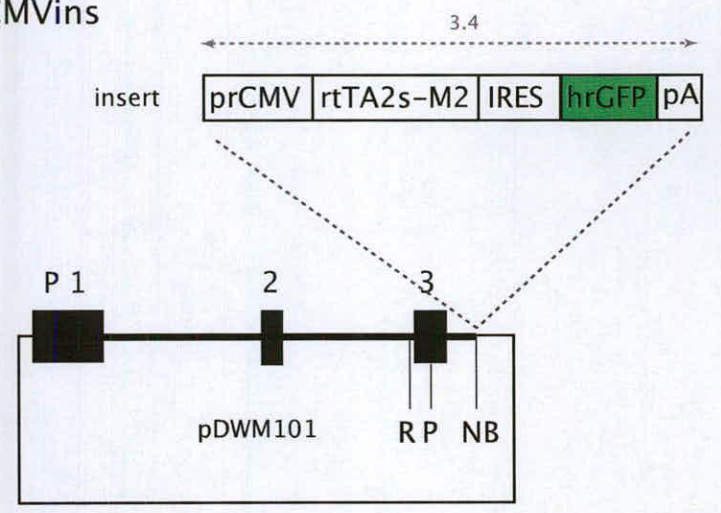
Not1, Mlu1 digest and ligation of sequences into pSKB1 replacement vector

Not1, BamH1 digest and ligation of sequences into pDWM101 insertion vector

d) p5'hprrtCMVrep



e) p5'hprrtCMVins



eGFP. The codon frequencies of this reporter have been altered to resemble those found in mammalian genes (hence the term ‘humanised’).

A unique *XhoI* site in the rtTA2<sup>s</sup>-M2-encoding plasmid pUHrT 62-1 was first abolished and replaced with a *NotI* site (Figure 2.3a). The CMV-rtTA2<sup>s</sup>-M2 fragment was then removed by *NotI* plus *BamHI* digestion and ligated into *NotI/BamHI* digested pIRES-hrGFP-1a (Figure 2.3b). The *BamHI* site downstream of the rtTA2<sup>s</sup>-M2 was subsequently removed by digesting pCMVrtTA2<sup>s</sup>-M2ihrGFP with *BamHI* and making the fragment ends flush by incubation with klenow polymerase and dNTPs before religating the plasmid. A *NotI* plus *MluI* digest of the resulting vector yielded the CMV-rtTA-IRES-hrGFP fragment, which was subsequently inserted into *NotI/MluI* digested pSKB1 to create the final replacement targeting vector p5’hp<sup>r</sup>tCMV<sup>rep</sup> (Figure 2.3d).

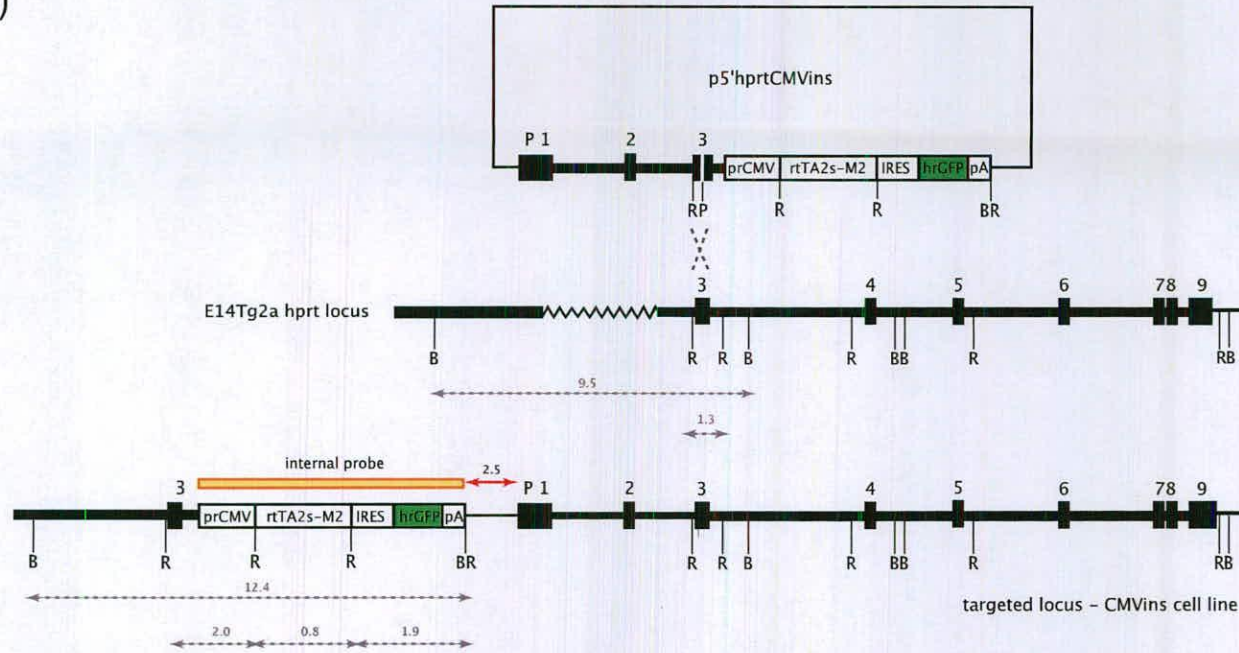
Plasmid pCMVrtTA2<sup>s</sup>-M2ihrGFP (Figure 2.3c) was further manipulated to replace the *MluI* site with a *BamHI* site. Digestion with *NotI* plus *BamHI* yielded the CMV-rtTA2<sup>s</sup>-M2-IRES-hrGFP sequence, which was then ligated into the polylinker of modified pDWM101 (the fragment was inserted into *NotI/BamHI* digested vector), to create the p5’hp<sup>r</sup>tCMV<sup>ins</sup> targeting vector (Figure 2.3e).

## 2.5 The targeting of p5’hp<sup>r</sup>tCMV<sup>rep</sup> and p5’hp<sup>r</sup>tCMV<sup>ins</sup> into E14TG2a cells

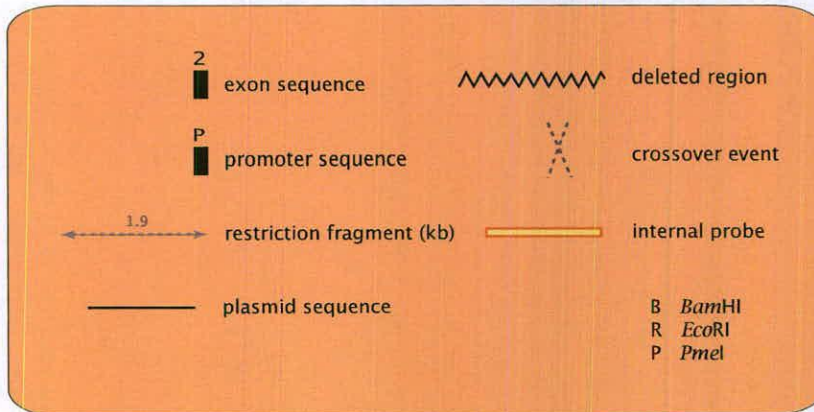
Both targeting vectors were linearised with *PmeI* restriction endonuclease and electroporated into E14TG2a cells. After electroporation, the cells were replated, and the next day seeded out in HAT selection medium. After 6-9 days of selection, HAT resistant colonies were picked and three clones from each transfection were expanded, after which genomic DNA was extracted and digested with *BamHI* or *EcoRI* prior to analysis by Southern blot hybridisation.

Figure 2.4a depicts the predicted targeting of the endogenous *hp<sup>r</sup>t* locus with p5’hp<sup>r</sup>tCMV<sup>ins</sup>, and the expected fragment sizes yielded after digestion with *BamHI* and *EcoRI*. Sequence encoding CMV-rtTA2<sup>s</sup>-M2-IRES-hrGFP was used as an internal probe

a)



Not drawn to scale



b)

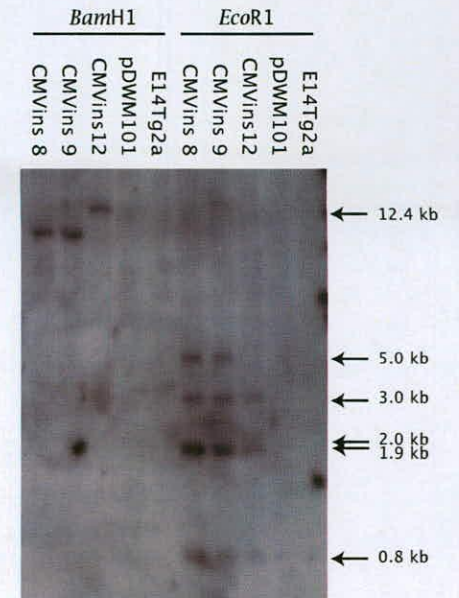


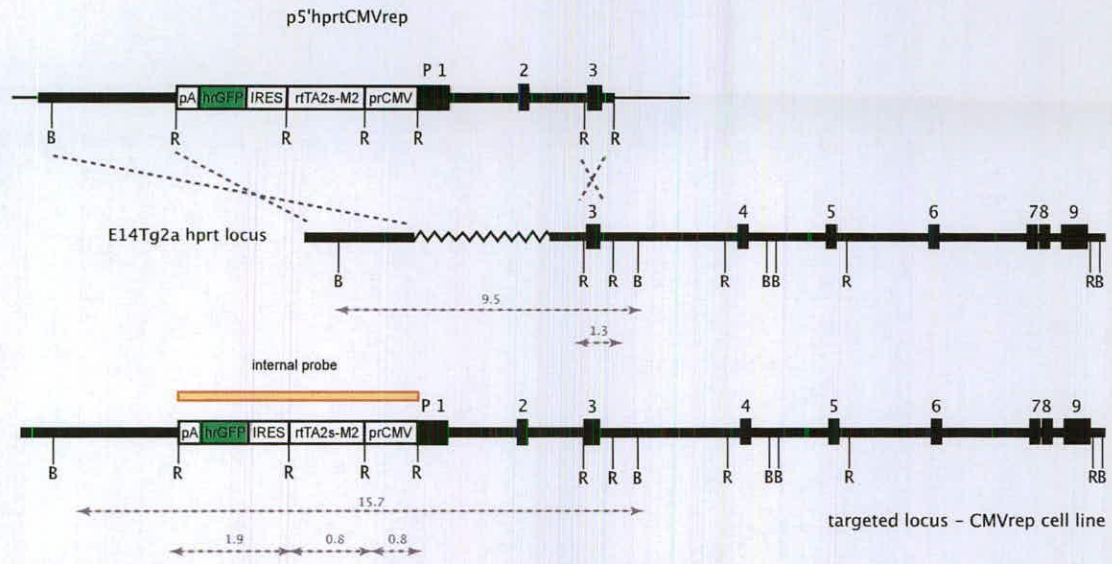
Figure 2.4 Targeting of p5'hprtCMVins to the *hprt* locus  
 a) p5'hprtCMVins was linearised with *PmeI* restriction endonuclease and electroporated into E14TG2a cells. A correctly targeted *hprt* allele is shown with predicted fragment sizes after digestion with *Bam*HI and *Eco*RI. The distance separating the *hprt* promoter and exogenous sequences is 2.5 kb as indicated in red. b) After electroporation and subsequent selection in HAT medium, the genomic DNA from HAT resistant colonies was digested with *Bam*HI and *Eco*RI and analysed by Southern blot hybridization. The sequence encoding CMV-rtTA2s-M2-IRES-hrGFP was used as an internal probe (as shown in a). The pDWM101 lane represents a cell line targeted with the vector alone. pdp: partial digest product

in the analysis of clones by Southern blot hybridisation. For a correctly targeted clone a 12.4 kb fragment was predicted after digestion with *Bam*HI and detection by hybridization to the vector-specific internal probe. Only clone 12 showed the band of predicted size, with clones 8 and 9 exhibiting a smaller sized band of about 9 kb in size (as determined by comparison to *Hind*III digested lambda marker, not shown), Figure 2.4b. The band sizes predicted from *Eco*RI digest were 2.0, 1.9 and 0.8 kb. Again, only clone 12 exhibited all three of these bands, however 0.8 and 1.9 kb bands appeared in the lanes of clones 8 and 9. The 2.0 kb band was absent in the digests of clones 8 and 9 and a band of around 5 kb was detected, which was not found in the digest of clone 12. The presence of unexpected band sizes found in both the *Bam*HI and *Eco*RI digests of clones 8 and 9 suggest that the integration event in these clones involved an alternative and unpredicted configuration of exogenous sequence at the *hprt* locus. A band of about 3 kb in size was seen in the lanes of all three clones in the *Eco*RI digest. This could possibly be a partial digest product of the 0.8 and 1.9 kb *Eco*RI fragments.

Figure 2.5b depicts the targeting of p5'*hprt*CMV<sup>rep</sup> into E14TG2a cells. Correctly targeted integrants yielded a 15.7 kb band upon *Bam*HI digest, and 1.9 and 0.8 kb bands upon digestion with *Eco*RI. All three clones analysed showed the predicted bands after *Bam*HI and *Eco*RI digest and subsequent hybridisation to the internal probe. Once again a 3.0 kb band was seen after *Eco*RI digest of the clones, which probably represents a partial digest product.

The targeting efficiency of the replacement vector was approximately 10 fold higher than that of the insertion vector, with the efficiency for p5'*hprt*CMV<sup>ins</sup> approximately  $10^{-6}$  HAT resistant colonies generated per cell electroporated with vector DNA, and for p5'*hprt*CMV<sup>rep</sup> approximately  $10^{-5}$ . The difference in targeting efficiency is most likely due to the replacement vector encoding longer stretches of homology to the endogenous *hprt* locus and also the fact that it is comprised of sequence isogenic with E14TG2a cells.

a)



Not drawn to scale

|  |     |                           |   |                 |
|--|-----|---------------------------|---|-----------------|
|  | 2   | exon sequence             |   | deleted region  |
|  | P   | promoter sequence         |   | crossover event |
|  | 1.9 | restriction fragment (kb) |   | internal probe  |
|  |     | plasmid sequence          |   |                 |
|  |     |                           | B | <i>Bam</i> H1   |
|  |     |                           | R | <i>Eco</i> R1   |

b)

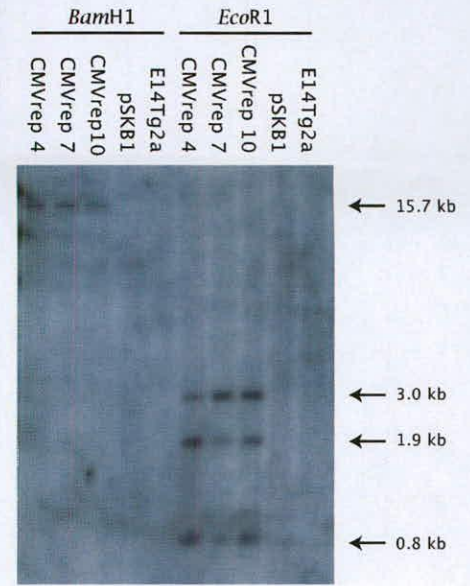


Figure 2.5 Targeting of p5'hpertCMVrep to the *hpert* locus  
 a) p5'hpertCMVrep was linearised with *Pme*I restriction endonuclease and electroporated into E14TG2a cells. A correctly targeted *hpert* allele is shown with predicted fragment sizes after digestion with *Bam*HI and *Eco*RI. b) After electroporation and subsequent selection in HAT medium, the genomic DNA from HAT resistant colonies was digested with *Bam*HI and *Eco*RI and analysed by Southern blot hybridization. The sequence encoding CMV-rtTA2s-M2-IRES-hrGFP was used as an internal probe (as shown in a). The pSKB1 lane represents a cell line targeted with the vector alone. pdp: partial digest product.

## 2.6 Analysis of rtTA2<sup>s</sup>-M2 expression by RT-PCR

The six targeted clones, CMVins 8, 9 and 12, and CMVrep 4, 7 and 10 were analysed for expression of the rtTA2<sup>s</sup>-M2 transcript (Figure 2.6a). RNA was extracted from each cell line and quantified before being used as a template for the production of cDNA by incubation with Superscript II reverse transcriptase (Invitrogen), oligoT primers (that bind to the polyA tracts of messenger RNA), and dNTPs. The cDNA derived from this reaction was then analysed for the presence of rtTA2<sup>s</sup>-M2 sequences by using a primer set designed to amplify a 700 base pair fragment of rtTA2<sup>s</sup>-M2. E14TG2a cDNA was used as a negative control. A 700bp band was detected for all clones analysed, with no fragments amplified in the minus reverse transcriptase controls (data not shown) or E14TG2a control. All targeted clones therefore expressed the rtTA2<sup>s</sup>-M2 transcript.

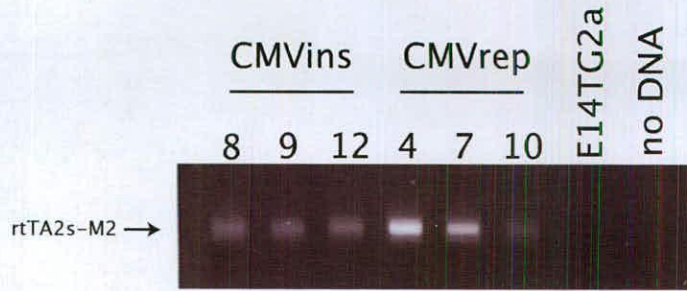
## 2.7 Analysis of rtTA2<sup>s</sup>-M2 function by transient luciferase assay

Following on from the detection of the rtTA2<sup>s</sup>-M2 by RT-PCR, the production of functional rtTA2<sup>s</sup>-M2 protein was determined by transiently transfecting all cell lines with the reporter plasmid pBI-L which encoded a TRE and linked firefly luciferase gene, followed by the assay of luciferase levels.

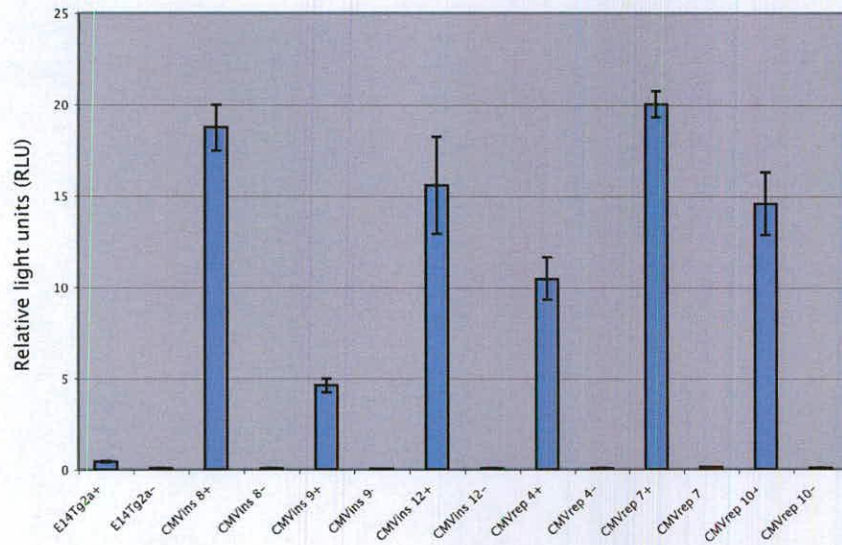
The targeted clones were transfected with pBI-L and pREN (a *renilla* luciferase-expressing plasmid which acted as an internal control for variations in transfection efficiency; see materials and methods). The following day 1 µg dox was added to each cell line (with some cultures from each cell line remaining free of dox), and 48 hours later the cells were lysed and analysed for the expression of the luciferase reporters.

All clones were shown to induce the expression of firefly luciferase in the presence of dox (Figure 2.6b), although the degree of expression varied between the cell lines. Negligible basal levels of luciferase expression were detected in the absence of dox. The fold induction of luciferase is shown graphically in Figure 2.6c. It can be seen that whilst some clones induced luciferase to a higher extent than others, there were no obvious

a)



b)



c)

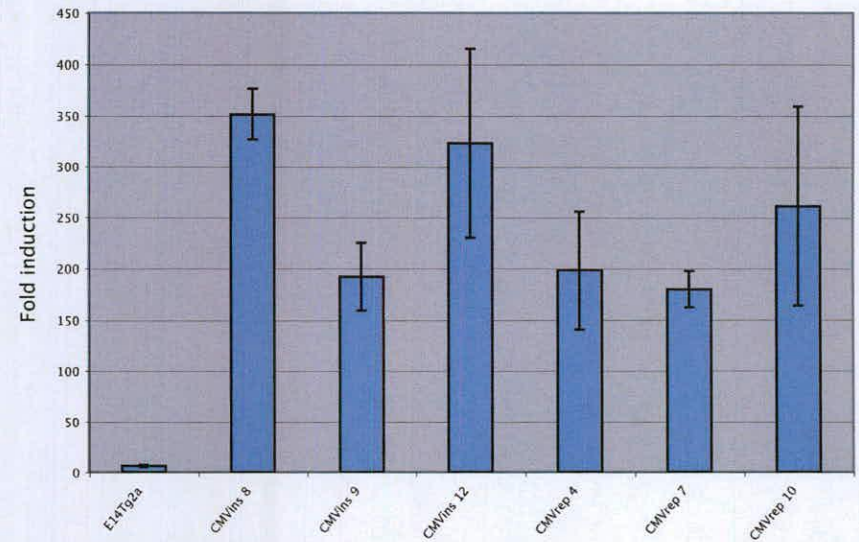


Figure 2.6 Analysis of rtTA2s-M2 expression. a) RT-PCR of the six targeted cell lines using primers specific for rtTA2s-M2. A 700bp band was detected in all targeted clones. b) The cell lines were transiently transfected with a reporter plasmid encoding a TRE and linked firefly luciferase sequence, and the following day some cells of each transfected clone were treated with 1ug/ml dox whilst others were left untreated. 48 hours later the cells were assayed for luciferase expression. c) Representation of the same experiment as fold induction of luciferase for each clone. All transfections were normalised by cotransfection with a renilla luciferase-expressing plasmid. All transfections were carried out in triplicate.

differences in the levels of induction between the CMVins group and the CMVrep group of clones.

## **2.8 Flow cytometry of targeted clones**

Flow cytometry was used to analyse the expression of rtTA2<sup>s</sup>-M2 at a cellular level. As described in section 2.3, the rtTA2<sup>s</sup>-M2 ORF was linked to an IRES-hrGFP sequence to create a bicistronic element. Once targeted, expression of hrGFP should therefore parallel that of rtTA2<sup>s</sup>-M2, and thus acts as a reporter of transactivator expression.

The analysis of all the targeted cell lines by flow cytometry is shown in Figure 2.7. E14TG2a acted as a negative control in order to establish the hrGFP positive region, and the percentage of cells falling within this positive region is shown in the bottom right-hand corner of each plot. Although all the clones analysed exhibited some degree of hrGFP fluorescence and therefore rtTA2<sup>s</sup>-M2 expression, only small populations of cells were positive for hrGFP. In addition those cells that were positive, were only marginally so, as determined by the very slight shift along the x-axis. It would appear that the CMVins cell lines contain a higher number of hrGFP positive cells on average, when compared to the CMVrep cell lines, however more clones would need to be derived and analysed in order to establish if this is a consistent trend.

In order to produce a system which exhibits population-wide induction of a transgene in the presence of dox, it is necessary that rtTA2<sup>s</sup>-M2 is expressed in every cell within a population, and in addition, at levels sufficiently high enough to induce homogenous expression of the transgene. RT-PCR and luciferase analysis of the CMVins and CMVrep clones showed that rtTA2<sup>s</sup>-M2 was expressed and active in all clones analysed, however flow cytometry revealed that rtTA2<sup>s</sup>-M2 expression was very mosaic and at low levels. It was therefore deemed necessary to utilise another promoter to drive the expression of rtTA2<sup>s</sup>-M2.

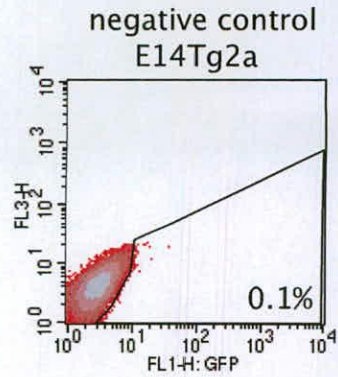
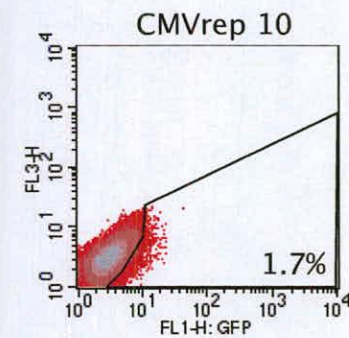
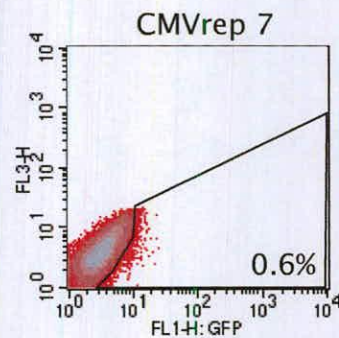
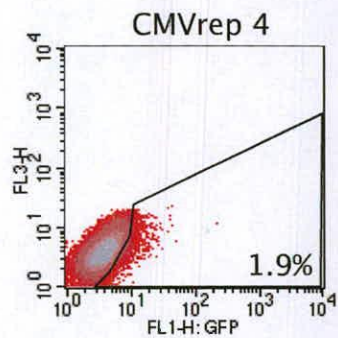
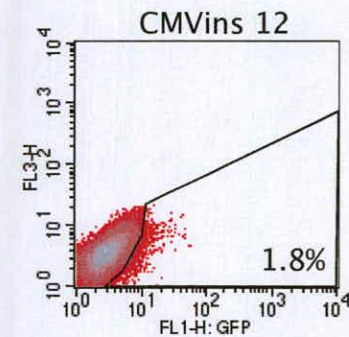
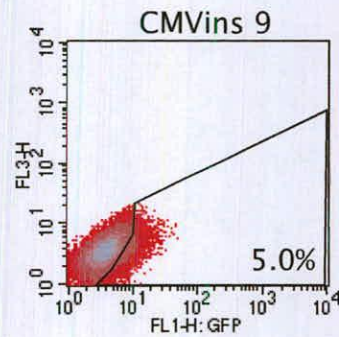
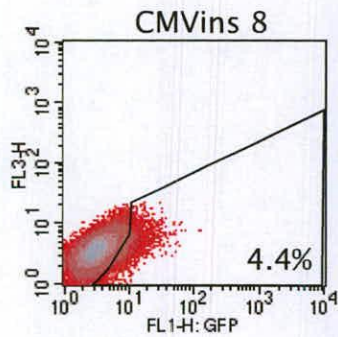


Figure 2.7 Flow cytometry of targeted cell lines. All six cell lines were analysed for hrGFP expression. hrGFP fluorescence is shown on the x-axis, with the y-axis redundant. E14TG2a was used as a negative control to determine the hrGFP positive region (this region is shown). The percentage of hrGFP positive cells falling within this region is indicated for each clone.



## 2.9 The CAG promoter to drive expression of the reverse transactivator

The CAG promoter is a composite promoter comprised of viral and eukaryotic sequences. In addition to a CMV enhancer and chicken  $\beta$ -actin promoter, it encodes a chimeric intron (derived from chicken  $\beta$ -actin, and rabbit  $\beta$ -globin introns) and a splice acceptor sequence. Transgenes placed downstream of this regulatory element have been reported to be expressed at high levels in ES cells (Chung *et al.*, 2002, Niwa *et al.*, 2000). The CAG element was therefore chosen as the second candidate promoter to drive expression of rtTA2<sup>S</sup>-M2.

In order to link the CAG promoter sequence to the transactivator, plasmid CAGGS-S was digested with *Eco*RI plus *Bam*HI and the rtTA2<sup>S</sup>-M2-IRES-hrGFP-fragment yielded from the digestion of pCMV rtTA2<sup>S</sup>-M2hrGFP with the same restriction enzymes was ligated into the digested pCAGGS-S backbone to generate CAG-rtTA2<sup>S</sup>-M2hrGFP (Figure 2.8a).

A *Spe*I/*Bam*HI double digest of plasmid CAG-rtTA2<sup>S</sup>-M2hrGFP excised the CAG-rtTA2<sup>S</sup>-M2-IRES-hrGFP-pA fragment. After the 5' protruding ends were made flush with klenow polymerase this fragment was inserted into *Not*I digested, pSKB1 (which had also been treated with klenow polymerase), to create the targeting vector p5'hprrtCAG<sup>rep</sup>. The orientation of the inserted sequences was determined by restriction digest. The final p5'hprrtCAG<sup>rep</sup> targeting vector contained the CAG-rtTA2<sup>S</sup>-M2-IRES-hrGFP-pA sequence in an opposite orientation to that of the 5' to 3' transcriptional orientation of genomic *hprrt* sequence encoded within same the vector (Figure 2.8c). This orientation is identical to that of the CMV-rtTA2<sup>S</sup>-M2-IRES-hrGFP sequences contained within the p5'hprrtCMV<sup>rep</sup> targeting vector used in the previous electroporation.

Due to the increased targeting efficiency of the pSKB1 vector over that of the pDWM101 vector, and also the possibility of targeted sequences becoming excised after use of the pDWM101 vector (see discussion), only pSKB1 was used for the targeting of CAG encoding exogenous sequences to the 5' *hprrt* locus.

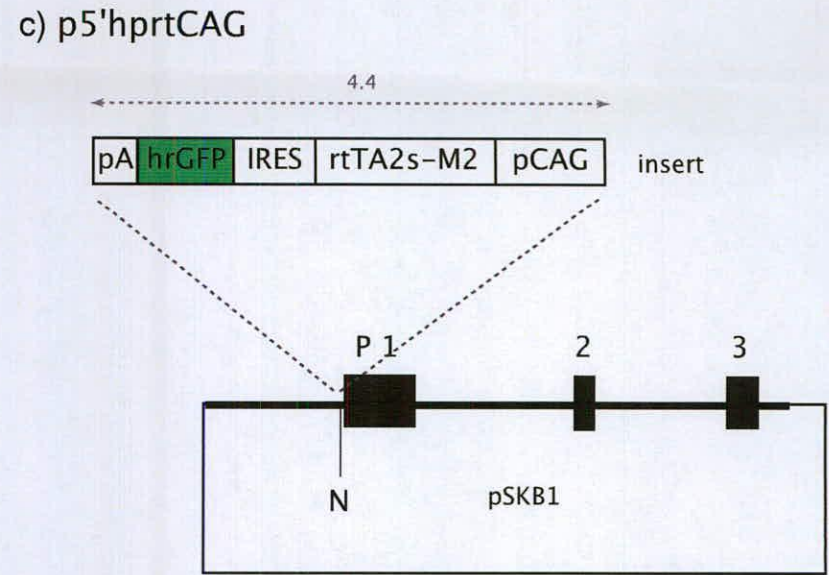
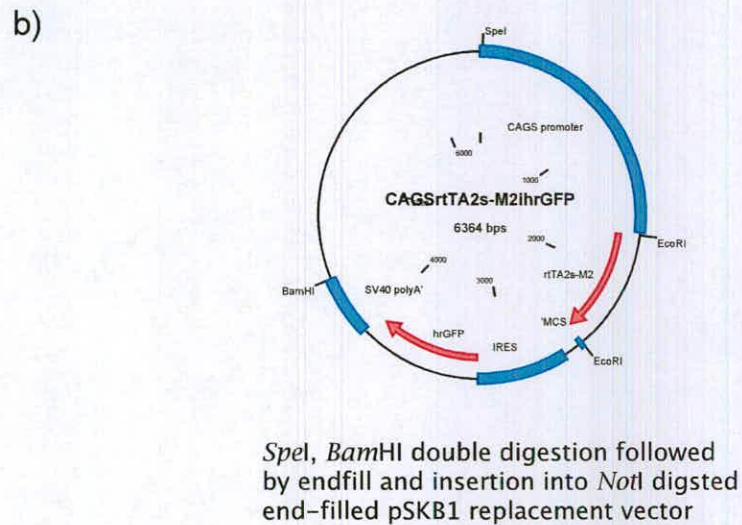
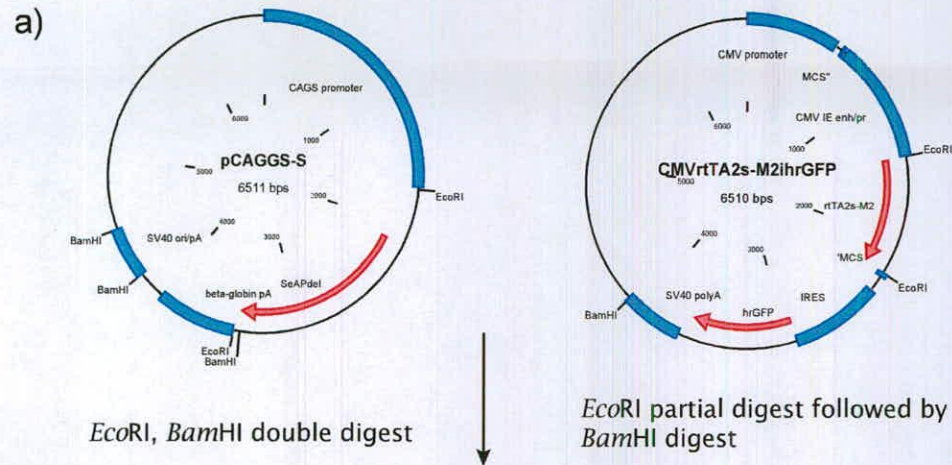


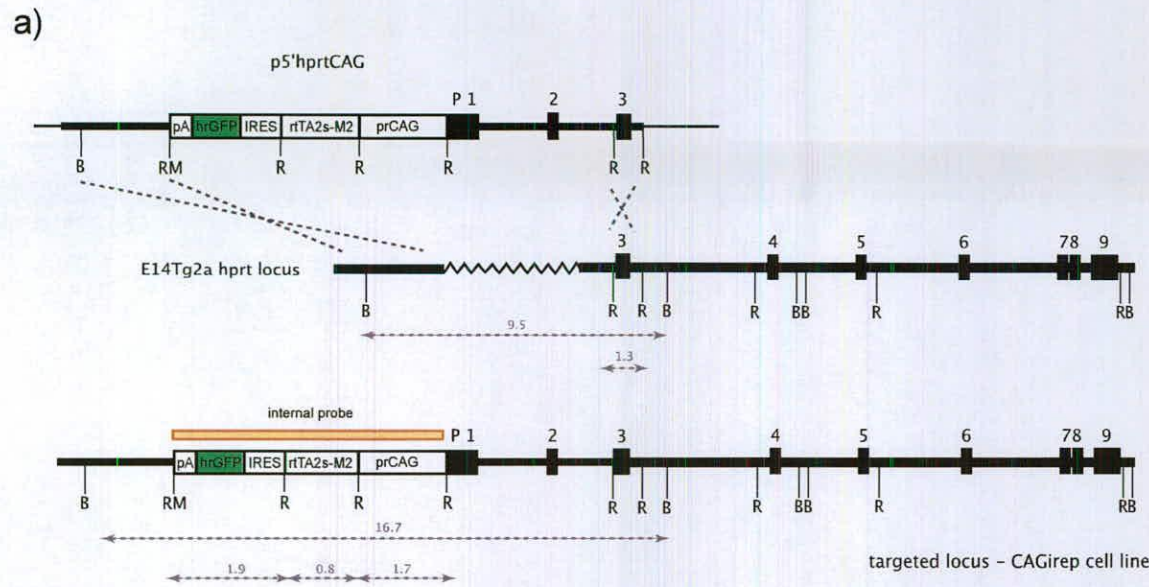
Figure 2.8 Cloning strategy for constructing the CAG-rtTA2s-M2-IRES-hrGFP-pA fragment and subsequent insertion into the pSKB1 targeting vector. a) Partial digest of pCMV-tTA2s-M2hrGFP with *EcoRI*, followed by digestion with *BamHI* yielded the tTA2s-M2-IRES-hrGFP-pA sequence, which was then inserted into *EcoRI/BamHI* restricted pCAGGS. b) *SpeI/BamHI* double digestion of the resulting plasmid yielded the CAG-rtTA2s-M2-IRES-hrGFP-pA fragment. The 3' recessed ends of this fragment were made flush with klenow polymerase and then inserted into *NotI* digested, pSKB1 replacement vector (the ends of this vector had also been made flush by klenow polymerase treatment), to create the p5'hrprtCAG targeting vector

The targeting vector was linearised with *PmeI* and electroporated into E14TG2a cells (Figure 2.9a). The next day the cells were plated out in HAT selection. After 6-9 days colonies were picked and eight clones were expanded and analysed by Southern blot hybridisation, using the CAG-rtTA2<sup>s</sup>-M2-IRES-hrGFP-pA sequence as an internal probe. The fragment sizes predicted from *BamHI* and *EcoRI* digestion of genomic DNA from correctly targeted clones are shown in Figure 2.9a. Digestion with *BamHI* was expected to yield a 16.7 kb band, and indeed this fragment was observed for all clones except clone CAGrepA3 (Figure 2.9b). Correctly targeted clones were expected to yield 1.9, 0.8 and 1.7 kb bands upon *EcoRI* digestion and subsequent hybridisation. Surprisingly only the 0.8 kb band was visible. The reason for the absence of the expected 1.9 and 1.7 kb bands has yet to be determined. Once again no bands were seen in the lane of CAGrepA3 suggesting that the endogenous *hprt* locus had not been targeted in this line.

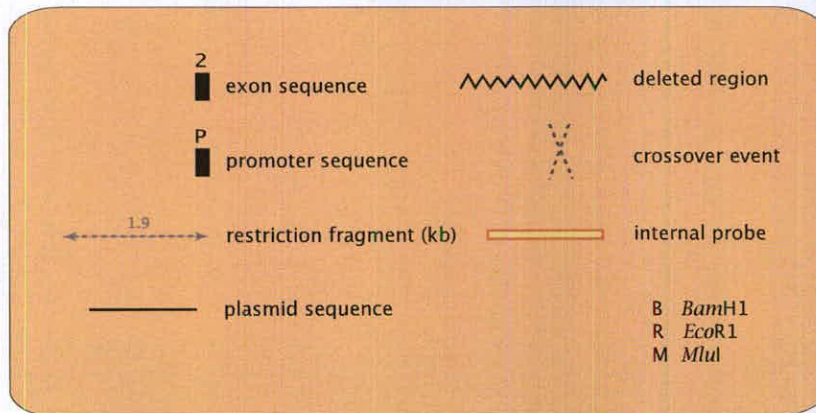
## 2.10 Flow cytometry of the CAGrep cell lines

The analysis of the CMVins and CMVrep clones by flow cytometry proved to be the most definitive test of rtTA2<sup>s</sup>-M2 expression. Four correctly targeted CAGrep cell lines were therefore analysed for expression of hrGFP by flow cytometry.

The outcome of this analysis is shown in Figure 2.10. In marked contrast to the previous CMVins and CMVrep cell lines, all four CAGrep clones analysed showed a definite and relatively homogenous shift along the x-axis, with the vast majority of cells possessing a clear hrGFP positive phenotype. The proportion of hrGFP positive cells between cell lines was remarkably consistent, with values of over 98% for each clone. This result establishes the CAG sequence as a suitable choice of promoter to drive expression of rtTA2<sup>s</sup>-M2 at the 5' region of *hprt*.



Not drawn to scale



b)

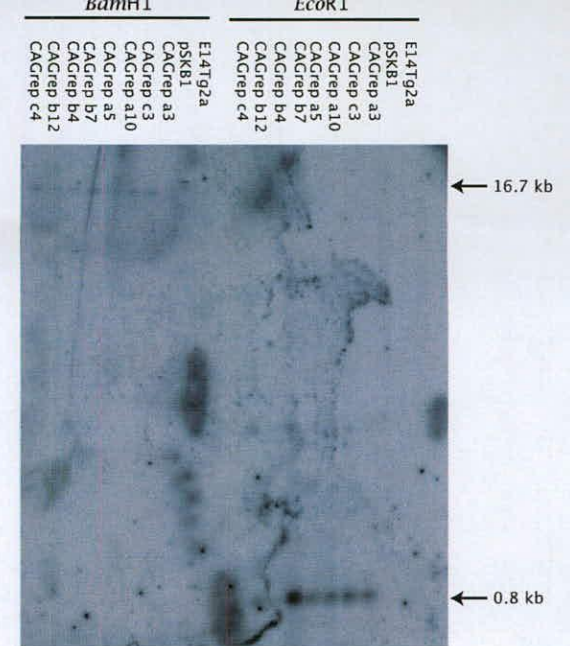


Figure 2.9 Targeting of p5'hprrCAG to *hprr* a) p5'hprrCAG was linearised with *Pme*I restriction endonuclease and electroporated into E14TG2a cells. A correctly targeted *hprr* allele is shown with predicted fragment sizes after digestion with *Bam*HI and *Eco*RI. b) After electroporation and subsequent selection in HAT medium, the genomic DNA from HAT resistant colonies was digested with *Bam*HI and *Eco*RI and analysed by Southern blot hybridization. The sequence encoding CAG-rTA2s-M2-IRES-hrGFP was used as an internal probe (as shown in 2.9a). The pSKB1 lane represents a cell line targeted with the vector alone.

