# A NOVEL APPROACH FOR REGULATED TRANSGENE EXPRESSION IN MAMMALIAN CELLS

By

Rachel Brough

A thesis submitted to the University of London in candidature for the degree of Doctor of Philosophy

Gene Targeting Group

MRC Clinical Sciences Centre

Imperial College School of Medicine

2005

# Acknowledgements

I would firstly like to thank my supervisor, Dr Andy Porter, for his patience, guidance and encouragement throughout this project and the invaluable advice he gave both during the PhD and whilst writing this thesis.

I would also like to thank all the other members of the Gene Targeting Group, both past and present, for making it such an enjoyable place to work and for all their advice and support. In particular, many thanks to Adam for encouraging my caffeine addiction over the last three years and organising lab surfing expeditions. Also, thank you Debbie for all the stimulating conversations about flapjacks, dieting and exercise. Thank you Roz, for demonstrating that you can do anything when you put your mind to it; and Chadwick, for providing me with daily entertainment through your ever changing ailments! Thank you, Ke, for making lab life so much fun by making me laugh every day. Finally, I would also like to show appreciation to Rajani, Nadeene, Emma and Sheng for all their help over the last three years.

I am also very grateful to Rolf Sprengel for donating the 'improved' transactivator construct and Adam Carpenter for providing the HT2 cells. In addition, many thanks to Maria Jasin for donating the pDR-NEO and pCMV-I-SCE1 constructs.

Thank you to all my friends at the CSC, especially the other students in the same year as myself (Adam, Fiona, Helen, Liz, Rebecca). In particular, Rebecca and Fiona for providing many mad nights out when we first arrived in London and Liz for friendship, gossip and daily lunch breaks discussing intelligent issues!

Thanks also to all my friends outside the CSC, especially those involved in the Serpentine Running Club, you have all helped make the last few years in London fun and exciting with endless trips away and regular nights out. Thank you also to my university friends who have kept in touch and for making me jealous with stories about their fancy jobs and designer outfits.

Finally, I would like to thank all of my family, in particular my Dad and Justin for their encouragement and generous support (financially and emotionally). I could not have done it without you both.

# Abstract

A novel 'screen and insert' strategy to improve the ease, efficiency and reproducibility with which tightly regulated transgenes can be generated is tested in human cell lines. The strategy involves firstly identifying and characterising by flow cytometry those rare clones whose integration site allows optimal regulation of a reporter gene (encoding the Green Fluorescent Protein, GFP) by the tetracycline (Tet-OFF) inducible system. *Cre*-mediated site-specific recombination between two mutant *loxP* sites (*lox71* and *lox66*) is then used to insert any gene of interest (for testing purposes the luciferase ORF was used) downstream of the tightly regulated promoter in such clones. To establish the 'screen and insert' approach in cell culture the human fibrosarcoma cell line, HT1080, and the telomerase immortalised retinal epithelial cell line, hTERT-RPE1, were used.

The low frequency of *Cre*-mediated insertion (~$1/10^5$ of transfected cells) requires a method of selection for isolating clones and two methods are described. The first (System One) involves insertion of a gene of interest (GOI) which is linked to an IRES-hygromycin cassette to enable selection in hygromycin for the insertion event. A limitation of this strategy, however, is that the GOI must be expressed whilst selection takes place and this may be undesirable for certain GOIs (e.g. when a gene's product is toxic). The second method (System Two) avoids this limitation by taking advantage of *flp*-mediated site-specific excision to delay GOI expression until selection for recombination is complete. The utility of the validated 'screen and insert' approach is demonstrated by use of the *RAD52* gene (involved in homologous recombination and DNA repair) and the *I-SceI* endonuclease as GOIs. The resulting cell lines will be used to characterise the deleterious effects of *RAD52* and *I-SceI* expression on genome stability and cell viability.

# Abbreviations

| | |
|---|---|
| aa | amino-acid |
| ATP | adenosine triphosphate |
| BAC | bacterial artificial chromosome |
| bp | base-pair |
| CMV | cytomegalovirus |
| d2EGFP | destabilised enhanced green fluorescent protein |
| dCTP | deoxycytosine triphosphate |
| DMEM | Dulbecco's modified eagle's medium |
| DMSO | dimethyl sulphoxide |
| DNA | deoxyribonucleic acid |
| Dox | doxycycline |
| ds | double strand |
| DSB | double strand breaks |
| dsRNA | double stranded RNA |
| *E. coli* | *Escherichia coli* |
| EDTA | ethylediaminetra-acetic acid |
| EGFP | enhanced green fluorescent protein |
| ES | embryonic stem |
| FACS | Fluorescence-activated cell sorting |
| FCS | foetal calf serum |
| G418 | gentamicin |
| GFP | green fluorescent protein |
| GOI | gene of interest |
| *gpt* | guanine phosphoribosyltransferase |
| HAT | hypoxanthine, aminopterin and thymidine |
| HR | homologous recombination |
| HRP | horseradish peroxidase |
| htTA | humanised tetracycline transactivator |
| *hygro* | *hygromycin* |
| IPTG | isopropyl-beta-D-thiogalactopyranoside |

| | |
|---|---|
| IRES | internal ribosomal entry site |
| ISRE | interferon stimulated response element |
| Kb | kilobase-pair |
| kDa | kilodalton |
| KI | knock-in |
| KO | knock-out |
| LacO | lac operator |
| LacR | lac repressor |
| LIF | leukaemia inhibitory factor |
| luc | luciferase |
| MCS | multiple cloning site |
| MPA | mycophenolic acid |
| mRNA | messenger ribonucleic acid |
| murA | muristerone A |
| NEAA | non-essential amino acid |
| *neo* | neomycin phosphotransferase |
| NHEJ | non-homologous end joining |
| NLS | nuclear localisation signal |
| nt | nucleotide |
| $OD_{260}$ | optical density at 260nm |
| oligo | oligonucleotide |
| ORF | open reading frame |
| PCR | polymerase chain reaction |
| polyA | poly adenylation |
| *puro* | *puromycine* |
| RNA | ribonucleic acid |
| RNAi | RNA interference |
| rpm | revolutions per minute |
| rtTA | reverse tetracycline transactivator |
| RXR | retinoid X receptor |
| SDS | sodium dodecyl sulphate |
| shRNA | short hairpin RNA |
| siRNA | small interfering RNA |
| ss | single stranded |

| | |
|---|---|
| TAE | tris, acetate, EDTA buffer |
| Tet | tetracycline |
| TetO | tetracycline operator |
| TRE | tetracycline response element |
| tTA | tetracycline transactivator |
| 6-TG | 6-thioguanine |
| 6-TX | 6-thioxanthine |
| USP | ultraspiracle |
| v/v | volume for volume |
| wt | wild-type |
| w/v | weight for volume |
| X | xanthine |
| *zeo* | *zeocin* |