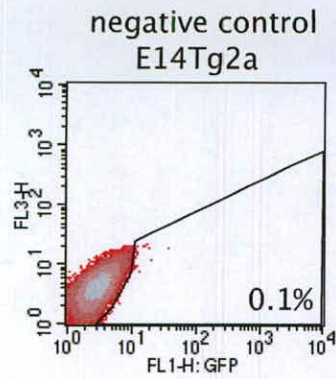
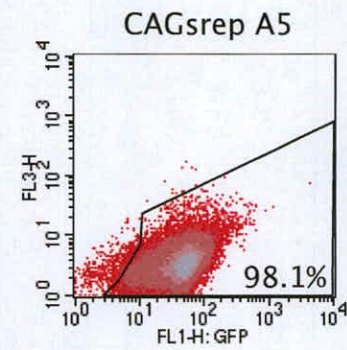
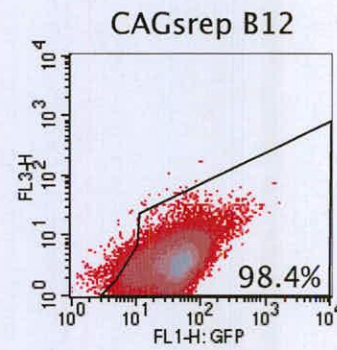
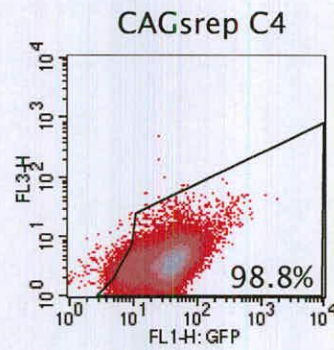
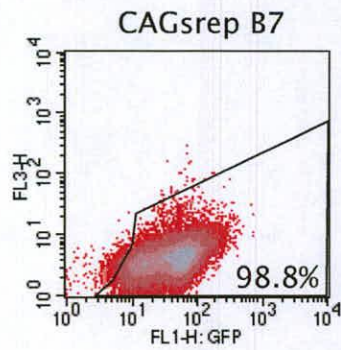


Figure 2.10 Analysis of targeted cell lines by flow cytometry. Four CAGrep clones were analysed for hrGFP expression. hrGFP fluorescence is shown on the x-axis, with the y-axis redundant. E14TG2a was used as a negative control to determine the hrGFP positive region (this region is shown). The percentage of hrGFP positive cells falling within this region is indicated for each clone.



## 2.11 Discussion

### **The targeting of rtTA2<sup>s</sup>-M2 to the 5' region of *hpert* using the pDWM101 and pSKB1 correction vectors**

Two types of targeting vector were used to independently integrate the reverse transactivator and linked sequences into the 5' region of *hpert*. The main reason for utilising two vector types was the different mode of integration for each; pDWM101 integrates as an insertion vector with one cross-over event taking place between the endogenous *hpert* locus and homologous vector sequences, whereas pSKB1 integrates as a replacement vector, via two cross-over events (Figure 2.2). As a consequence, the integrated transgenic sequences were predicted to be positioned at different distances from the *hpert* promoter. Use of the pDWM101 targeting vector resulted in the integration of transgenic sequences 2.5 kilobases upstream of the *hpert* promoter, whilst the pSKB1 vector integrated sequences immediately upstream of the promoter. Whilst both vectors have previously been used to target transgenic sequences to the *hpert* locus (Bronson *et al.*, 1996), previous studies have not included the analysis of transgene expression on a single cell basis. As it is essential that the rtTA2<sup>s</sup>-M2 transactivator is expressed throughout the ES cell population in order for the tet-system to be utilised effectively, it was decided to use both vectors to independently target the reverse transactivator sequence, with the view that one vector may give rise to superior rtTA2<sup>s</sup>-M2 expression. The rtTA2<sup>s</sup>-M2 ORF was linked to a fluorescent reporter gene via an IRES sequence to enable the detection of rtTA2<sup>s</sup>-M2 levels indirectly by the flow cytometric analysis of hrGFP expression.

No substantial difference in expression of hrGFP was observed in cell lines derived from the targeting of the CMV-rtTA2<sup>s</sup>-M2-IRES-hrGFP encoding sequence, using the two different vectors. As described previously, the correct integration of the pDWM101 vector results in the duplication of exon 3 (in the region upstream of the inserted exogenous sequences). This duplication event could potentially lead to the excision of

the integrated exogenous sequences via homologous recombination between the duplicated regions. Although the continued presence of HAT medium could select against this event, such a situation would not allow the further targeting of the TRE into the 3' region of *hprt*, as this requires the selection of an HPRT negative phenotype (to be described in the next chapter). Due to these concerns the sequence encoding the CAG promoter, and linked transactivator, was targeted using only the pSKB1 targeting vector. An additional advantage to using this vector was the increased efficiency of targeting over the pDWM101 insertion vector.

### **The cytomegalovirus immediate early enhancer and promoter to drive the expression of rtTA2<sup>s</sup>-M2**

It was found that despite the initial promising results of detecting the rtTA2<sup>s</sup>-M2 transcript by RT-PCR, and the induction of luciferase upon transient transfection of a TRE and linked luciferase reporter plasmid into the newly targeted clones (upon the addition of dox), flow cytometry analysis revealed that expression of hrGFP (and therefore rtTA2<sup>s</sup>-M2) was very low and mosaic, with most cells possessing a hrGFP negative phenotype. This observation was very surprising as it has been reported that the CMV promoter and enhancer sequence is able to direct high expression of transgenes both in numerous cell lines (Muller *et al.*, 1990; Keating *et al.*, 1990) and in transgenic animals, (Furth *et al.*, 1991).

It has previously been described that the transient transfection of a plasmid encoding an eGFP ORF directly linked to a CMV promoter gave rise to eGFP fluorescence in over 50% of the population in several ES cell lines, as determined by flow cytometry, (the authors state that under optimal conditions 50%-80% transfection efficiencies could be achieved; Ward & Stern 2002). However an earlier paper by an independent group revealed evidence that the transient transfection of a plasmid encoding an IRES-hrGFP sequence, linked to a CMV promoter did not result in hrGFP fluorescence in D3 or J1 ES cells, or indeed in ES cells that underwent differentiation by embryoid body formation (Chung *et al.*, 2002).

The conflicting results of these two studies can possibly be explained by the fact that the former study utilises a plasmid in which the fluorescent reporter sequence is directly linked to the CMV promoter, whereas the hrGFP reporter sequence used by the latter group is linked to the CMV promoter via an IRES. It has been shown that IRES-dependent gene expression is significantly lower than that of cap-dependent gene expression, and therefore it is possible that translation of the hrGFP sequence was too low for hrGFP fluorescence to be detected by flow cytometry (Mizuguchi *et al.*, 2000). In addition it is likely that the transfection efficiency achieved by Chung *et al* was very low, owing to a suboptimal protocol for transfecting ES cells. In the study by Chung *et al*, adherent cells were transfected with DNA (using Lipofectamine™ 2000), whereas Ward and Stern transfected cells in suspension, (just after trypsinisation), using the same reagent. This author has found the overall transfection efficiency is increased using the latter protocol.

In light of these studies it is possible to explain the poor expression of hrGFP in CMVrtTA2<sup>s</sup>-M2rep and CMVrtTA2<sup>s</sup>-M2ins cells. The targeting of exogenous sequences usually results in the integration of just one copy of sequence, (in contrast several copies of plasmid can reside in a single cell after transient transfection). In addition the strategy implemented in this study utilises an IRES sequence upstream of the hrGFP reporter, which as stated previously probably results in lower expression of the IRES-dependent second gene (when compared to the upstream cap-dependent gene). A previous study has also shown that when compared to the *apolipoprotein (apo)* locus, the targeting of transgenes to the 5' region of *hpvt* gives rise to relatively low levels of expression (Hatada *et al.*, 1999). Therefore it is possible that in the context of the *hpvt* locus, and at single copy number, the CMV promoter is not able to drive high level expression of the integrated transgenic sequences.

It is clear from the study of the CMVrep and CMVins cell lines that all clones exhibit some degree of mosaicism with regards to the expression of the hrGFP reporter. It is possible that mosaicism is characteristic of CMV-driven transgene expression in ES cells.

In the previous studies described (Chung *et al.*, 2002, Ward & Stern., 2002), expression of the transgenes involved the transient transfection of cells, and therefore cannot give any insight into the potential mosaic nature of this promoter in ES cells. There have however been previous reports that the CMV promoter gives rise to mosaic transgene expression in certain tissues *in vivo* (Furth *et al.* 1994).

It is interesting to note that although all clones were able to induce the expression of a TRE-linked firefly luciferase reporter in the presence of dox (after transient transfection of the reporter plasmid), in reality it was only a very small proportion of the population that had the ability to induce luciferase. This analysis highlights how assays involving the detection of the luciferase reporter can often be deceptive. As detection involves the lysis of the cells prior to analysis, one does not gain any insight into whether luciferase levels are indicative of population-wide effects, or the result of induction in only a small sub-population of cells. In addition, the luciferase assay is based on an enzymatic reaction (that is the conversion of luciferin substrate into oxyluciferin by luciferase) and is therefore a sensitive assay that can amplify low levels of luciferase. This provides an explanation of how populations that contain few rtTA2<sup>s</sup>-M2-expressing cells are able to induce luciferase reporter to reasonable levels (Figure 2.6b and 2.6c).

### **The CAG promoter to drive the expression of rtTA2<sup>s</sup>-M2**

The second candidate promoter chosen to drive expression of rtTA2<sup>s</sup>-M2 was the CAG promoter. Several studies which involved the placement of transgene sequences under the regulation of the CAG promoter reported high levels of transgene expression, (Nathwani *et al.*, 2001; Niwa *et al.*, 1991; Miyazaki *et al.*, 1989). Robust transgene expression driven by the CAG promoter has also been observed in studies undertaken in ES cells (Niwa *et al.*, 2000., Chambers *et al.*, 2003).

The CAG promoter is a composite promoter which contains both eukaryotic and viral sequences. The capacity of this promoter to drive high levels of transgene expression most likely results from the presence of numerous enhancer elements encoded within its

sequence. The promoter was developed over several years, with each modification increasing the strength of the promoter. It was first observed that inserting the 3' region of the second intron of a rabbit  $\beta$ -globin gene (including the splice acceptor and 5' part of the third exon), into the first intron of the chicken  $\beta$ -actin promoter, increased the expression of a downstream lacZ transgene (Miyazaki *et al.*, 1989). This promoter, termed the AG promoter ('AG' stands for actin and globin), was then further modified by the addition of a CMV immediate-early enhancer element to form the CAG promoter (Niwa *et al.*, 1991). It is thought that both the first intron of the  $\beta$ -actin gene, and the 3' region of the rabbit  $\beta$ -globin gene contain enhancer-like activity, in addition to the CMV sequence. The presence of an intron in the CAG regulatory sequence may also enhance transcript stability as there is evidence that splicing itself is a prerequisite for stable RNA formation in eukaryotic cells (Hamer and Leder., 1979).

In contrast to the CMV promoter, when the CAG promoter was used to drive the expression of rtTA2<sup>s</sup>-M2 and hrGFP reporter sequences, over 98% of cells exhibited a hrGFP positive phenotype (Figure 2.10). The hrGFP levels approached two log shifts as represented on the two dimensional density plots. Whilst the majority of targeted cells fell within a discrete population of hrGFP positive cells (the lighter the areas on the plot indicate the areas of higher cell density), there did appear to be a small population of cells which expressed lower levels of fluorescence. This could indicate a slight mosaicism in transgene expression upon regulation by the CAG promoter, or alternatively epigenetic differences in the 5' region of *hprt* between cells, which affect the levels of transgene expression.

In summary the CAG promoter is clearly a superior promoter in driving transgene expression at the 5' *hprt* locus in ES cells. In contrast to the CMV promoter, high and fairly uniform levels of transgene fluorescence were detected. As it has previously been shown that the CMV promoter is active in ES cells (Ward and Stern., 2002), the poor transgene expression observed for the CMVins and CMVrep lines was most likely due to a combination of factors, including the single copy insertion of sequences, the presence of a IRES sequence upstream of the hrGFP ORF and possible low potential of the region

of chromatin targeted to give rise to high levels of transgene expression. All these restrictions are however overcome with utilisation of the CAG promoter. This is probably due to the presence of numerous enhancer sequences present within the promoter that enable the strong expression of the targeted sequences.

With the establishment of ES cell lines that robustly express rtTA2<sup>s</sup>-M2, the next step was to target the TRE and linked transgene sequences into the 3' region of the *hprt* locus. This will be described in the next chapter.

## Chapter 3

### Integration of the transactivator-response element into the 3' coding region of *hprt*

#### 3.1 Introduction

In the previous chapter it was described how homogenous expression of the reverse transactivator (rtTA2<sup>s</sup>-M2) was achieved by juxtaposing the rtTA2<sup>s</sup>-M2 open reading frame to a CAG promoter sequence and targeting these exogenous sequences to the 5' region of the *hprt* gene in E14TG2a cells. This chapter details the targeting of the transactivator-response element (TRE) and linked DsRED2 reporter into the coding region of *hprt*, by utilising an already existing replacement vector designed to target the 3' end of the gene. This positions the TRE at a distance of approximately 30 kb from the upstream transactivator and linked promoter. Analysis of the induction of DsRED2 expression upon addition of dox is also described.

### 3.2 Construction of the 3'hpTREDsRED2F and 3'hpTREDsRED2R targeting vectors and electroporation into CAGrepB7 cells

Plasmid BigTlucTREDsRED2 (kind gift from Barry Heavey) contains a bidirectional TRE which is juxtaposed to a luciferase open reading frame in one direction and a DsRED2 open reading frame in the other (Figure 3.1a). The TRE sequence, linked reporter genes, and corresponding polyA sequence were extracted as an *AscI*, *ScaI* fragment from plasmid BigTlucTREDsRED2, and the protruding fragment ends were made flush by incubation with klenow polymerase and dNTPs. The M42 targeting vector (gift of M. Richardson and A.J.H.Smith, unpublished work), is a plasmid containing 9kb of *hpT* genomic sequence (spanning introns 4 to exon 9), and encodes a *NotI* site within exon 6 (which is used to insert transgenic sequences). This vector was digested with *NotI* and subjected to klenow treatment. The M42 vector and fragment sequences were then ligated (Figure 3.1b and 3.1c). The resulting blunt-ended ligation produced two vector products. Plasmid 3'hpTREDsRED2F contained the DsRED2 sequence in the same 5' to 3' orientation as that of the genomic *hpT* coding sequence present in the M42 vector (Figure 3.1b), and plasmid 3'hpTREDsRED2R contained the DsRED2 sequence in the opposite orientation (Figure 3.1c).

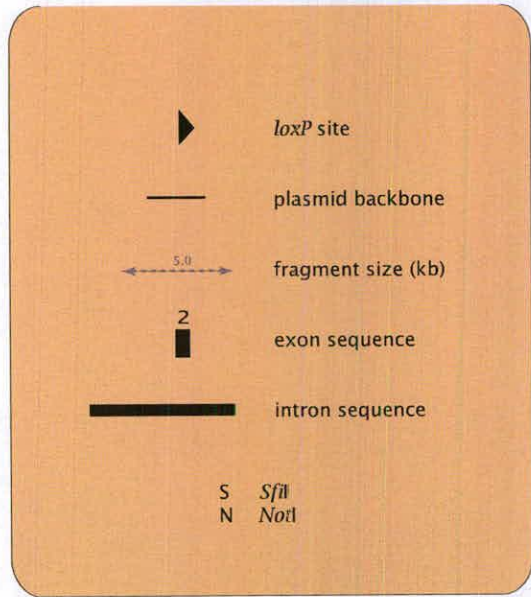
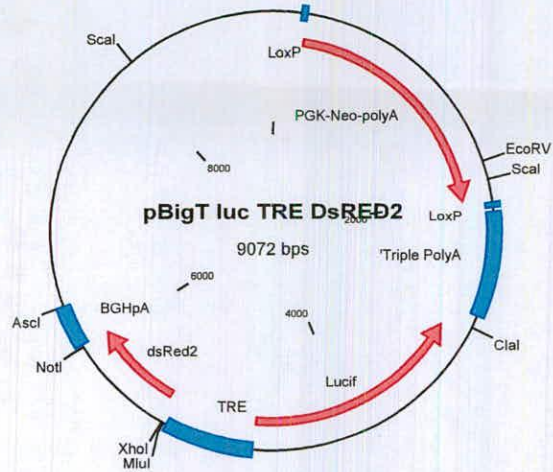
Both the 3'hpTREDsRED2F and 3'hpTREDsRED2R targeting vectors were linearised with *SfiI* restriction endonuclease. This restriction enzyme cuts within each vector at a unique site located at one end of the cloned insert. The linearised vectors were then electroporated into CAGrepB7 cells, after which they were replated and passaged non-selectively for 7 days. After a week, 6-thioguanine (6-TG) selection was added to the transfected cells to select for those cells that exhibited a loss of HPRT activity (HPRT catalyses the conversion of 6-TG into a toxic metabolite that results in cell death). 6-TG resistance was indicative of a correct targeting event due to the disruption of the *hpT* ORF after integration, Figure 3.2a and 3.2b.

6-TG resistant colonies were picked, expanded and analysed by Southern blot hybridisation to confirm correct integration of the TRE-containing sequence. Figures

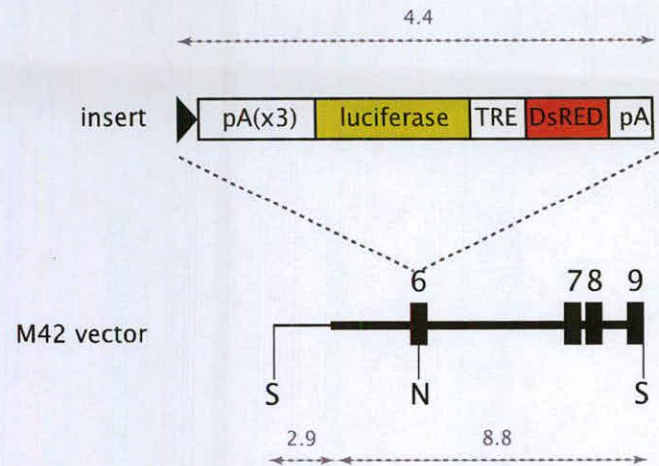
Figure 3.1 The cloning of the TRE and linked reporter sequences into the M42 vector

a) Plasmid BigTlucTREdsRED2 was digested with *AscI* and *ScaI* to excise the TRE and linked reporter sequences. The protruding ends of the TRE fragment were made flush by treatment with klenow polymerase. The sequences were then inserted into *NotI* digested M42 targeting vector (which had also been treated with klenow polymerase) to create the vectors b) 3'hprtTREdsRED2F and c) 3'hprtTREdsRED2R that harboured the inserted sequences in opposite orientations. Both vectors are depicted in their linearised forms (after digestion with *SfiI* restriction endonuclease).

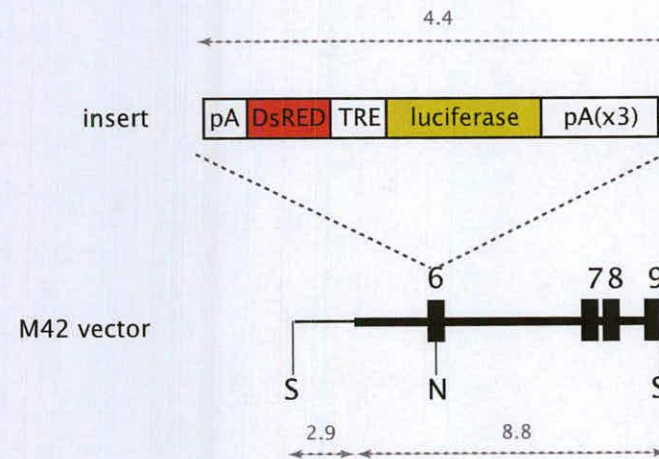
a)



b) 3'hprtTREDsRED2F



c) 3'hprtTREDsRED2R



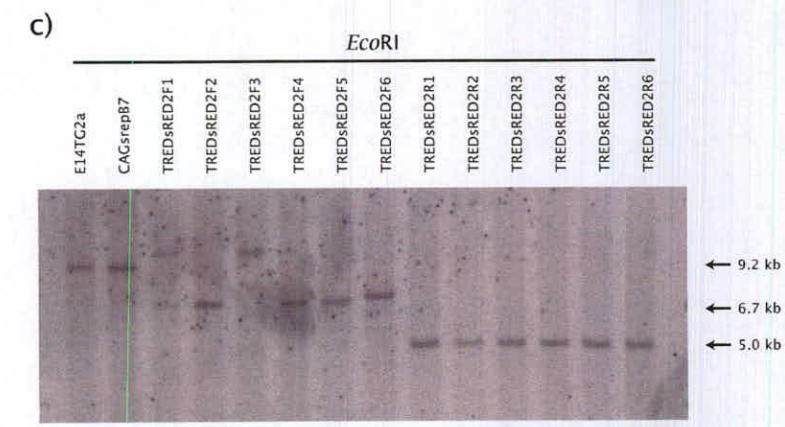
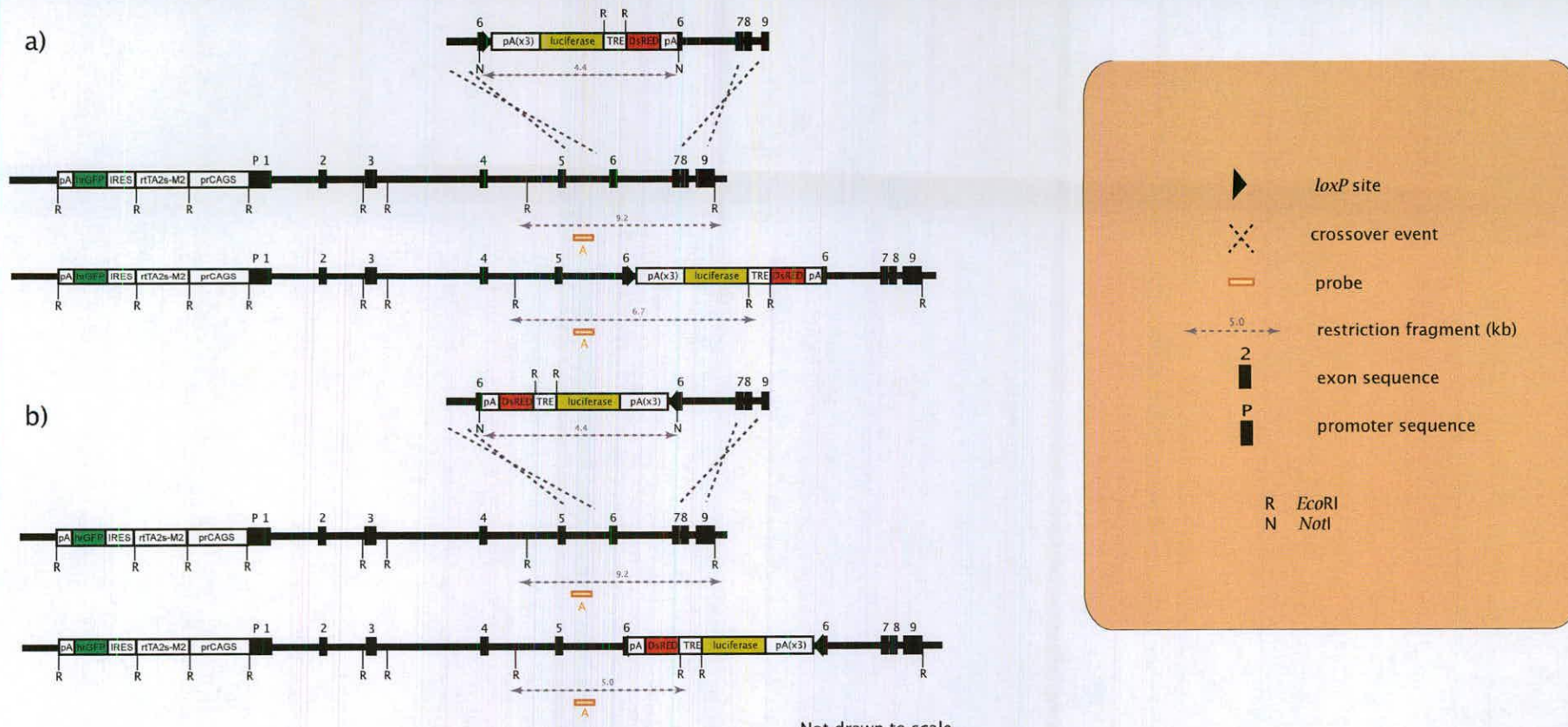


Figure 3.2 The targeting vectors 3'hprtTREDsRED2F a) and 3'hprtTREDsRED2R b) were electroporated into CAGrepB7 cells. The expected restriction fragments yielded after digestion with *EcoRI* and hybridisation with probe A are shown. c) After electroporation and subsequent selection in 6-TG supplemented media, genomic DNA from 6-TG resistant colonies was digested with *EcoRI* and analysed by Southern blot hybridisation.

3.2a and 3.2b depict the sizes of fragments expected after digestion of genomic DNA with *EcoRI* restriction endonuclease, and hybridisation with a probe (probe A), corresponding to a region within intron 5 of the *hprt* gene.

*EcoRI* digestion of the parental CAGrepB7 cell line was expected to yield a 9.2 kb fragment after hybridisation to probe A (Figure 3.2a). A 9.2 kb band was indeed present in the Southern blot analysis of E14TG2a and CAGrepB7 cells (Figure 3.2c). It was predicted that the correct targeting of the 3'hprtTREDsRED2F vector into CAGrepB7 cells would yield a smaller band of 6.7 kb upon *EcoRI* digestion and probe hybridisation. Clones 2, 4 and 5 from the TREDsRED2F cell-lines exhibited this 6.7 kb band. Clones TREDsRED2F1 and TREDsRED2F3 yielded a band that was larger than the wildtype 9.2 kb fragment. Owing to the fact their loci were no longer wild type, and that both clones exhibited a 6-TG resistant phenotype it can be concluded that an integration event had occurred. However the nature of this integration event was not readily apparent. Both clones underwent digestion with *EcoRI* (data not shown) and indeed ihprtDsRED2F1, (but not clone ihprtDsRED2F3) was able to express DsRED2 in the presence of dox, as discussed at a later point (Figure 3.3). It is therefore possible that the nature of the integration event was different for each of these clones.

*EcoRI* digestion of clone TREDsRED2F6 yielded a band of around 7 kb, which was larger than the expected fragment by about 0.3 kb. This extra sequence could possibly be due to a duplication event that occurred within the vector prior to integration, however the aberrant nature of this recombination event has not been fully characterised.

The correct targeting of 3'hprtTREDsRED2R was expected to yield a 5 kb fragment after digestion of genomic DNA with *EcoRI* and hybridisation to probe a . All the clones analysed exhibited this size of fragment, and therefore were concluded to be correctly targeted.

### 3.3 Flow cytometry of ihprtDsRED2 cell lines

Six ihprtDsRED2F and six ihprtDsRED2R clones were analysed for hrGFP and DsRED2 fluorescence by flow cytometry, after culture in the presence and absence of 1 µg/ml dox (dox was added for 48 hours), the results of which are shown in figures 3.3 and 3.4. The first thing to note is that although leakiness of transgene expression in the absence of dox has previously been reported to be a problem when implementing the tet system (Izumi and Gilbert., 1999; Pacheco *et al.*, 2000; Qu *et al.*, 2004), virtually no transgene expression could be detected in any of the clones analysed in the absence of dox (the DsRED2 positive gate was set whereby the population of cells from the parental CAGrepB7 line falling within the gate was less than 1%). It was therefore concluded that the targeting of the TRE element to the 3'*hpert* locus did not render the TRE open to activation by nearby enhancer elements, which would have resulted in high levels of basal transcription.

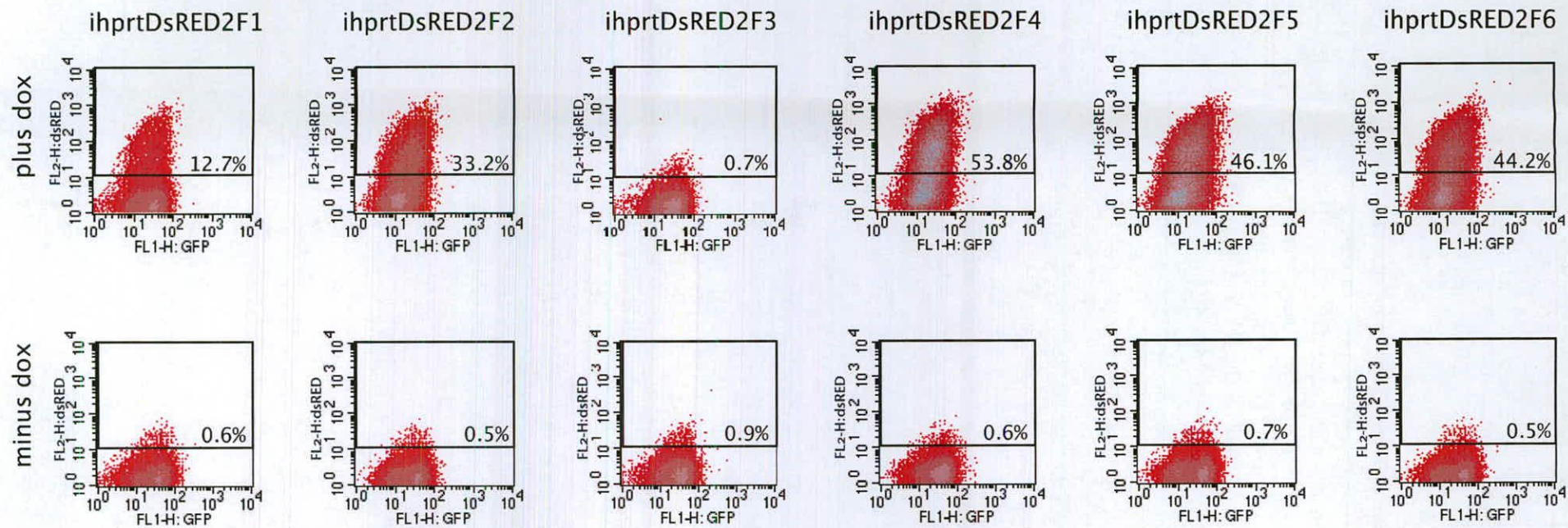
It was readily apparent that upon the addition of doxycycline all clones (bar ihprtDsRED2F3 which was shown to contain an aberrant targeting event by Southern blot analysis – Figure 3.1), exhibited expression of the DsRED2 reporter (Figure 3.3a and 3.4a). However the percentage of cells positive for DsRED2 fluorescence varied from clone to clone with values ranging from 12.7% (ihprtDsRED2F1) to 75.2% (ihprtDsRED2R1), with all clones exhibiting heterogeneity in DsRED2 expression. There did not appear to be any significant difference in fluorescence levels between the ihprtDsRED2F and ihprtDsRED2R cell lines.

One intriguing observation was that the addition of dox resulted in a decrease in the median hrGFP fluorescence value for every clone analysed. This effect is most clearly viewed in the histograms in figure 3.3b and 3.4b which depict hrGFP fluorescence in the presence and absence of 1 µg/ml dox. The reason for this negative shift in hrGFP fluorescence has not yet been determined. However, as this phenomenon was present in every clone it suggests that activity of the TRE at this locus was somehow detrimental to the expression of the transactivator from the 5' region of the *hpert* locus.



Figure 3.3 Six ihprtDsRED2F clones were cultured in the presence and absence of 1  $\mu\text{g/ml}$  dox for 48 hours and analysed by flow cytometry. a) Two dimensional flow cytometry plots for each cell line in the presence and absence of dox. The x-axis depicts hrGFP fluorescence and the y-axis DsRED2 fluorescence. The proportion of DsRED2 positive cells within each cell population is shown. b) Histograms depicting hrGFP fluorescence in the presence and absence of dox, with median hrGFP fluorescence stated.

a)



b)

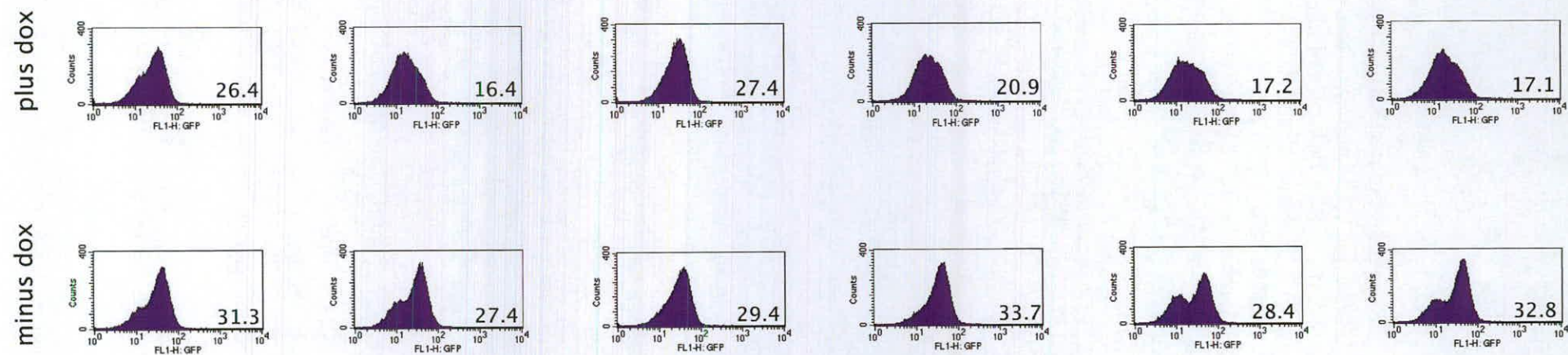
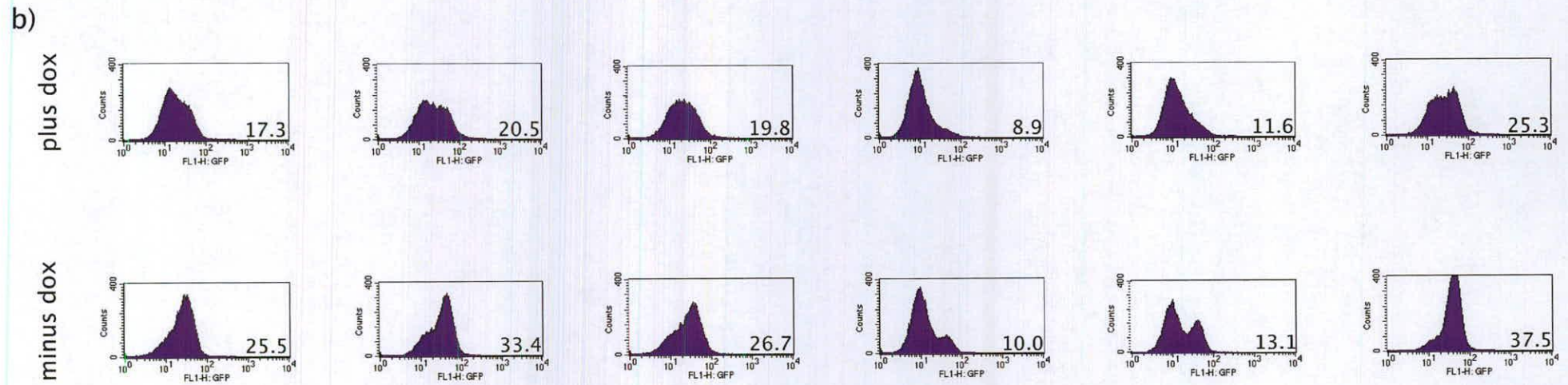
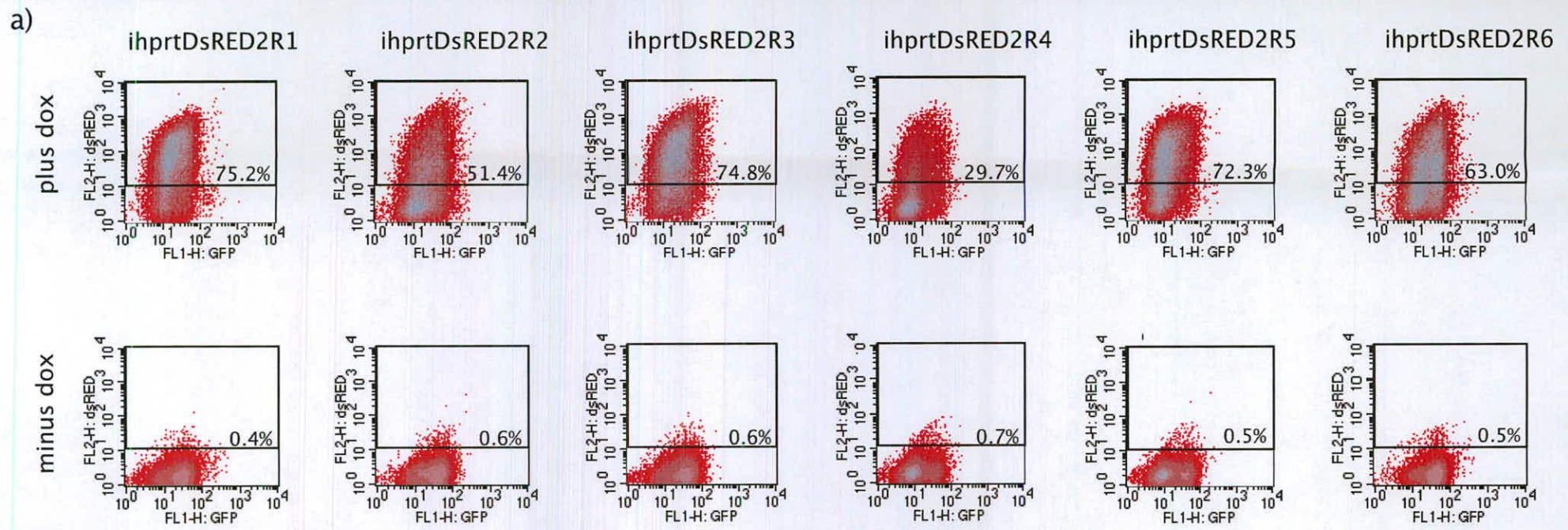


Figure 3.4 Six ihprtDsRED2R clones were cultured in the presence and absence of 1  $\mu\text{g/ml}$  dox for 48 hours and analysed by flow cytometry. a) Two dimensional flow cytometry plots for each cell line in the presence and absence of dox. The x-axis depicts hrGFP fluorescence and the y-axis DsRED2 fluorescence. The proportion of DsRED2 positive cells within each cell population is shown. b) Histograms depicting hrGFP fluorescence in the presence and absence of dox, with the median hrGFP fluorescence stated.



It was evident that some cell lines in the absence of dox contained two populations of cells that exhibited differential hrGFP fluorescence (the most extreme cases were ihprtDsRED2R4 and ihprtDsRED2R5 minus dox – Figure 3.4). It is possible that these cell lines underwent partial epigenetic silencing at their *hprt* promoter region during their derivation, resulting in a sub-population of cells expressing lower levels of rtTA2<sup>S</sup>-M2 and hrGFP protein.

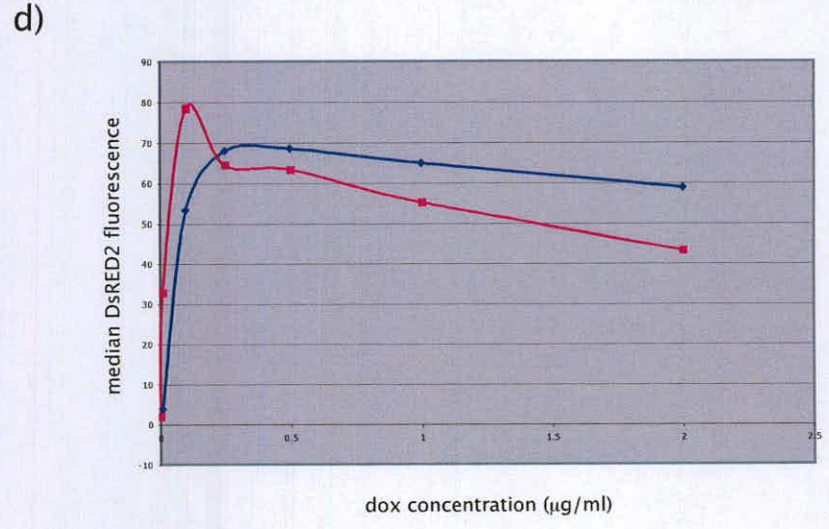
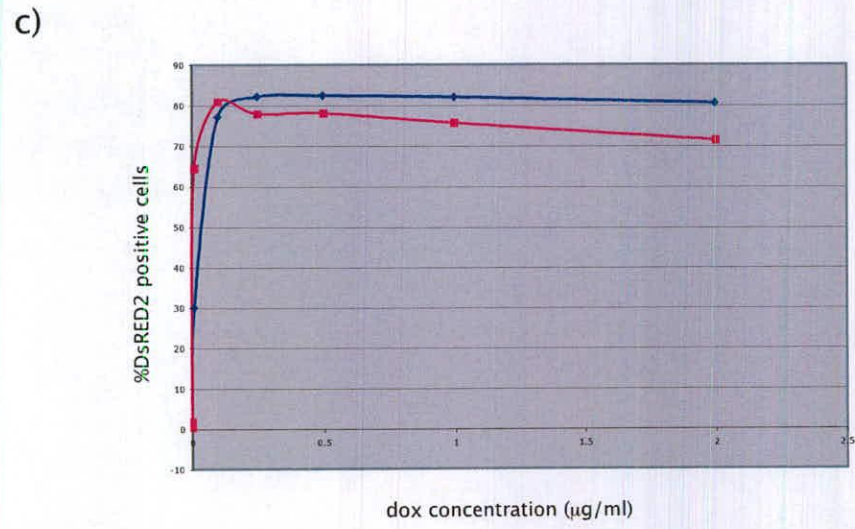
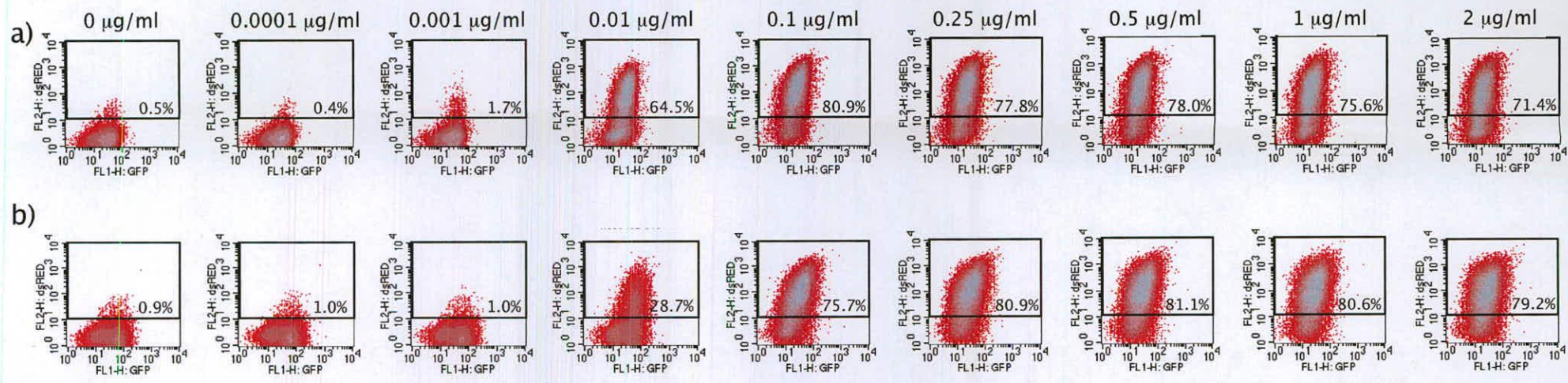
### **3.4 DsRED2 induction in a gradient of concentrations of dox.**

The ihprtDsRED2R3 and ihprtDsRED2F4 cell lines were cultured in a gradient of dox concentrations for 48 hours, to determine the concentrations of dox that were necessary to induce DsRED2 expression. Both clones were then analysed by flow cytometry and the population of DsRED2 expressing cells (Figure 3.5a,b and c) and median DsRED2 fluorescence (Figure 3.5d) at each concentration was determined (Figure 3.5).