# Table of Contents

# List of Figures

## Chapter One

## Chapter Three

## Chapter Four



## Chapter Five

## Chapter Six

## Chapter Seven

## Chapter Eight

# List of Tables

## Chapter One



## Chapter Two



## Chapter Three



## Chapter Four

## Chapter Five



## Chapter Six



## Chapter Seven



## Chapter Eight

# Chapter One

# Introduction

## 1.1 Gene function studies in the post-genomic era

### 1.1.1 The post-genomic challenge

Successful completion of various genome sequencing projects has made it possible for scientific researchers to identify almost all the genes in several organisms (e.g. *Saccharomyces cerevisiae* [Goffeau *et al.*, 1997], *Caenorhabditis elegans* [Abbot *et al.*, 1998], *Drosophila melanogaster* [Adams *et al.*, 2000], *Homo sapiens* [Lander *et al.*, 2001; Venter *et al.*, 2001] and *Mus musculus* [Gregory et al., 2002]). Assigning a biological function to each of these genes now represents a major task, which will involve both forward (classical) and reverse genetic approaches (Figure 1.1). Consequently, in the post-genomic era there is a growing requirement for efficient technology for genetic analyses. In this thesis, a system is developed to facilitate gene function studies by allowing the simpler generation of tightly regulated transgenes in mammalian cells.

### 1.1.2 A brief history of genetic technology

*Model systems facilitate the study of genetics*

One of the most pronounced biological advances that facilitated gene function studies was the development of model systems and these include bacteria, yeast, *Drosophila*, embryonic stem (ES) cells and somatic cells. Using bacteria as a model system for gene function studies is adequate for examining general traits (such as DNA repair); however their applications are limited by their simple prokaryotic genome. The yeast genome was the first eukaryotic model organism to be fully sequenced and the

functions of more than half of its genes are already known (Aylon and Kupeic, 2004). The rapid progress of gene function studies in the yeast has been greatly facilitated by its small haploid genome and the ease with which genetic modifications can be executed. Also, yeast are very easy to grow and their meiotic onset can be controlled, generating a simple system for mutational phenotypic analysis. Many biological mechanisms have been conserved throughout evolution and numerous gene function studies obtained from yeast are applicable to the homologues found in higher eukaryotic organisms. For studying more complex traits (e.g. development, immunity and brain function), however, a more appropriate system is clearly required. *Drosophila* have provided a more advanced eukaryotic system for gene function studies (e.g. in development), however when researching mammalian specific traits (such as immunity) they cannot be utilised.

Murine embryonic stem (ES) cells are derived from pluripotent, noncommitted cells of the inner cell mass of pre-implantation blastocysts and remain undifferentiated in the appropriate culture conditions (either using feeder cells or by the addition of the leukaemia inhibitory factor (LIF); Pease and Williams, 1990). The advantage of using ES cells as a model system for gene function studies is that they can be used to generate transgenic mice after the desired genetic manipulation has taken place. The generation of transgenic mice, however, can be costly and time consuming and is not always necessary for analysing some biological functions.

Mammalian somatic cell culture has provided a valuable tool for the analysis of simpler cellular phenotypes, without the need for complex animal models, such as the examination of genes involved in DNA repair (Yanez and Porter, 2002), cell cycle (Ji *et al.*, 2005), immortalisation (Guerrero *et al.*, 1984), transcription (Cheng *et al.*, 2001), apoptosis (Gil *et al.*, 2004) and many more. The use of human somatic cells is particularly relevant for the study of human gene function, in cases where the function of the mouse or yeast homologue (if there is one) may be different.

*Forward 'classical' genetics for gene function studies*

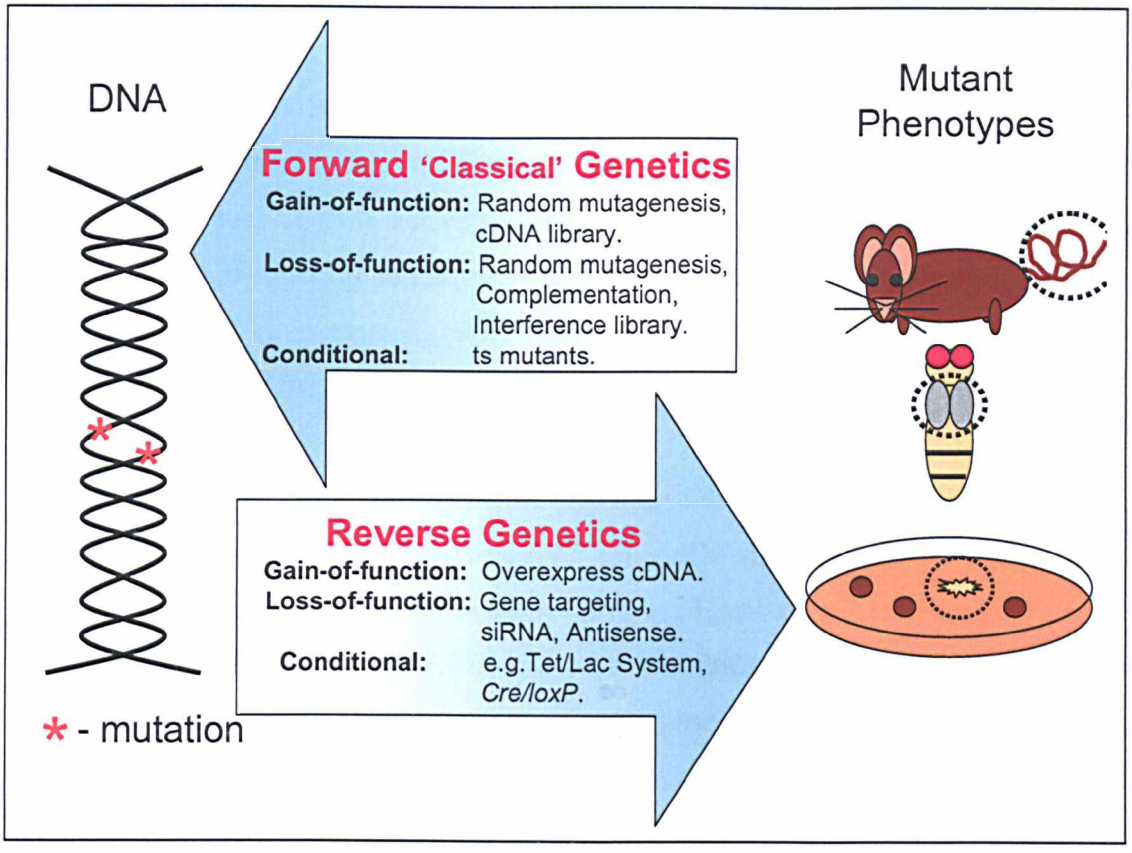



**Figure 1.1 A general outline of typical gene function studies:** The bottom part of the figure illustrates the direction of forward 'classical' genetics which begins with a mutant phenotype of interest and identifies the gene bearing the mutation. The top part of the figure illustrates the direction of reverse genetics which usually begins with a gene of interest which is manipulated in such a way to investigate its phenotype. A number of example methods are listed for achieving gain-of-function, loss-of-function and conditional gene expression for both forward and reverse approaches.