Significant induction of the DsRED2 reporter was only observed at concentrations of 0.01 µg/ml doxycycline and above, and maximal induction of reporter gene expression in both lines was achieved at around 0.1 µg/ml. Even though the majority of the cell population could be induced to express DsRED2 at 0.1 µg/ml dox the expression was heterogenous in nature, with some cells expressing high levels of the reporter whilst others cells remained negative. Uniform expression of DsRED2 was not achieved by increasing the concentration of dox, indeed the median DsRED2 fluorescence started to decline at concentrations above 0.1 µg/ml for ihprtDsRED2R3 and 0.25 µg/ml for ihprtDsRED2F4 (Figure 3.5d).

### **3.5 Sorting of the DsRED2 positive cells**

It was hypothesised that the heterogenous nature of DsRED2 expression may be due to the presence of a sub-population of cells within the clonal population unable to express the reporter (for example due to epigenetic silencing mechanisms), with the remaining cells able to activate transgene expression to varying extents. If this was the case it was



◆ ihprtDsRED2R3  
 ■ ihprtDsRED2F4

Figure 3.5 a) ihprtDsRED2R3 and b) ihprtDsRED2F4 were cultured in a gradient of dox concentrations for 48 hours. Two dimensional density plots are shown with the percentage of DsRED2 positive cells indicated. The x-axis depicts hrGFP fluorescence, and the y-axis DsRED2 fluorescence. c) Graphical representation of the percentage of DsRED2 positive cells in each cell line at each concentration of dox. d) Graphical representation of the median DsRED2 fluorescence at each dox concentration.

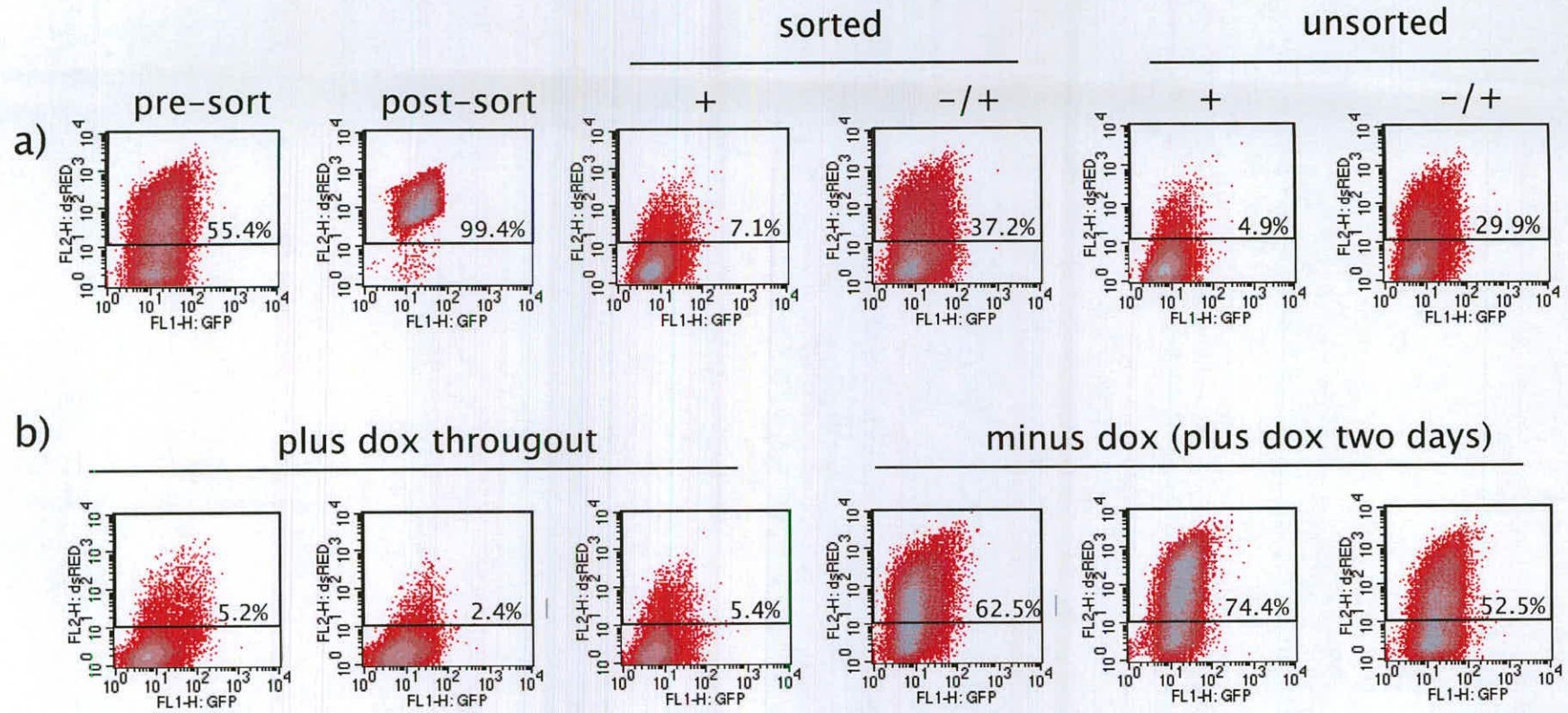


Figure 3.6 The sorting of high DsRED2-expressing cells. a) Two-dimensional flow cytometry plots depicting DsRED2 (y-axis) and hrGFP fluorescence (x-axis) for sorted and unsorted cell populations, (+) cultured in dox for 6 days, (-/+ ) cultured in the absence of dox for 4 days after which dox was added for the two remaining days. b) Single cell sorting. Flow cytometry plots of 6 clones each derived from the single cell sorting of DsRED2 positive cells, either maintained in dox for 8 days or cultured in the absence of dox for 6 days with the addition of dox for the remaining 2 days.

reasoned that the sorting of high DsRED2 expressing cells from the rest of the heterogenous population may lead to the establishment of a population of cells that were able to reproducibly activate DsRED2 expression to high and homogenous levels.

The ihprtDsREDF6 cell line was therefore cultured in 1 µg/ml dox for 48 hours, after which the DsRED2 positive cells were sorted from the negative population and cultured separately. Half of the highly expressing DsRED2 cells were maintained in dox for 6 days (these will be referred to as sorted + cells), whilst the other half were cultured without dox for 4 days with dox subsequently added for the remaining two days preceding analysis (these will be referred to as sorted -/+ cells). An unsorted population of cells was also cultured in parallel under identical conditions (unsorted + and unsorted -/+ respectively). All cells were analysed at day 6 (Figure 3.6a)

It was observed that the DsRED2 expression of the sorted + cells could not be maintained after six days in dox. Indeed, surprisingly the cell population of sorted + was mostly negative for DsRED2, with significantly lower fluorescence levels than those possessed by the initial pre-sort population of cells. There was no apparent difference in the levels of DsRED2 induction between the sorted + and the unsorted + cells.

In a parallel experiment, single DsRED2 positive cells were sorted from the ihprtDsRED2F6 population into the wells of a 96 well plate. These cells were either maintained in dox for 8 days (sorted +) or maintained in the absence of dox, with the addition of dox two days before flow cytometric analysis (sorted -/+). Six clones are shown in figure 3.6b, the profiles of which are representative of all clones analysed. It can be seen that clones cultured in the presence of dox were all largely negative, whereas the clones cultured in the absence of dox could be induced to express DsRED2 levels comparable to those of the ihprtDsRED2F6 presort. It could therefore be concluded that high DsRED2 levels could not be maintained after the sorting of high DsRED2-expressing cells.

### 3.6 Induction of DsRED2 after different periods of culture in dox

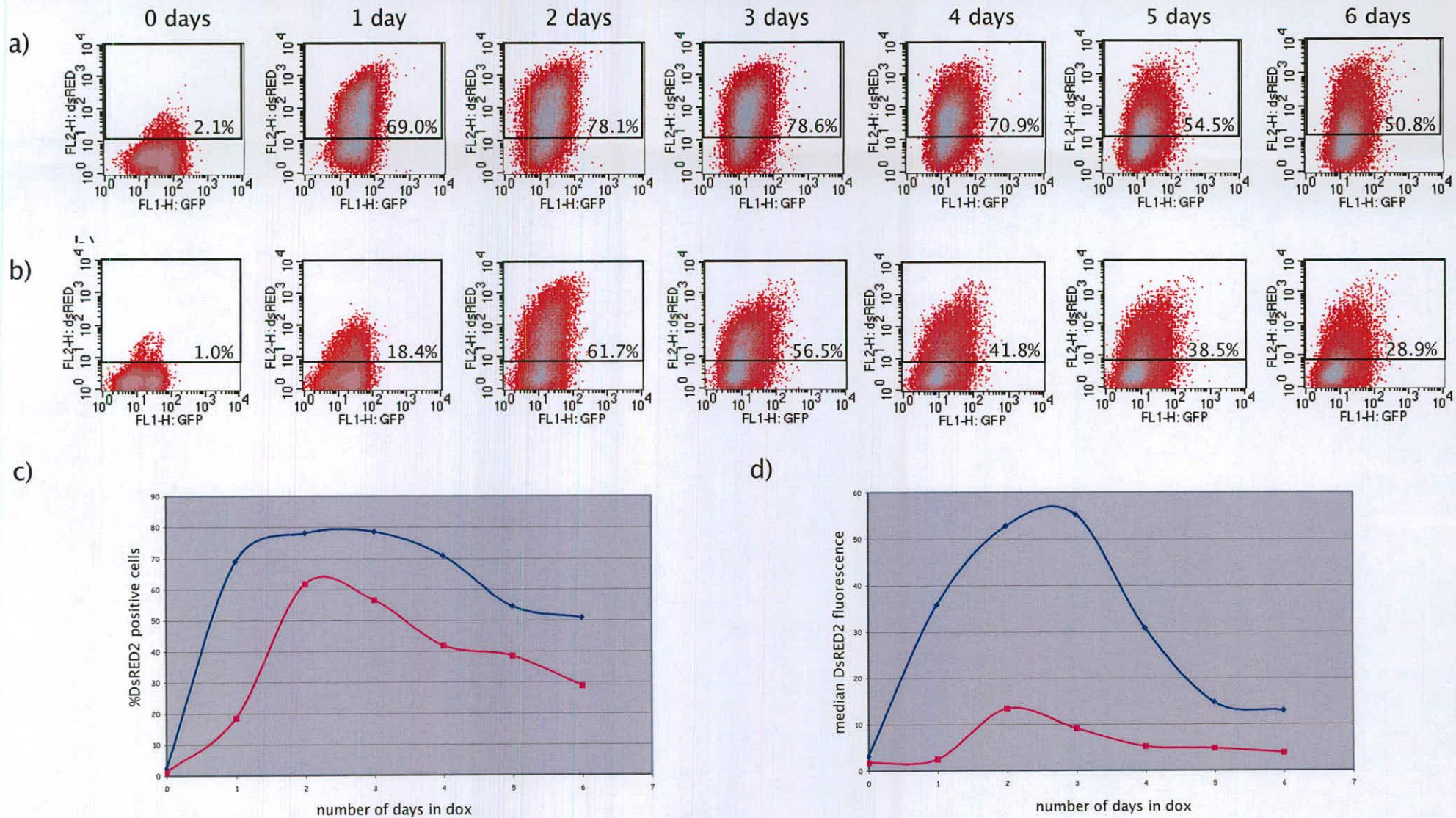
It was clear from the cell sorting experiment that the propagation of ihprtDsRED2 cells in dox for long periods of time was detrimental to DsRED2 expression. This was investigated further by culturing the ihprtDsREDF4 and ihprtDsREDR2 cell lines in 1µg/ml dox for 0 to 6 days (Figure 3.7). After 6 days each timepoint was analysed for DsRED2 fluorescence by flow cytometry.

It can be observed from figure 3.7 that maximal induction of the DsRED2 transgene was achieved after 2 days in dox (78.1% and 61.7% of cells were positive for DsRED2, ihprtDsREDF4 and ihprtDsREDR2 respectively). At day 3 this high expression was maintained in the ihprtDsREDF4 cell line whereas the ihprtDsREDR2 cell line exhibited a slight reduction in DsRED2 fluorescence.

After day 3 the two cell lines exhibited a decline, in both in the percentage of cells that remained positive and in the median DsRED2 fluorescence recorded (Figure 3.7c and 3.7d). At day 6 a large proportion of the cell population was negative for DsRED2 in both clones. Maintenance of the clones in dox for periods of longer than 3 days therefore appears to be detrimental to DsRED2 expression, which correlates with the inability to maintain DsRED2 reporter expression in the previous sorting experiment.

### 3.7 Discussion

It has been shown previously that homogenous inducible expression of fluorescent reporter transgenes can be achieved using the tet-system (Izumi and Gilbert *et al.*, 1999, Qu *et al.*, 2004). In both studies, the reporter sequences were linked to the TRE, and the expression of these transgenes was analysed by flow cytometry upon the addition of dox to the cell culture.



- ◆ ihprtDsRED2R3
- ihprtDsRED2F4

Figure 3.7 a) ihprtDsRED2F4 and b) ihprtDsRED2R2 were cultured in 1  $\mu\text{g/ml}$  dox for 0 to 6 days. Two dimensional density plots are shown with the percentage of DsRED2 positive cells indicated. The x-axis depicts hrGFP fluorescence, and the y-axis DsRED2 fluorescence. c) Graphical representation of the percentage of DsRED2 positive cells for each cell line at each timepoint. d) Graphical representation of the median DsRED2 fluorescence.

The main problem encountered by these authors, and indeed an issue that has been widely cited by users of the tet-system (Shockett *et al.*, 1995, Meyer-Ficca *et al.*, 2004) is the problem of high basal expression of the transgene in the absence of dox. This has been attributed to the residual affinity of the rtTA for the TRE in the absence of dox in those situations where the original rtTA has been used, and also to activation of the TRE by nearby enhancer-like elements upon random integration of the TRE element (Izumi and Gilbert *et al.*, 1999, Qu *et al.*, 2004). In light of this evidence a strategy was devised to target both the reverse transactivator and the response element to the *hprt* locus, a housekeeping gene which is expressed in all tissues (Bronson *et al.*, 1996). The novel rtTA<sup>2<sup>s</sup></sup>-M2 reverse transactivator, which is an updated version of the original rtTA, has no residual affinity for the TRE in the absence of dox, and was targeted to 5' *hprt* as described in chapter 2. Once this parental cell line had been established, the TRE element (linked to DsRED2) was targeted into the *hprt* open reading frame, at a distance of 30 kilobases from the CAG promoter driving rtTA<sup>2<sup>s</sup></sup>-M2 expression, and at a considerable distance from any known enhancer (thus it was reasoned that the integration of the TRE at this location would not result in high basal levels of transgene transcription in the absence of dox). This targeting strategy was therefore designed to assess whether or not *hprt* was a viable locus for implementing transgene expression via the tet-system. After the derivation of the ihprtDsRED2 cell lines, they were analysed for transgene induction in the presence and absence of dox.

It was clear upon analysis of the ihprtDsRED2 cells lines derived, that background activity of the TRE element in the absence of dox was not a problem encountered with this strategy. However, an issue of concern was the heterogenous expression of DsRED2, which was observed in every clone analysed. For the tet-system to be used optimally, it is necessary that homogenous expression of the transgene in question is observed. There are a number of possible explanations for this heterogeneity. Firstly one must question the possibility of toxicity effects of DsRED2 in ES cells. Secondly the integrated TRE may not be fully accessible to the rtTA<sup>2<sup>s</sup></sup>-M2, or indeed it may be the case that the levels of rtTA<sup>2<sup>s</sup></sup>-M2 expressed are too low to activate the TRE-linked fluorescent reporter maximally. There is also the question of mosaicism of transgene

expression due to epigenetic silencing of the locus. These will all be discussed in further detail below.

### **DsRED2 toxicity**

The DsRED2 reporter gene, a variant of the original DsRED1 exhibits improved solubility and reduced tendency to form aggregates within transfected cells. Toxicity in mammalian cells was reported to be a problem with DsRED1, however, owing to the existence of several mutations within the ORF of the novel fluorescent protein, DsRED2 shows improved ability to be tolerated *in vivo* (manufacturers product information).

In the event of possible DsRED2 toxicity to ES cells, there would be a selective pressure on cells within a population to express low amounts of DsRED2. Whilst this would seem to be a plausible explanation for the heterogenous profiles viewed for the ihprtDsRED2 clones, it is known from independent experiments with the tet system (Fisher, Heavey and Cunningham, unpublished data) that high levels of DsRED2 can be tolerated in ES cells for at least three days without there being any obvious signs in reduction of DsRED2 levels. Therefore the heterogeneity viewed here cannot be attributed to possible toxicity of DsRED2 to ES cells.

### **Accessibility of the TRE to rtTA2<sup>S</sup>-M2**

#### **i) Is the TRE located in a region of chromatin which is suboptimal for rtTA2<sup>S</sup>-M2 binding?**

The random integration of the TRE and linked luciferase reporter into cell lines which overexpress tet-transactivators has been shown to give rise to varying levels of reporter induction depending on where the exogenous sequences integrate into the genome (Gossen *et al.*, 1992; Kistner *et al.*, 1996). Whilst some clones exhibit high basal levels of transgene expression in the absence of inducer, presumably as a result of integration of the TRE close to nearby enhancers, other clones do not exhibit any luciferase expression

in the presence of inducer. One reason for this could be the integration of the TRE and linked reporter into a region devoid of any transcriptional activity, such as areas of heterochromatin.

Other clones reportedly generated from the random integration of the TRE exhibit poor expression of the linked luciferase reporter in the presence of inducer. In this case it can be imagined that whilst the TRE has inserted into a locus permissive for transcription, the chromatin configuration is not completely supportive of maximal transgene expression. In light of this observation it may be hypothesised that the heterogenous expression exhibited by the *ihprtDsRED2* cell lines may be the result of integration into a locus that was not permissive for maximal activation of the TRE. As the element was integrated into exon 6 of the *hpert* gene however, and *hpert* is a housekeeping gene which is ubiquitously expressed, it seems intuitive that this locus would possess a fairly open chromatin configuration, leading one to presume that this hypothesis is unlikely to be true.

## **ii) Read-through of *hpert* mRNA resulting in inhibition of transgene transcription initiation**

In a previous study in which exogenous sequences were targeted to the coding region of *hpert* it was reported that there was read-through from the *hpert* promoter (Shaw-White *et al.*, 1993). The authors placed a *LacZ* reporter gene downstream of a combination promoter which contained a herpes simplex virus, thymidine kinase (*HSVtk*) promoter plus a modified polyoma enhancer, and targeted the sequences into exon 3 of the *hpert* gene in ES cells (a neomycin selection gene was also contained within the construct). *LacZ* expression was observed in 40 out of 45 of the targeted clones analysed, and mosaic expression was seen in all the positive clones. Read-through transcripts that originated from the *hpert* promoter and that contained the targeted exogenous sequences could be detected by Northern blot. The authors concluded that the mosaicism in *LacZ* expression was due to 'transcriptional interference' that occurred as the transcriptase complex passed through the exogenous promoter, inhibiting initiation of transcription from this promoter. However in a parallel experiment, random integration of the promoter and linked *LacZ*

sequence failed to give rise to any uniformly blue *LacZ* positive lines and it is therefore unclear if the heterogenous expression was indeed due to ‘transcriptional interference’ as a result of read-through, or if the composite promoter and enhancer sequence was unable to drive homogenous expression of transgenes in ES cells. It is also possible that methylation of the exogenous promoter and *LacZ* sequences resulted in mosaicism in expression of the *LacZ* transgene (Chevalier-Mariette *et al.*, 2003).

Although the argument put forward by Shaw-White *et al.*, that the observed mosaicism was due to *hprt* read-through from the endogenous promoter could not be substantiated from their evidence presented, it was clear that read-through did indeed occur, as transcripts originating from the *hprt* promoter region, containing the exogenous sequences were detected by Northern blot hybridisation. It is very likely that this phenomenon also occurs in the *ihprtDsRED2* cell lines established, although this has yet to be confirmed by Northern analysis. In addition this form of ‘promoter occlusion’ arising from transcriptional read-through has previously been documented in HeLa cells in a study where two tandem HIV-1 promoters were integrated into the genome (Greger *et al.*, 1998). It was found that the transcriptional activation of the upstream promoter reduced transcription from the downstream promoter, the mechanism of which was found to be due to the occlusion of Sp1 transcription factors from the downstream promoter as determined by *in vivo* footprinting. This reduction in downstream promoter activity was abolished when a transcriptional terminator sequence was placed between the two promoters and thus it was hypothesised that the movement of the polymerase complex over the downstream promoter was responsible for the promoter occlusion.

In light of this evidence it is a possible that the non-uniform expression of the DsRED2 reporter observed in the *ihprtDsRED2* lines can be attributed to the occlusion of rtTA2<sup>S</sup>-M2 binding to the TRE element, as the polymerase complex passes over the TRE sequence. This would result in sub-optimal levels of expression of the DsRED2 reporter and could give rise to the heterogenous expression of reporter observed.

## Heterogenous expression of DsRED2 as a result of sub-optimal levels of rtTA2<sup>S</sup>-M2

It has been reported in the literature that low internal concentrations of transactivator give rise to heterogenous expression of TRE-linked transgenes when using the tet-system. Izumi and Gilbert conducted a study whereby they linked the transactivator tTA, to a hygromycin antibiotic sequence via an IRES and randomly integrated the construct into cells. Stable transformants expressing tTA were derived by placing the transfected cells in hygromycin selection. It was found that upon stable integration of the TRE element and linked GFP reporter, a number of clones were established which exhibited homogenous levels of the reporter throughout the population upon withdrawal of inducer, and this was reasoned to be due to the initial selection for cells expressing high levels of tTA (using the IRES hygromycin strategy). This was in contrast to earlier tTA-expressing cell lines, which were not established by selecting for high concentrations of tTA, and therefore only ever gave rise to heterogenous reporter expression when stably transfected with the TRE-GFP reporter construct. In these experiments, the tTA itself was not linked to a reporter that could be used to directly compare the levels of transactivator expression between the different cell types created, and therefore the absolute levels of tTA produced were never established. It is however logical to presume that maximal expression of TRE-linked transcript is in part dependant on internal concentrations of rtTA2<sup>S</sup>-M2. One might expect that if the internal concentration of rtTA2<sup>S</sup>-M2 was limiting then some of the *tetO* sequences within the TRE would remain unoccupied and this would reduce the frequency at which DsRED2 would be transcribed (due to the sub-optimal recruitment of transcriptional machinery). This raises the issue as to whether or not limiting internal concentrations of rtTA2<sup>S</sup>-M2 were responsible for the heterogenous profiles of the ihprtDsRED2 cells.

Whilst further experimentation into the implementation of the tet system (described in the next chapter) suggest that the levels of rtTA2<sup>S</sup>-M2 expressed in the ihprtDsRED2 cells (as determined by analysis of hrGFP levels in the absence of dox), should be sufficient to activate transgene expression at high and homogenous levels, the addition of dox inducer appears to decrease intracellular rtTA2<sup>S</sup>-M2 levels (reasons for this are discussed below).

It is therefore possible that the heterogeneity in DsRED2 expression observed could be the result of sub-optimal levels of transactivator that are unable to activate maximal transgene transcription.

### **The negative shift in hrGFP expression upon addition of inducer**

The reason for the observed negative shift in hrGFP upon the application of dox inducer to ihprtDsRED2 cells is unclear. Such a shift was observed in every correctly targeted cell line analysed and was in some cases so pronounced that the differences in hrGFP levels could be clearly visualised on the two dimensional flow cytometry plots.

The addition of inducer leads to the formation of a functional transactivator (dox bound rtTA2<sup>S</sup>-M2), which in turn leads to activation of the TRE as a result of transactivator binding and subsequent polymerase recruitment. It therefore appears that the activation of the TRE-DsRED2 transcriptional unit is detrimental to the transcription of the bicistronic rtTA2<sup>S</sup>-M2 and hrGFP sequences. Transcriptional interference between two transcriptional units has previously been shown to occur when independently regulated genes are positioned in close proximity (Eszterhas *et al.*, 2002). In this study GFP and YFP fluorescent reporters each possessing their own promoter and terminator sequences, were positioned in a construct (with less than 100 base pairs separating the two units), in divergent, convergent and tandem orientations. When compared to the targeting of single fluorescent reporter genes to two independent loci in MEL cells, it was found that the expression of both transgenes exhibited reduced expression when targeted in the same construct at both loci analysed. This was the case for all orientations analysed. When the promoter region of one reporter gene was mutated so as to be rendered inactive, it was found that the levels of expression of the remaining gene were comparable to those seen when only one transcriptional unit was present. This suggested that interference between the two promoter regions of the transgenes hindered the expression of both fluorescent markers.

In the study described, both transcriptional units were placed in close proximity. However, it is not unreasonable to assume that the CAG-rtTA2<sup>S</sup>-M2 and TRE-DsRED2 transcriptional units, even though positioned approximately 30 kilobases apart, exert some influence over each other. If one thinks in terms of the three dimensional structure of chromatin, in real distances the two units may not be positioned that far apart. This can be exemplified by the fact that some endogenous enhancers are known to act over considerable distances, through the 'looping' of intervening chromatin which enables the enhancer element to interact with its target gene promoter.

A recent study in *Saccharomyces cerevisiae* showed that the promoter regions of some genes interact with the polyA addition signal sequence contained in the same transcriptional unit. Transcriptional run-on analysis and chromosome conformation capture techniques led to the conclusion that a 'gene loop' structure juxtaposed the promoter and terminator sequences in the two genes analysed (O'Sullivan *et al.*, 2004). Although such structures have yet to be discovered in mammalian cells it seems unlikely that such configurations are confined to yeast.

If such looping configurations between promoter and terminator sequences are present in mammalian cells it is possible that the integration of a transcriptional unit into the ORF of an endogenous gene (as is the case in the targeting of the TRE to endogenous *hprt*), may disrupt the three-dimensional structure of the chromatin region, which in turn may influence the transcription of nearby genes. Although not enough is yet known about the complex chromatin structures involved in transcription to be able to predict the consequences of targeting exogenous elements to particular loci, the studies mentioned above highlight the complexity of interactions within and between transcriptional units, and offer a possible explanation for the observed shift in hrGFP upon addition of inducer.

### **Potential silencing of rtTA2-M2 and TRE upon integration into the *hprt* locus**

The rtTA2<sup>S</sup>-M2 and TRE sequences both contain viral and prokaryotic elements, and it therefore might be envisaged that host cell defence mechanisms, such as RNAi mediated

pathways, may act to silence the activity of these units. Possible mechanisms of silencing could include the cleavage of the rtTA2<sup>S</sup>-M2 mRNA, or the epigenetic silencing of the transactivator or TRE sequences at the chromatin level, (Schramke *et al.*, 2003). If this silencing effect was not to occur in all the cells within the population it may be expected to give rise to heterogenous expression of DsRED2 reporter as observed.

In order to determine whether the non-uniform expression of DsRED2 was due to inheritable differences between the highly-expressing DsRED2 cells and the cells negative for DsRED2 in the presence of dox (due to chromatin based epigenetic silencing mechanisms), the highly-expressing DsRED2 positive cells from clone ihprtDsRED2F6 were sorted and cultured for 6 days in dox (Figure 3.6a). It was predicted that this population of cells would either become heterogenous for DsRED2 expression – indicating that all cells within the population had the potential to express DsRED2, or would remain a discrete population of high DsRED2-expressers – suggesting that there were sub-populations that contained different inheritable transgene activation potentials. Surprisingly however, after 6 days in dox, most of the cell population became negative for DsRED2 expression. Although this proved that the high DsRED2-expressing state in the presence of dox was not permanent, it was not clear why the majority of the cells should become negative for DsRED2. This effect was investigated further in a timecourse experiment.

### **Reduction in DsRED2 expression upon increasing time in dox**

Two clones were analysed for their DsRED2 expression after varying periods of culture in dox. It was clear that between 2 to 3 days (depending on the clone in question), DsRED2 expression peaked and thereafter declined to near negative levels (Figure 3.7). At first it may appear that either the application of dox, or prolonged DsRED2 expression is toxic to ES cells. However independent experiments using the tet-system show that both high levels of DsRED2, and concentrations of up to 1 µg/ml dox can be tolerated by ES cells (Fisher, Heavey and Cunningham, unpublished data), which suggest that the

decline in DsRED2 observed in the timecourse experiment is unlikely to be due to dox or DsRED2 toxicity.

The outcome of this experiment highlights the experimental limitations of using this strategy to overexpress cDNAs in ES cells, as the period of time over which a transgene can be expressed is limited by the rapid decline in expression levels. It would therefore only be applicable in situations where factors need to be overexpressed for about 3-4 days, and in addition in situations where heterogeneity in transgene expression would not pose a problem (for example in the inducible expression of secreted factors).

### **DsRED2 expression in a gradient of dox concentrations**

Previous studies have reported that the tet-system can be used to obtain a range of transgene expression levels by varying the concentration of inducer added to the cell population (Gossen *et al.*, 1992). These studies however involved the use of a luciferase assay which was dependent on the lysis of the cell population and therefore the results obtained offered no insight into transgene expression at a cellular level. It is possible that the intermediate concentrations of luciferase observed upon culture with lower concentrations of inducer could have been due to the absence of luciferase expression in a proportion of the cell population. With these technical limitations in mind, it was decided to use a fluorescent reporter in the establishment of the tet-system in order to analyse transgene expression at a cellular level by flow cytometry.

ES cells were cultured in a gradient of concentrations of dox, over a period of 48 hours. It was found that increasing the concentration of dox did not lead to an incremental increase in the levels of the DsRED2 reporter (Figure 3.5a and 3.5b). Rather a bimodal, all-or-nothing effect was more apparent (this is most obvious in Figure 3.5a 0.01 µg/ml), with a sub-population of cells expressing DsRED2 to a high level while a large proportion of the population remained negative for DsRED2 expression. This result did not support the claim that the tet-system can be used to express varying levels of transgene.

Interestingly the median levels of DsRED2 declined above approximately 0.25 µg/ml dox (Figure 3.5d), and it therefore appears that high concentrations of dox lead to sub-optimal DsRED2 expression. Again independent experiments suggest that this is not due to toxic effects of DsRED2 expression, or dox addition (Fisher, Heavey and Cunningham, unpublished data), and the reason for the decline in transgene expression levels over concentrations of 0.25 µg/ml dox remain to be determined.

After characterisation of the ihprtDsRED2 cell lines it can be concluded that the strategy of targeting both components of the tet-system to the *hprt* allele in ES cells can give rise to expression of the DsRED2 reporter in the presence of dox, with no basal transgene expression visible in the absence of dox. There is however need for improvement both in terms of levels of reporter expression and uniformity of expression across the whole population. It was therefore decided to implement an alternative strategy in order to address these issues.

## Chapter 4

### A biallelic approach to implementing the tet-system

#### 4.1 Introduction

The previous chapter described how the transactivator-response element (TRE) and linked reporter genes were targeted into the 3' coding region of *hprt* in CAGrepB7 cells. The CAGrepB7 cell line expressed homogenous constitutive levels of the reverse transactivator rtTA2<sup>s</sup>-M2, however the subsequent integration of the response-element into these cells did not result in the anticipated homogenous expression of TRE-linked DsRED2 reporter. In light of this result an alternative approach was implemented, with the aim of achieving population-wide induction of transgene expression.

From the previous targeting of cDNAs described in chapter 2, it has been shown that the 5' proximal region of *hprt* in E14TG2a cells is permissive for the constitutive expression of integrated transgenes. In an alternative approach to targeting both elements of the tet-

system to the *hpri* locus the transactivator-responsive element (TRE) was targeted to the 5' region of *hpri*, in a cell line that expressed high and homogenous levels of the rtTA2<sup>s</sup>-M2 transgene at an unlinked locus. The cell line used was R26-rtTA2<sup>s</sup>-M21F, which expressed the reverse transactivator from the endogenous ROSA26 promoter, (this cell line was a kind gift from Barry Heavey, for information regarding its construction please refer to Figure 4.1).

#### **4.2 Construction of the 5'hpriTREDsRED2F and 5'hpriTREDsRED2R targeting vectors and electroporation into R26- rtTA2<sup>s</sup>-M2 1F cells.**

The bidirectional TRE sequence and linked luciferase and DsRED2 reporter sequences used in the construction of the 5'hpriTREDsRED2 targeting vectors (which are represented in Figure 4.2 in their linear form), were identical to those used in the construction of the 3'hpriTREDsRED2 targeting vectors detailed in chapter 3 (Figure 3.1). Plasmid BigTlucTREDsRED2 was digested with *AscI*, plus *ScaI* restriction endonucleases (Figure 4.2a), and the TRE-encoding fragment was incubated with klenow polymerase to fill-in the recessed 3' fragment ends. The fragment was then ligated into *NotI* digested and blunted (by klenow end-fill) pSKB1 targeting vector. Two ligation products were produced, one in which the DsRED2 cDNA coding sequence was in the same 5' to 3' orientation as the genomic *hpri* coding sequence present in the vector (Figure 4.2b) and one in which the DsRED2 cDNA sequence was in the opposite orientation (Figure 4.2c). These two targeting vectors were termed 5'hpriTREDsRED2F and 5'hpriTREDsRED2R respectively.

After linearisation of the two vectors with *PmeI* endonuclease, 5'hpriTREDsRED2F and 5'hpriTREDsRED2R were transfected into R26- rtTA2<sup>s</sup>-M21F cells via electroporation. The R26- rtTA2<sup>s</sup>-M21F cells were derived from the E14TG2a cell line and were therefore HPRT negative. The targeting of the correctional vectors to *hpri* resulted in the restoration of the *hpri* ORF (Figure 4.3a and 4.3b), and conferred an HPRT positive phenotype upon the cells. The day after electroporation, transfected cells were plated out in HAT selection media. HAT resistant colonies were then picked and expanded, after

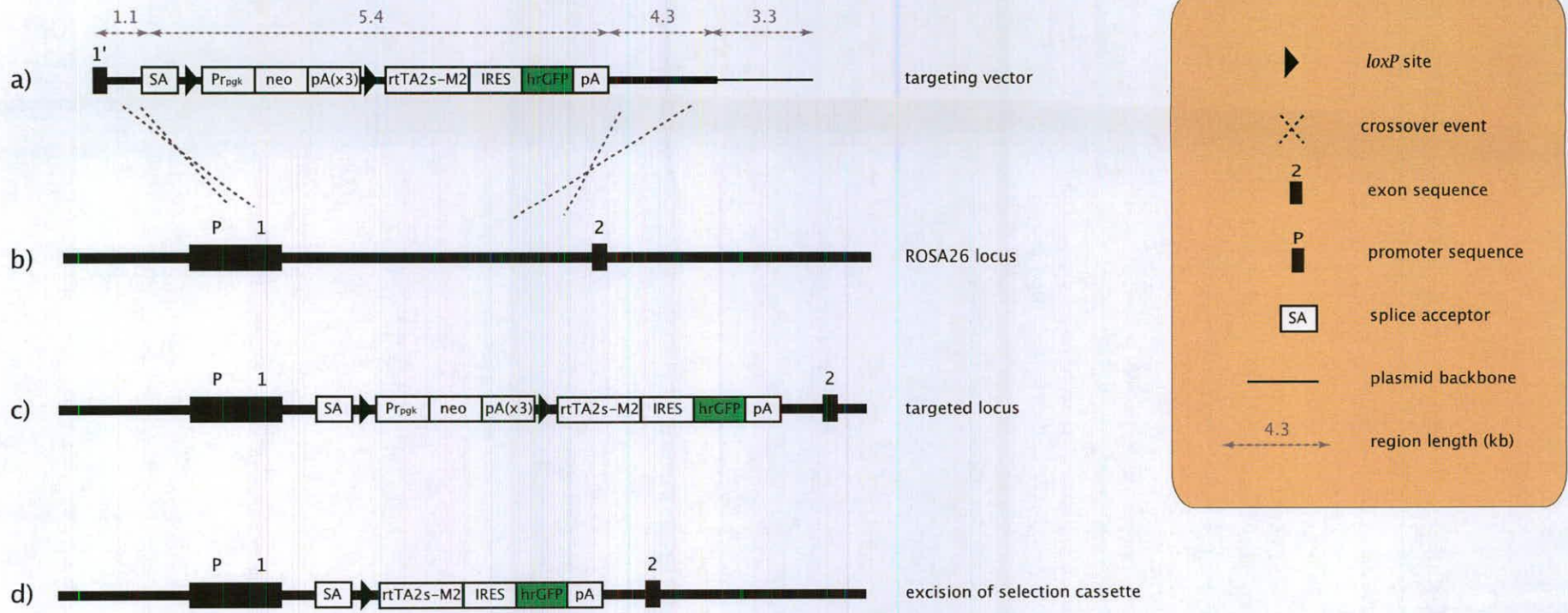
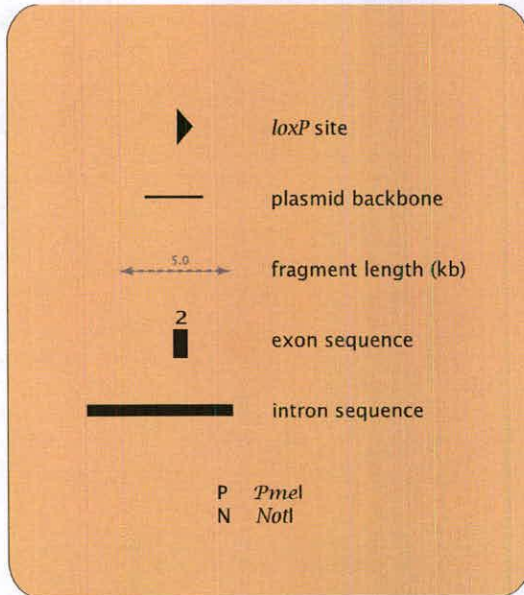
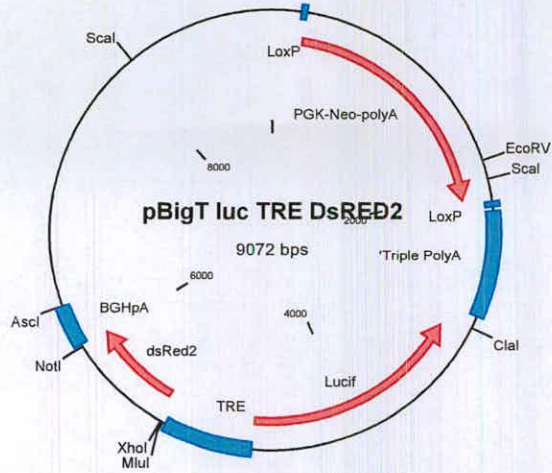


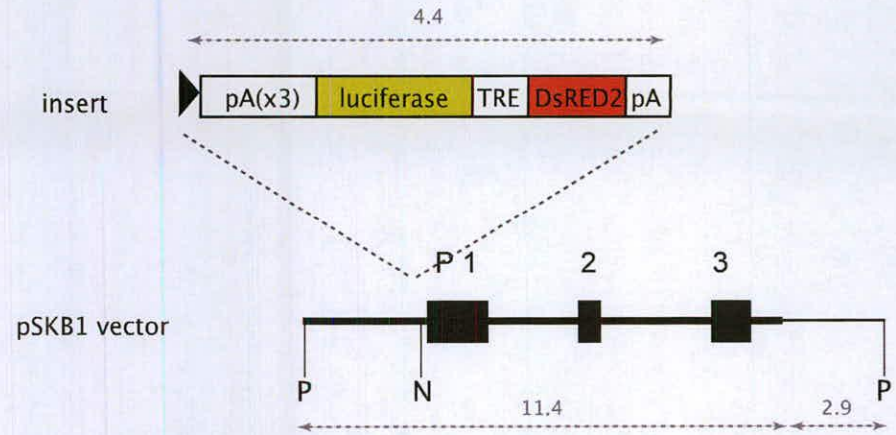
Figure 4.1 The targeting of the reverse transactivator to the ROSA26 locus a) The ROSA26 targeting vector contained a splice acceptor site, a selection cassette and rtTA2s-M2 linked to an IRES-hrGFP sequence and was electroporated into E14TG2a cells to target the endogenous ROSA26 locus (shown in b). c) Transfected cells were then placed in G418 selection after which G418 resistant clones were analysed by Southern blotting to identify clones with a correctly targeted ROSA26 allele. The presence of an exogenous splice acceptor (SA), positioned downstream of the endogenous non-coding exon 1 resulted in the transcription of the neomycin resistance gene only, as the triple polyA, pA(x3) sequence inhibited expression of the reverse transactivator. d) Transfection of targeted cells with a Cre-recombinase expressing plasmid resulted in site-specific recombination between the *loxP* sites flanking the selection cassette, and excision of the intervening sequences. This led to the expression of the reverse transactivator and hrGFP. Note that although three endogenous transcripts are known to arise from the ROSA26 locus, for simplicity only the exons of one of the transcripts have been shown. Abbreviations; Pr<sub>pgk</sub>, pgk promoter; 1', homology to the 3' sequence of exon 1.

Figure 4.2 Cloning of the TRE and linked reporter sequences into the pSKB1 vector. a) Plasmid BigTlucTREDSRED2 was digested with *ScaI* plus *AscI* after which the 3' recessed ends were made flush by treatment with klenow polymerase. The fragment was inserted into *NotI* digested pSKB1 vector in both orientations (the vector had also been treated with klenow polymerase after *NotI* digestion to end-fill the protruding ends), to give rise to the 5'hprtTREDSRED2F targeting vector b) and the 5'hprtTREDSRED2R targeting vector c). Note that the vectors are represented in their linearised forms (site of linearization is *PmeI*).

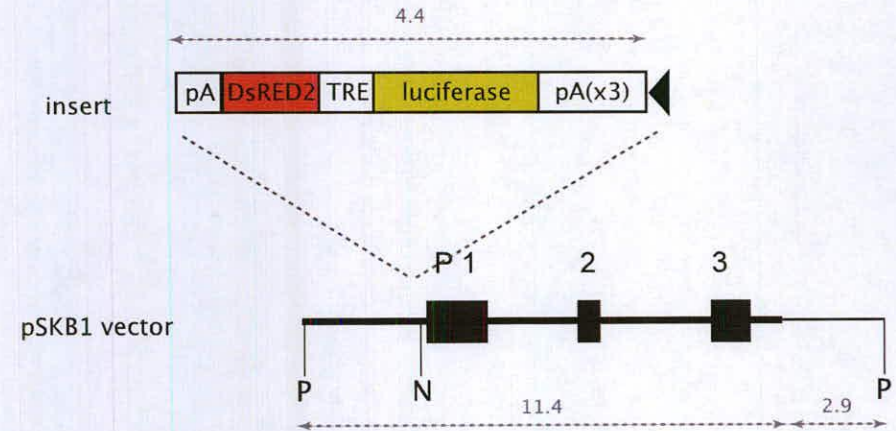
a)



b) 5'hprtTREDsRED2F



c) 5'hprtTREDsRED2R



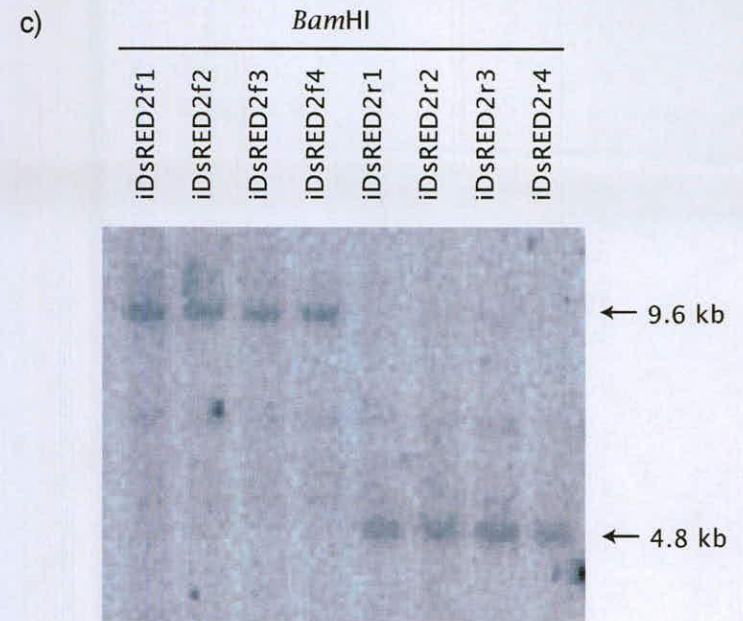
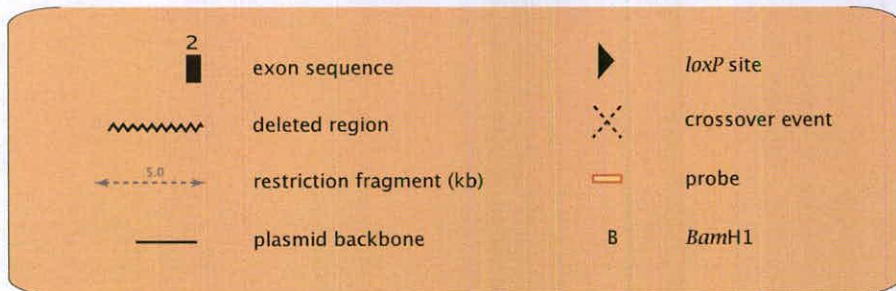
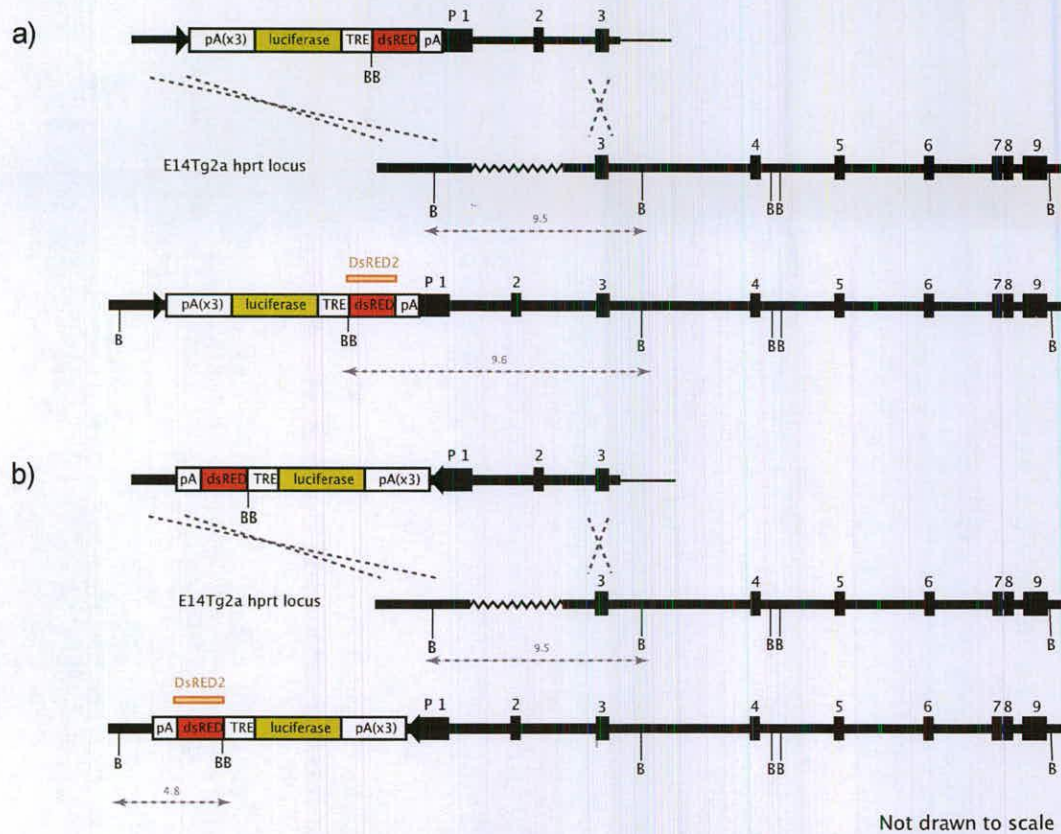


Figure 4.3. The targeting vectors 5'hprtDsRED2F a) and 5'hprtDsRED2R b) were electroporated into the R26-rtTA2s-M2 1F cell line. The expected restriction fragments yielded after digestion with *Bam*H1 and hybridisation with a probe encoding the DsRED2 sequence are shown. c) After electroporation and subsequent selection in HAT medium, genomic DNA from HAT resistant colonies was digested with *Bam*H1 and analysed by Southern blot hybridisation.

which genomic DNA was extracted and analysed by Southern blot hybridisation to determine the nature of the integration event. Figure 4.3a and 4.3b depict the fragment sizes predicted in correctly targeted clones after digestion with *Bam*HI and subsequent hybridisation to a probe encoding the DsRED2 sequence. Eight cell lines were analysed (Figure 4.3c), four cell lines from the 5'hpTREDsRED2F and four from the 5'hpTREDsRED2R vector transfection (the resulting cell lines were termed iDsRED2F1-4 and iDsRED2R1-4 respectively). The fragment size predicted after *Bam*HI digestion and hybridisation with the DsRED2 probe was 9.6 kb for the iDsRED2F clones and 4.8 kb for the iDsRED2R clones. All eight clones exhibited the correct fragment sizes and it was therefore concluded that all clones had been correctly targeted.

#### **4.3 Flow cytometry analysis of iDsRED2 cell lines in the presence and absence of dox**

The four iDsRED2F and four iDsRED2R cell lines were plated in 1 µg/ml doxycycline for two days and then analysed by flow cytometry for expression of the DsRED2 reporter transgene (Figures 4.4 and 4.5). Whereas all the ihpTREDsRED2 cell lines described in the previous chapter exhibited very heterogenous transgene expression throughout the clonal cell populations, the expression of DsRED2 in the iDsRED2 cell lines was far more homogenous, with the majority of cells becoming positive for DsRED2 expression upon the addition of inducer. All clones exhibited DsRED2 fluorescence in over 90% of their population (with the exception of iDsRED2R4 in which 88.8% of the population was positive for DsRED2). In the absence of dox no 'leaky' expression of DsRED2 was observed in any of the cell lines. These observations establish the 5' region of *hpT* as a locus suitable for TRE integration, as high expression levels of linked reporter are achieved in the presence of doxycycline inducer yet no background expression is observed in the absence of inducer. The orientation of DsRED2 upon integration into the genomic 5'*hpT* locus did not appear to have any influence upon the extent of DsRED2 transgene expression, as there was no discernible difference between the profiles of the iDsRED2F clones and the iDsRED2R clones.

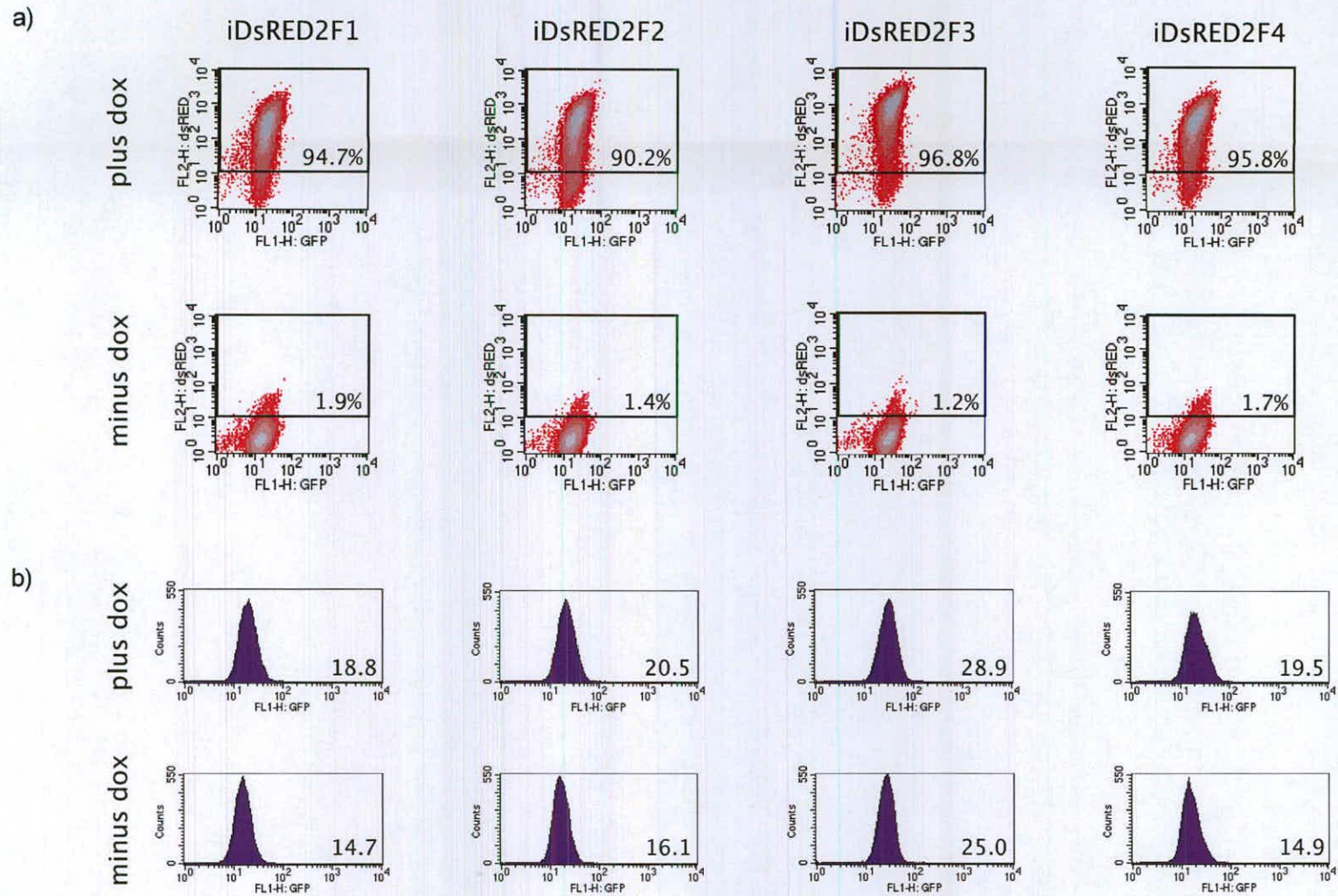


Figure 4.4. Four iDsRED2F cell lines were cultured in the presence and absence of 1 ug/ml dox for two days and analysed by flow cytometry. a) Two dimensional density plots with DsRED2 fluorescence along the y-axis and hrGFP fluorescence along the x-axis. The percentage of DsRED2 positive cells is indicated. b) Histograms depicting hrGFP fluorescence. The median hrGFP fluorescence is shown.

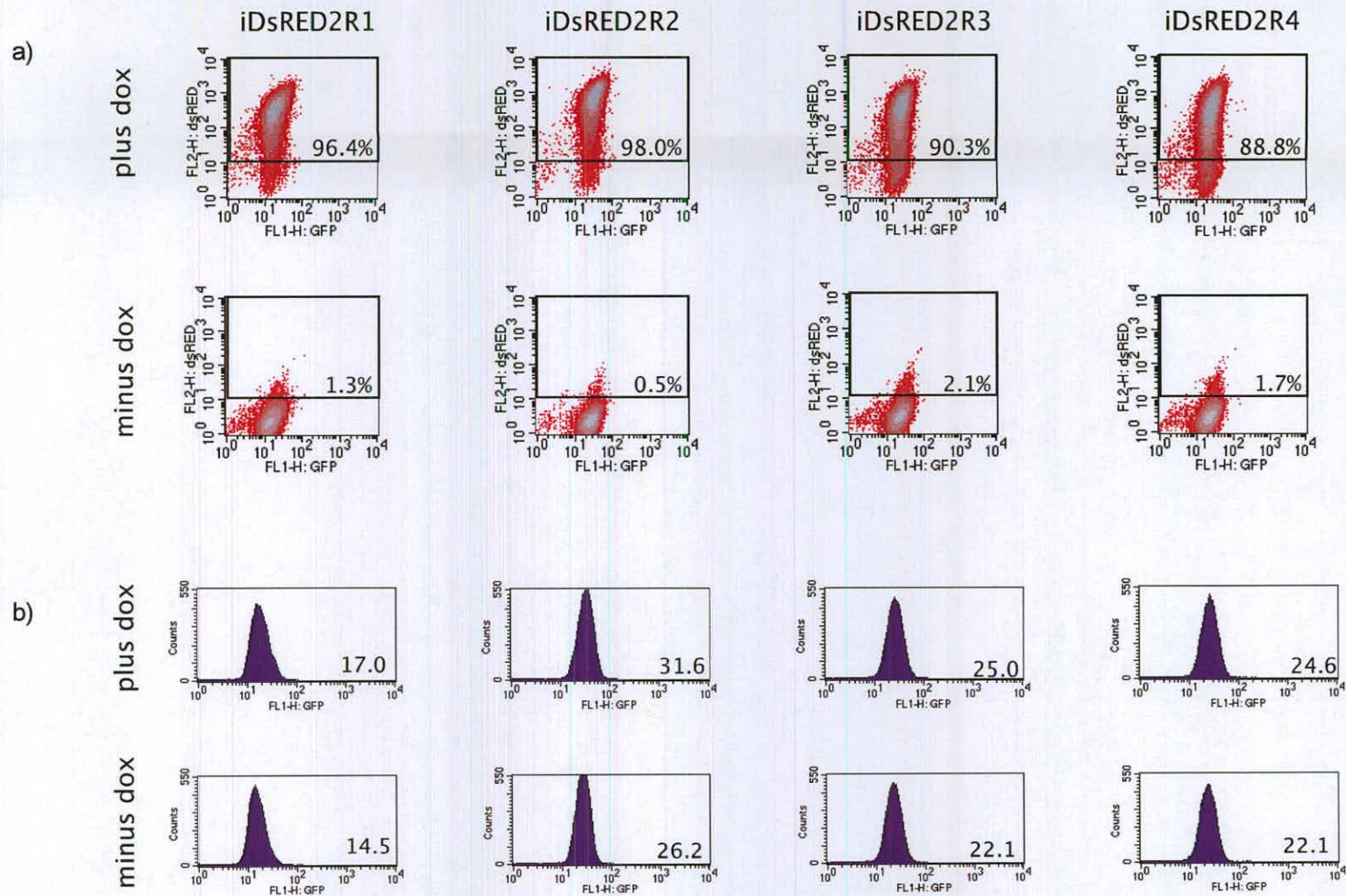


Figure 4.5. Four iDsRED2R cell lines were cultured in the presence and absence of 1  $\mu\text{g/ml}$  dox for two days and analysed by flow cytometry. a) Two dimensional density plots with DsRED2 fluorescence along the y-axis and hrGFP fluorescence along the x-axis. The percentage of DsRED2 positive cells is indicated. b) Histograms depicting hrGFP fluorescence. The median hrGFP fluorescence is shown.

Figure 4.4b and 4.5b depict the hrGFP levels detected for each clone in the presence and absence of dox. It was observed that in contrast to the ihprtDsRED2 cell lines characterised in Chapter 3, which exhibit a negative shift in hrGFP fluorescence upon the addition of dox inducer, the iDsRED2 cell lines analysed exhibit a slight increase in hrGFP fluorescence when dox was added to the cells. This slight shift can be explained by the leakage of DsRED2 fluorescence into the FL1 channel (through which hrGFP is detected) at the compensation settings used during the acquisition of the data. The fact that a negative shift in hrGFP fluorescence was not seen when dox is added to the iDsRED2 cell lines indicates that this is a phenomenon that is restricted to ihprtDsRED2 cell lines.

#### **4.4 Induction of DsRED2 in a gradient of doxycycline concentrations**

Four randomly selected iDsRED2 cell lines were cultured in a range of concentrations of dox for two days after which they were analysed by flow cytometry (Figure 4.6). All the cell lines were shown to respond in a similar fashion to increasing levels of doxycycline. At 0.25  $\mu\text{g/ml}$  dox approximately half the cells from each cell line acquired a DsRED2 positive phenotype. At 0.5  $\mu\text{g/ml}$  dox the majority of cells became DsRED2 positive, and at 1  $\mu\text{g/ml}$  each cell line exhibited DsRED2 fluorescence in over 90% of their population. The highest levels of fluorescence were observed at 2  $\mu\text{g/ml}$  dox.

Interestingly at intermediate dox concentrations (such as 0.25 and 0.5  $\mu\text{g/ml}$  dox), rather than the majority of the cell population expressing intermediate DsRED2 fluorescence levels, there was more of a heterogenous spread of DsRED2-expressing cells along the y-axis. This heterogeneity seemed to have an 'all-or-nothing' element to it as cells predominantly fell into either a relatively high DsRED2 positive population or a DsRED2 negative population (as indicated by the two open circles in the 0.5  $\mu\text{g/ml}$  iDsREDR3 density plot).

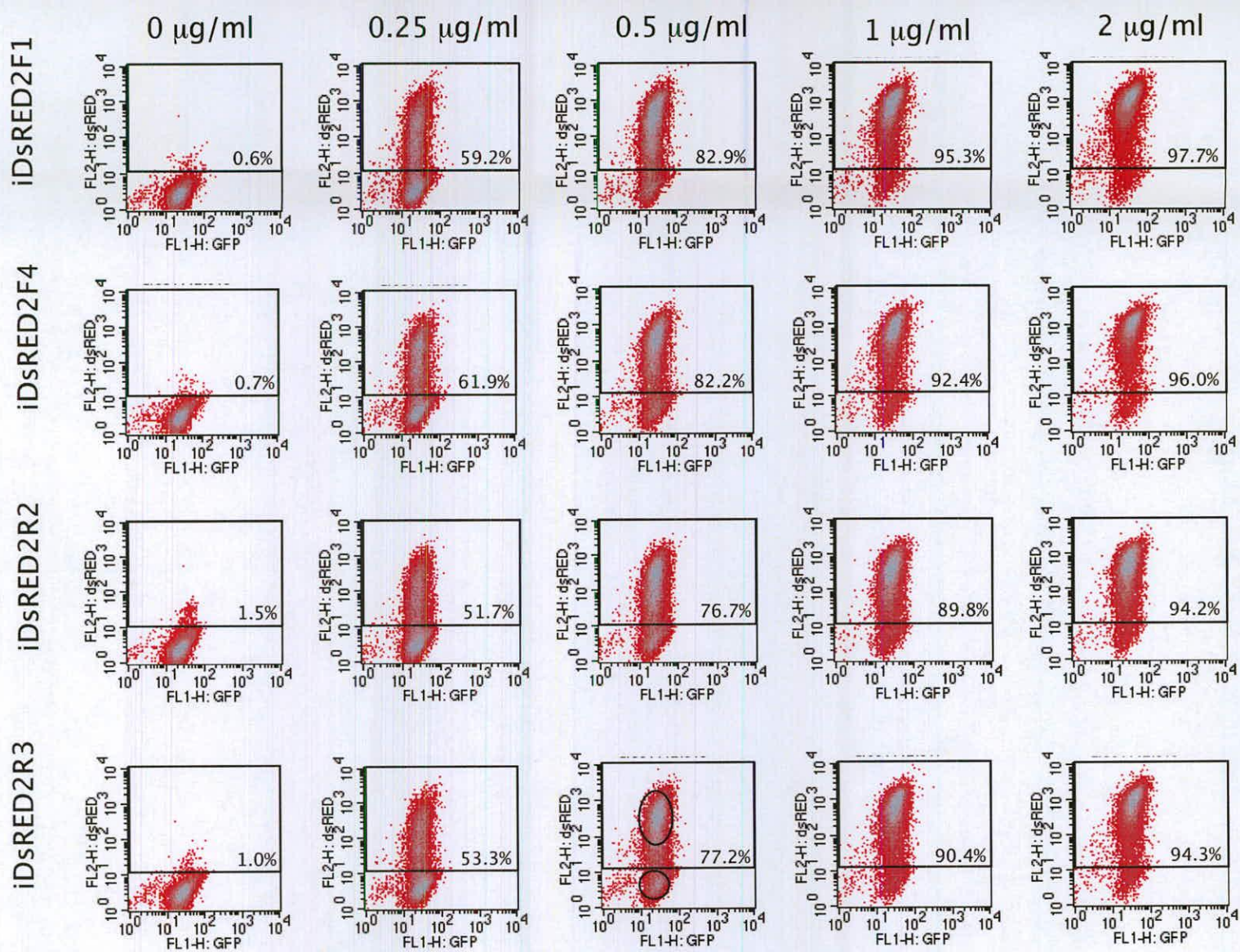


Figure 4.6. Four cell lines were cultured in different concentrations of dox for two days. Two dimensional density plots show DsRED2 fluorescence on the y-axis and hrGFP fluorescence on the x-axis. The percentage of DsRED2 positive cells is indicated. The open circles in iDsRED2R3 at 0.5 µg/ml highlight two discrete populations of cells.

The median DsRED2 fluorescence values are represented graphically in figure 4.7b. It is clear that whilst the percentages of DsRED2-expressing cells did not vary widely from clone to clone (Figure 4.7a), the median fluorescence values were not so consistent between the four cell lines with some clones expressing higher levels of DsRED2 than others. This has implications experimentally, as it is not only the percentage of cells that over-express the desired transgene that is crucial, but also the absolute levels of transgene produced.

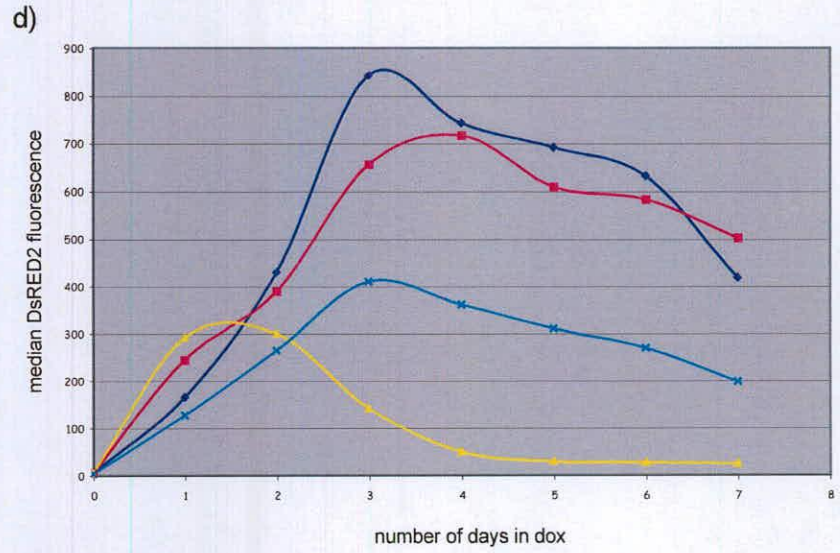
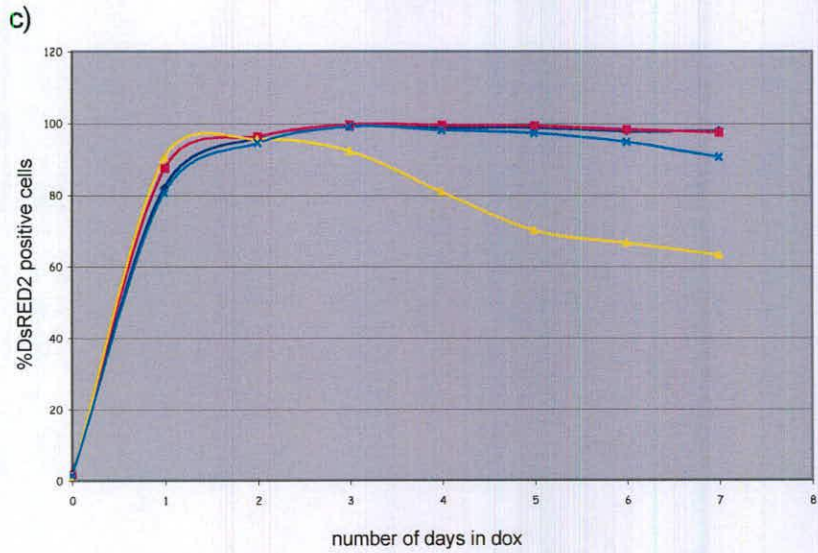
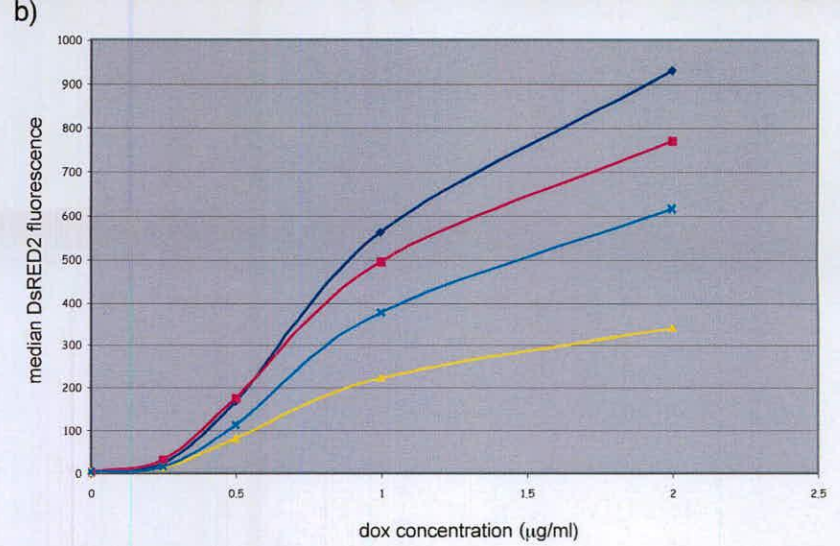
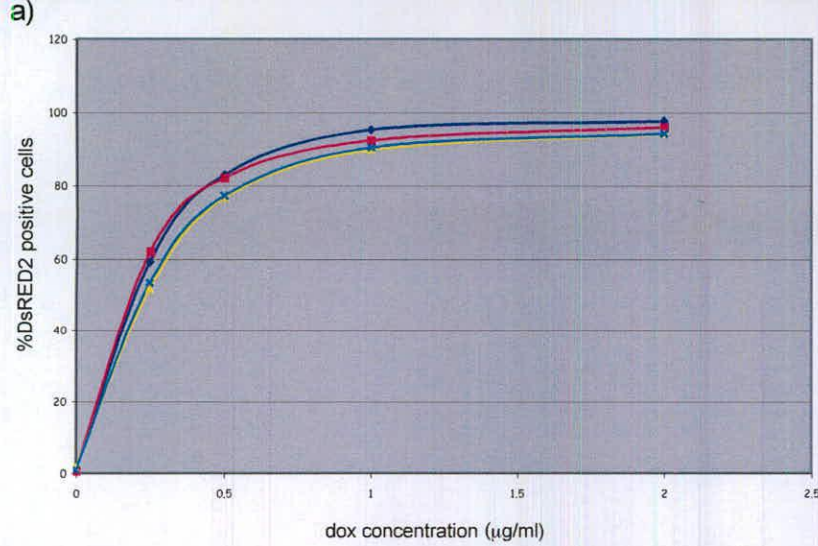
#### **4.5 Induction of DsRED2 after different periods of culture in dox**

The four cell lines were cultured in 1 µg/ml doxycycline over a period of 0-7 days after which time they were analysed for DsRED2 fluorescence by flow cytometry. Two-dimensional density plots for each cell line at each time-point are shown in figure 4.8. It was observed that at day 3, the iDsRED2F1, iDsRED2F4 and iDsRED2R3 cell lines reached their maximum value with regards to the percentage of DsRED2 positive cells (also viewed in Figure 4.7c). After this time these cell lines continued to exhibit DsRED2 in over 90% of their cell populations, however the median DsRED2 fluorescence levels gradually started to decline. Although the iDsRED2R2 cell line was correctly targeted (Figure 4.2), the two-dimensional density profile of this cell line did not follow those of the other cell lines during the time-course experiment. Expression of DsRED2 was observed to peak at day 2, after which time it rapidly declined (Figure 4.7d).

#### **4.6 Discussion**

##### **Induction of DsRED2 in the iDsRED2 cell lines**

The strategy of targeting the rtTA2<sup>S</sup>-M2 to the ROSA26 locus and the TRE to the 5'*hpvt* region gives a far superior readout of transgene expression when compared to the previous approach of targeting both elements to the *hpvt* locus. Whereas the ihprtDsRED2 cell lines exhibited heterogenous expression of the DsRED2 reporter upon



- ◆ iDsRED2F1
- iDsRED2F4
- ▲ iDsRED2R2
- × iDsRED2R3

Figure 4.7 a) Scatter plot depicting percentage of DsRED2 positive cells in each of the four clones at different concentrations of dox. b) Scatter plot of median DsRED2 fluorescence versus dox concentration. c) Scatter plot of percentage of DsRED2 positive cells after different lengths of time in dox. d) Scatter plot of median DsRED2 fluorescence after different periods in dox.

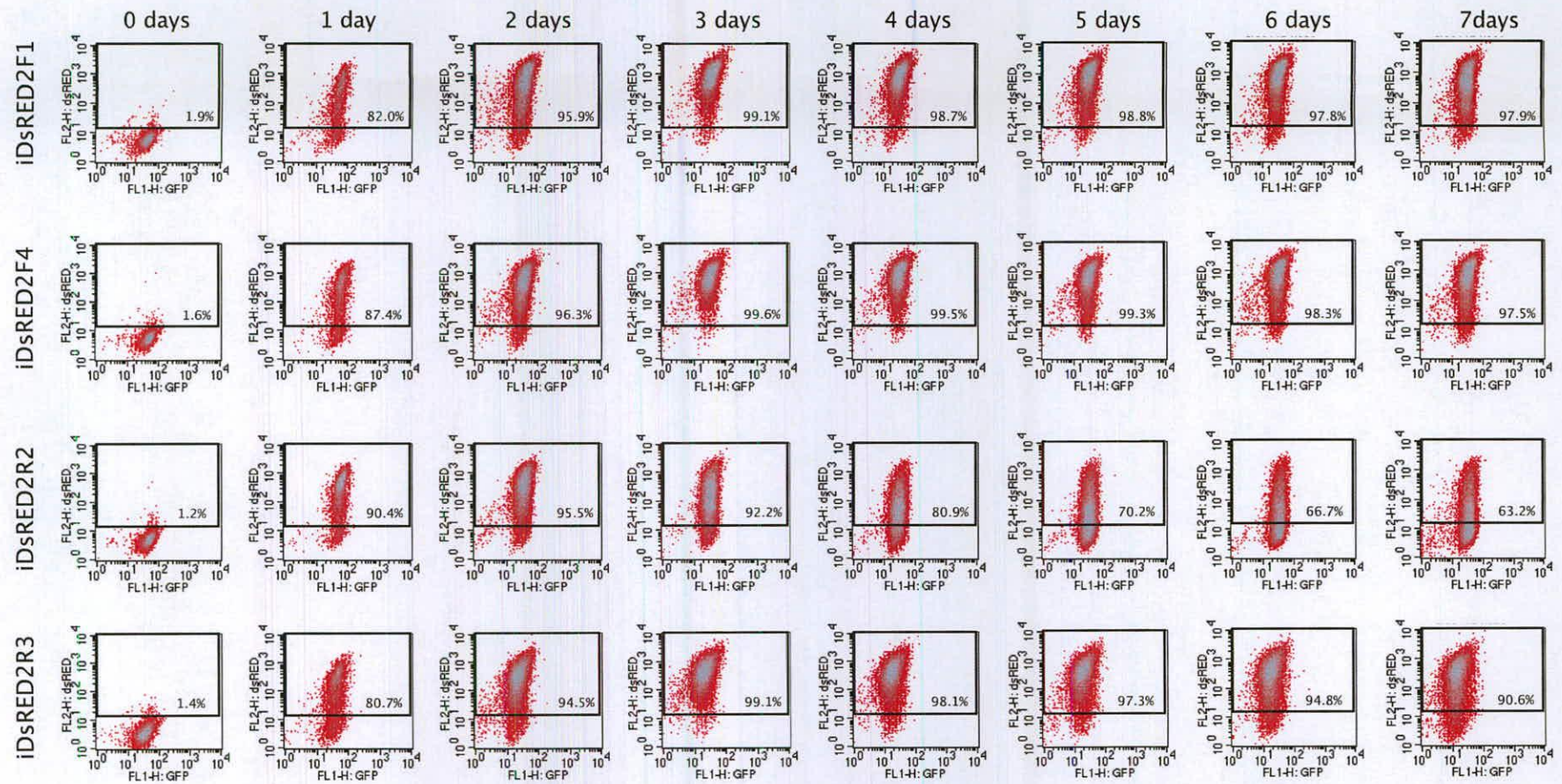


Figure 4.8. Four iDsRED2 lines were cultured in dox for different periods of time. Two dimensional density plots show DsRED2 fluorescence (y-axis) and hrGFP fluorescence (x-axis) at each time-point. The percentage of DsRED2 positive cells is indicated.

the addition of dox, the iDsRED2 cell lines described here expressed DsRED2 at high levels and in a relatively homogenous manner. A similar approach has been carried out by two independent groups (Wutz *et al.*, 2002, Kyba *et al.*, 2002). In both studies the original reverse transactivator rtTA was placed under the control of the endogenous ROSA26 promoter and the TRE targeted to the 5' region of *hprt*. It was found that upon addition of 1 µg/ml dox to ES cell lines harbouring tet-inducible GFP, 94% of the cells became positive for transgene expression (Kyba *et al.*, 2002). This result correlates with the inductions described here.

### **Transgene induction in a gradient of dox concentrations**

The iDsRED2 cell-lines appeared to exhibit some degree of intermediate level of reporter expression. With progressively decreasing concentrations of dox the highest DsRED2-expressing cells could be observed to shift down the y-axis (Figure 4.6; 2 µg/ml to 0.5 µg/ml dox). However at concentrations of 0.5 µg/ml dox and below, an 'all or nothing' effect with regards to levels of DsRED2 expression appeared, where the cells either fell into a relatively high DsRED2-expressing population or a negative DsRED2 population. The tet-system has been previously described as one in which the experimenter is able to manipulate levels of transgene expression by varying the concentration of inducer (Gossen *et al.*, 1995; Urlinger *et al.*, 2000). Although as stated previously this observation was originally based on analysis of luciferase transgene expression, intermediate transgene expression levels have been reported using a rtTA-regulated fluorescent reporter in primary mouse myoblasts, (Kringstein *et al.*, 2000). An independent study, however, reported that rather than population-wide low levels of fluorescent reporter at low concentrations of dox, an increase in the population of negative cells was found (Qu *et al.*, 2004). This study was carried out in ML20 cells using the rtTA2<sup>S</sup>-M2 version of the transactivator. It is therefore possible that the nature of transgene expression at intermediate concentrations of inducer may vary between cell types.

## Analysis of DsRED2 expression after time in dox

The optimum culture period in dox necessary to maximally stimulate DsRED2 reporter expression was 3 days for all of the cell lines analysed bar iDsRED2R2. Whilst the proportion of DsRED2 positive cells in these clones remained constant from 3 to 7 days, it should be noted that there was a gradual decrease in the median value of fluorescence. It is possible that high levels of DsRED2 interfere with ES cell metabolism and therefore over a period of time there is a gradual selection for cell that express lower levels of this fluorescent reporter. Alternatively the concentration of dox used (1 µg/ml) could be slightly toxic to ES cells over long periods of time. If this was the case one would expect a gradual selection for cells able to maintain lower internal concentrations of dox, for example those able to efflux the antibiotic more readily. An indirect effect of this selection pressure would be the accumulation of lower DsRED2 expressing cells as low internal concentrations of dox will not result in maximal reporter transcription.

Experimentally it can be seen that high levels of transgene can be maintained for up to 7 days, which is in contrast to the ihprtDsRED2 cell lines that exhibited near negative transgene expression after 6 days in dox. Although there is a slight downregulation of DsRED2 after 3 days, the majority of the cells still express high levels of the reporter at day 7, and this strategy is therefore able to confer robust, high levels of transgene expression over time (Figure 4.8). The exception to this is the iDsRED2R2 cell line which exhibited a sharp decline in DsRED2 expression after 2 days. It is unclear why this cell line should respond differently to prolonged culture in dox, however it acts as a future caution to check that high expression of a tet-regulated transgene is maintained over the induction time required by the experimental application.