Classical, or forward genetics, involves the mapping and identification of a gene responsible for a particular phenotype (Figure 1.1). Prior to the development of recombinant DNA technology harsh techniques were typically used, such as exposure to X-rays, UV light or chemicals, to induce mutations in the whole genome of cells (most often bacteria, yeast, fungi and *Drosophila*) to generate different mutant phenotypes for analysis (Struhl, 1983). This method does not allow control over

which genes and how many genes are affected by the mutagen and often results in tedious mating procedures and the characterisation of numerous progeny. Desirable mutant phenotypes were then identified by genetic screens (where large populations are visually checked for possible mutant phenotypes) or genetic selection (where only mutant cells survive in otherwise toxic conditions), after which the gene causing the phenotype would be identified. Typically, such methods often result in loss-of-function mutant phenotypes, which describe a mutation that generates a reduction in or a complete loss-of-gene activity when compared to the wild type. Less often, however, gain-of-function mutations are achieved, in which a cell whose mutant phenotype results from extra (hypermorph) or novel (neomorph) gene activity that is not characteristic of the wild type (Griffiths *et al.*, 2005).

A limitation of the random mutagenesis approach is the difficulty in isolating mutants involved in processes that are essential for the cell's viability. For example, when carrying out a screen for genes involved in the control of cell-division cycle (*cdc* mutants) in the yeast strains *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, it was expected that mutations of such genes would compromise the cell's survival. A screen was therefore carried out for mutations that led to the generation of conditional temperature-sensitive *cdc* mutants in which cells demonstrated a wild-type phenotype at normal temperatures and a mutant phenotype at higher temperatures (Jimenez *et al.*, 1990). Such mutants resulted from amino acid changes that lead to deleterious protein shape alterations at elevated temperatures. The mutants generated from this type of screen have enabled researchers to define many of the proteins that regulate the highly programmed progression through the cell cycle. Comparative genomics has shown that the same genes are at work in the cell cycle of humans, and that many of these genes are defective in cancers (Wassmann and Benezra, 2001).

The development of gene cloning techniques, plasmid vectors and transformation procedures in bacteria (*Escherichia coli*) and yeast (*S. cerecisiae* and *S. pombe*) in the 1970s (Struhl *et al.*, 1976) facilitated the development of complementation experiments for the isolation of genes altered in certain recessive mutant phenotypes. This involved transforming a mutant host cell line with wild-type DNA and isolating the plasmid DNA from cells in which the wild type phenotype is restored (Struhl, 1984). In addition, the development of recombinant DNA technology for applications

in mammalian cells, which includes various DNA vectors (e.g. plasmid, BACS and viral) and techniques for their transfection into the host cell's nucleus (e.g. electroporation, lipofection and infection), has produced the gene transfer methods for mammalian gene function studies. Gene function studies in higher eukaryotes, however, are more complicated than in yeast, due to their diploid genomes and DNA alterations in both alleles are often required to generate a mutant phenotype.

Despite the availability of the entire human genome sequence, forward genetic approaches are still adopted today which deploy recombinant DNA techniques for the identification of novel genes involved in certain cellular processes. Genome wide cDNA libraries (for gene overexpression) or RNA interference (RNAi) libraries (for gene silencing) are commonly used to screen for mammalian genes involved in generating a particular phenotype (e.g. cell cycle, senescence (Gil *et al.*, 2005)). After this, the cDNA or RNAi whose effect generated the desired phenotype can be isolated and its genomic location can be mapped using one of many sequence databases now available (e.g. http://www.ncbi.nlm.nih.gov/BLAST/). In addition, gene trap approaches are still commonly used for indiscriminately identifying genes involved in specific cellular processes and typically involves randomly integrating a promoterless reporter gene into the genome to identify genes whose expression is limited, for example, to a particular tissue (e.g. the discovery of *Arkadia* gene, Niederlander *et al.*, 2001; for a review describing further current forward genetic approaches see Branda and Dymecki, 2004). Although classical forward genetics is still being successfully executed today, the development of genome sequencing techniques and recombinant DNA technology has also made possible reverse genetics, the principal approach currently used for gene function studies.

*Reverse genetics for analysing gene function*

Reverse genetic analysis begins with a DNA sequence and then attempts to elucidate its biological function (Figure 1.1), by mutational analysis, gene silencing (loss-of-function) or alternatively by expressing an extra copy of the gene within the cell for overexpression studies (gain-of-function). The development of gene cloning and transformation techniques in the 1970s allowed genetic manipulations (by enzymatic

or chemical treatment) to be performed on specific DNA sequences in the test-tube, rather than involving the entire intact organism. In yeast, the phenotypic consequences of the mutation (typically loss-of-function) could be assessed upon the reintroduction of the DNA into the host organism (Struhl, 1983). This approach was possible by the high efficiency of integration of the mutated DNA into the host cell's genome at their original chromosomal location by a process known as gene targeting (Figure 1.2; Hinnen *et al.,* 1978). Gene targeting utilises the host cell's ability to undergo homologous recombination, which is a cellular process required for accurate DNA repair (discussed further in section 1.1.2). Gain-of-function studies in yeast (and bacteria) did not require the integration of DNA into the genome as gene expression from episomes was typically used.

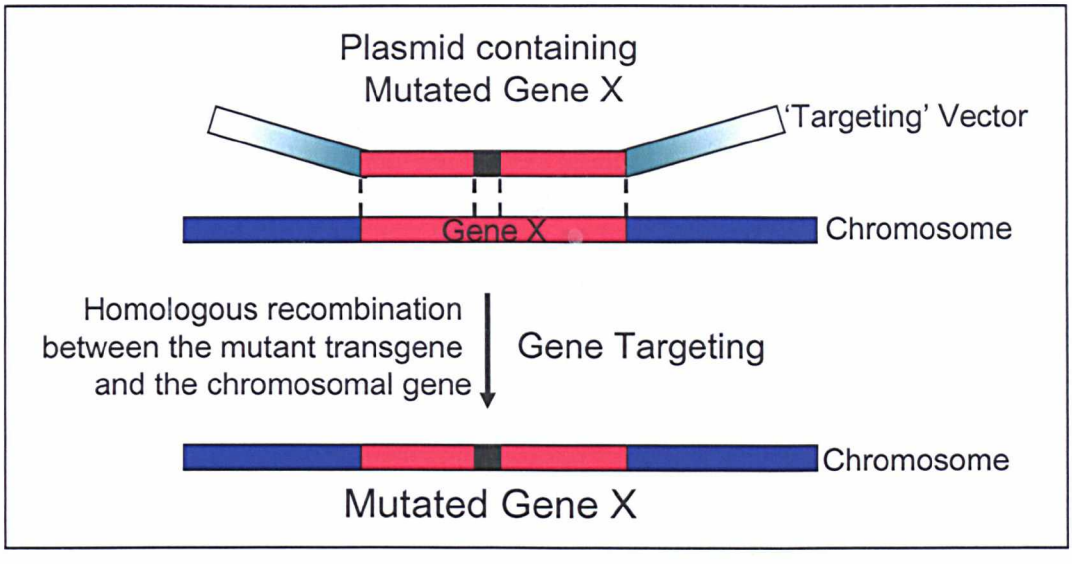



**Figure 1.2 The basic molecular event in gene targeting:** A 'target' vector containing sequences homologous to the target gene but with a mutated segment of DNA is introduced into a cell. Homologous recombination between the 'targeting' vector and the target gene results in a gene replacement reaction, incorporating the mutant segment into the host cells 'wild type' gene.