In conclusion, the strategy of targeting the transactivator to the ROSA26 locus and the TRE to the 5' region of the *hprt* gene, established an inducible system that could be used to manipulate the expression of a DsRED2 reporter in a tight and predictable manner. Due to the success of this strategy, this approach was applied to investigate the involvement of the mitogen-activated protein kinase (MAPK) signalling pathway during

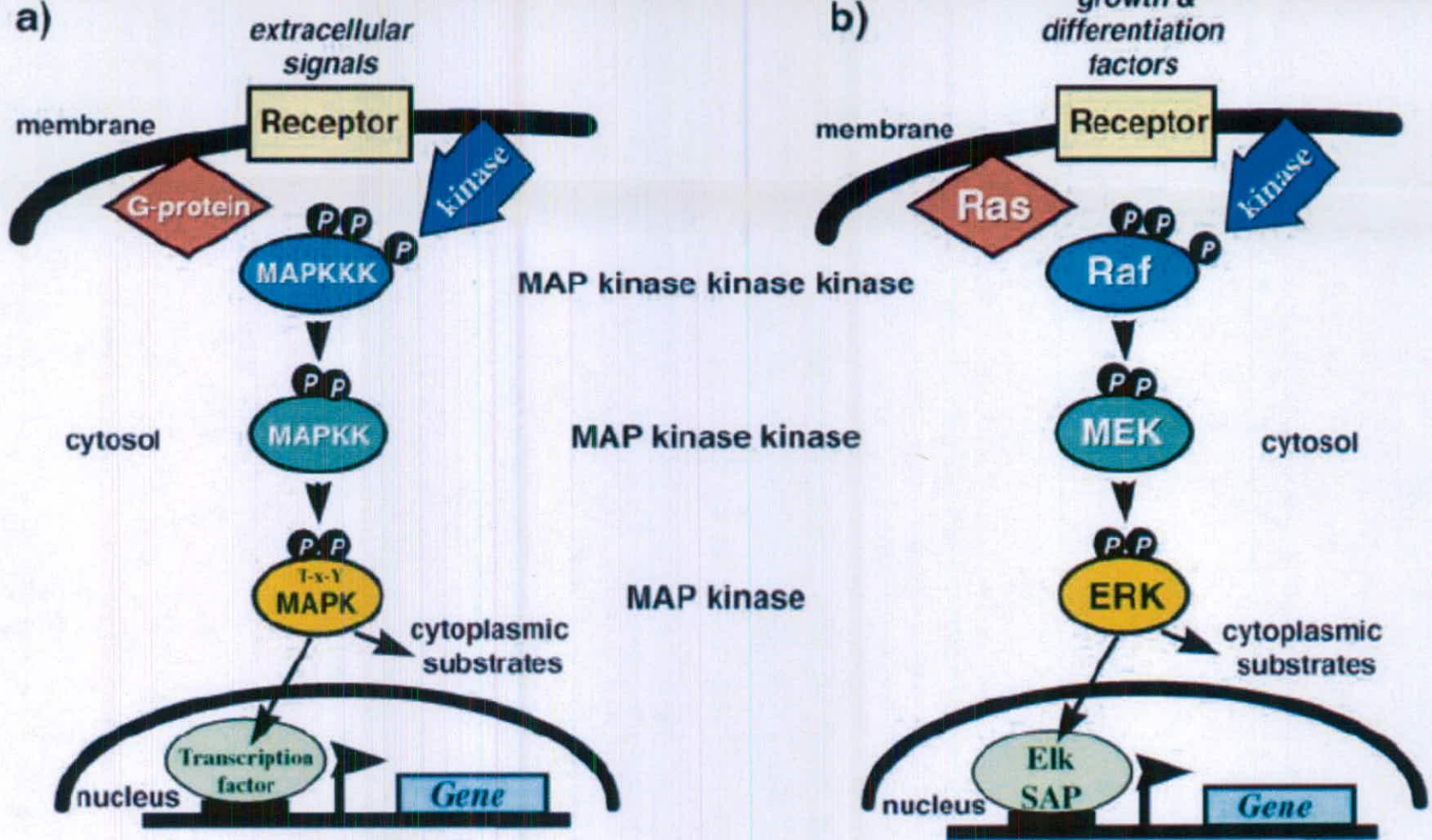
the *in vitro* differentiation of ES cells into neural progenitor cells. This is described in the following chapter.

## Chapter 5

The establishment of inducible MEK1<sub>act</sub> expression in ES cells for investigation into the involvement of the MAPK pathway in neural induction

### 5.1 Introduction

The mitogen-activated protein kinase (MAPK) pathway is a highly conserved signalling pathway that controls fundamental cellular processes such as proliferation, differentiation and cell survival. It is comprised of a series of protein kinase modules that are activated via phosphorylation by upstream kinases, and in turn activate downstream substrates by phosphorylation. The classical MAPK pathway is the extracellular-signal-regulated (ERK) kinase pathway, which comprises Raf, MEK1 (MAPK/ERK kinase), MEK 2 (MEK 1 and MEK 2 are collectively referred to as MEK1/2), and ERK1/2 proteins (see figure 5.1 for a detailed description of this pathway). The MAPK/ERK pathway can be activated by stimulation of many different receptor types including tyrosine kinase (Blume-Jensen *et al.*, 2001), G-protein coupled (Belcheva and Coscia., 2001), and integrin receptors (Barberis *et al.*, 2000). In ES cells the MAPK/ERK pathway acts to



Kolch, Biochem. J. 2000

Figure 5.1 The MAPK/ERK pathway. Five mammalian MAPK pathways have been discovered, with each pathway possessing distinct biological functions. a) A general schematic of MAPK pathway signalling. b) The extracellular signal-regulated kinase (ERK) pathway. The pathway is shown with its three basic kinase modules Raf, MEK1/2 and ERK1/2. Receptor activation leads to the activation of the G-protein ras which recruits Raf to the membrane by binding its N-terminal regulatory domain. Once localised to the membrane Raf becomes activated by phosphorylation and in turn activates downstream MEK1/2 by phosphorylation of specific serine residues. MEK1/2 is a dual specificity kinase which activates its substrate ERK1/2 by phosphorylation of threonine and tyrosine residues located within an activation loop. This action is highly specific as no additional substrates have been found for MEK1/2. Upon activation, ERK1/2 dissociates from MEK1/2 and phosphorylates downstream nuclear and cytoplasmic substrates on serine or threonine residues (active ERK1/2 is able to translocate to the nucleus).

impair self-renewal, and inhibition of MEK signalling inhibits the differentiation of ES cells in embryoid bodies (Burdon *et al.*, 1999).

It has recently been shown that the *in vitro* differentiation of ES cells into neural progenitor cells is dependent on FGF signalling in serum-free conditions (Ying *et al.*, 2003). The pathway acting downstream of FGF receptor (FGFR) signalling has yet to be determined, although the MAPK/ERK pathway is a strong candidate as it is known to regulate the onset of neural development in invertebrate embryos (Kim *et al.*, 2001, Hudson *et al.*, 2003). In addition, recent studies in chick suggest that it functions downstream of FGFR signalling in higher vertebrates to regulate neural development and is also required for the later differentiation of the neural plate (Eblaghie *et al.*, 2003). A proposed mechanism of action is the inhibitory phosphorylation of the Smad1 linker region by Erk1/2, leading to the exclusion of Smad1 from the nucleus and attenuation in BMP signalling (Pera *et al.*, 2003), a state which is characteristic of neural development in the frog and chick. In the chick, early neural genes can be induced by FGF signalling even in the presence of BMP4 (Streit *et al.*, 2000), suggesting that FGF regulates multiple aspects of neural induction.

The recent development of a serum-free adherent monoculture protocol that gives rise to the efficient conversion of ES cells into neuroectodermal precursors offers an *in vitro* model of early neural development and therefore enables investigation into the involvement of the MAPK/ERK pathway in neural induction (Ying *et al.*, 2003). Having established a robust strategy for implementing inducible gene expression in ES cells it was decided to use this system to overexpress an activated component of the MAPK/ERK pathway during the *in vitro* differentiation of ES cells into neural progenitor cells. The prediction was that activation of MAPK/ERK signalling would either lead to an increase in the proportion of neural progenitors obtained after ES cells differentiation (the standard differentiation procedure does result in 100% conversion), or in a more speedy acquisition of the neural precursor fate. Such an outcome would provide evidence that the MAPK/ERK pathway is involved in the conversion of ES cells into neural

progenitor cells, and support the findings that it functions to regulate neural development in higher vertebrates (Eblaghie *et al.*, 2003).

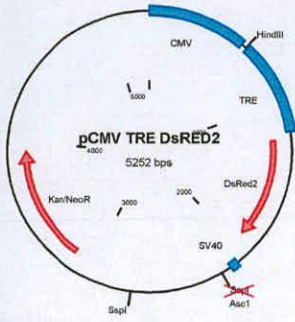
A mutant form of MEK1 (hereon referred to as MEK<sub>act</sub>) has been developed which is able to stimulate the ERK pathway independently of upstream Raf activity (Brunet *et al.*, 1994). The mutant MEK protein is 'activated' due to the substitution of serine residues at positions 218 and 222 (S218/S222) with aspartic acid residues. Both S218 and S222 are phosphorylated by Raf in the MAPK/ERK cascade, and the negatively charged aspartic acid residues that are substituted for S218 and S222 in the mutant MEK<sub>act</sub> mimic this process and thus render MEK<sub>act</sub> constitutively active. The cDNA encoding MEK<sub>act</sub> was placed under the control of the tet-responsive promoter in ES cells, with the aim of investigating the effects (on neural acquisition) of activating the ERK pathway at different timepoints during the differentiation of ES cells into neural progenitors using the serum-free monoculture differentiation protocol. As the expression of tet-regulated transgenes can be inhibited after initial overexpression, this offers the option of transiently expressing MEK<sub>act</sub> during different stages of the differentiation procedure, in order to investigate those windows during differentiation in which the activation of MAPK/ERK pathway has the most pronounced effect on neuroectoderm fate acquisition. Transient overexpression of MEK<sub>act</sub> during the initial stages of differentiation also allows effects on neural induction to be separated from possible mitogenic effects of the MAPK pathway on differentiated cell types. In addition, the pharmacological inhibition of the FGF receptors in ES cells and subsequent induction of MEK<sub>act</sub> enables investigation into the sufficiency of the MAPK/ERK pathway in neural specification.

## **5.2 Construction of the 5'hprtTREMEK<sup>act</sup> and 5'hprtTREMEK<sup>wt</sup> targeting vectors**

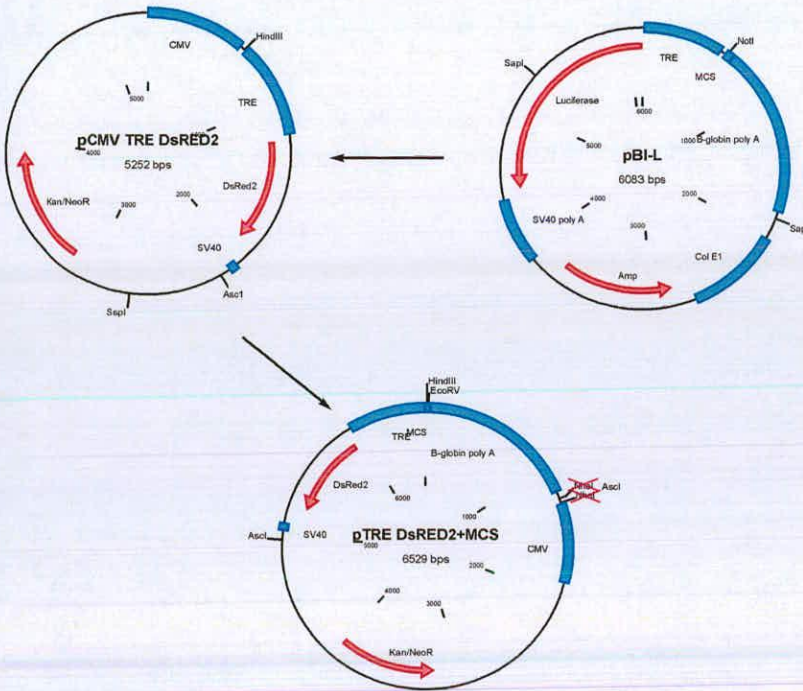
In the event of inserting TRE-linked cDNAs into the pSKB1 targeting vector, it was first necessary to create an intermediate plasmid, (TREDsRED2+MCS(*AscI*), Figure 5.2c) that contained a multiple cloning (MCS) site downstream of the TRE in order to facilitate cloning. This plasmid contained two *AscI* sites flanking the TRE and juxtaposed

Figure 5.2 Juxtaposition of HA-MEK to the TRE sequence. a) Plasmid CMVTREsRED2 was partially digested with *SspI* and subsequently incubated with klenow polymerase and dNTPs to blunt the protruding fragment ends. An oligonucleotide sequence encoding an *AscI* restriction site was then inserted into this site. b) Plasmid BI-L was digested with *NotI* and *SapI* restriction endonucleases to extract the fragment encoding the multiple cloning site (MCS) and the  $\beta$ -globin poly A sequence. This fragment was incubated with klenow polymerase and dNTPs to make the protruding fragment ends flush, after which it was ligated into the modified CMVTREsRED2 plasmid at the *HindIII* site (in preparation for this ligation the plasmid was also 'end-filled' with klenow polymerase and dNTPs). The resulting plasmid TREsRED2+MCS was digested with *NheI* restriction endonuclease, and the overhanging ends were blunted with klenow polymerase, before an oligonucleotide sequence encoding an *AscI* restriction site was inserted. c) Plasmid ECE-HA-MEKact and plasmid ECE-HA-MEKwt were digested with *HindIII* and *XbaI* restriction endonucleases to extract the HA-MEKact and HA-MEKwt encoding fragments (as both plasmids yield *HindIII/XbaI* fragments of the same size only one plasmid is shown here). These fragments were then inserted into pTREsRED2+MCS(asc1) which had also previously been digested with *HindIII* and *XbaI* restriction endonucleases. The resulting plasmid, HA-MEKTREsRED2 contained the MEK and the DsRED2 open reading frames in divergent orientations. Digestion of the plasmid with *AscI* yielded the TRE and linked transgene sequences (see next figure).

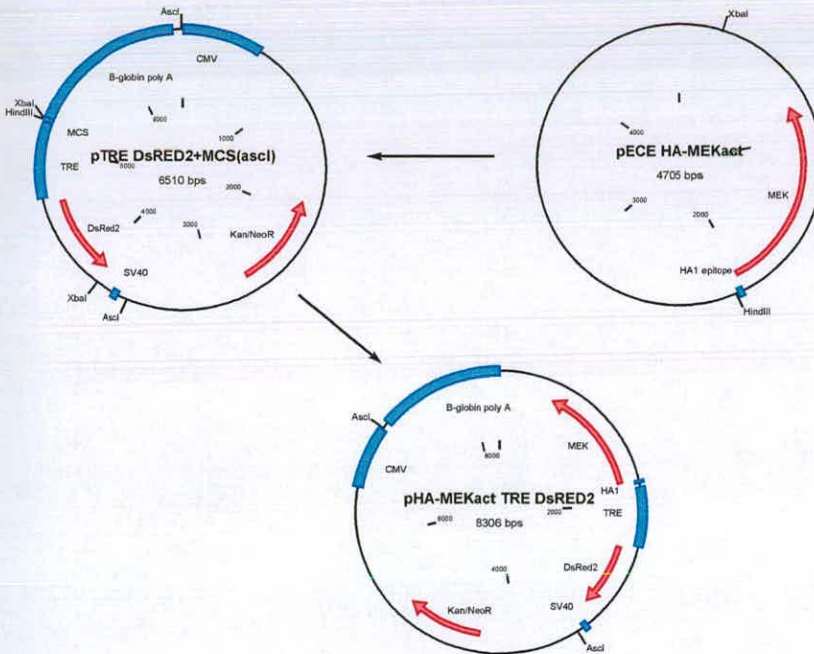
a)



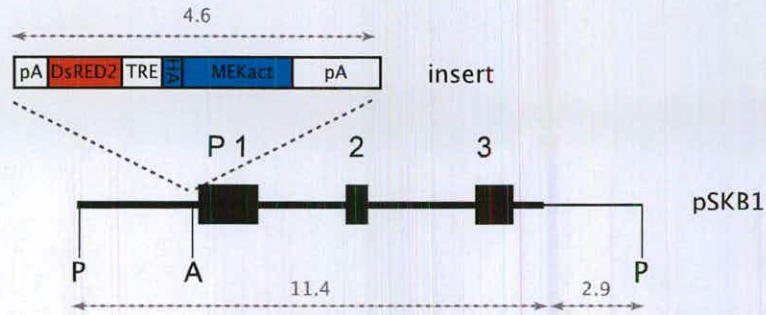
b)



c)



a) 5'hprtTREMEKact



b) 5'hprtTREMEKwt

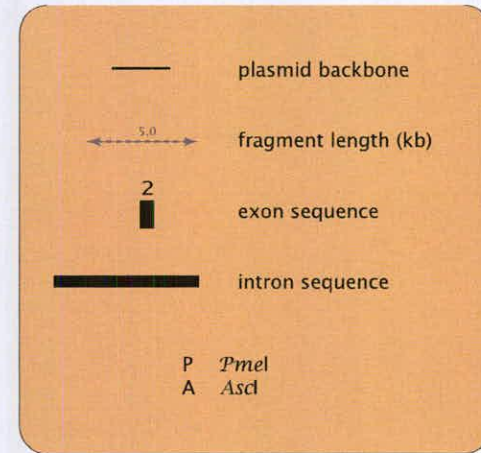
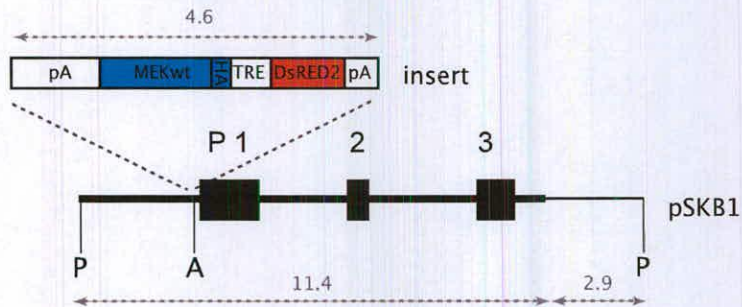


Figure 5.3 Creation of the 5'hprtTREMEKact and 5'hprtTREMEKwt targeting vectors. The TRE-containing fragment yielded from *Ascl* digest of plasmids MEKactTREDSRED2 and MEKwtTREDSRED2 (see previous figure), were inserted into the pSKB1 targeting vector immediately upstream of the promoter sequence (the pSKB1 vector had previously been modified to replace the *NotI* insertion site with an *Ascl* encoding oligonucleotide sequence to facilitate the cloning of these elements). The final targeting vectors were termed 5'hprtTREMEKact a) and 5'hprtTREMEKwt b) respectively.

sequences to be inserted into the pSKB1 targeting vector. This enabled the straightforward insertion of the transgenic sequences as an *AscI* fragment into an *AscI* site positioned just upstream of the genomic *hprt* promoter in the pSKB1 vector sequence.

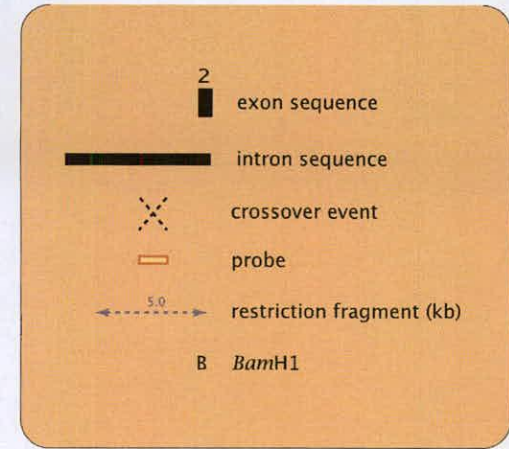
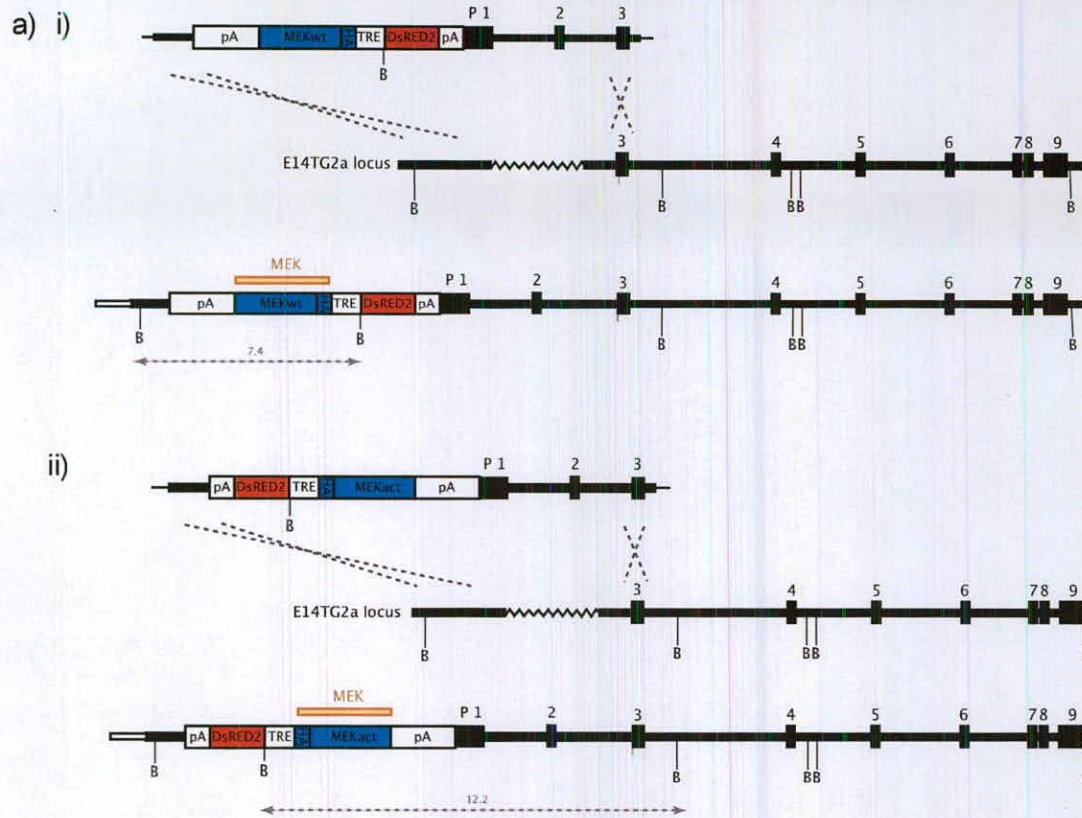
The first step in the creation of this intermediate plasmid was to change the *SspI* site downstream of the SV40 polyA sequence in plasmid CMVTREDSRED2 to an *AscI* site. This was achieved by partial digestion of the plasmid with *SspI*, followed by end-fill of the protruding ends by incubation with klenow polymerase and dNTPS, and subsequent ligation with an oligonucleotide encoding an *AscI* restriction site (Figure 5.2a). The MCS and  $\beta$ -globin polyA sequences from plasmid BI-L were then inserted into pCMVTREDSRED2 in the opposite orientation to that of the DsRED2 sequence, (Figure 5.2b). This was carried out by digestion of pBI-L with *NotI* and *SapI* restriction endonucleases and concomitant digestion of CMVTREDSRED2 with *HindIII* enzyme. Both the CMVTREDSRED2 plasmid and the extracted fragment were treated with klenow polymerase and dNTPs to blunt the protruding ends, after which the two DNAs were ligated. This resulted in the production of pTREDSRED2+MCS (Figure 5.2b). Digestion of this plasmid with *NheI* restriction endonuclease, followed by klenow end-fill and insertion into this site of an oligonucleotide encoding an *AscI* restriction site resulted in the production of pTREDSRED2+MCS(*AscI*) (Figure 5.2c).

In addition to creating cell lines harbouring an inducible form of active MEK, cell lines expressing inducible wild-type MEK were also created to act as control cell lines. The active MEK and wt MEK sequences were previously fused to an HA-epitope (Pages *et al.*, 1994), and will be referred to as HA-MEK<sub>act</sub> and HA-MEK<sub>wt</sub> respectively. HA-MEK<sub>act</sub> and HA-MEK<sub>wt</sub> were extracted as *HindIII/XbaI* fragments from plasmids ECE-HA-MEK<sub>act</sub> and ECE-HA-MEK<sub>wt</sub> respectively (kind gifts from Gilles Pages). These DNA fragments were inserted next to the TRE in *HindIII/XbaI* digested pTREDSRED2+MCS(*AscI*) (Figure 5.2c). Digestion of these resulting plasmids with *AscI* yielded the TRE and linked cDNAs, with respective polyA sequences, which were then inserted into *AscI* digested pSKB1 targeting vector (Figure 5.3a and 5.3b), to produce 5'*hprt*TREMEK<sup>act</sup> and 5'*hprt*TREMEK<sup>wt</sup> respectively. Note that the MEK<sub>act</sub> and

MEK<sub>wt</sub> sequences were present in opposite orientations within the pSKB1 vector sequences, (this was due to the ability of the cloned fragment to insert in either orientation as a consequence of identical cohesive ends). As it has previously been shown that the orientation of transgenes targeted to the region 5' of *hprt* has no effect on their expression levels (chapter 4) this was considered to be of no significance to the subsequent experiments.

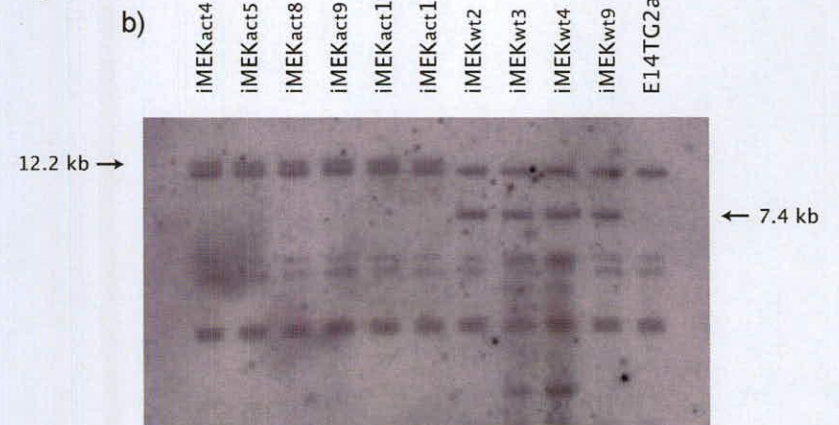
### 5.3 The targeting of the vectors to the 5' region of *hprt*

After the linearisation of 5'*hprt*TREMEK<sup>act</sup> and 5'*hprt*TREMEK<sup>wt</sup> with *Pme*I restriction endonuclease, both vectors were electroporated into R26-rtTA2<sup>s</sup>-M21F ES cells. The day after electroporation the cells were plated out in HAT selective medium. After approximately 10 days, HAT resistant colonies were picked and expanded, and genomic DNA was extracted for analysis by southern hybridisation. Figure 5.4a depicts the expected integration even upon correct targeting of the locus by a double recombination event between the correction vector and the endogenous *hprt* locus, with the predicted fragment sizes detected after digestion with *Bam*HI and hybridisation to a probe encoding the MEK sequence shown. The resulting blot shows that all iMEK<sup>act</sup> clones (those targeted with the 5'*hprt*TREMEK<sup>act</sup> vector), and all iMEK<sup>wt</sup> clones (those targeted with the 5'*hprt*TREMEK<sup>wt</sup> vector) exhibited the expected fragment sizes (12.2 kb and 7.4 kb for iMEK<sup>act</sup> and iMEK<sup>wt</sup> cell lines respectively). It should be noted that the size of band expected for the two types of cell line are different owing to the opposite orientation of MEK<sup>act</sup> and MEK<sup>wt</sup> in their respective targeting vectors. The other bands visible on the Southern autoradiograph are presumably a result of the MEK probe hybridising to endogenous *mek*, as such bands are also visible on the E14TG2a control. The two clones iMEK<sub>wt</sub>3 and iMEK<sub>wt</sub>4 exhibit an additional band of around 2.5 kb, which cannot be accounted for.



Not drawn to scale

Figure 5.4 a) i) The targeting vectors 5'hprtTREMekact and ii) 5'hprtTREMekwt were electroporated into R26rtTA2s-M21F cells. The expected restriction fragments detected after digestion with *Bam*H1 and hybridisation to a MEK sequence-specific probe are shown. b) Ten clones derived from these two electroporations were analysed by Southern hybridisation. E14TG2a, unelectroporated cells.



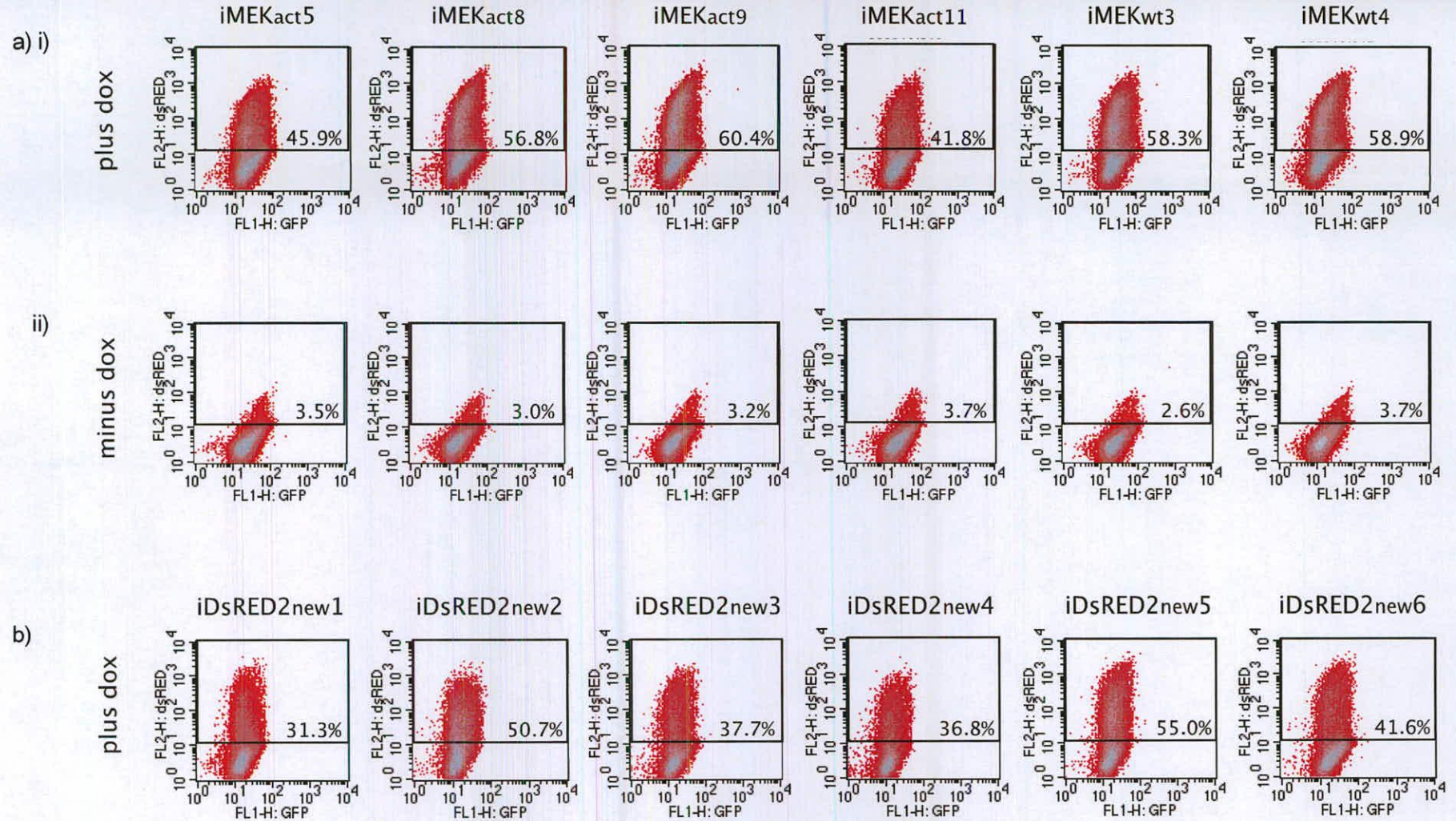


Figure 5.5 a) Four iMEKact and two iMEKwt clones were cultured in the presence i) and absence ii) of 2  $\mu\text{g/ml}$  dox for three days after which time they were analysed by flow cytometry. The resulting two-dimensional plots are shown. b) R26rtTA2s-M21F cells were targeted with 5'hrpTREDsRED2F vector and analysed by flow cytometry after three days in 2  $\mu\text{g/ml}$  dox. For all plots hrGFP fluorescence is shown on the x-axis and DsRED2 fluorescence is shown on the y-axis.

#### **5.4 Flow cytometry in the presence and absence of dox**

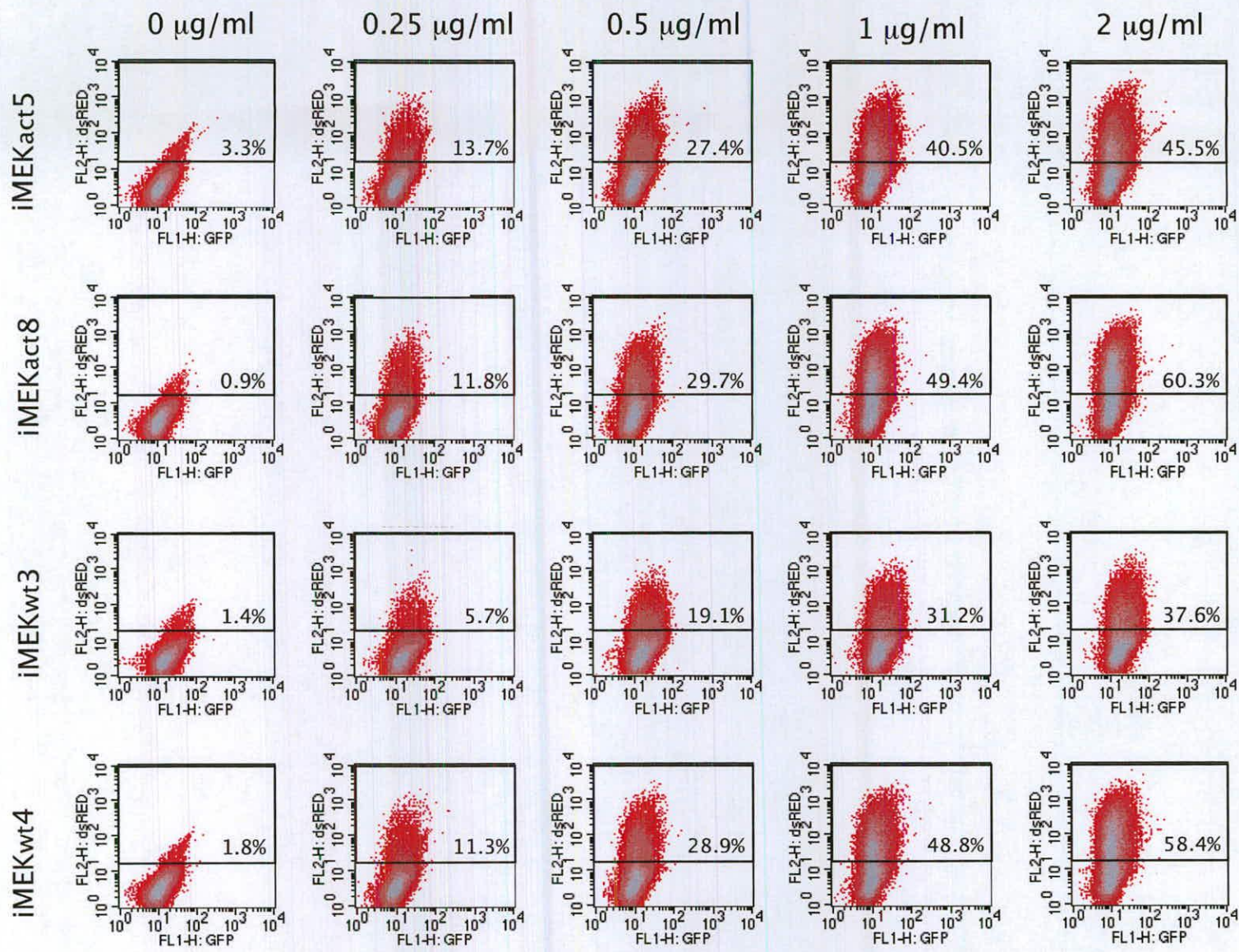
The cell lines were cultured in the presence (Figure 5.5ai) and absence (Figure 5.5aii) of 2 µg/ml dox for 3 days (as previously this was shown to be the optimal time for maximal induction of DsRED2 in the iDsRED2 cell lines as described in Chapter 4), after which time they were analysed by flow cytometry. Unexpectedly the levels of DsRED2 fluorescence induced were not as high as those exhibited by the iDsRED2 cell lines described in the previous chapter. Induction of the DsRED2 reporter was seen in the presence of dox in every cell line analysed, with negligible basal levels seen in the absence of dox. There was however, a large population of cells negative for DsRED2 fluorescence observed in every clone cultured in dox, although the size of this negative population varied between cell lines. The remainder of the population expressed DsRED2 to varying extents with the profile of expression appearing very heterogenous.

#### **5.5 Induction in a gradient of dox concentrations and timecourse**

Two iMEKact and two iMEKwt clones were chosen at random to culture in a gradient of dox concentrations for three days after which time they were analysed by flow cytometry (Figure 5.6 and Figures 5.8a and 5.8b). All clones induced increasing numbers of DsRED2 positive cells with increasing concentrations of dox (Figure 5.6), and also increasing levels of DsRED2 fluorescence (Figure 5.8b). The highest inductions were seen at 2 µg/ml dox, however this concentration was not sufficient to induce high and homogenous levels of DsRED2. Indeed DsRED2 expression was heterogenous at all concentrations. The overall levels of transgene expression were relatively poor when compared to the iDsRED2 cell lines from the previous chapter.