When similar approaches were carried out using mammalian cells, it soon became clear that DNA rarely integrates into the mammalian genome by homologous recombination, but by a method known as random integration into non-specific, indiscriminate chromosomal sites (Wigler and Axel, 1979). This feature has greatly facilitated gene overexpression studies in mammalian cells to allow the phenotype of gene up-regulation to be analysed (gain-of-function). This method of gene transfer

was greatly facilitated by the development of selection and screening approaches to enrich for DNA integration events. These included the development of drug resistance cassettes such as neomycin and hygromycin (Santerre *et al.*, 1984) and positive-negative marker genes; for instance the guanine phosphoribosyltransferase (*gpt*) gene (Lupton *et al.*, 1991). In addition, the use of reporter genes such as GFP, luciferase and β-galactosidase (lacZ) have generated a means of analysing gene transfer and expression in a highly sensitive and non-destructive manner (Welsh and Kay, 1997).

The reduced homologous recombination efficiency observed for mammalian cells has greatly limited their applications for gene alteration experiments by gene targeting (loss-of-function). As a result, alternative strategies for analysing loss-of-gene function in mammalian cells have been developed and these include the use of antisense oligonucleotides (Pierce *et al.*, 2003) and RNA interference (RNAi), which will be described in more detail later in this chapter (see section 1.1.3). Surprisingly, reduced gene targeting frequencies are not observed for mouse embryonic stem (ES) cells; as they appear to undergo homologous recombination with relatively high efficiency, when compared to other mammalian cells (Arbones *et al.*, 1994). In fact, many knockout mice (1000s) have already been generated for loss-of-gene function analyses.

Both gain-of-function and loss-of-function techniques cannot, however, be applied to those GOIs whose expressions are toxic or essential for the cell's survival. Conditional gene expression systems are therefore necessary for such gene function studies as they allow the control of a gene's expression in a temporal and quantitative manner. The most commonly deployed inducible systems are those generated from regulatory elements of *E. coli* (Tet and Lac Operon; discussed later, see section 1.2.4), as they provide an efficient circuit for expression in mammalian cells, which can also be utilised for conditional gene expression in transgenic animals. The remainder of this Chapter examines the major advances in reverse genetic approaches for studying gene function (both constitutively and conditionally), which have taken place over the last few decades. In addition, the limitations of constitutive gene function studies will be discussed and improved conditional gene expression systems will be considered.

## 1.1.3 Gene targeting approaches for studying gene function

*Standard homologous recombination-based gene alteration strategies*

As mentioned earlier, gene targeting is a technique used to make specific alterations to the genome of living cells (Muller *et al.*, 1999) and involves homologous recombination between a chromosomal 'target' locus and an exogenous DNA molecule, called a targeting vector (Figure 1.2). Gene targeting relies on a cell's potential to carry out efficient homologous recombination, a mechanism utilised by all cells to accurately repair DNA double strand breaks, and it is a process which involves two DNA molecules of similar sequence interacting before undergoing an exchange of genetic information (Smithies *et al.*, 1985). As mentioned above, the principles of gene targeting were established in yeast the 1970s, as these cells demonstrate an elevated frequency of homologous recombination (and therefore gene targeting) when compared to random integration (Struhl, 1983). Random integration describes the integration of the target vector into any chromosomal position other than the 'target' locus. Gene targeting in mammalian somatic cells is limited by a low gene targeting frequency and a high efficiency of random integration (Yanez and Porter, 1998). As a result gene targeting was initially used in mammalian cells to disrupt selectable genes. The hypoxanthine phosphoribosyltransferase (HPRT) gene is an example of such a selectable gene. HPRT codes for an enzyme that catalyses the metabolic salvage of the purine bases hypoxanthine and guanine and inactivation of the HPRT gene can be selected for in 6-thioguanine (Thomas and Capecchi, 1987).

As well as gene function studies, gene targeting also has a potential application in gene therapy, for the accurate correction of inherited gene disorders. Gene targeting is currently the only procedure that can produce specific gene alterations in the genome of eukaryotic cells (Yanez and Porter, 1998). Unlike gene augmentation approaches for gene therapy, which involves the random integration of a functional therapeutic gene (Anderson, 1984), gene targeting can correct both loss-of and gain-of-function mutations. In addition, the gene correction step is permanent and under the control of its own endogenous promoter following homologous recombination. Unfortunately, the risks associated with gene augmentation, which leads to the

generation of additional mutations within the genome upon its integration, still apply to gene targeting experiments due to a high efficiency of random integration. Dramatic improvements in gene targeting efficiencies will, therefore, be required if gene therapy is going to be delivered safely for both *ex vivo* (e.g. somatic stem cells) and *in vivo* (patients) applications.

A number of factors have been shown to influence gene targeting frequencies and these include: an increased length of homology between the target locus and the targeting vector (Hasty *et al.,* 1991; Deng and Capecchi, 1992, Yanez and Porter, 1999b); the use of isogenic DNA (te Riele *et al.,* 1992); and the choice of target locus. The latter is most likely a result of the varying accessibility of the chromatin to the recombination proteins or repair template. Despite slight increases in the targeting frequency in mammalian cells, it still remains low; with approximately 1 targeted clone produced per $10^{-5}$–$10^{-7}$ cells transfected (the absolute targeting frequency), this is compared to random integration where DNA integrates into the genome in 1 cell per $10^{-3}$–$10^{-5}$ cells transfected (Yanez and Porter, 1998). Targeting constructs, therefore, often contain a selection cassette to allow a simpler means of isolating targeted clones. Oliver Smithies was the first to achieve gene targeting of a non-selectable endogenous gene (Beta-globin) in mammalian cells using a gene coding for neomycin resistance (Smithies *et al,* 1985).

Two types of targeting vectors are commonly used: replacement and insertion vectors (Capecchi, 1987; Thomas and Capacchi, 1987; Hasty and Bradley, 1993). Replacement vectors require a double crossover event and act by disrupting the GOI by integration of a positive selection cassette to disrupt its ORF (Figure 1.3A). Replacement vectors are linearised in such a way so that the arms of homology remain collinear with the target sequence and none of the vector sequence is integrated upon gene targeting (Mansour *et al.*, 1988). Insertion vectors are similar to replacement vectors although the positive selectable marker can be located in either the homologous region or the vector backbone (Figure 1.3B). Insertion vectors are linearised within the region of homology and gene targeting results in insertion of the entire construct (including the vector sequences), generating a partial duplication of the targeted gene.

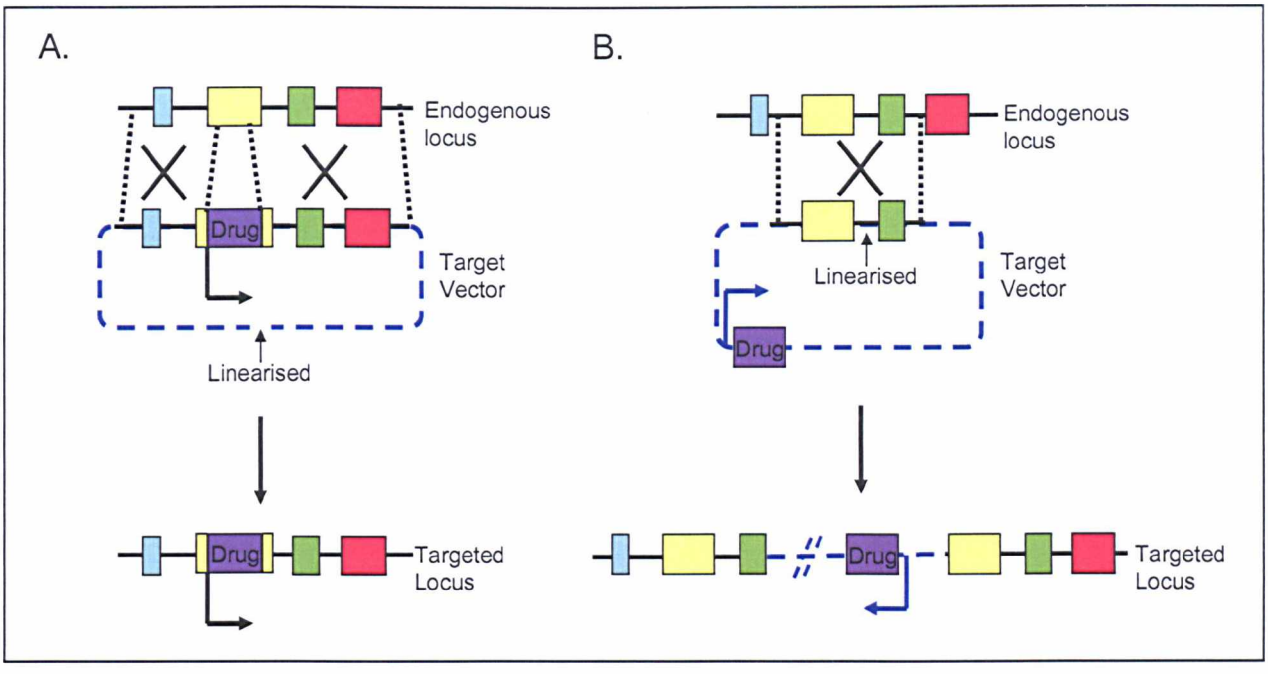



**Figure 1.3 Gene Targeting Vectors: A**. Diagram represents a hypothetical gene containing four exons (coloured boxes) which is targeted by a replacement vector. The targeting event removes the second exon from the hypothetical gene and renders the encoded protein non-functional. Insertion events are selected for in the appropriate drug. **B.** Represents the same hypothetical gene as before, which is targeted by an insertion vector. The targeting event leads to partial duplication of the gene and the function of the gene product is altered. Insertion events are selected for in the appropriate drug.

Gene targeting, using replacement and insertion vector strategies, can appear inefficient in somatic cells as it often requires the laborious screening of numerous clones due to the generation of very few targeted events when compared to random integrants (varies from 1:100 – 1:10,000; Yanez and Porter, 1998). A more efficient type of targeting vector was subsequently developed, called a promoter-trap vector, which consists of a promoterless reporter gene whose activity depends upon targeted insertion downstream of the GOIs transcriptional machinery (Figure 1.4; Sedivy and Sharp, 1989; Friedrich and Soriano, 1991; Itziaki and Porter, 1991). This method reduces the number of random integrants isolated after the selection process and it can only be applied to target genes that are normally expressed in transfected cells.

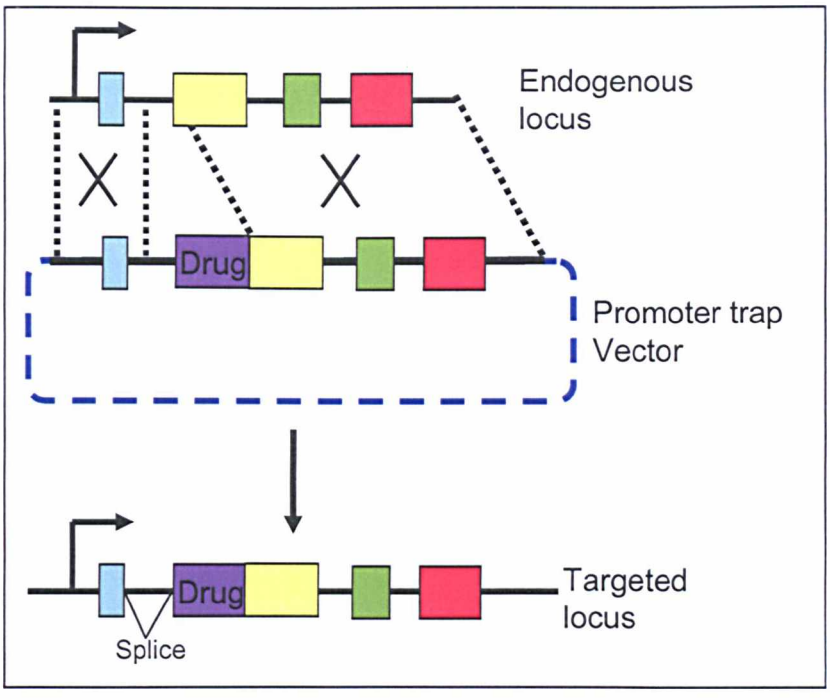



**Figure 1.4 The promoter-trap targeting vector:** The diagram represents the same hypothetical gene as Figure 1.3, which is targeted by a promoter trap vector. The targeting event leads to disruption of the second exon (in yellow) and this event renders the encoded protein non-functional. The expression of the drug resistance cassette is dependant upon gene targeting, downstream of the GOIs transcriptional machinery.

The above methods typically generate cells with a null target allele and therefore do not easily permit subtle sequence alterations (such as point mutations), to allow the effects of a particular disease mutation to be examined, the analysis of the biological function of a particular protein domain or targeted gene correction. It would be desirable to make such alterations to genomic DNA without using or leaving selectable marker cassettes and excess bacterial backbone sequences at the target locus. The reason for this is that these sequences can potentially interfere with expression of neighbouring endogenous genes, which makes it difficult to assign a phenotype to the target gene and is certainly unsuitable for gene therapy applications (Hug *et al.*, 1996; Pham *et al.*, 1996; Rigby, 1997). Adaptations of both replacement and insertion vectors (Figure 1.3A and B) can be used to generate subtle mutations without leftover bacterial DNA; however, two rounds of homologous recombination are required to achieve this (e.g. in-and-out [Hasty *et al.*, 1991] and double replacement strategy [Askew *et al.*, 1993]).