The four cell lines were then cultured in 2 µg/ml dox for periods of between 0 and 7 days (Figures 5.7, and 5.8c and 5.8d). It was observed that after two days the percentage of DsRED2 expressing cells and median fluorescence levels remained relatively constant until day 5 when levels started to decline. An exception to this was iMEKwt4 which

Figure 5.6 Two iMEKact and two iMEKwt clones were cultured in a gradient of dox concentrations for 3 days. Two dimensional flow cytometry plots are shown with the percentage of DsRED2 positive cell indicated. HrGFP fluorescence is shown on the x-axis and DsRED2 fluorescence on the y-axis



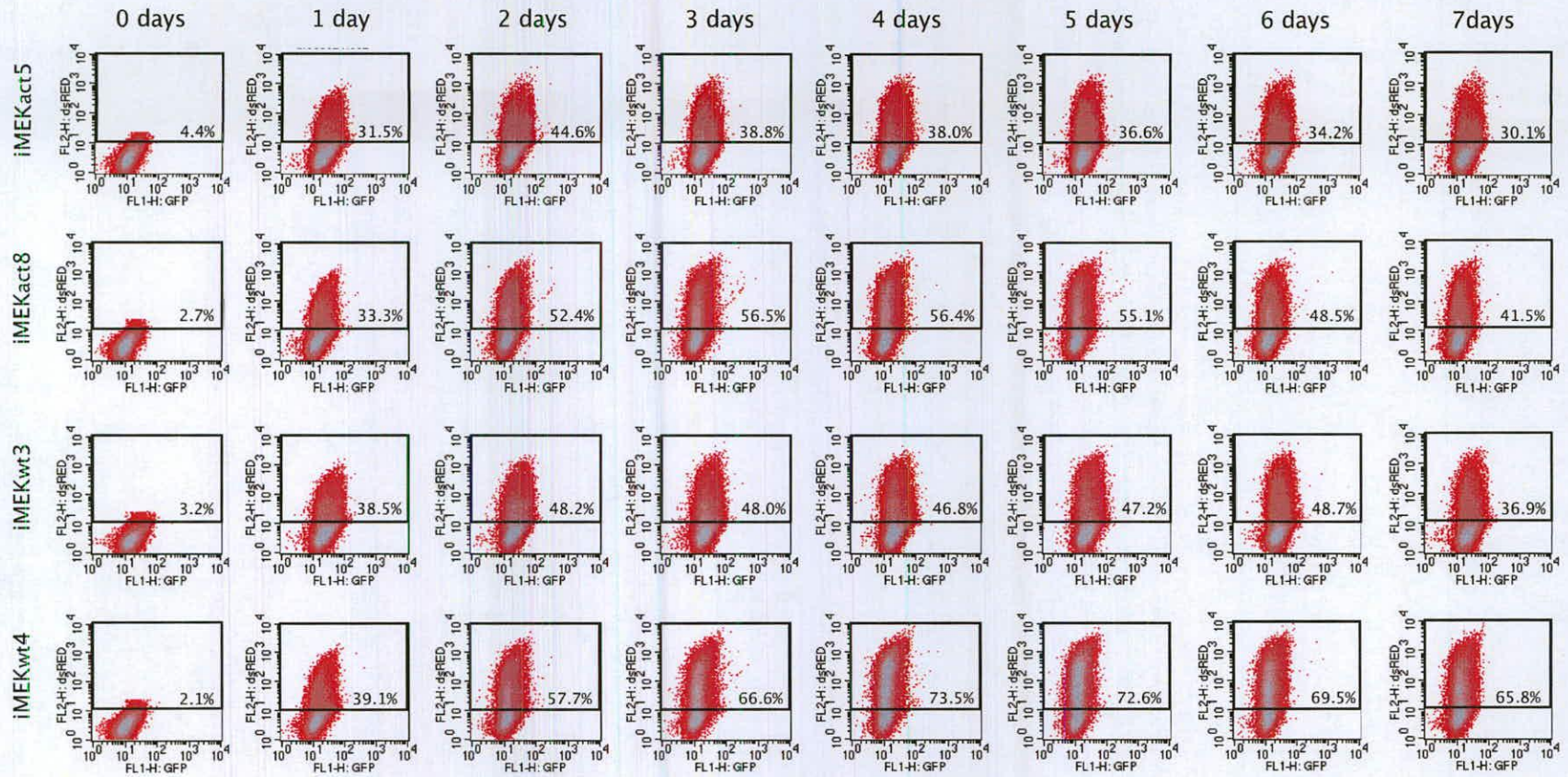
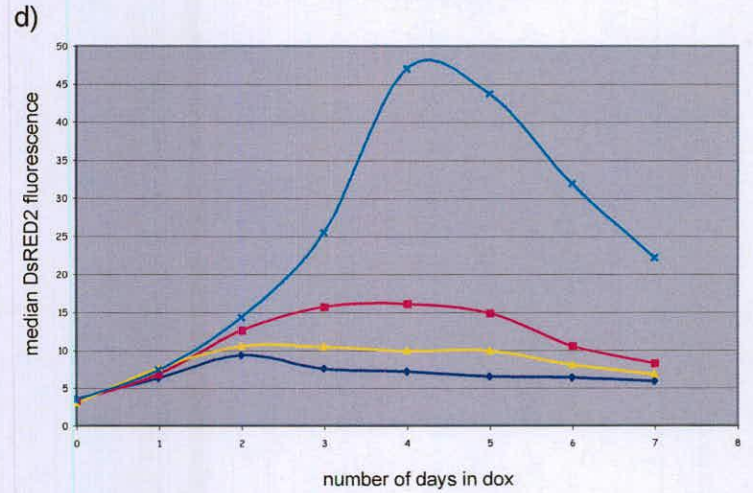
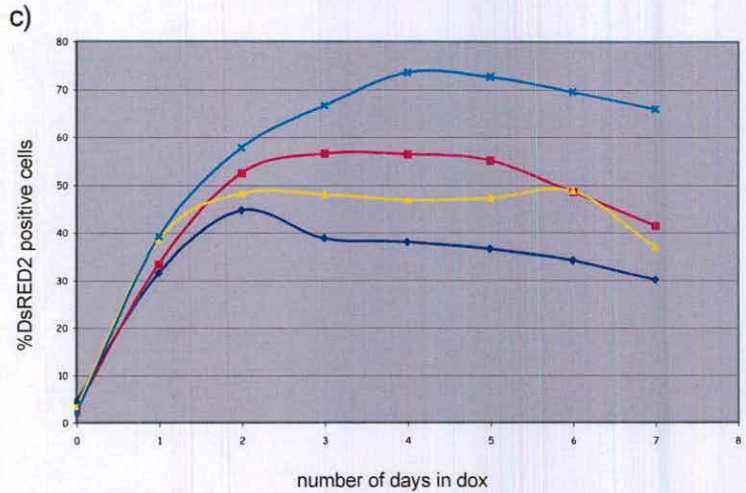
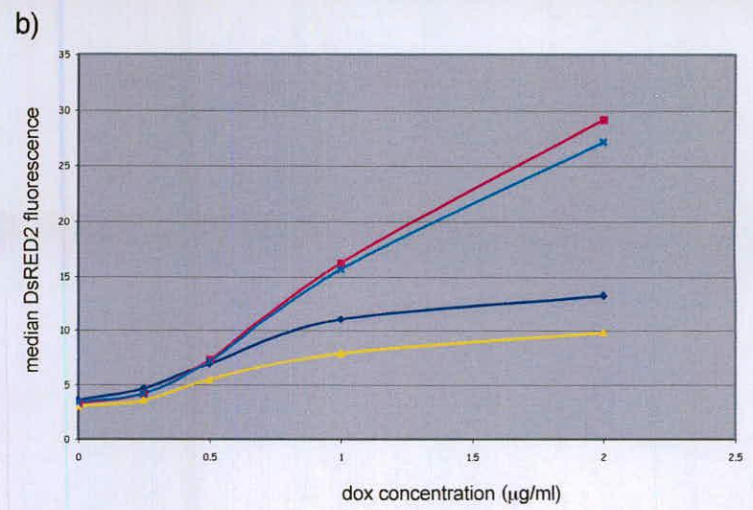
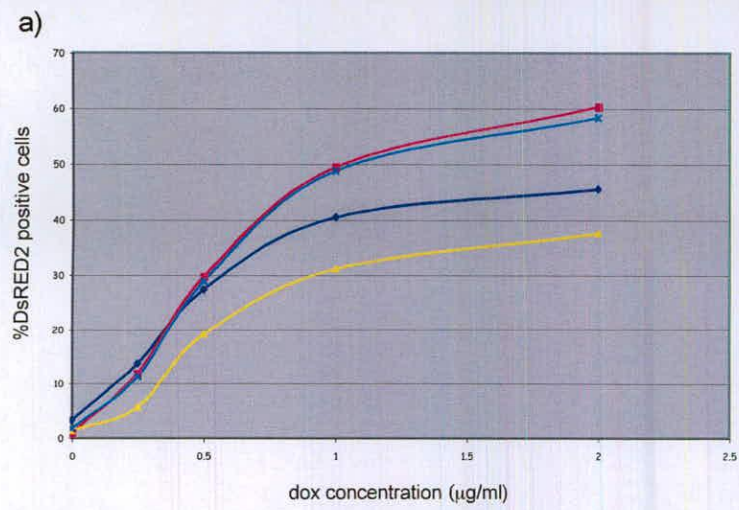


Figure 5.7 Two iMEKact and two iMEKwt clones were cultured in 2  $\mu\text{g/ml}$  dox for 0 – 7 days and then analysed by flow cytometry. Two dimensional flow cytometry plots are shown with hrGFP fluorescence on the x-axis and DsRED2 fluorescence on the y-axis



- ◆ iMEKact5
- iMEKact8
- ▲ iMEKwt3
- × iMEKwt4

Figure 5.8 a) Scatter plot depicting percentage of DsRED2 positive cells in each of the four clones at different concentrations of dox. b) Scatter plot of median DsRED2 fluorescence versus dox concentration. c) Scatter plot of percentage of DsRED2 positive cells after different lengths of time in dox. d) Scatter plot of median DsRED2 fluorescence after number of days in dox.

exhibited the highest levels of DsRED2 fluorescence at day 4 (Figure 5.8d), after which time the levels dramatically declined.

## 5.6 Retargeting R26-rtTA2<sup>s</sup>-M21F cells with iDsRED2 vectors

The iMEKact and iMEKwt cell lines exhibited low and heterogenous expression of DsRED2 reporter when compared to the iDsRED2 clones of the previous chapter. Both experimental strategies involved targeting the TRE and linked cDNAs to the 5' region of *hprt* in R26-rtTA2<sup>s</sup>-M21F cells. However, the establishment of the iMEKact and iMEKwt cell lines involved the creation of new intermediate TRE-linked plasmids for the cloning procedure, one such consequence of which was that the DsRED2 ORF was positioned closer to the TRE sequence in the 5'*hprt*TREMEK vectors than in the 5'*hprt*TREDSRED2 vectors.

In an effort to determine whether the differences in transgene expression between the iMEK and iDsRED2 cell lines was due to subtle differences in the positioning of the transgene open reading frames in relation to the TRE, the R26-rtTA2<sup>s</sup>-M21F parental cells that were used in the establishment of the iMEK cell lines were thawed and electroporated with the 5'*hprt*TREDSRED2F targeting vector (see Figure 4.2). After selection in HAT medium for about 10 days, colonies were picked, expanded and cultured in dox inducer for three days. Flow cytometry was then used to detect levels of DsRED2 fluorescence. The resulting two-dimensional density plots are shown in Figure 5.5b. It is clear that these iDsRED2new clones all expressed poor levels of the DsRED2 reporter and therefore it appears that the suboptimal expression of inducible transgene was not a result of vector manipulation but rather an issue of the parental cell line.

## 5.7 Immunohistochemistry of iMEKact and iMEKwt cell lines

Having previously confirmed that the iDsRED2 reporter was inducibly expressed, the cell lines were analysed for the expression of their exogenous MEK proteins. The cell-lines iMEKact8 and iMEKwt4 were cultured in dox for three days, after which time they were

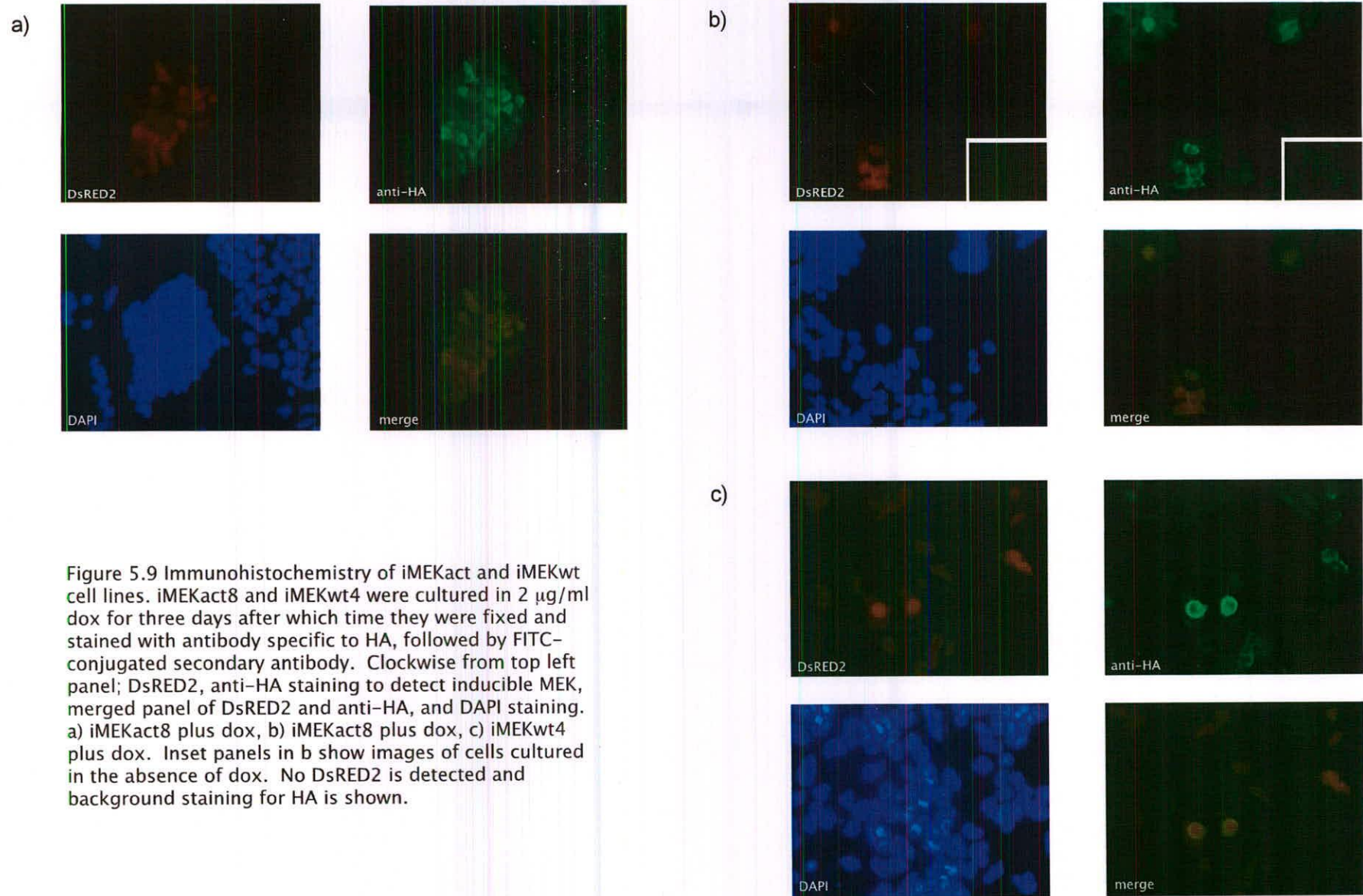


Figure 5.9 Immunohistochemistry of iMEKact and iMEKwt cell lines. iMEKact8 and iMEKwt4 were cultured in 2  $\mu\text{g/ml}$  dox for three days after which time they were fixed and stained with antibody specific to HA, followed by FITC-conjugated secondary antibody. Clockwise from top left panel; DsRED2, anti-HA staining to detect inducible MEK, merged panel of DsRED2 and anti-HA, and DAPI staining. a) iMEKact8 plus dox, b) iMEKact8 plus dox, c) iMEKwt4 plus dox. Inset panels in b show images of cells cultured in the absence of dox. No DsRED2 is detected and background staining for HA is shown.

fixed and analysed for MEK expression by immunohistochemistry and subsequent microscopy (as inducible MEK was fused to an HA epitope, immunostaining with an antibody specific to HA detected only inducible MEK, not endogenous MEK). Both iMEKact8 and iMEKwt4 expressed their respective HA-MEK fusion proteins (Figure 5.9). It was shown that the inducible MEK protein co-localised with DsRED2 in the ES cells treated with dox (merged panels of the anti-HA and DsRED2 fields). As MEK is a cytoplasmic protein and DsRED2 is both nuclear and cytoplasmic, a yellow halo around the nucleus of cells expressing the inducible proteins was seen in the merged views (most clearly viewed in figures 5.9b and 5.9c). This confirmed that the MEK fusion protein was induced in the presence of dox, and also that the TRE was truly bidirectional, as cells that highly expressed inducible MEK also expressed DsRED2.

## **5.8 Immunoblotting of iMEKact and iMEKwt cell lines**

In order to analyse activation of the MAPK/ERK pathway by MEKact, the iMEKact8 and iMEKwt4 ES cell lines were cultured in the presence or absence of dox and then analysed by immunoblot for phosphorylated ERK1/2 levels after stimulation with either LIF, FGF4, or in the absence of stimulation. The detection of elevated phospho-ERK1/2 (P-ERK) levels in the iMEKact8 cell line in comparison to those levels of the iMEKwt4 clone cultured in identical conditions, would indicate activation of the MAPK/ERK pathway by the exogenous active MEK. This would enable investigation into the involvement of this pathway during neural induction using a monoculture differentiation protocol, (Ying *et al.*, 2003).

The MAPK/ERK pathway is permanently activated by factors in standard cell culture conditions (such as growth factors present in serum and the addition of LIF cytokine to cell culture media). Before lysis of the cells for p-ERK1/2 detection, it was therefore necessary to culture the cells overnight in the absence of serum and LIF, in order to obtain basal levels of ERK1/2 phosphorylation, (basal phospho-ERK1/2 are routinely detected, after culture in the absence of serum and LIF, probably as a result of intercellular FGF and integrin signalling). It was predicted that the iMEKact8 cell line

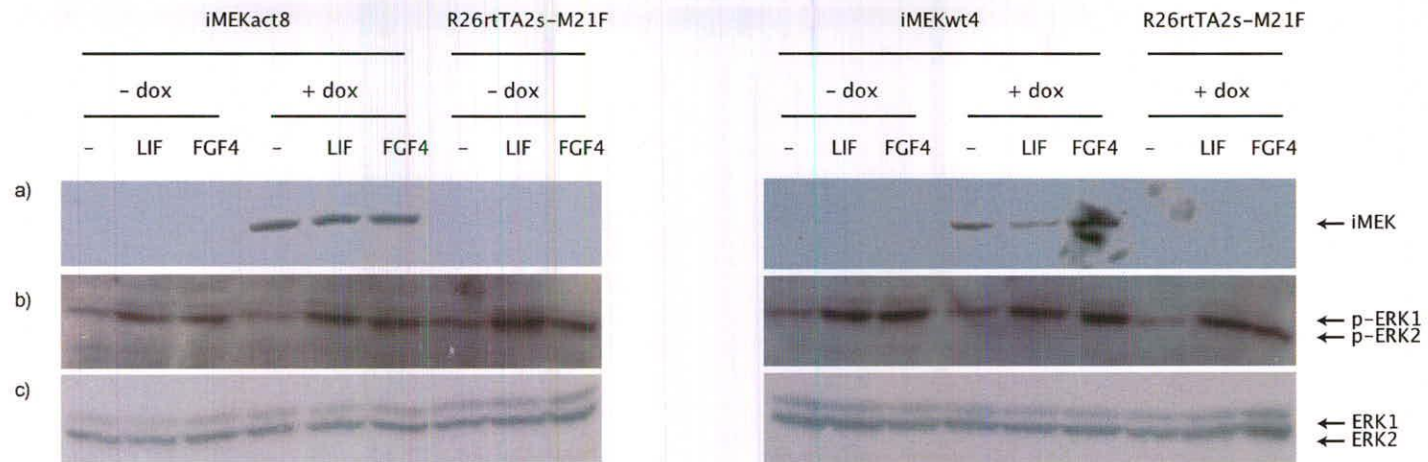


Figure 5.10 Immunoblotting of targeted cell lines. iMEKact 8, iMEKwt4 and R26rtTA2s-M21F cells were plated out at the same density and cultured in the presence or absence of 2  $\mu\text{g/ml}$  dox for 72 hours. After 60 hours all cells were serum starved overnight (with dox added to the appropriate cells). At 72 hours cells were either left untreated or stimulated with LIF or FGF4 for 15 minutes, after which time they were washed once with cold PBS and then lysed in SDS lysis buffer. Samples were sonicated and one fifth of the total lysate was run on a 10% Tris-HCl gels and transferred onto PVDF membrane overnight. Membranes were then blocked and subsequently probed with antibody to a) HA, b) phospho-ERK1/2 (p-ERK1/2) or c) ERK1/2. After washing, the membranes were probed with secondary antibody and protein levels were detected by chemiluminescent assay. iMEK; inducible MEK

cultured in the presence of dox (iMEKact8 + dox) would exhibit higher basal levels of phospho-ERK1/2 in comparison to those exhibited by iMEKwt4 (also cultured in the presence of dox - iMEKwt4 + dox) owing to the action of activated MEK. Cells were also stimulated with LIF and FGF4 (which activates the MAPK/ERK pathway through cell surface receptors), before lysis to investigate whether p-ERK levels were elevated in the cell line overexpressing MEKact when compared to the MEKwt expressing line. This would indicate increased activation of the pathway by MEKact.

Both cell lines and in addition, the parental R26-rtTA2<sup>s</sup>-M21F line (this line acted as a control for the possibility of dox addition leading to ERK1/2 phosphorylation), were cultured in the presence or absence of 2 µg/ml dox for three days, in standard self-renewal conditions. Approximately 12 hours before lysis, medium lacking serum and LIF was added to the cell cultures (dox was also added to those cells that had previously been cultured in the presence of dox). The next morning cells were either left untreated or stimulated for 15 minutes with medium containing LIF or FGF4, after which time they were lysed, sonicated and then analysed by immunoblotting with antibodies to HA, p-ERK1/2 and ERK1/2 (Figure 5.10).

Both iMEKact8 and iMEKwt4 were shown to express exogenous MEK, (MEKact and MEKwt respectively), in the presence of dox as determined by probing with antibody to the HA epitope (Figure 5.10a). No inducible MEK was detected in the absence of dox, which supports the flow cytometry analysis that showed lack of DsRED2 fluorescence in the absence of dox. In the absence of LIF or FGF4 stimulation, it was found that basal p-ERK1/2 levels were of a similar magnitude in all cell lines, irrespective of culture in the presence or absence of dox (Figure 5.10b). MEKact is a constitutively active form of MEK which has been previously been shown to phosphorylate and activate ERK1/2 in the absence of cytokine or growth factor stimulation (Pages *et al.*, 1994), therefore the induction of MEKact was predicted to enhance basal p-ERK levels in the absence of LIF and FGF4 stimulation. The observed levels however, were comparable to those detected upon culture in the absence of dox (iMEKact8 – dox), and those seen in the iMEKwt4 cell line cultured in the presence of dox (iMEKwt4 + dox), (it should be noted that

overexpression of MEKwt would not be expected to constitutively activate the MAPK-ERK pathway). These results therefore precluded investigation into the involvement of the MAPK/ERK pathway during monoculture serum-free differentiation, as it was clear that overexpression of MEKact in the iMEKact8 cell line did not lead to increased activation of the pathway in the absence of stimulating conditions.

The stimulation of all cell lines with LIF and FGF4 lead to increased pERK1/2 levels as expected. However expression of MEKact did not lead to elevated pERK1/2 levels when compared to iMEKact8-dox (+LIF or +FGF4), or iMEKwt4 (+LIF or +FGF4). These observations suggest that in stimulating conditions MEKact expression did not lead to increased activation of the MAPK/ERK pathway.

## 5.9 Discussion

### Low expression levels of DsRED2 in iMEKact and iMEKwt cell lines

The previous chapter described the targeting of a TRE and linked DsRED2 reporter to the 5' *hprt* locus in the R26-rtTA2<sup>s</sup>-M21F cell line, and the resulting high levels of DsRED2 fluorescence obtained upon induction of the correctly targeted lines with dox. An identical strategy was used to establish the inducible expression of MEKact and MEKwt proteins in ES cells (again using DsRED2 as a reporter transcribed from a bidirectional TRE), however analysis of the newly targeted cell lines after three days in dox revealed that the intensities of DsRED2 fluorescence observed were much lower than those observed in the previous iDsRED2 cell lines (compare figure 5.3 to figure 4.4). Also the induction was very heterogenous in nature, with a large population of cells negative for DsRED2 fluorescence in every clone analysed in the presence of dox.

To exclude the possibility that the reason for poor induction was due to differences between the two vectors used in the establishment of the iDsRED2 and iMEK cell lines, identical passage R26-rtTA2<sup>s</sup>-M21F cells that were used to establish the iMEK clones were transfected with the 5'*hprt*TREDsRED2R vector. Had the poor expression levels

observed with the iMEK clones been a problem intrinsic to the targeted transgenic sequences, one would have expected the targeting of R26-rtTA2<sup>s</sup>-M21F cells with the 5' hprtTREDsRED2R vector to give rise to discrete, high levels of DsRED2 as described in chapter 4. However the newly targeted iDsRED2new clones all exhibited transgene expression that was reminiscent of those obtained with the iMEK clones, and therefore the hypothesis of poor induction being an issue of transgenic sequence configuration can be discounted.

This suggests that the poor induction observed was due to the R26-rtTA2<sup>s</sup>-M21F cells themselves. One possibility is that the R26-rtTA2<sup>s</sup>-M21F cells used in the electroporations that gave rise to the iMEK clones, expressed lower levels of transactivator and therefore were not able to drive high levels of transgene expression in the presence of dox. Analysis of both cell lines simultaneously by flow cytometry however, uncovered no difference between the hrGFP fluorescence levels of the iDsRED2 and iMEK cell lines (data not shown), and therefore it is unlikely that poor induction levels are the result of low transactivator levels.

An alternative explanation is that the epigenetic state of the 5' *hprt* region is different in the iDsRED and iMEK cell lines. The *hprt* promoter is located within a CpG island that is methylated during X inactivation. E14TG2a cells are male and therefore contain only one (transcriptionally active) X-chromosome, however it is possible that this region of chromatin is prone to transcriptional silencing. If there is a gradual accumulation of methylation with time in cell culture, this region may be expected to be particularly susceptible to silencing as it contains a high proportion of CpG dinucleotides. Indeed it has previously been shown that the transcriptional repression of the human HPRT gene is highly correlated with the methylation of just three specific CpG nucleotides located within the HPRT promoter (Chen *et al.*, 2001). There may therefore be the risk of methylation spreading to the TRE sequence which is targeted to this CpG rich area. It is however difficult to correlate this with the fact that the targeting of 5' hprtTREMEdact into low passage cells (which are unlikely have accumulated many methyl-CpG dinucleotides) still gives rise to the poor induction of the DsRED2 transgene (data not

shown). The reason for the differences in transgene expression therefore still remains to be determined.

### **Dox titration and timecourse experiments**

The four iMEK cell lines cultured in a gradient of 0-2  $\mu\text{g/ml}$  dox exhibited an increase in both the proportion of DsRED2 positive cells and an increase in median DsRED2 fluorescence levels with increasing concentration of dox (Figure 5.8). This parallels the increase in DsRED2 observed in the iDsRED2 cell lines in the previous chapter, however the overall percentages of DsRED2 positive cells and median DsRED2 values were lower for the iMEK cell lines for reasons which have been commented upon above.

In the timecourse experiment the four lines were cultured in 2  $\mu\text{g/ml}$  dox for between 0 and 7 days. It was observed that for most of the cell lines DsRED2 expression (in terms of proportion of DsRED2 positive cells and median fluorescence) peaked at around two to three days and remained relatively constant until about 5 days, after which expression gradually declined (Figure 5.8). The exception to this was the iMEKwt4 cell line which exhibited maximal DsRED2 expression at day 4 after which median fluorescence levels rapidly declined. This pattern of expression of DsRED2 with time is reminiscent to that of the iDsRED2 cell lines. The iMEKwt4 cell line expressed the highest levels of DsRED2 out of the four clones analysed, and this suggests that the reason for the observed decline in transgene levels may be linked to the initial levels of transgene expression, with the highest DsRED2 expressing clones exhibiting the greatest rate of decline in DsRED2 fluorescence. This phenomenon is therefore likely to be due to a toxic effect of DsRED2 expression in ES cells, with the cell lines exhibiting the highest levels of DsRED2 being subject to greater selection pressures for lower DsRED2 levels, and therefore exhibiting the most pronounced reductions in transgene levels. If the decline was in fact due to toxicity related to dox administration as discussed in the previous chapter, one might expect that all cell lines would be subjected to the same selection pressure and therefore exhibit the same decrease in DsRED2 levels with time (as the selection pressure would be to decrease the internal concentration of dox, for

example by selecting for cells which exhibit increased efflux of the antibiotic). As the iMEK cell lines express very low levels to start with this would probably result in the cell lines becoming completely negative for DsRED2 with time. Whilst DsRED2 levels do decline in the iMEK cell lines, fluorescence is still observed, even after 7 days, and therefore it is more likely that the reason for the reduction DsRED2 levels with time, is linked to DsRED2 toxicity rather than to dox toxicity.

### **The influence of MEKact on the MAPK pathway**

The MAPK signalling pathway plays vital roles in cellular proliferation, differentiation and cell cycle regulation of somatic cells. The overexpression of an activated form of MEK has previously been utilised to investigate the involvement of this pathway in these diverse cellular processes, (Cowley *et al.*, 1994; Brunet *et al.*, 1994; Montesano *et al.*, 1999). Cowley *et al* showed that the overexpression of a constitutively activated form of MEK1 was sufficient to both drive the differentiation of PC12 cells and to transform the NIH 3T3 cell lines, and Montesano *et al* reported that overexpression of activated MEK conferred an invasive phenotype upon MDCK epithelial cells.

In an attempt to investigate the involvement of the MAPK pathway in the differentiation of ES cells into the neuroectoderm lineage using serum free adherent monoculture conditions, an activated form of MEK (MEK<sub>act</sub>) was inducibly expressed in ES cells using the tet-system. The aim was to overexpress MEKact during monoculture differentiation at different timepoints in order to determine if activation of the MAPK/ERK pathway led to increased acquisition of a neuroectodermal cell fate. The use of a bidirectional TRE was used to drive the transcription of a DsRED2 reporter in one direction and the *mek* ORF transgene in the other. Upon analysing the expression of the MEKact and MEKwt proteins by immunohistochemistry both proteins were found to be expressed in their respective cell lines upon the addition of dox (Figure 5.9), however in correlation to the flow cytometry analysis of these cell lines, this induction was found to be mosaic. As anticipated the inducible MEK and DsRED2 proteins were shown to be

expressed in the same cells (as observed in the merged panels in figure 5.9), showing that the TRE was indeed functioning in a bidirectional manner.

Investigation into the involvement of the MAPK/ERK pathway during the differentiation of ES cells to neural progenitor cells required that the cell line iMEKact8 overexpressed MEKact to levels high enough to activate the ERK pathway in serum-free conditions (as indicated by ERK1/2 phosphorylation), when induced by dox addition. MEKact expression was induced in the iMEKact8 cell line and levels of p-ERK1/2 were analysed by immunoblot and compared to those detected when iMEKact8 was cultured in the absence of dox and also to the levels produced by the iMEKwt4 cell line cultured in identical conditions. The iMEKact8 and iMEKwt4 ES cell lines were therefore cultured in the presence or absence of dox for three days and analysed for p-ERK1/2 levels by immunoblot. 12 hours before lysis, the cells were treated with medium lacking serum or LIF, in order to obtain basal levels of p-ERK. Just prior to lysis cells were either left untreated, or stimulated with either LIF or FGF4. It was found that even though MEKact was clearly expressed in the presence of dox, this did not result in increased activation of the ERK pathway in either unstimulated or stimulating conditions.

The lack of increased MAPK/ERK signalling may result from insufficient levels of expression of MEK<sub>act</sub>, unable to give rise to significant activation of ERK1/2. MAPK signalling is however highly regulative, and several inhibitors of this pathway have been identified, such as Pyst1/MKP3, a dual specificity phosphatase that acts on MAPK (Groom *et al.*, 1996), and Spred, which acts on Raf, (Wakioka *et al.*, 2001). Upon overexpression of MEK<sub>act</sub> in ES cells it is therefore possible that ERK pathway activation is 'dampened' by the activation of downstream inhibitors of signalling such as Pyst1/MKP3. Indeed studies in chick show that activation of the MAPK pathway in the developing embryo (by the implantation of FGF-soaked beads into extra-embryonic epiblast) induce Pyst1/MKP3 expression after 4 hours (Eblaghie *et al.*, 2003). In addition the electroporation of a plasmid encoding an activated form of MEK into the chick neural tube has been shown to induce the expression of Pyst1/MKP3, (Smith *et al.*, 2005). It currently still remains to be determined if the lack of ERK1/2 activation observed in the

iMEKact8 ES cell line resulted from the upregulation of negative inhibitors of ERK signalling, or if MEKact was expressed at levels insufficient to bring about significant phosphorylation of ERK1/2. However it is clear that higher levels of MEKact are needed in order to activate the MAPK/ERK pathway. Alternative strategies of tet-system implementation proposed to give rise to high levels of inducible cDNA expression will be discussed in the next chapter.

## Concluding notes

The tet-system is currently one of the most popular inducible gene expression systems in use today. The aim of this project was to develop a strategy for implementing the tet-system in ES cells, which would give rise to robust inducible expression of cDNAs to use in the study of pathways involved in the *in vitro* differentiation of these cells. Two strategies have been described. The first involved the targeting of both elements of the system to the *hpri* locus, the transactivator to the 5' proximal region of *hpri*, and the TRE into exon 6 of the open reading frame of *hpri* to create ihpriDsRED2 cell lines. Addition of doxycycline inducer to the established cell lines resulted in the heterogenous expression of the DsRED2 reporter which instigated investigation into alternative strategies.

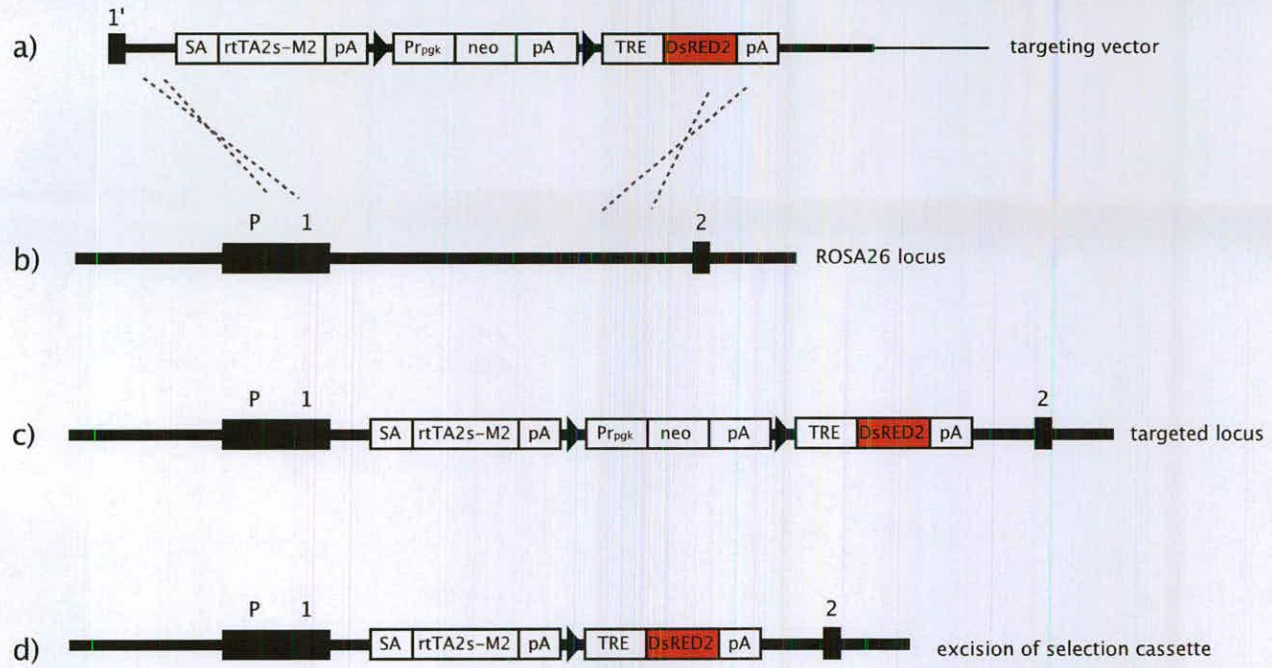
The second strategy involved the use of the previously established R26-rtTA2<sup>s</sup>-M21F cell line. This cell line expressed the rtTA2<sup>s</sup>-M2 transactivator from the endogenous ROSA26 promoter. The targeting of the TRE element to the 5' region of *hpri* gave rise to the iDsRED2 cell lines, which upon addition of dox, expressed the DsRED2 reporter in a high and homogenous manner. Further attempts to inducibly express an activated form of MEK in the same experimental design however, resulted in cell lines that expressed the MEK and DsRED transgenes at low and heterogenous levels. The reasons for the poor expression exhibited by the iMEK cell lines still remains to be determined, however it is most likely a problem of the R26-rtTA2<sup>s</sup>-M21F cell-line, as control experiments

suggest it was not due to differences in vector sequences (see Chapter 5, discussion section).

### **Improved strategies for establishing the tet-system in ES cells**

In order to manipulate transgene expression in a predictable manner, new strategies of implementation need to be investigated. Independent work in which the rtTA2<sup>s</sup>-M2 and TRE sequences were targeted to the endogenous ROSA26 locus within a single construct, gave rise to high levels of transgene expression (Heavey and Cunningham, personal communication – Figure 6.1). Although previous studies involving the integration of transactivator and TRE elements encoded within a single element, have reported that high basal levels of transgene expression can be a problem (Hofmann *et al.*, 1996, Markusic *et al.*, 2005), in the ‘single-allele’ approach mentioned above there was no detection of the DsRED2 reporter gene expression in the absence of inducer.

Further adaptation of this strategy into a ‘cassette-exchange’ approach would facilitate the derivation of cell-lines with regulatable transgenes (Figure 6.2). A hypothetical scheme is proposed in Figure 2.1. The first step would involve targeting rtTA2<sup>s</sup>-M2 to the ROSA26 locus in a construct that also encodes a selection cassette (encoding neomycin and HSV thymidine kinase selection transgenes), flanked by heterospecific *loxP* sites, (Figure 6.2a and 6.2b). Correct integrants would be isolated by selection in G418-containing media and subsequent analysis of genomic DNA. This step only needs to be carried out once as the correctly targeted cell line acts as the parental cell line in all future cassette-exchange transfections, as will be described. The exchange vector (Figure 6.2c) would encode the TRE and linked cDNA (and also a reporter gene linked via an IRES if desired). These sequences would also be flanked by heterospecific *loxP* sites. Upon transfection of the exchange vector and a plasmid expressing Cre-recombinase into the rtTA2<sup>s</sup>-M2-expressing parental cell line, site-specific recombination between the homospecific *loxP* sites present in the targeted ROSA26 locus and the exchange vector would result in the exchange of the selection cassette (from the ROSA26 locus), with the TRE and linked sequences (from the exchange vector) as shown in Figure 6.2d. Such an



- loxP site
- crossover event
- 2
- 1'
- P
- SA
- plasmid backbone

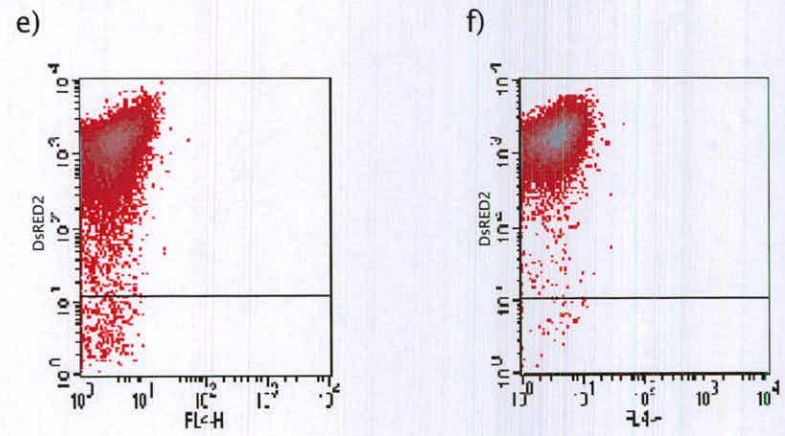


Figure 6.1 The 'single-allele' approach, targeting both the rtTA2s-M2 and TRE in a single construct to the ROSA26 locus (unpublished work done by Heavey and Cunningham). a and b) The targeting construct (here shown in its linearised form) was electroporated into E14TG2a cells c) Correctly targeted clones were isolated by selection in G418-containing media and subsequent analysis by Southern blot hybridisation. d) Subclone of the above, derived by transfection with a plasmid expressing Cre-recombinase. Clones with excised selection cassettes were determined by Southern blot hybridisation. e) Addition of 1 µg/ml dox for two days to clones containing the selection cassette (as shown in c) resulted in high and homogenous expression of the DsRED2 transgene. f) Addition of 1 µg/ml dox for two days to clones in which the selection cassette had been excised (as shown in d) resulted in even higher transgene expression. The y-axis represents DsRED2 fluorescence with the x-axis redundant. Not drawn to scale.

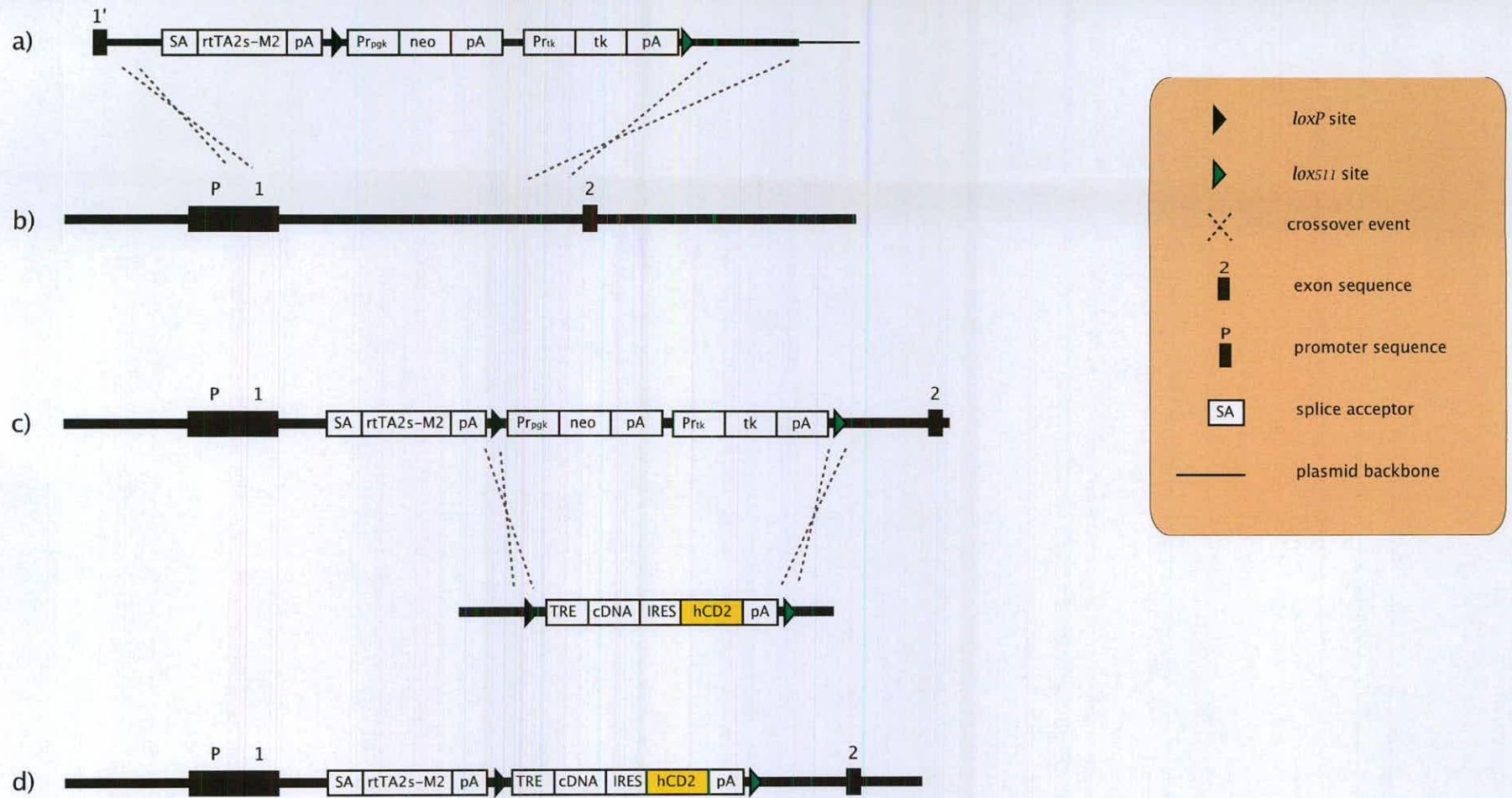


Figure 6.2 Proposed implementation of 'cassette-exchange' strategy. a) The rtTA2s-M2 encoding construct is targeted to the endogenous ROSA26 locus shown in b. c) Upon isolation of correctly targeted clones (by selection in G418-containing media and analysis by Southern blot hybridisation) cells are transfected with the exchange vector encoding a TRE sequence linked to the desired cDNA. A reporter (in this case human CD2 antigen) is linked to the cDNA sequence via an IRES. A plasmid that encodes Cre-recombinase is also transfected. d) Site-specific-recombination between the two loxP sites and two lox511 sites results in exchange of the targeted selection cassette with the TRE encoding sequences (here only the ROSA26 locus is shown). Clones undergoing this exchange can be selected in gancyclovir (as an exchange event results in the removal of the HSVtk ORF).

exchange event could be selected for using ganciclovir (Mansour *et al.*, 1988), as any cells harbouring 'un-exchanged' loci will not survive this drug owing to the presence of the HSV *tk* transgene (Figure 6.2d).