Apart from the relatively high efficiency of homologous recombination in ES cells facilitating the production of numerous knock-out mice, most other applications of gene targeting in mammalian cells (including the analysis of gene function and therapeutic repair) are limited by the low frequency of homologous recombination and understanding how this can be safely increased is currently under examination by many researchers (Vasquez *et al.*, 2001). In addition, understanding why mammalian cells choose to randomly integrate exogenous DNA rather than undergo homologous recombination, like in the yeast and bacterial systems, and how this frequency can be reduced to encourage gene targeting represents an major intellectual and technological challenge.

*Manipulating gene expression as a method for improving gene targeting efficiencies*

The chicken B cell line DT40 is widely used for gene targeting experiments as they exhibit low levels of random integration and therefore screening for targeted events is much simpler than for somatic mammalian cells. The relative targeting frequency (homologous recombination/random integration X 100) of this cell line represents between 10-100% of drug selected clones (Buerstedde and Takeda, 1991), compared to approximately 1% in mammalian cell types (Yanez and Porter, 1998), therefore generating a much simpler system for gene function studies. A comparison of the absolute targeting frequency (homologous recombination/cells transfected X 100) between mammalian somatic cells ($\sim10^{-5}$-$10^{-7}$; Yanez and Porter, 1998) and DT40 cells ($\sim10^{-6}$-$10^{-7}$; Beurstedde and Takeda, 1991) confirmed that DT40s do not exhibit an enhanced homologous recombination capacity but rather a reduced random integration efficiency. More recently, the lymphoid human cell line, DG75, was also shown to have an elevated relative gene targeting frequency approaching 69% of drug selected clones (Feederle *et al.*, 2004).

The existence of these naturally occurring recombinogenic cell lines suggest that it should be possible to manipulate the expression levels of the relevant factors of other cell types to mimic this behaviour transiently during a gene targeting event. A strategy for improving the gene targeting efficiency in somatic cells is the manipulation of expression of genes involved in DNA repair. Candidate genes are

some of those involved in homologous recombination and they are called the *RAD52* epistasis group of genes (including *RAD51* (plus the vertebrate homologues of *RAD51: DCMI, XRCC2, XRCC3, RAD51B, RAD51C* and *RAD51D)*, *RAD52, RAD54, RAD55, RAD57, MRE11, RAD50,* and *XRS2*) (For review see Vasquez *et al.*, 2001; Figure 1.5A).

It is expected that elevated expression of one or a combination of these genes may enhance gene targeting frequencies. For example, stable overexpression of the *RAD51* gene, which forms filaments around single-stranded DNA and catalyses the strand exchange reaction of homologous recombination (Figure 1.5A), in the human fibrosarcoma cell line, HT1080, resulted in a 2-3-fold increase in gene targeting frequencies (Yanez and Porter, 1999a). Gene targeting is reduced by 50-fold in *RAD51B* deficient DT40 cells (Lim and Hastys, 1996); and *XRCC3* deficient hamster cells demonstrate a 25-fold reduction in the homology-directed repair of DNA double-strand breaks (Pierce *et al.*, 1999). The overexpression of these *RAD51* vertebrate homologues are therefore additional potential candidates for increasing gene targeting frequencies.

The *RAD54* gene, which is involved in chromatin remodelling and strand invasion (Figure 1.5A), is also thought to be a potential candidate gene as its disruption in ES cells (Essers *et al.*, 1997) and chicken DT40 cells (Bezzubova *et al.*, 1997) results in a decrease in gene targeting (a 100-fold decrease in DT40 cells). In addition, overexpression of yeast *RAD54* has recently been shown to stimulate gene targeting efficiencies by an average of 27-fold in *Arabidopsis* (Shaked *et al.*, 2005). These results suggest that the *RAD54* gene may be a limiting factor for efficient gene targeting. *RAD52* is also involved in the DNA strand transfer reaction (Figure 1.5A) and has been shown to be essential for homologous recombination in *S. cerevisiae*, therefore suggesting *RAD52* as a candidate gene for elevating gene targeting frequencies in mammalian cells. The generation of *RAD52* null DT40 (Rijkers *et al.*, 1998) and ES (Yamaguchi-Iwai *et al.*, 1998) cells resulted in only a modest 2-fold reduction in homology directed repair. Alternatively, overexpression of the human *RAD52* gene in HT1080 cells actually inhibited gene targeting (Yanez and Porter, 2002), therefore suggesting a loss-of-functional activity of the *RAD52* gene in influencing homologous recombination in vertebrate cells. The latter hypothesis was

further illustrated by the demonstration that gene targeting could be stimulated 37-fold when the yeast *RAD52* gene was overexpressed in HeLa cells (Primio *et al.*, 2005).