The construction of an exchange vector with a multiple cloning site downstream of the TRE sequence would facilitate the efficient cloning of cDNAs into the vector. This is in contrast to the attempted integration of cDNA sequences into large stretches of vector sequence, as would be the problem if all elements were encoded within the same vector. Multiple cell-lines containing conditional genes could be derived with relative ease once the initial parental rtTA<sup>s</sup>-M2-expressing cell line had been established (Figure 6.2a and 6.2b). Figure 6.2 depicts the use of a reporter gene downstream of the TRE and cDNA. As expression of DsRED2 at high levels was found to be slightly toxic to ES cells here the human CD2 marker is shown as an alternative.

A similar strategy of cassette-mediated exchange at the ROSA26 locus has previously been reported (Masui *et al.*, 2005). It was observed however that after about five passages, transgene induction started to decline. This may result from intolerance of the cells for the reverse transactivator as the original rtTA was used. Such a rapid decline in DsRED2 was not seen with the aforementioned single-allele approach (Heavey and Cunningham, personal communication).

### **Experimental approaches to investigating the involvement of the MAPK/ERK signalling pathway in the differentiation of ES cells into neural progenitor cells**

The utilisation of the 46C ES cell-line that harbours an eGFP reporter sequence under the regulation of the endogenous *Sox1* promoter would greatly aid the investigation of pathways involved in neural induction (Ying *et al.*, 2003). *Sox1* is the earliest known marker of neuroectoderm and the proportion 46C ES cells that acquire this cell fate during differentiation can be readily analysed by eGFP fluorescence. The establishment of the tet-exchange system in this cell line would therefore provide a quantitative and qualitative read-out of neural induction. In addition to the overexpression of MEK<sub>act</sub>, the

tet-regulated expression of other factors involved in the activation, or inhibition of the MAPK/ERK pathway would provide complementary information about the involvement of ERK signalling during the *in vitro* differentiation of ES cells. The production of such inducible cell lines could be efficiently established using the tet-exchange approach. The MAPK phosphatases (MKPs) which negatively regulate the pathway by dephosphorylating ERK1 and 2, would be interesting candidates. Indeed it would be possible to determine if the block in ES cell differentiation in serum-free monolayer conditions, resulting from blockade of FGFR signalling (Ying *et al.*, 2003), was due entirely to inhibition of the MAPK/ERK pathway or if other candidate signalling pathways such as Phosphatidylinositol-3-OH (PI3K) and Phospholipase C (PLC) mediated pathways play a role. If the overexpression of MKPs did result in the inhibition of neuroectoderm specification these induced cells could then be analysed to determine if they remained in a pluripotent ES cell state or if they were restricted in developmental potential.

An advantage to establishing the inducible regulation of these cDNAs in ES cells is that the engineered cell lines can be used to investigate the MAPK/ERK pathway in multiple lineages by using alternative *in vitro* differentiation protocols. Indeed ERK activation has also been shown to be important for the induction of mesoderm (Li *et al.*, 2003). The tet-regulated cell lines could also be used to establish conditional expression of cDNAs *in vivo*. It should be noted however that the restriction of transgene expression to a particular tissue would require the re-establishment of the tet-exchange system under a tissue-specific promoter (as the ROSA26 promoter described above is ubiquitously expressed *in vivo*).

In conclusion, the work undertaken in this project describes two strategies that can be used to achieve inducible gene expression in ES cells using the tet-system, with the later strategy (utilising both the ROSA26 and *hprt* loci), giving rise to superior expression of the inducible DsRED2 transgene. The approaches outlined above highlight possible ways in which tet-inducible expression of transgenes can be further improved, both in terms of reproducibility in expression and in ease of inducible cDNA establishment.

## **Materials and Methods**

### **7.1.1 Molecular Biology Materials**

All chemicals were of analytical grade and supplied by Sigma-Aldrich Co, BDH Laboratory Supplies or Fisher Scientific UK Limited. All bacterial media constituents were supplied by DIFCO laboratories and electrophoresis grade agarose was supplied by Invitrogen Life Technologies. Synthetic oligonucleotides were supplied by Eurogentec EGT Group. Radioisotopes were supplied by Amersham Biosciences and X-ray film by Eastman Kodak Company. Positively charged nylon membrane was supplied by Roche Diagnostics GmbH and Sequi-Blot™ PVDF membrane by Bio-Rad Laboratories. Superscript™ First-Strand Synthesis System for RT-PCR was supplied by Invitrogen Life Technologies and the Dual Luciferase® Reporter Assay System by Promega UK. RNeasy® mini kits, Plasmid Maxi kits, Qiaprep® Spin Miniprep kits and Qiaquick® Gel Extraction kits were all supplied by Qiagen.

Clone Manager Professional Suite 7, version 7.11 was used for DNA analysis.

### 7.1.2 Enzymes

All enzymes were obtained from New England Biolabs except for the following: Klenow polymerase, Taq polymerase, and T4 DNA ligase from Roche Diagnostics GmbH and Alkaline Phosphate from Boehringer Mannheim GmbH.

### 7.1.3 Plasmids

Commercially available

*pBI-L* - BD Biosciences, Clontech,

*pIRES-hrGFP-1a* - Stratagene

Constructed/provided

Please refer to appendix 1

### 7.1.4 Probes for Southern blot hybridisation

| Probe                    | Plasmid                   | Digest              | Fragment size (bp) |
|--------------------------|---------------------------|---------------------|--------------------|
| CMV-rtTA2s-M2-IRESihrGFP | pCMVrtTA2s-M2-IRESihrGFP  | <i>NotI/BamHI</i>   | 3431               |
| CAG-rtTA2s-M2-IRESihrGFP | pCAG-rtTA2s-M2-IRESihrGFP | <i>SpeI/BamHI</i>   | 4380               |
| a                        | M19                       | <i>SpeI/BglII</i>   | 543                |
| DsRED2                   | pCMVTREDSRED2             | <i>NotI/BamHI</i>   | 697                |
| MEK                      | pECE-HA-MEK               | <i>HindIII/XbaI</i> | 1814               |

### 7.1.5 Primers

| Primer                         | Sequence              |
|--------------------------------|-----------------------|
| rtTA2 <sup>S</sup> -M2 forward | agactggacaagagcaaaagt |
| rtTA2 <sup>S</sup> -M2 reverse | gcatgtcaagggtcaaaatcg |
| HA-MEK1 *                      | ggcgtcaaagcaacatggca  |
| HA-MEK2 *                      | tgtgcataatcttgtgettc  |

\* sequencing primers

### 7.1.6 Bacterial strains

*XL-1 blue* - Stratagene

*DH5 $\alpha$*  - Invitrogen Life Technologies

*Stbl2<sup>TM</sup>* - Invitrogen Life Technologies

### 7.1.7 Bacterial growth media

*Luria Broth (LB)*; 5 LB pellets (DIFCO) added to 1 litre dH<sub>2</sub>O and autoclaved

*Luria Broth based agar plates*; 15 g agar (DIFCO) added to 1 litre LB

### 7.1.8 Antibiotics

| Antibiotic   | Stock concentration          | Working concentration |
|--------------|------------------------------|-----------------------|
| Ampicillin   | 50 mg/ml in H <sub>2</sub> O | 50 $\mu$ g/ml         |
| Kanamycin    | 25 mg/ml in H <sub>2</sub> O | 25 $\mu$ g/ml         |
| Tetracycline | 6.25 mg/ml in 100% EtOH      | 12.5 $\mu$ g/ml       |
| Doxycycline  | 1 mg/ml in H <sub>2</sub> O  | 1 $\mu$ g/ml *        |

\* For use as inducer of transgene expression using the tet-system (all other antibiotics were used for bacterial selection during cloning procedures)

### 7.1.9 Standard stock solutions and buffers

All solutions were prepared with dH<sub>2</sub>O

*TE Buffer*; 10 mM Tris.HCl (pH 8.0), 0.1 mM EDTA (pH 8.0)

*20x SSC (per 4 litres)*; 701.3 g NaCl, 352 g tri-sodium citrate, adjusted to pH 7.0 with HCl

*PBS (per 2 litres)*; 14.7 g NaCl, 4.7 g Na<sub>2</sub>HPO<sub>4</sub>; 2.6 g NaH<sub>2</sub>PO<sub>4</sub>.2H<sub>2</sub>O, adjusted to pH 7.1

*50x TAE (per 1 litre)*; 242 g Tris base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA (pH 8.0).

*DEPC-treated dH<sub>2</sub>O for RNA preparations*; 0.1% DEPC in dH<sub>2</sub>O. This solution was prepared by overnight incubation at 37°C before being autoclaved, and then allowed to cool to room temperature before use.

### 7.2.1 Molecular Biology Methods

Unless stated otherwise all centrifugations were carried out in a Heraeus Sepatech benchtop centrifuge

#### 7.2.2 Preperation of nucleic acids

*Small scale preparation of plasmid DNA*; A single colony was inoculated into 2 ml of LB broth supplemented with 50 µg/ml ampicillin. The culture was incubated at 37°C overnight, with shaking. The next day the culture was centrifuged at 13000 rpm for 3 minutes to form a pellet. The supernatant was then removed and the plasmid DNA harvested using the Qiaprep® Spin miniprep system as according to the manufacturers instructions.

*Large scale preparation of plasmid DNA;* A single colony was inoculated into 2 ml of LB broth supplemented with 50 µg/ml ampicillin for 6-8 hours. This culture was then used to inoculate 250 ml LB supplemented with 50 µg/ml ampicillin. This culture was incubated at 37°C overnight, with shaking. The next day the culture was centrifuged at 6000 rpm for 15 minutes using a Sorvall® GSA rotor in a Sorvall® RC5C centrifuge. Plasmid DNA was then harvested using the Qiagen Plasmid Maxi Kit as according to the manufacturers instructions.

*Preparation of RNA;* For RNA extraction  $5 \times 10^6$  cells were washed twice in PBS after which they were incubated with trypsin (see tissue culture materials and methods) and then spun at 1200 rpm for 5 minutes in a MSE Mistral 1000 centrifuge. The cells were then lysed and RNA extracted using the RNeasy® mini kit as according to manufacturers instructions.

### **7.2.3 General cloning methods**

*Nucleic acid digestion by restriction enzymes;* 5 units of the appropriate restriction enzyme was used to digest 1 µg of DNA at the optimal temperature for enzyme action for 1 hour. Reaction mixes included the DNA to be digested, restriction enzyme, appropriate buffer (supplied with the enzyme), dH<sub>2</sub>O and BSA when needed.

*Blunting of DNA fragment ends;* When necessary 3'recessed ends were made flush by incubation with NEB klenow polymerase. 5 µl of 500 µM dNTPs and 1 µl of (2U/µl) klenow polymerase was added to previously digested DNA (that had not been precipitated and therefore still contained restriction enzyme buffer, the volume was 50µl), and incubated for 5 minutes at room temperature. The reaction was stopped by addition of EDTA to a final concentration of 5 mM and heating the mix at 75°C for 10 minutes.

*Dephosphorylation of linearised vectors;* An equivalent volume of TE was added to previously digested DNA (that had not been precipitated and therefore still contained

restriction enzyme buffer, the volume was 50 $\mu$ l). Then 1  $\mu$ l of alkaline phosphatase was added and the reaction mix was incubated at 37°C for 15 minutes. The reaction was stopped by addition of EDTA to a final concentration of 5 mM and heating the mix at 75°C for 10 minutes.

*Phenol/chloroform extraction of reaction mixes;* Phenol/chloroform extraction was routinely used to remove protein from reaction mixes in preparation for ligation reactions. An equal volume of phenol/chloroform mix was added the reaction mix (with the phenol/chloroform at a ratio of 1:1), for example 50  $\mu$ l of phenol and 50  $\mu$ l of chloroform was added to a 100  $\mu$ l reaction mix. The mix was then vortexed and centrifuged at 13000 rpm for 5 minutes. The upper aqueous phase was then removed into a fresh tube.

*Ethanol precipitation of nucleic acids;* DNA was precipitated from solution by addition of one-tenth volume 3 M sodium acetate (pH 8.0) and 2 volumes of ice-cold 100% ethanol. The nucleic acid was then incubated at -20°C for 20 minutes after which it was centrifuged for 10 minutes at 13000 rpm. The supernatant was discarded and 100  $\mu$ l of 70% ethanol was added to the pellet and the tube inverted several times. The pellet was centrifuged for a further 2 minutes at 13000 rpm after which the ethanol was removed and the pellet air-dried for 10 minutes. The DNA was then resuspended in an appropriate volume of TE buffer.

*Gel purification of DNA fragments;* Fragments of digested DNA were resolved by agarose gel electrophoresis. Gels were visualised under UV light (340nm wavelength), and the appropriate band excised with a scapel blade. DNA was extracted from the gel slice using Qiaquick® Gel Extraction Kit as according to the manufacturers instructions.

*Ligation of DNA;* Ligations were carried out in 20  $\mu$ l reaction volumes using three fold excess of insert to vector. Appropriate amounts of vector and insert were incubated

with 2  $\mu$ l T4 DNA ligase and 2  $\mu$ l of T4 ligase buffer at overnight at 16°C (for a blunt-ended ligation), or for 2 hours at 16°C (for the ligation of fragments with cohesive ends).

#### **7.2.4 Preparation of chemically-competent cells**

10 ml LB broth supplemented with 12.5  $\mu$ g/ml tetracycline was inoculated with an *E.coli* colony, and incubated at 37°C with shaking overnight. The next day 250 ml LB broth was inoculated with 5 ml of the overnight culture and incubated at 37°C with shaking. Once the optical density of 0.5 had been achieved (absorbance was measured at 600nm), the culture was divided into six, 50 ml Corning tubes and centrifuged at 4000 rpm in a MSE Centraur 2 centrifuge for 10 minutes at 4°C. The supernatant was then discarded, and 20 mls of chilled 50 mM calcium chloride was added to each tube and the pellet resuspended with a plastic Pasteur. After 15 minutes on ice, the culture was centrifuged at 4000 rpm for 10 minutes at 4°C (MSE Centraur 2 centrifuge). The supernatant was discarded and 4 mls of freezing mix was added to each tube (0.1 M KCl, 0.05 M CaCl<sub>2</sub>, 0.01 M CH<sub>3</sub>COOK, 10% glycerol). The culture was once again resuspended with a plastic Pasteur pipette and aliquoted into 1.5 ml screw capped tubes.

#### **7.2.5 Transformation of chemically-competent XL-1 *Escherichia coli***

Chemically-competent *E.coli* were thawed on ice and dispensed into 100  $\mu$ l aliquots in pre-chilled 14 ml polypropylene Falcon tubes. 10  $\mu$ l ligation mix was then added to 100  $\mu$ l volume of bacterial cells and mixed by flicking the tube. The mix was left on ice for 20 minutes. The bacterial/ligation mix was then incubated at 37°C for 2 minutes after which 1 ml prewarmed (to 37°C) LB broth was added to the tube. After incubation at 37°C with shaking for 45 minutes 50  $\mu$ l of the mix was plated out on LB agar plates containing appropriate selection (for sticky-ended ligation), or all of the mix was centrifuged for 2 minutes at 13000 and resuspended in 100  $\mu$ l LB broth before being plated out (for blunt-ended ligations). The selection plates were then incubated at 37°C overnight.

### **7.2.6 Transformation of chemically-competent recombinant Stbl2<sup>TM</sup> *Escherichia coli***

Chemically-competent *E. coli* were thawed on ice and dispensed into 100 µl aliquots in pre-chilled 14 ml polypropylene Falcon tubes. 10 µl ligation mix was then added to 100 µl volume of bacterial cells and mixed by flicking the tube. The mix was left on ice for 20 minutes. The bacterial/ligation mix was then incubated at 42°C for 30 seconds, followed by 2 minutes on ice. 1 ml of SOC medium was then added to the tube after which the culture was incubated at 32°C for 1 hour with shaking. 50 µl of the mix was plated out on LB agar plates containing appropriate selection (for sticky-ended ligation) or all of the mix was centrifuged for 2 minutes at 13000 and resuspended in 100 µl LB broth before being plated out (for blunt-ended ligations). The selection plates were then incubated at 32°C overnight.

*SOC medium*; 5 g/l Yeast extract, 20 g/l tryptone, 10 mM NaCl<sub>2</sub>, 2.5 mM KCl. This solution is autoclaved after which the following compounds are added, 20 mM Glucose, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>.

### **7.2.7 Screening of bacterial colonies – Colony lifts**

When high incidence of vector religation were observed in ligations involving compatible ends (at both 5' and 3' of the vector and insert fragment ends), or in blunt-ended ligations, the ligation of insert into the vector backbone was screened using colony lifts. An 8.2 cm Nytran nylon membrane (Schleicher & Schuell) was placed on top of the agar plate and orientation marks were made with a needle in both the membrane and agar. The membrane was then placed colony side up on Whatmann 3MM paper, presoaked in 2xSSC, 5% SDS for 5 minutes. The membrane and Whatmann 3MM paper was then microwaved for 2-3 minutes until the colonies had a shiny appearance. The membrane was then placed in 2xSSC to wet and then transferred to the probe mix which contained the appropriate radiolabelled probe. The agar plates were let at room temperature

overnight to allow the colonies to regrow. The next day the membrane was rinsed in 6xSSC, 0.1% SDS at room temperature, followed by two 10 minute washes in 6xSSC, 0.1% SDS, after which they were analysed by autoradiography.

*1 litre of 2x hybridisation solution;* 0.5 g sodium pyrophosphate, 500 ml (20x) SSC, 200 mls (50x) Denhardt's solution and 100 mg yeast tRNA are added to 800 mls dH<sub>2</sub>O. This solution is stirred and then made up to 1 litre after which it is filtered and can be stored at 4°C.

*Probe mix;* 50 mls (2x) hybridisation solution is added to 50mls formamide. The radiolabelled probe is then added.

*Denhardt's solution (50x);* 5 g Ficoll, 5 g polyvinylpyrrolidone, 5 g BSA (pentax fraction V) is added to 450 ml dH<sub>2</sub>O, stirred and then made up to 500 mls. This is then filtered, aliquoted into 50 ml Falcon tubes and stored at 4°C.

### **7.2.8 Lysis of ES cells, extraction and digestion of genomic DNA**

500 µl cell lysis buffer was added to each confluent well of a six well plate and incubated overnight at 37°C for complete lysis to take place. The next day the lysate for each well was transferred to 1.5ml eppendorfs and 500 µl isopropanol was added to each tube. The tubes were inverted several times until the genomic DNA appeared precipitated. The DNA was then removed with a yellow tip and transferred into fresh eppendorf tubes containing 1 ml 70% ethanol. These tubes were inverted several times and the precipitated DNA once again transferred to fresh eppendorf tubes containing 1 ml 70% ethanol. The tubes were centrifuged at 13000 for 10 minutes after which the 70% ethanol was removed and the pellets air-dried for 10 minutes. The DNA was then resuspended in 200 µl TE and transferred to 55°C block for to facilitate the resuspension process. Tubes were flicked every hour for 4 hours and then left at room temperature overnight. The next day once the DNA was quantified and digests were set up in 25 µl reactions (5 µg DNA, 2.5 µl (10x) buffer, 2.5 µl (10)x spermidine at 30mM, 2 µl enzyme made up to 25

µl with dH<sub>2</sub>O) and incubated at the appropriate temperature. After 3 hours the reactions were 'spiked' with an additional 2 µl of enzyme and incubated at the appropriate temperature overnight.

*Cell lysis buffer*; 100 mM Tris, 5 mM EDTA, 200 mM NaCl, 2% SDS. At pH 8.5.  
Proteinase K added just before use too 100 µg/ml

### **7.2.9 Electrophoresis of genomic DNA and transfer onto nylon membrane**

The next day the digested genomic fragments were resolved by gel electrophoresis. The digests were run alongside *Hind*III digested lambda DNA which acted as a marker in the future determination of the size of probe hybridised fragments (this marker is not shown in data chapter figures), after which the gel was stained with ethidium bromide for 30 minutes. The genomic DNA within the gel was fragmented with short wave UV light for 45 seconds after which the gel was placed in denaturing solution (87.66 g NaCl, 20 g NaOH made up to 1 litre in dH<sub>2</sub>O), for two 20 minutes washes with rocking. After rinsing the gel in dH<sub>2</sub>O it was then placed in neutralising solution (175.32 g NaCl, 60.55 g Tris, made up to 1 litre after adding HCl to pH 8), for two 30 minute washes.

For transfer of DNA onto Hybond-N+ nylon membrane by capillary action the gel was placed on a wick of Whatmann 3MM paper soaked in 20x SSC with the ends of the wick in a reservoir of 20xSSC. The nylon membrane was then placed onto of the agarose gel followed by 4 sheets of 20xSSC soaked Whatmann 3MM paper. Saran wrap was placed around the edges of the gel to cover all areas of the wick that were not in contact with the gel. A stack of paper towels were then placed on top of the top sheets of soaked 3MM paper followed by a glass plate and a 500 g weight. The transfer assembly was left overnight after which the membrane was removed and baked at 120°C for 1 hour to cross-link the DNA to the membrane.

### **7.2.10 Random-primed <sup>32</sup>P-labelling of oligonucleotides**

<sup>32</sup>P-labelled probes were produced from 25 ng DNA using Rediprime II Random Prime Labelling System (Amersham Biosciences) as according to the manufacturers instructions.

### **7.2.11 Hybridisation of Southern blot membrane**

For the hybridisation of radiolabelled probe to the nylon membrane, the membrane was prehybridised in Church and Gilbert buffer for 1 hour at 65°C. The radiolabelled probe was then boiled for 5 minutes after which it was placed on ice for 5 minutes. It was then added to the membrane in fresh Church and Gilbert buffer and incubated overnight at 65°C. The following day the membrane was rinsed in 2xSSC followed by two 10 minute washes in 2xSSC, 0,1% SDS at 65°C. After another 20 minute wash in 0.5xSSC, 0.1% SDS also at 65°C the membrane was covered in Saran wrap and exposed onto film.

*Church and Gilbert buffer*; For 100 mls, 50mls (1M) anhydrous Na<sub>2</sub>HPO<sub>4</sub> pH 7.2, 35 mls 20% SDS, 2 mls 0.5 M EDTA, 13 mls H<sub>2</sub>O.

### **7.2.12 Reverse transcription**

First-strand cDNA synthesis was performed on 50 ng of total RNA using the SuperScript™ First-Strand Synthesis System for RT-PCR kit (Invitrogen) as according to the manufacturers instructions. An additional DNase step was performed on the total RNA to remove any traces of genomic DNA. The RNase-Free DNase Set kit was used as according to the manufacturers instructions.

### **7.2.13 Polymerase Chain Reaction (PCR)**

Each PCR reaction contained a final concentration of the following; 200 μM dNTP mix, 1 μM of each primer, 100 ng template DNA, appropriate dilution of PCR reaction buffer

(containing  $MgCl_2$ ) and 1 unit of Taq DNA polymerase. For amplification of a 0.5 kb fragment the following program was cycled 25 times: 94°C; 30 seconds,  $T_m$ ; 60 seconds, 72°C 45 seconds, where  $T_m$  is the annealing temperature of the primer set.

#### **7.2.14 Antibody staining of ES cells**

Cell monolayers were washed twice in PBS and then fixed for 20 minutes at room temperature in 4% PFA (4% PFA in PBS). After a further two washes in PBST (PBS, 0.1% Triton X-10), blocking solution (3% goat serum, 1% BSA in PBST) was added for 30 minutes at room temperature. The appropriate dilution of primary antibody was then added to the blocking solution and the cells left overnight at 4°C with rocking. The next day after three, 15 minute washes in PBST, secondary antibody at the appropriate dilution was added to the monolayer in blocking solution. The cells were then wrapped in aluminium foil incubated for 2-3 hours after which they underwent three, 5 minute washes in PBST.

#### **7.2.15 Immunoblotting by Western protocol**

ES cells were lysed in 100  $\mu$ l of SDS sample buffer (for 1 ml; 0.5 ml Laemmli solution, 0.4 ml dH<sub>2</sub>O, 0.1 ml (1 M) DTT, 10  $\mu$ l (0.5 mM) PMSF protease inhibitor). Cells were scraped off the plastic and pipetted into 1.5 ml eppendorf tubes. Genomic DNA was fragmented by sonication and samples boiled for 5 minutes at 100°C. Bromophenol blue was added to each sample and 20  $\mu$ l of protein was resolved on 10% Tris-HCl BioRad gels at 20 mA in SDS running buffer. The protein was then transferred onto PVDF membrane (Roche) using BioRad gel transfer apparatus. PVDF membrane was activated in methanol after which it was soaked in transfer buffer and positioned on top of the gel. 2 sheets of Whatmann 3MM paper were soaked in transfer buffer and placed on top of the PVDF membrane, followed by a transfer buffer soaked sponge. Two more sheets of Whatmann 3MM paper and sponge (all pre-soaked in transfer buffer) were positioned on the other side of the gel and the stack placed in the transfer apparatus such that the

membrane was positioned nearest the anode in relation to the Tris-HCl gel. Transfer was carried out at 30 V overnight at 4°C.

The next day the PVDF membrane was blocked overnight in 10% milk powder in TBST at 4°C. The following day the membrane was placed in 5% milk solution (5% w/v milk powder in TBST) containing the appropriate dilution of primary antibody, this and all subsequent steps were done at room temperature. After incubation for an hour at room temperature the membrane was washed for four, 15 minute washes of TBST before being incubated in secondary ECL<sup>TM</sup> peroxidase labelled antibody. After 1 hour the membrane was once again washed for four, 15 minute washes in TBST followed by incubated for 5 minutes with Supersignal® West Pico Chemiluminescent Substrate. The membrane was then covered in Saran wrap and exposed to film.

*Laemmli solution 2x*; 4% SDS, 20% glycerol, 120 mM Tris (pH 6.8). Can be stored at room temperature.

*10x SDS running buffer*; 250 mM Tris, 1.92 M glycine, 1% SDS in 1 litre dH<sub>2</sub>O.

*Transfer buffer 1 litre*; 1g SDS, 5.8 g Tris, 29 g glycine in 800 mls dH<sub>2</sub>O. 200 ml methanol was added to bring the volume up to 1 litre.

TBST; 100 mM NaCl, 10 mM Tris pH 7.5, 0.1% Tween in dH<sub>2</sub>O.

### **7.2.16 Sequencing**

Automated sequencing was performed by ICAPB sequencing service using the DYEnamic<sup>TM</sup> ET Terminator cycle sequencing kit (Amersham Biosciences).

### **7.3.1 Tissue Culture**

All tissue culture flasks, plates and 10cm diameter Petri dishes were obtained from IWAKI. 50 ml and 15 ml tubes were supplied by Corning Incorporated. 20 ml universal tubes were supplied by Bibby Sterilin Ltd and Cryotube<sup>TM</sup> vials by Nunc. All

centrifugations were carried out in a MSE Mistral 1000 Centrifuge unless stated otherwise.

### 7.3.2 E14TG2a cell line

The ES cells used in all experiments were E14TG2a cells, a feeder-independent cell-line which relies on an exogenous source of leukocyte inhibitory factor (LIF). These cells are karyotypically male and were derived from the E14 cell line by selection for a HPRT negative phenotype (in 6-thioguanine selection) as described in (Hooper *et al.*, 1987).

### 7.3.3 General tissue culture solutions

*ES cell culture medium (per 550 ml)*; The following constituents were added to 500 ml 1x Glasgow Minimum Essential Media (GIBCO); 200 mM glutamine and 100 mM sodium pyruvate (GIBCO), 0.1 M 2-Mercaptoethanol (Sigma), foetal calf serum (GIBCO), 5.5 ml 1X MEM non-essential amino acids (GIBCO), 550 µl LIF (1000x stock – see below)

*Leukocyte inhibitory factor (LIF)*; LIF was prepared by D.Colby at the ISCR by transient expression of human LIF expression plasmids in COS-7 cells using the method described in (Smith *et al.*, 1991). Serial dilutions of the supernatant were tested on ES cells plated in 24-well plates. A 100-fold higher dilution than the minimal dilution required to keep ES cells undifferentiated was typically used. Serum batches were tested for their ability to sustain the growth, differentiation and viability of ES cells grown at clonal density in the presence and absence of LIF.

*Trypsin solution*; 250 mg trypsin (Difco) and 372 mg EDTA disodium salt (Sigma) was dissolved in 1 litre of PBS. 10 ml chicken serum (Flow Labs) was then added and the solution filter sterilised. The trypsin was stored in 20 ml aliquots at  $-20^{\circ}\text{C}$ .

*1% gelatin*; 1 g of gelatin was dissolved in 100 ml H<sub>2</sub>O. The solution was then autoclaved and stored in 20 ml aliquots at 4°C. For a working solution of 0.1%, 10 ml of 1% gelatin was added to 90ml PBS.

*0.1M 2-mercaptoethanol*; 100 ml 2-mercaptoethanol (Sigma) was added to 14.1 ml PBS. This was stored for up to a month at 4°C.

*2x Freezing solution*; 4 ml of DMSO added to 16 ml of cell culture medium and filter sterilised. The solution was stored for up to a month at 4°C.

*Quench*; An equal volume of serum was added to ES cell culture medium

#### **7.4.1 Growth conditions during ES cell culture**

Methods for routine culture of ES cells are based on those described in Smith, *et al.* 1991. Briefly, ES cells were maintained in a humidified incubator (Heraeus) at 7.5 % CO<sub>2</sub> and 37°C. All ES cell manipulations were carried out in a laminar flow sterile hood (ICN, flow). All solutions were warmed in a 37°C water bath prior to use. Tissue culture grade flasks and dishes (Corning) were gelatin coated (0.1 % in PBS, Sigma) and then aspirated before use. ES cell manipulations methods varied depending on whether the cells were being grown in flasks and dishes or in 96 well plates, usually following clonal selection of ES cells. This section describes both methods.

#### **7.4.2 Passage, expansion and freezing of ES cells**

ES cell culture medium was changed every other day and cells were most commonly passaged at a confluence of 80-90% coverage of the culture vessel surface. Once this confluence was reached, cells were washed twice in PBS after which trypsin was added to the monolayer (approximately 1 ml/25cm<sup>2</sup>) and the cells incubated at 37°C for approximately 2-3 minutes. After this time the flask was tapped gently and then inspected to ensure that the cells were completely dislodged from the culture vessel

surface. If the cells were not appropriately dislodged the flask was returned to the incubator. The trypsinisation process did not proceed for longer than 5 minutes.

Once dislodged the cells were made into a single-cell suspension with pipetting and then transferred to a universal. The flask was rinsed with an equivalent volume of ES cell culture medium and this was also transferred to the universal to inhibit the trypsinisation process. The cells were centrifuged at 1200 rpm for 3 minutes, and the cells resuspended in new ES cell culture medium and passaged into a new culture vessel at appropriate ratios (between 1:2 and 1:6), depending on experimental requirements.

For freezing of ES cells, after centrifugation cells were resuspended in an appropriate amount of ES cell culture medium followed by the equivalent volume of 2x freezing solution (2x freezing solution; 20% DMSO ES cell culture medium), after which the solution was aliquoted into 1ml cryovials. For example for a well of a 6 well plate, after trypsinisation and centrifugation cells were resuspended in 500  $\mu$ l of ES cell culture medium. 500  $\mu$ l of 2x freezing solution was then added and cells mixed gently. This cell-suspension was then transferred into a cryovial and place at  $-80^{\circ}\text{C}$ . The next day, (and no later than 2 months) the vial was transferred to liquid nitrogen for long term storage.

#### **7.4.3 ES cell passage, expansion and freezing in 96 well plate format (1:5 passage)**

The cell monolayers in each well of the 96 well plate were washed twice in PBS using a multichannel pipette. 30  $\mu$ l of trypsin was added to each well and the cells incubated as described above. Once the cells were dislodged from the surface of the plate they were made in a single cell suspension by pipetting. 70  $\mu$ l of Quench media (1:1 serum:ES cell culture medium), was then added to each well to inhibit the trypsinisation process. 20  $\mu$ l was then added to a pre-prepared 96 well plate (gelatinised and 180  $\mu$ l ES cell culture medium per well), for 1:5 passage. The remaining 80  $\mu$ l of ES cell suspension was often frozen at this point (addition of 80  $\mu$ l 2x freezing solution and storage at  $-80^{\circ}\text{C}$ ) and

appropriate wells thawed upon determination of correct targeting by Southern blot hybridisation.

#### **7.4.4 Revival of frozen ES cells**

The cells were retrieved from the  $-80^{\circ}\text{C}$  freezer or liquid nitrogen storage and kept on dry ice until the point of thawing. The vial was held in the  $37^{\circ}\text{C}$  waterbath until the liquid began to thaw and then the cells were transferred to a universal and centrifuged at 1000rpm for 3 minutes. The cells were then resuspended in ES cell medium and transferred into an appropriate tissue culture vessel.

#### **7.4.5 Revival of frozen ES clones from a 96 well plate**

The 96 well plate was retrieved from the  $-80^{\circ}\text{C}$  freezer and kept on dry ice until the point of thawing. 100  $\mu\text{l}$  ES cell culture media was added to the wells that were to be thawed. When the cells began to thaw the suspension was transferred into the preprepared well (gelatinised and containing ES cell culture media) of a 24 well plate.

#### **7.4.6 Vector preparation of plasmid DNA for targeting experiments**

75  $\mu\text{g}$  of targeting vector DNA was linearised with the appropriate restriction endonuclease, and the protein subsequently removed from the digest by phenol/chloroform extraction. The DNA was precipitated in one-tenth volume of sodium acetate and two volumes of 100% ethanol for 20 minutes at  $-20^{\circ}\text{C}$ , and then pelleted by centrifugation at 13000 rpm for 10 minutes. 1 ml of 70% ethanol was pipetted into the eppendorf tube after which it was transferred to the laminar flow hood. The 70% ethanol was removed by aspiration and the pellet air-dried for 5 minutes. The linearised DNA was then resuspended in 100  $\mu\text{l}$  PBS.

#### 7.4.7 Transfection of ES cells by electroporation protocol

Cells were grown to confluency in two 150 cm<sup>2</sup> tissue culture flasks. The cells were washed twice in PBS and then trypsinised as described above. After centrifugation at 1000 rpm for 5 minutes cells were pooled and resuspended in a total of 10 mls of PBS and then counted. The volume of cell suspension containing  $1 \times 10^8$  cells was transferred into a fresh universal and pelleted by centrifugation. This pellet was then resuspended in PBS to give a total volume of 500  $\mu$ l. 100  $\mu$ l of PBS containing the resuspended linearised plasmid DNA (see section) was added to the cells and carefully mixed by re-pipetting. The cells were then transferred to a 0.4 cm electroporation cuvette and capped. The cuvette was pulsed in a BioRad Gene Pulsar unit set a 3  $\mu$ F, 0.8 kV. Cells were left for 10 minutes to recover and then transferred to a 20 ml universal containing 10ml ES cell culture media. This cell suspension was then transferred into two preprepared 150 cm<sup>2</sup> flasks (gelatinised and containing 45mls medium). Cells were incubated for 24 hours before clonal selection was added.

#### 7.4.8 HAT selection of targeted E14TG2a ES cells

The targeting of *hprt* correctional vectors into E14TG2a cells yielded HPRT positive clones owing to the reconstruction of a full *hprt* ORF upon correct integration of vector sequences (as they encoded the *hprt* promoter sequence, and exons 1, 2, which are absent in E14TG2a cell owing to a deletion within the 5' region of the gene). HAT selection medium, selects the HPRT positive clones. The day after electroporation appropriate numbers of 10 cm diameter culture dishes were prepared by gelatinisation, and addition of 9 mls of ES cell culture and 200  $\mu$ l (50x) HAT selection. Cells were washed in PBS and trypsinised as described above. After pelleting by centrifugation at 1200 rpm for 5 minutes the cells were pooled and resuspended in 10mls ES culture media, after which live cells were counted. The volume of cell suspension was adjusted to obtain a suspension of  $5 \times 10^6$  cells/ml, repelleting the cells if necessary. Serial dilutions were made with culture media, e.g  $5 \times 10^6$  cells,  $2.5 \times 10^6$  cells,  $1.25 \times 10^6$  cells, and 1 ml aliquots

of each dilution were added to the prepared 10 cm culture dishes. Selection was applied for 10 days before ES cell colonies were picked.

*HAT selection medium*; Effective concentration in ES cells Hypoxanthine 0.1 mM, Aminopterin 0.4  $\mu$ M, Thymidine 16  $\mu$ M

#### **7.4.9 6-TG selection of targeted ES cells**

6-Thioguanine selection was used to select for HPRT negative clones after the targeting of the M42 vector to the 3' region of *hprt* (this targeting event resulted in the disruption of the *hprt* ORF). Cells were passaged when needed for 7 days after electroporation to allow sufficient time for any HPRT protein present before electroporation to be degraded. Cells were then plated out at a clonal density in prepared 10 cm diameter culture dishes containing 6-TG (final concentration 10 mM) in ES cell media. Selection was applied for 6 days before ES cell colonies were picked.

#### **7.4.10 Picking ES cell colonies**

After the clonal expansion of ES cells in appropriate selective media, visible, discrete colonies of cells were observed. To isolate the colonies and expand for DNA analysis and storage, the 10 cm diameter plate was washed once in PBS, after which 5mls of PBS was added. Using a pipette, the colonies were isolated and transferred into single wells of a 96 well plate.

#### **7.4.11 Transient transfection of ES cells by Lipofectamine<sup>TM</sup> protocol**

For all transient transfections  $1 \times 10^6$  cells were washed in PBS, trypsinised and pelleted, after which they were resuspended in 2 mls ES cell culture medium and transferred into a gelatinised well of a 6 well plate and placed in the 37°C incubator. 3  $\mu$ g of DNA (to be transfected) was resuspended in 250  $\mu$ l of ES cell culture media lacking serum and in a separate eppendorf tube 3  $\mu$ l Lipofectamine<sup>TM</sup> 2000 was mixed in another 250  $\mu$ l ES cell

culture media lacking serum. After 5 minutes incubation at room temperature the lipofectamine<sup>TM</sup> 2000 and DNA containing media was mixed by gentle pipetting and incubated for a further 20 minutes at room temperature, after which it was added to the cell culture and mixed by gentle rocking of the plate.