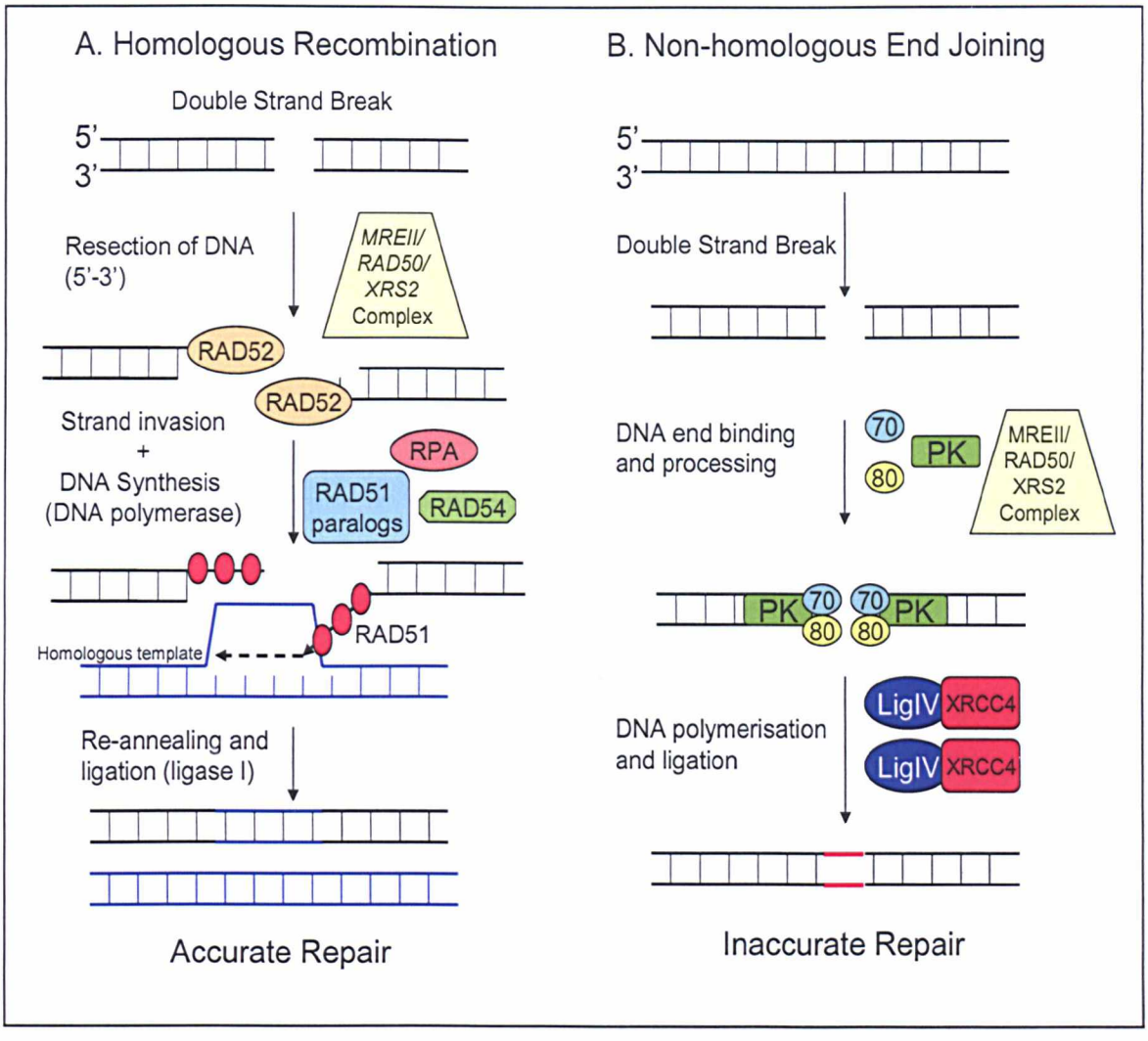



**Figure 1.5 Homologous and non-homologous DNA repair: A.** A representation of DNA repair by homologous recombination (illustrated by yeast homologues). After a double strand break, it is repaired faithfully by the *RAD52* epistasis group of genes. **B.** A representation of DNA repair by non-homologous end-joining (illustrated by the yeast homologues). After a double strand break the DNA is repaired inaccurately. For a more detailed description see Aylon *et al.* (2004).

Conversely, down-regulating genes thought to inhibit homologous recombination or promote non-homologous end joining (a cellular process for inaccurately repairing DNA double strand breaks) are also considered to be potential candidates for increasing gene targeting efficiencies. Genes that have been associated with non-homologous end joining include *BLM, ku70, ku80, XRCC4, RAD50, XRS2* and *Mre11*

(Figure 1.5B). A candidate gene is one found to be mutated in Bloom's syndrome, *BLM* (Ellis *et al.*, 1995). This gene functions as a helicase in eukaryotic cells to suppress holliday junctions (an intermediate of homologous recombination) that develop inappropriately, for example at a stalled replication fork (Karrow *et al.*, 2000). In fact, gene targeting frequencies are elevated by an average of 3-fold in *BLM*-deficient DT40 cells (Wang *et al.*, 2000) and more recently it was shown to have a similar effect in yeast and HT1080 cells (Langston *et al.*, 2005). In yeast, a double down-regulation of both the *BLM* homologue (*SGS1*) and the exonuclease gene (*Exo1*) resulted in a 15-fold stimulation of gene targeting compared to a 4-fold enhancement when each gene was down-regulated independently (Stafa *et al.*, 2005). Similarly, combined down-regulation of yeast *BLM* (*SGS1*) and *ku70* (a helicase that associates at DNA ends for the recruitment of proteins associated with non-homologous end-joining; (Figure 1.5B [Yael and Kupiec, 2004]) resulted in an 8-fold stimulation of gene targeting over wild-type, compared to a 4-fold enhancement when each gene was silenced independently (Yamana *et al.*, 2005). Finally, the use of poly (ADP-ribose) polymerase (PARP) inhibitors, which hampers the non-homologous recombination pathway (Waldman and Waldman, 1990), has been shown to stimulate gene targeting by up to 8-fold in mouse fibroblasts (Semionov *et al.*, 2003).

Although these increases in gene targeting, summarised in Table 1.1, are encouraging they are also very modest. Much larger enhancements of homologous recombination or reductions in the frequency of random integration will be essential to stimulate gene targeting enough so that the desirable (targeted) clones can be isolated using non-selectable conditions. This will facilitate the precise alteration of a gene for specific gene function studies, disease mutation analysis and eventually therapeutic applications. It may simply be a matter of finding the correct combination of proteins to overexpress and down-regulate. It is predicted that permanent stimulation of homologous recombination in mammalian cells is likely to compromise cell viability, as uncontrolled recombination reactions between the many repeat sequences present in the complex genome may occur. A transient method of delivery during a gene targeting experiment would expect to achieve the desired affects without effecting cell survival.

**Table 1.1 The effects manipulating various genes expression on gene targeting efficiencies:** A summary of effects on gene targeting when various genes were either up- or down-regulation. Stimulations in gene targeting were calculated as the fold expression above endogenous levels. HR, homologous recombination; NHR, non-homologous recombination.

| **Overexpression** | | | | |
|---|---|---|---|---|
| **Gene** | **Role** | **Cell Line** | **Stimulation** | **Reference** |
| Human *RAD51* | HR | Human (HT1080) | 2-3 | Yanez & Porter, 1999 |
| Hamster RAD51 | HR | Mouse ES cells | 2-4 | Dominguez-Bendala *et al.*, 2003 |
| Human *RAD52* | HR | Human (HT1080) | 0 | Yanez & Porter, 2002 |
| Yeast *RAD52* | HR | Human (HeLa) | 15-37 | Primio *et al.*, 2005 |
| Yeast *RAD54* | HR | *Arabidopsis* (GV3101) | 27 | Shaked *et al.*, 2005 |
| **Gene Suppression** | | | | |
| **Gene** | **Role** | **Cell Line** | **Stimulation** | **Reference** |
| Yeast *BLM* | NHR | Yeast | 4.4 | Stafa *et al.*, 2005 |
| Yeast *ExoI* | NHR | Yeast | 4.9 | Stafa *et al.*, 2005 |
| Yeast *BLM / ExoI* | NHR | Yeast | 15.7 | Stafa *et al.*, 2005 |
| Yeast *Ku70* | NHR | Yeast | 4 | Yamana *et al.*, 2005 |
| Yeast Ku70 / BLM | NHR | Yeast | 8 | Yamana *et al.*, 2005 |
| PARP | - | Mouse (Ltk) | 8 | Semionov *et al.*, 2003 |

*Nuclease-induced double-strand breaks elevate gene targeting frequencies*

The generation of double strand breaks in the genome of mammalian cells has been shown to stimulate homologous recombination by 1000-fold (Jasin, 1996) and even 5000-fold (Cohen-Tannoudjii *et* al., 1998; Donoho *et al.*, 1998). After a double strand break has been induced, the host cell's recombination machinery exploits a homologous template (e.g. targeting vector) to repair the break faithfully, by DNA

strand exchange (Johnson and Jasin, 2001). The use of rare cutting endonuclease systems in mammalian cells has allowed the affects of double strand breaks on gene targeting and other forms of homologous recombination to be studied. The most commonly used endonuclease is the I-*SceI* protein from the yeast strain *S. cerevisiae* (Jasin, 1996), where it functions to copy specific introns from an intron-containing gene to an intronless version of the same gene, termed intron-encoded homing (Dujon, 1989). I-*SceI* recognises an 18bp sequence that is thought to be absent from most mammalian genomes (Colleaux et al., 1988; Jasin, 1996). This characteristic is important as it enables manipulation at a specific pre-integrated recognition site to facilitate elevated gene targeting frequencies. Although no I-*SceI* recognition sites are thought to be present in the mammalian genome, it has been reported that the endonuclease may recognise sites that slightly deviate from its 18bp recognition sequence (Colleaux, 1988). Non-specific cleavage of genomic DNA (i.e. at cryptic sites) may lead to mutagenic consequences.