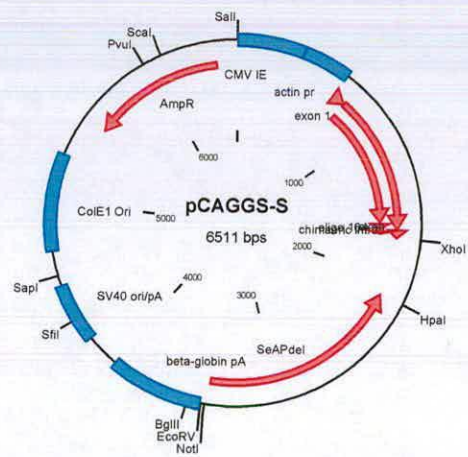
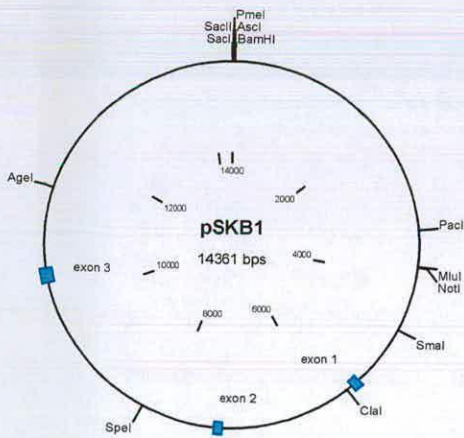
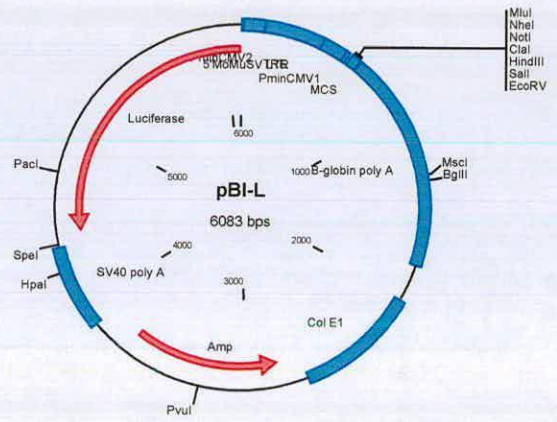
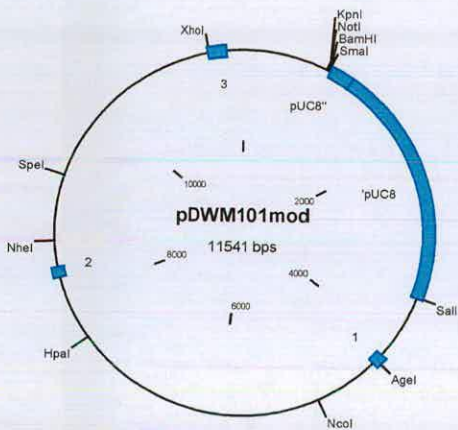
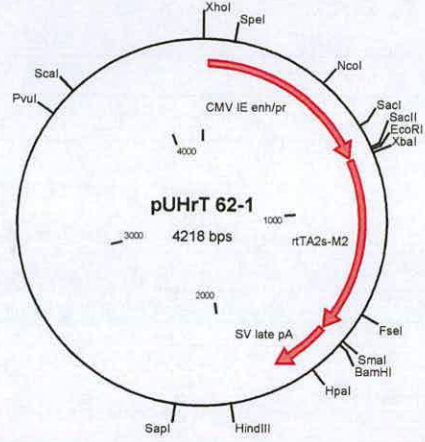
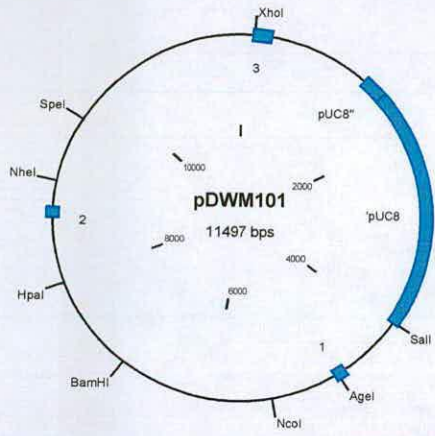
#### **7.4.12 Luciferase analysis**

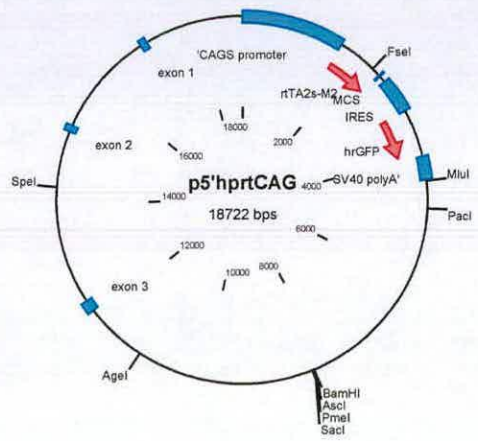
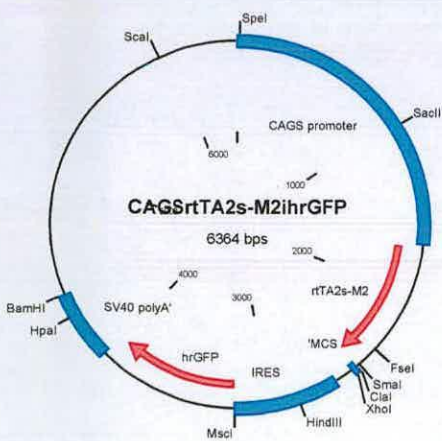
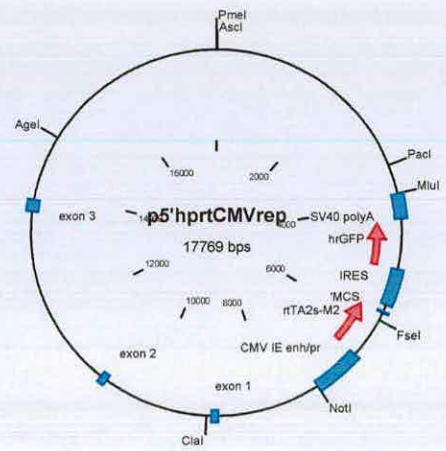
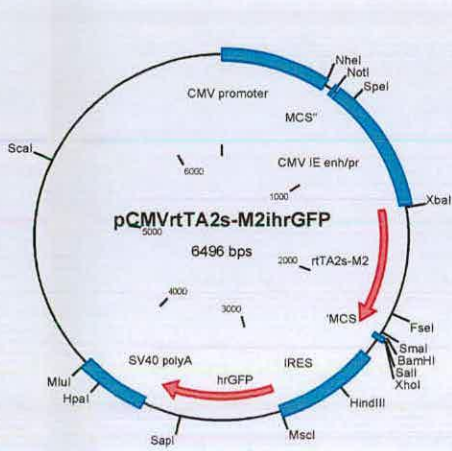
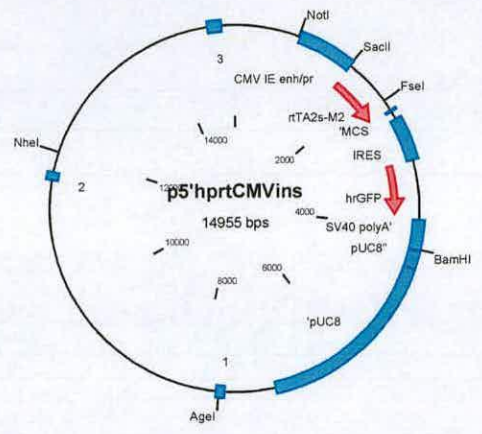
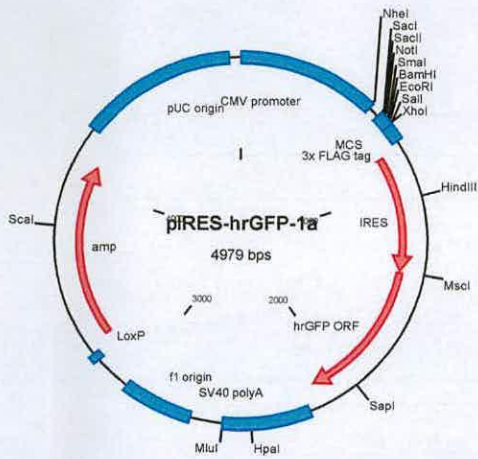
Used in the analysis of the CMVins and CMVrep cell lines. Transient transfections were carried out in the wells of a 6 well plate.  $1 \times 10^6$  cells were cotransfected with pBI-L (reporter plasmid containing TRE-linked firefly-luciferase) and ptk-renilla (encoding renilla luciferase linked to an HSV thymidine kinase promoter – this acted as an internal control for the transfection) using Lipofectamine<sup>TM</sup> 2000 as described above. Two transfections were carried out for each cell line. The next day dox (1  $\mu\text{g}/\text{ml}$ ) was added to one of the wells. 48 hours after the addition of dox cell monolayers were lysed and analysed for firefly luciferase and renilla luciferase using the Dual-Luciferase<sup>®</sup> Reporter Assay System and a plate-reading luminometer (Mediators PhL). Deviations from the manufactures protocol are as follows: cells were lysed in 250  $\mu\text{l}$  of passive lysis buffer instead of the recommended 500  $\mu\text{l}$ , and 50  $\mu\text{l}$  of LARII and Stop & Glo<sup>®</sup> reagent were used instead of 100  $\mu\text{l}$ .

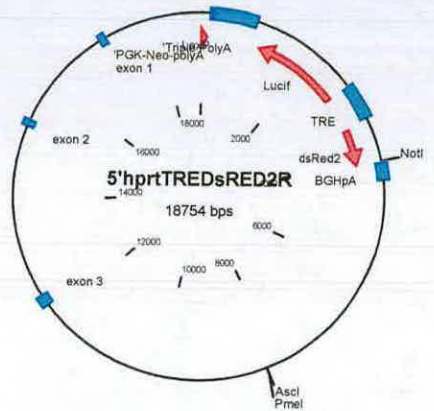
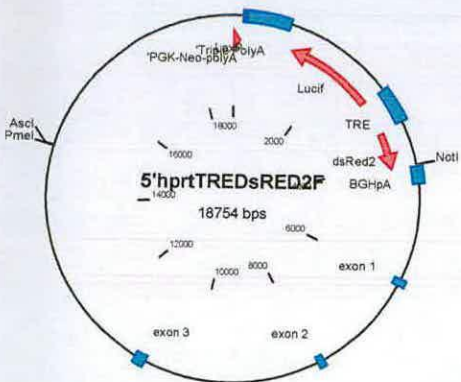
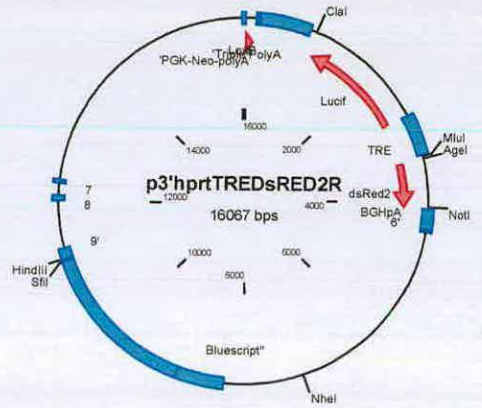
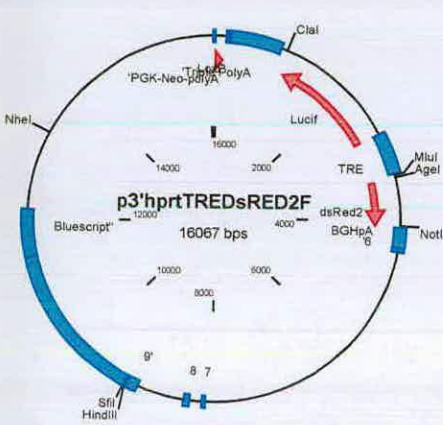
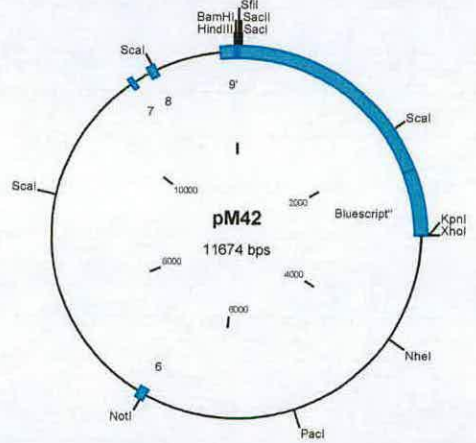
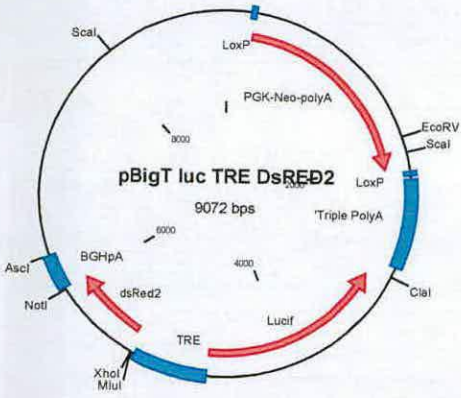
#### **7.4.13 Preperation of cells for immunoblotting**

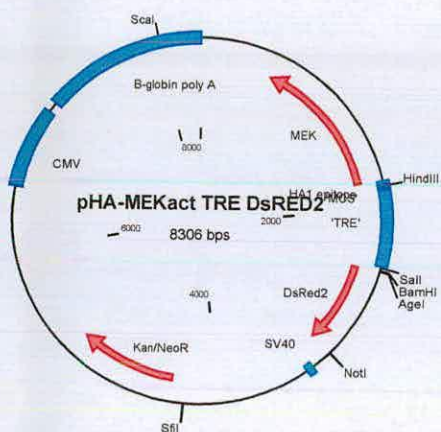
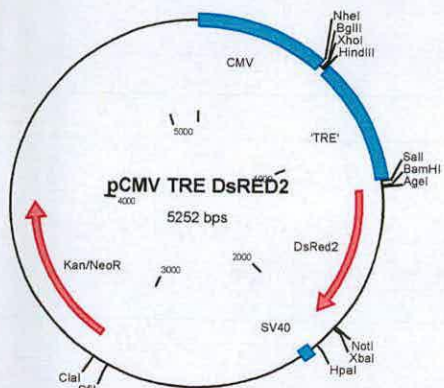
$1 \times 10^6$  cells were plated out into well of a 6 well plate in the presence or absence of 1  $\mu\text{g}/\text{ml}$  dox as according to experimental procedure. After 60 hours the cells were placed in media lacking serum or LIF cytokine (serum starved), and incubated overnight. The next morning cells were either left untreated, or stimulated with LIF cytokine, at a 1:1000 dilution (see above) or FGF4 (recombinant human – R&D systems) to a final concentration of 25  $\text{ng}/\mu\text{l}$  for 15 minutes, after which they were washed once in cold PBS and lysed in 100  $\mu\text{l}$  Laemmli's buffer (see 7.2.15)

# Appendix I









## Bibliography

- Akagi, K., Kanai, M., Saya, H., Koze, T. and Berns, A.** (2001). A novel tetracycline-dependent transactivator with E2F4 transcriptional activation domain. *Nucleic Acids Res* **29**, E23.
- Albanese, C., Reutens, A. T., Bouzahzah, B., Fu, M., D'Amico, M., Link, T., Nicholson, R., Depinho, R. A. and Pestell, R. G.** (2000). Sustained mammary gland-directed, ponasterone A-inducible expression in transgenic mice. *Faseb J* **14**, 877-84.
- Barberis, L., Wary, K. K., Fiucci, G., Liu, F., Hirsch, E., Brancaccio, M., Altruda, F., Tarone, G. and Giancotti, F. G.** (2000). Distinct roles of the adaptor protein Shc and focal adhesion kinase in integrin signaling to ERK. *J Biol Chem* **275**, 36532-40.
- Baron, U., Gossen, M. and Bujard, H.** (1997). Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucleic Acids Res* **25**, 2723-9.
- Baron, U., Schnappinger, D., Helbl, V., Gossen, M., Hillen, W. and Bujard, H.** (1999). Generation of conditional mutants in higher eukaryotes by switching between the expression of two genes. *Proc Natl Acad Sci U S A* **96**, 1013-8.
- Belcheva, M. M. and Coscia, C. J.** (2002). Diversity of G protein-coupled receptor signaling pathways to ERK/MAP kinase. *Neurosignals* **11**, 34-44.
- Belteki, G., Haigh, J., Kabacs, N., Haigh, K., Sison, K., Costantini, F., Whitsett, J., Quaggin, S. E. and Nagy, A.** (2005). Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction. *Nucleic Acids Res* **33**, e51.
- Berger, S. L., Pina, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L., Triezenberg, S. J. and Guarente, L.** (1992). Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* **70**, 251-65.
- Blanpain, C., Lowry, W. E., Geoghegan, A., Polak, L. and Fuchs, E.** (2004). Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* **118**, 635-48.
- Blau, H. M. and Rossi, F. M.** (1999). Tet B or not tet B: advances in tetracycline-inducible gene expression. *Proc Natl Acad Sci U S A* **96**, 797-9.
- Blume-Jensen, P. and Hunter, T.** (2001). Oncogenic kinase signalling. *Nature* **411**, 355-65.

- Bond, C. T., Sprengel, R., Bissonnette, J. M., Kaufmann, W. A., Pribnow, D., Neelands, T., Storck, T., Baetscher, M., Jerecic, J., Maylie, J. et al.** (2000). Respiration and parturition affected by conditional overexpression of the Ca<sup>2+</sup>-activated K<sup>+</sup> channel subunit, SK3. *Science* **289**, 1942-6.
- Brinster, R. L., Chen, H. Y., Trumbauer, M., Senear, A. W., Warren, R. and Palmiter, R. D.** (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* **27**, 223-31.
- Bronson, S. K., Plaehn, E. G., Kluckman, K. D., Hagaman, J. R., Maeda, N. and Smithies, O.** (1996). Single-copy transgenic mice with chosen-site integration. *Proc Natl Acad Sci U S A* **93**, 9067-72.
- Brunet, A., Pages, G. and Pouyssegur, J.** (1994). Constitutively active mutants of MAP kinase kinase (MEK1) induce growth factor-relaxation and oncogenicity when expressed in fibroblasts. *Oncogene* **9**, 3379-87.
- Burdon, T., Smith, A. and Savatier, P.** (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol* **12**, 432-8.
- Burdon, T., Stracey, C., Chambers, I., Nichols, J. and Smith, A.** (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev Biol* **210**, 30-43.
- Cao, T., Longley, M. A., Wang, X. J. and Roop, D. R.** (2001). An inducible mouse model for epidermolysis bullosa simplex: implications for gene therapy. *J Cell Biol* **152**, 651-6.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. and Smith, A.** (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-55.
- Chambers, I. and Smith, A.** (2004). Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* **23**, 7150-60.
- Chen, C., Yang, M. C. and Yang, T. P.** (2001). Evidence that silencing of the HPRT promoter by DNA methylation is mediated by critical CpG sites. *J Biol Chem* **276**, 320-8.
- Chevalier-Mariette, C., Henry, I., Montfort, L., Capgras, S., Forlani, S., Muschler, J. and Nicolas, J. F.** (2003). CpG content affects gene silencing in mice: evidence from novel transgenes. *Genome Biol* **4**, R53.
- Chin, L., Tam, A., Pomerantz, J., Wong, M., Holash, J., Bardeesy, N., Shen, Q., O'Hagan, R., Pantginis, J., Zhou, H. et al.** (1999). Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468-72.

- Chung, S., Andersson, T., Sonntag, K. C., Bjorklund, L., Isacson, O. and Kim, K. S.** (2002). Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines. *Stem Cells* **20**, 139-45.
- Chung, S., Sonntag, K. C., Andersson, T., Bjorklund, L. M., Park, J. J., Kim, D. W., Kang, U. J., Isacson, O. and Kim, K. S.** (2002). Genetic engineering of mouse embryonic stem cells by Nurr1 enhances differentiation and maturation into dopaminergic neurons. *Eur J Neurosci* **16**, 1829-38.
- Corson, L. B., Yamanaka, Y., Lai, K. M. and Rossant, J.** (2003). Spatial and temporal patterns of ERK signaling during mouse embryogenesis. *Development* **130**, 4527-37.
- Cowley, S., Paterson, H., Kemp, P. and Marshall, C. J.** (1994). Activation of MAP kinase kinase is necessary and sufficient for PC12 differentiation and for transformation of NIH 3T3 cells. *Cell* **77**, 841-52.
- Cronin, C. A., Gluba, W. and Scrable, H.** (2001). The lac operator-repressor system is functional in the mouse. *Genes Dev* **15**, 1506-17.
- Dai, Y., Vaught, T. D., Boone, J., Chen, S. H., Phelps, C. J., Ball, S., Monahan, J. A., Jobst, P. M., McCreath, K. J., Lamborn, A. E. et al.** (2002). Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. *Nat Biotechnol* **20**, 251-5.
- Dancz, C. E., Haraga, A., Portnoy, D. A. and Higgins, D. E.** (2002). Inducible control of virulence gene expression in *Listeria monocytogenes*: temporal requirement of listeriolysin O during intracellular infection. *J Bacteriol* **184**, 5935-45.
- Denning, C., Burl, S., Ainslie, A., Bracken, J., Dinnyes, A., Fletcher, J., King, T., Ritchie, M., Ritchie, W. A., Rollo, M. et al.** (2001). Deletion of the alpha(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep. *Nat Biotechnol* **19**, 559-62.
- Deuschle, U., Meyer, W. K. and Thiesen, H. J.** (1995). Tetracycline-reversible silencing of eukaryotic promoters. *Mol Cell Biol* **15**, 1907-14.
- Eblaghie, M. C., Lunn, J. S., Dickinson, R. J., Munsterberg, A. E., Sanz-Ezquerro, J. J., Farrell, E. R., Mathers, J., Keyse, S. M., Storey, K. and Tickle, C.** (2003). Negative feedback regulation of FGF signaling levels by Pyst1/MKP3 in chick embryos. *Curr Biol* **13**, 1009-18.
- el Marjou, F., Janssen, K. P., Chang, B. H., Li, M., Hindie, V., Chan, L., Louvard, D., Chambon, P., Metzger, D. and Robine, S.** (2004). Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* **39**, 186-93.

- Eszterhas, S. K., Bouhassira, E. E., Martin, D. I. and Fiering, S.** (2002). Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol Cell Biol* **22**, 469-79.
- Evans, M. J. and Kaufman, M. H.** (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-6.
- Felsher, D. W. and Bishop, J. M.** (1999). Reversible tumorigenesis by MYC in hematopoietic lineages. *Mol Cell* **4**, 199-207.
- Freundlieb, S., Schirra-Muller, C. and Bujard, H.** (1999). A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *J Gene Med* **1**, 4-12.
- Furth, P. A., Hennighausen, L., Baker, C., Beatty, B. and Woychick, R.** (1991). The variability in activity of the universally expressed human cytomegalovirus immediate early gene 1 enhancer/promoter in transgenic mice. *Nucleic Acids Res* **19**, 6205-8.
- Furth, P. A., St Onge, L., Boger, H., Gruss, P., Gossen, M., Kistner, A., Bujard, H. and Hennighausen, L.** (1994). Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. *Proc Natl Acad Sci U S A* **91**, 9302-6.
- Gogos, J. A., Osborne, J., Nemes, A., Mendelsohn, M. and Axel, R.** (2000). Genetic ablation and restoration of the olfactory topographic map. *Cell* **103**, 609-20.
- Gossen, M. and Bujard, H.** (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **89**, 5547-51.
- Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W. and Bujard, H.** (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766-9.
- Greger, I. H., Demarchi, F., Giacca, M. and Proudfoot, N. J.** (1998). Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res* **26**, 1294-301.
- Grieshammer, U., Lewandoski, M., Pevette, D., Oppenheim, R. W. and Martin, G. R.** (1998). Muscle-specific cell ablation conditional upon Cre-mediated DNA recombination in transgenic mice leads to massive spinal and cranial motoneuron loss. *Dev Biol* **197**, 234-47.
- Gu, H., Marth, J. D., Orban, P. C., Mossmann, H. and Rajewsky, K.** (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103-6.

- Guo, Z. M., Xu, K., Yue, Y., Huang, B., Deng, X. Y., Zhong, N. Q., Hong, X., Chen, X. G. and Xiao, D.** (2005). Temporal control of Cre recombinase-mediated in vitro DNA recombination by Tet-on gene expression system. *Acta Biochim Biophys Sin (Shanghai)* **37**, 133-8.
- Hamer, D. H. and Leder, P.** (1979). Splicing and the formation of stable RNA. *Cell* **18**, 1299-302.
- Hatada, S., Kuziel, W., Smithies, O. and Maeda, N.** (1999). The influence of chromosomal location on the expression of two transgenes in mice. *J Biol Chem* **274**, 948-55.
- Heidt, A. B. and Black, B. L.** (2005). Transgenic mice that express Cre recombinase under control of a skeletal muscle-specific promoter from *mef2c*. *Genesis* **42**, 28-32.
- Hillen, W., Gatz, C., Altschmied, L., Schollmeier, K. and Meier, I.** (1983). Control of expression of the Tn10-encoded tetracycline resistance genes. Equilibrium and kinetic investigation of the regulatory reactions. *J Mol Biol* **169**, 707-21.
- Hofmann, A., Nolan, G. P. and Blau, H. M.** (1996). Rapid retroviral delivery of tetracycline-inducible genes in a single autoregulatory cassette. *Proc Natl Acad Sci U S A* **93**, 5185-90.
- Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M.** (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**, 292-5.
- Hoppe, U. C., Marban, E. and Johns, D. C.** (2000). Adenovirus-mediated inducible gene expression in vivo by a hybrid ecdysone receptor. *Mol Ther* **1**, 159-64.
- Hu, M. C. and Davidson, N.** (1987). The inducible lac operator-repressor system is functional in mammalian cells. *Cell* **48**, 555-66.
- Hudson, C., Darras, S., Caillol, D., Yasuo, H. and Lemaire, P.** (2003). A conserved role for the MEK signalling pathway in neural tissue specification and posteriorisation in the invertebrate chordate, the ascidian *Ciona intestinalis*. *Development* **130**, 147-59.
- Imai, T., Jiang, M., Chambon, P. and Metzger, D.** (2001). Impaired adipogenesis and lipolysis in the mouse upon selective ablation of the retinoid X receptor alpha mediated by a tamoxifen-inducible chimeric Cre recombinase (Cre-ERT2) in adipocytes. *Proc Natl Acad Sci U S A* **98**, 224-8.
- Imumi, M. and Gilbert, D. M.** (1999). Homogeneous tetracycline-regulatable gene expression in mammalian fibroblasts. *J Cell Biochem* **76**, 280-9.

- Keating, A., Horsfall, W., Hawley, R. G. and Toneguzzo, F.** (1990). Effect of different promoters on expression of genes introduced into hematopoietic and marrow stromal cells by electroporation. *Exp Hematol* **18**, 99-102.
- Keefer, C. L., Baldassarre, H., Keyston, R., Wang, B., Bhatia, B., Bilodeau, A. S., Zhou, J. F., Leduc, M., Downey, B. R., Lazaris, A. et al.** (2001). Generation of dwarf goat (*Capra hircus*) clones following nuclear transfer with transfected and nontransfected fetal fibroblasts and in vitro-matured oocytes. *Biol Reprod* **64**, 849-56.
- Khillan, J. S., Deen, K. C., Yu, S. H., Sweet, R. W., Rosenberg, M. and Westphal, H.** (1988). Gene transactivation mediated by the TAT gene of human immunodeficiency virus in transgenic mice. *Nucleic Acids Res* **16**, 1423-30.
- Kim, G. J. and Nishida, H.** (2001). Role of the FGF and MEK signaling pathway in the ascidian embryo. *Dev Growth Differ* **43**, 521-33.
- Kim, S. H., Moores, J. C., David, D., Respass, J. G., Jolly, D. J. and Friedmann, T.** (1986). The organization of the human HPRT gene. *Nucleic Acids Res* **14**, 3103-18.
- Kirby, S. L., Cook, D. N., Walton, W. and Smithies, O.** (1996). Proliferation of multipotent hematopoietic cells controlled by a truncated erythropoietin receptor transgene. *Proc Natl Acad Sci U S A* **93**, 9402-7.
- Kistner, A., Gossen, M., Zimmermann, F., Jerecic, J., Ullmer, C., Lubbert, H. and Bujard, H.** (1996). Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. *Proc Natl Acad Sci U S A* **93**, 10933-8.
- Klucher, K. M., Gerlach, M. J. and Daley, G. Q.** (1997). A novel method to isolate cells with conditional gene expression using fluorescence activated cell sorting. *Nucleic Acids Res* **25**, 4858-60.
- Kringstein, A. M., Rossi, F. M., Hofmann, A. and Blau, H. M.** (1998). Graded transcriptional response to different concentrations of a single transactivator. *Proc Natl Acad Sci U S A* **95**, 13670-5.
- Krisher, R. L., Gibbons, J. R. and Gwazdauskas, F. C.** (1995). Nuclear transfer in the bovine using microinjected donor embryos: assessment of development and deoxyribonucleic acid detection frequency. *J Dairy Sci* **78**, 1282-8.
- Kyba, M., Perlingeiro, R. C. and Daley, G. Q.** (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29-37.

- Lamartina, S., Silvi, L., Roscilli, G., Casimiro, D., Simon, A. J., Davies, M. E., Shiver, J. W., Rinaudo, C. D., Zampaglione, I., Fattori, E. et al.** (2003). Construction of an rtTA2(s)-m2/tts(kid)-based transcription regulatory switch that displays no basal activity, good inducibility, and high responsiveness to doxycycline in mice and non-human primates. *Mol Ther* **7**, 271-80.
- Lederer, T., Kintrup, M., Takahashi, M., Sum, P. E., Ellestad, G. A. and Hillen, W.** (1996). Tetracycline analogs affecting binding to Tn10-Encoded Tet repressor trigger the same mechanism of induction. *Biochemistry* **35**, 7439-46.
- Leone, D. P., Genoud, S., Atanasoski, S., Grausenburger, R., Berger, P., Metzger, D., Macklin, W. B., Chambon, P. and Suter, U.** (2003). Tamoxifen-inducible glia-specific Cre mice for somatic mutagenesis in oligodendrocytes and Schwann cells. *Mol Cell Neurosci* **22**, 430-40.
- Lewandoski, M., Sun, X. and Martin, G. R.** (2000). Fgf8 signalling from the AER is essential for normal limb development. *Nat Genet* **26**, 460-3.
- Li, X., Chen, Y., Scheele, S., Arman, E., Haffner-Krausz, R., Ekblom, P. and Lonai, P.** (2001). Fibroblast growth factor signaling and basement membrane assembly are connected during epithelial morphogenesis of the embryoid body. *J Cell Biol* **153**, 811-22.
- Magin, T. M., McEwan, C., Milne, M., Pow, A. M., Selfridge, J. and Melton, D. W.** (1992). A position- and orientation-dependent element in the first intron is required for expression of the mouse hprt gene in embryonic stem cells. *Gene* **122**, 289-96.
- Markusic, D., Oude-Elferink, R., Das, A. T., Berkhout, B. and Seppen, J.** (2005). Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promoter. *Nucleic Acids Res* **33**, e63.
- 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 U S A* **78**, 7634-8.
- Masui, S., Shimosato, D., Toyooka, Y., Yagi, R., Takahashi, K. and Niwa, H.** (2005). An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit. *Nucleic Acids Res* **33**, e43.
- Meyer-Ficca, M. L., Meyer, R. G., Kaiser, H., Brack, A. R., Kandolf, R. and Kupper, J. H.** (2004). Comparative analysis of inducible expression systems in transient transfection studies. *Anal Biochem* **334**, 9-19.
- Miyazaki, J., Takaki, S., Araki, K., Tashiro, F., Tominaga, A., Takatsu, K. and Yamamura, K.** (1989). Expression vector system based on the chicken beta-actin promoter directs efficient production of interleukin-5. *Gene* **79**, 269-77.

- Miyazaki, S., Yamato, E. and Miyazaki, J.** (2004). Regulated expression of pdx-1 promotes in vitro differentiation of insulin-producing cells from embryonic stem cells. *Diabetes* **53**, 1030-7.
- Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E. and Hayakawa, T.** (2000). IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Mol Ther* **1**, 376-82.
- Montesano, R., Soriano, J. V., Hosseini, G., Pepper, M. S. and Schramek, H.** (1999). Constitutively active mitogen-activated protein kinase kinase MEK1 disrupts morphogenesis and induces an invasive phenotype in Madin-Darby canine kidney epithelial cells. *Cell Growth Differ* **10**, 317-32.
- Muller, S. R., Sullivan, P. D., Clegg, D. O. and Feinstein, S. C.** (1990). Efficient transfection and expression of heterologous genes in PC12 cells. *DNA Cell Biol* **9**, 221-9.
- Nathwani, A. C., Davidoff, A., Hanawa, H., Zhou, J. F., Vanin, E. F. and Nienhuis, A. W.** (2001). Factors influencing in vivo transduction by recombinant adeno-associated viral vectors expressing the human factor IX cDNA. *Blood* **97**, 1258-65.
- Nguyen, H. G., Yu, G., Makitalo, M., Yang, D., Xie, H. X., Jones, M. R. and Ravid, K.** (2005). Conditional overexpression of transgenes in megakaryocytes and platelets in vivo. *Blood* **106**, 1559-64.
- Niwa, H., Miyazaki, J. and Smith, A. G.** (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-6.
- Niwa, H., Yamamura, K. and Miyazaki, J.** (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193-9.
- No, D., Yao, T. P. and Evans, R. M.** (1996). Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc Natl Acad Sci USA* **93**, 3346-51.
- O'Sullivan, J. M., Tan-Wong, S. M., Morillon, A., Lee, B., Coles, J., Mellor, J. and Proudfoot, N. J.** (2004). Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* **36**, 1014-8.
- Pacheco, T. R., Maxwell, F., Vasiloudes, P. E., Robinson, W. A., Norris, D. A. and Maxwell, I. H.** (2000). Positive tetracycline control of expression of p15INK4B from an Epstein-Barr autonomous plasmid in a human melanoma cell line. *Gene* **242**, 249-56.
- Pages, G., Brunet, A., L'Allemain, G. and Pouyssegur, J.** (1994). Constitutive mutant and putative regulatory serine phosphorylation site of mammalian MAP kinase kinase (MEK1). *Embo J* **13**, 3003-10.

- Pera, E. M., Ikeda, A., Eivers, E. and De Robertis, E. M.** (2003). Integration of IGF, FGF, and anti-BMP signals via Smad1 phosphorylation in neural induction. *Genes Dev* **17**, 3023-8.
- Pierson, T. M., Wang, Y., DeMayo, F. J., Matzuk, M. M., Tsai, S. Y. and Omalley, B. W.** (2000). Regulable expression of inhibin A in wild-type and inhibin alpha null mice. *Mol Endocrinol* **14**, 1075-85.
- Qu, Z., Thottassery, J. V., Van Ginkel, S., Manuvakhova, M., Westbrook, L., Roland-Lazenby, C., Hays, S. and Kern, F. G.** (2004). Homogeneity and long-term stability of tetracycline-regulated gene expression with low basal activity by using the rtTA2S-M2 transactivator and insulator-flanked reporter vectors. *Gene* **327**, 61-73.
- Redfern, C. H., Coward, P., Degtyarev, M. Y., Lee, E. K., Kwa, A. T., Hennighausen, L., Bujard, H., Fishman, G. I. and Conklin, B. R.** (1999). Conditional expression and signaling of a specifically designed Gi-coupled receptor in transgenic mice. *Nat Biotechnol* **17**, 165-9.
- Reid, L. H., Shesely, E. G., Kim, H. S. and Smithies, O.** (1991). Cotransformation and gene targeting in mouse embryonic stem cells. *Mol Cell Biol* **11**, 2769-77.
- Rideout, W. M., 3rd, Wakayama, T., Wutz, A., Eggan, K., Jackson-Grusby, L., Dausman, J., Yanagimachi, R. and Jaenisch, R.** (2000). Generation of mice from wild-type and targeted ES cells by nuclear cloning. *Nat Genet* **24**, 109-10.
- Rossi, F. M., Guicherit, O. M., Spicher, A., Kringstein, A. M., Fatyol, K., Blakely, B. T. and Blau, H. M.** (1998). Tetracycline-regulatable factors with distinct dimerization domains allow reversible growth inhibition by p16. *Nat Genet* **20**, 389-93.
- Rossi, F. M., Kringstein, A. M., Spicher, A., Guicherit, O. M. and Blau, H. M.** (2000). Transcriptional control: rheostat converted to on/off switch. *Mol Cell* **6**, 723-8.
- Saam, J. R. and Gordon, J. I.** (1999). Inducible gene knockouts in the small intestinal and colonic epithelium. *J Biol Chem* **274**, 38071-82.
- Saez, E., Nelson, M. C., Eshelman, B., Banayo, E., Koder, A., Cho, G. J. and Evans, R. M.** (2000). Identification of ligands and coligands for the ecdysone-regulated gene switch. *Proc Natl Acad Sci U S A* **97**, 14512-7.
- Sage, J., Miller, A. L., Perez-Mancera, P. A., Wysocki, J. M. and Jacks, T.** (2003). Acute mutation of retinoblastoma gene function is sufficient for cell cycle re-entry. *Nature* **424**, 223-8.
- Schramke, V. and Allshire, R.** (2003). Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. *Science* **301**, 1069-74.

- Schultze, N., Burki, Y., Lang, Y., Certa, U. and Bluethmann, H.** (1996). Efficient control of gene expression by single step integration of the tetracycline system in transgenic mice. *Nat Biotechnol* **14**, 499-503.
- Schwarze, S. R., Ho, A., Vocero-Akbani, A. and Dowdy, S. F.** (1999). In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science* **285**, 1569-72.
- Schwenk, F., Kuhn, R., Angrand, P. O., Rajewsky, K. and Stewart, A. F.** (1998). Temporally and spatially regulated somatic mutagenesis in mice. *Nucleic Acids Res* **26**, 1427-32.
- Shaw-White, J. R., Denko, N., Albers, L., Doetschman, T. C. and Stringer, J. R.** (1993). Expression of the lacZ gene targeted to the HPRT locus in embryonic stem cells and their derivatives. *Transgenic Res* **2**, 1-13.
- Shigehara, T., Zaragoza, C., Kitiyakara, C., Takahashi, H., Lu, H., Moeller, M., Holzman, L. B. and Kopp, J. B.** (2003). Inducible podocyte-specific gene expression in transgenic mice. *J Am Soc Nephrol* **14**, 1998-2003.
- Shockett, P., Difilippantonio, M., Hellman, N. and Schatz, D. G.** (1995). A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. *Proc Natl Acad Sci U S A* **92**, 6522-6.
- Sonntag, K. C., Simantov, R., Kim, K. S. and Isaacson, O.** (2004). Temporally induced Nurr1 can induce a non-neuronal dopaminergic cell type in embryonic stem cell differentiation. *Eur J Neurosci* **19**, 1141-52.
- Soriano, P.** (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21**, 70-1.
- Streit, A., Berliner, A. J., Papanayotou, C., Sirulnik, A. and Stern, C. D.** (2000). Initiation of neural induction by FGF signalling before gastrulation. *Nature* **406**, 74-8.
- Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. and Melton, D. W.** (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313-21.
- Trumpp, A., Depew, M. J., Rubenstein, J. L., Bishop, J. M. and Martin, G. R.** (1999). Cre-mediated gene inactivation demonstrates that FGF8 is required for cell survival and patterning of the first branchial arch. *Genes Dev* **13**, 3136-48.
- Tsai, S. Y., O'Malley, B. W., DeMayo, F. J., Wang, Y. and Chua, S. S.** (1998). A novel RU486 inducible system for the activation and repression of genes. *Adv Drug Deliv Rev* **30**, 23-31.

**Tsuda, H., Maynard-Currie, C. E., Reid, L. H., Yoshida, T., Edamura, K., Maeda, N., Smithies, O. and Jakobovits, A.** (1997). Inactivation of the mouse HPRT locus by a 203-bp retroposon insertion and a 55-kb gene-targeted deletion: establishment of new HPRT-deficient mouse embryonic stem cell lines. *Genomics* **42**, 413-21.

**Uhm, S. J., Chung, H. M., Kim, C., Shim, H., Kim, N. H., Lee, H. T. and Chung, K. S.** (2000). In vitro development of porcine enucleated oocytes reconstructed by the transfer of porcine fetal fibroblasts and cumulus cells. *Theriogenology* **54**, 559-70.

**Urlinger, S., Baron, U., Thellmann, M., Hasan, M. T., Bujard, H. and Hillen, W.** (2000). Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A* **97**, 7963-8.

**Vallier, L., Mancip, J., Markossian, S., Lukaszewicz, A., Dehay, C., Metzger, D., Chambon, P., Samarut, J. and Savatier, P.** (2001). An efficient system for conditional gene expression in embryonic stem cells and in their in vitro and in vivo differentiated derivatives. *Proc Natl Acad Sci U S A* **98**, 2467-72.

**Vooijs, M., Jonkers, J. and Berns, A.** (2001). A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent. *EMBO Rep* **2**, 292-7.

**Wang, X. J., Liefer, K. M., Tsai, S., O'Malley, B. W. and Roop, D. R.** (1999). Development of gene-switch transgenic mice that inducibly express transforming growth factor beta1 in the epidermis. *Proc Natl Acad Sci U S A* **96**, 8483-8.

**Ward, C. M. and Stern, P. L.** (2002). The human cytomegalovirus immediate-early promoter is transcriptionally active in undifferentiated mouse embryonic stem cells. *Stem Cells* **20**, 472-5.

**Weber, P., Cammas, F., Gerard, C., Metzger, D., Chambon, P., Losson, R. and Mark, M.** (2002). Germ cell expression of the transcriptional co-repressor TIF1beta is required for the maintenance of spermatogenesis in the mouse. *Development* **129**, 2329-37.

**Wolf, D. P., Mitalipov, S. and Norgren, R. B., Jr.** (2001). Nuclear transfer technology in mammalian cloning. *Arch Med Res* **32**, 609-13.

**Wutz, A., Rasmussen, T. P. and Jaenisch, R.** (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* **30**, 167-74.

**Yamamoto, A., Lucas, J. J. and Hen, R.** (2000). Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. *Cell* **101**, 57-66.

**Yarranton, G. T.** (1992). Inducible vectors for expression in mammalian cells. *Curr Opin Biotechnol* **3**, 506-11.

**Ying, Q. L., Stavridis, M., Griffiths, D., Li, M. and Smith, A.** (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* **21**, 183-6.

**Yusa, K., Horie, K., Kondoh, G., Kouno, M., Maeda, Y., Kinoshita, T. and Takeda, J.** (2004). Genome-wide phenotype analysis in ES cells by regulated disruption of Bloom's syndrome gene. *Nature* **429**, 896-9.

**Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., Kerr, W. G. and Soriano, P.** (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A* **94**, 3789-94.

**Zhao, B., Chua, S. S., Burcin, M. M., Reynolds, S. D., Stripp, B. R., Edwards, R. A., Finegold, M. J., Tsai, S. Y. and DeMayo, F. J.** (2001). Phenotypic consequences of lung-specific inducible expression of FGF-3. *Proc Natl Acad Sci U S A* **98**, 5898-903.

**Zhu, Z., Ma, B., Homer, R. J., Zheng, T. and Elias, J. A.** (2001). Use of the tetracycline-controlled transcriptional silencer (tTS) to eliminate transgene leak in inducible overexpression transgenic mice. *J Biol Chem* **276**, 25222-9.