A limitation of using such a system for encouraging gene targeting is that it is necessary to first carry out a targeting step to integrate the unique I-*SceI* recognition sequence into the target gene, prior to making the required genomic change. Integration of the recognition site typically takes place using a method of selection for the easy isolation of such clones; therefore, you can subsequently target your GOI using a selection-free targeting cassette at a much higher efficiency. This method allows accurate alterations to be made to the genome, which can mimic *in vivo* disease systems and avoid completely disrupting the gene with excess plasmid DNA. This approach would be particularly desirable in situations where a range of mutations at a given locus are required. This approach was successfully used to mutate the endogenous villin gene in mouse ES cells at a frequency 100-fold greater than conventional gene targeting (Cohen-Tannoudjii *et al.*, 1998). In addition, the frequency of gene targeting at the endogenous HPRT locus in ES cells was shown to be 5000-fold higher when stimulated using restriction at a preintegrated I-*SceI* site, when compared to cells containing no recognition site (Donoho *et al.*, 1998). A simplified illustration of the two-step I-*SceI* targeting approach deployed for each of these experiments is shown in Figure 1.6.

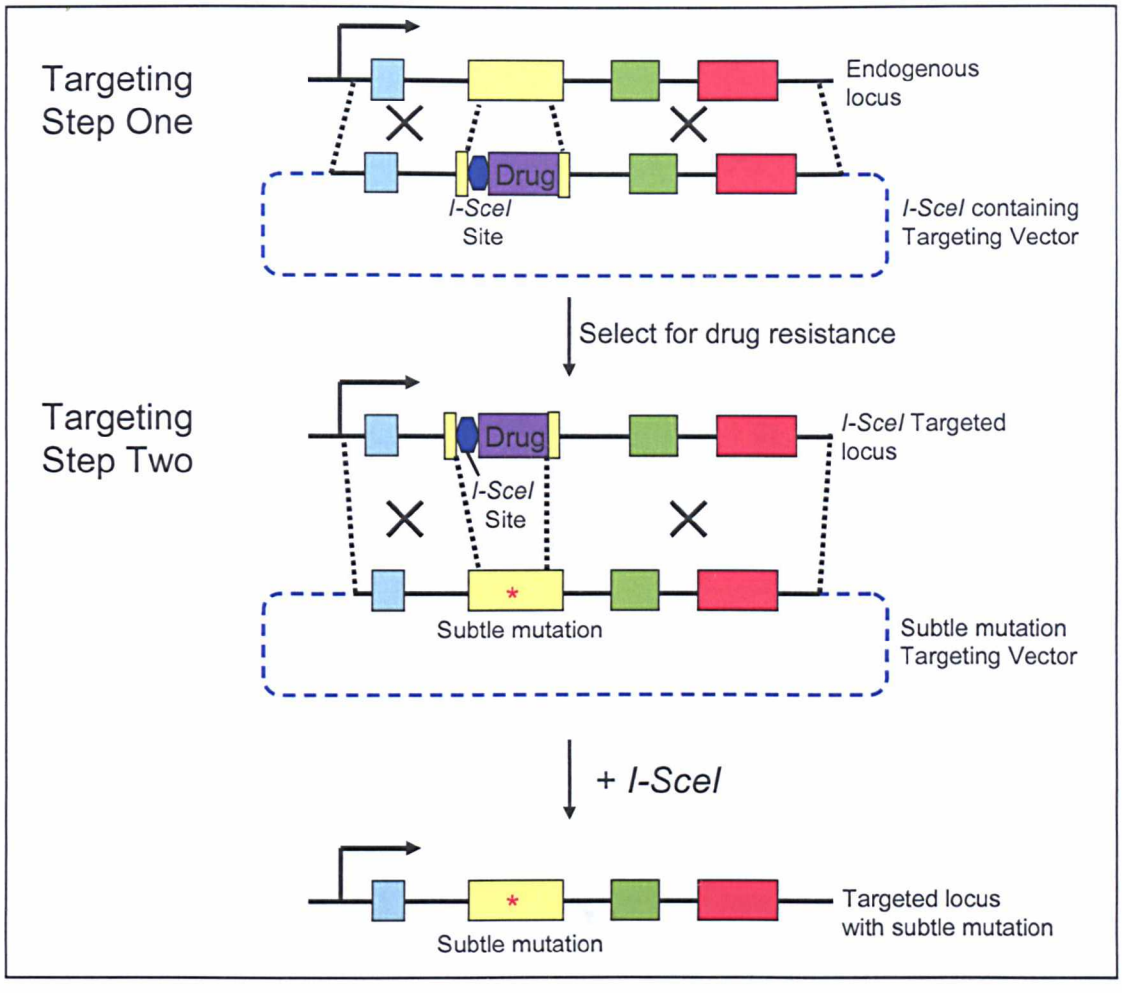



**Figure 1.6 Double strand break induced gene targeting:** Targeting step one: Firstly the I-*SceI* site must be targeted into the locus using a drug selection cassette, possibly a selectable marker gene (Drug). Targeting step two: A DNA double strand break can be induced at the target locus by expression of the I-*SceI* endonuclease. The targeting construct, containing the subtle mutation, is used by the cells as a template for the repair of the DNA break and the subtle mutation is then copied to the target locus.

For exonucleases to be more useful in stimulating gene targeting at any chosen target gene, a method for generating rare cutting endonucleases with any desired sequence specificity is required. This has recently been made possible by the generation of chimeric nucleases that consist of three zinc finger DNA binding domains linked to the *FokI* nuclease domain (Chandrasegaran and Smith, 1999). Each of the zinc fingers is designed so that it recognises and binds to a triplet of unique genomic sequence (Figure 1.7). The zinc finger endonucleases are designed in such a way that their site-specific binding to a unique 18bp target sequence brings together two *FokI* nuclease molecules and activates them to generate a double strand break (Figure 1.7; Jantz et al., 2004; Urnov *et al.*, 2005). Recently, successful gene targeting was

observed in up to 20% of non-selected K562 cells using this zinc finger endonuclease method (Urnov *et al.*, 2005). Although designer endonucleases are an exciting new prospect for the gene targeting field, they are still in their infancy and the techniques require further optimisation. In the future, however, it is predicted that chimeric endonucleases will provide a very useful tool for both gene function studies and gene therapy applications (Reviewed: Porteus and Carroll, 2005).

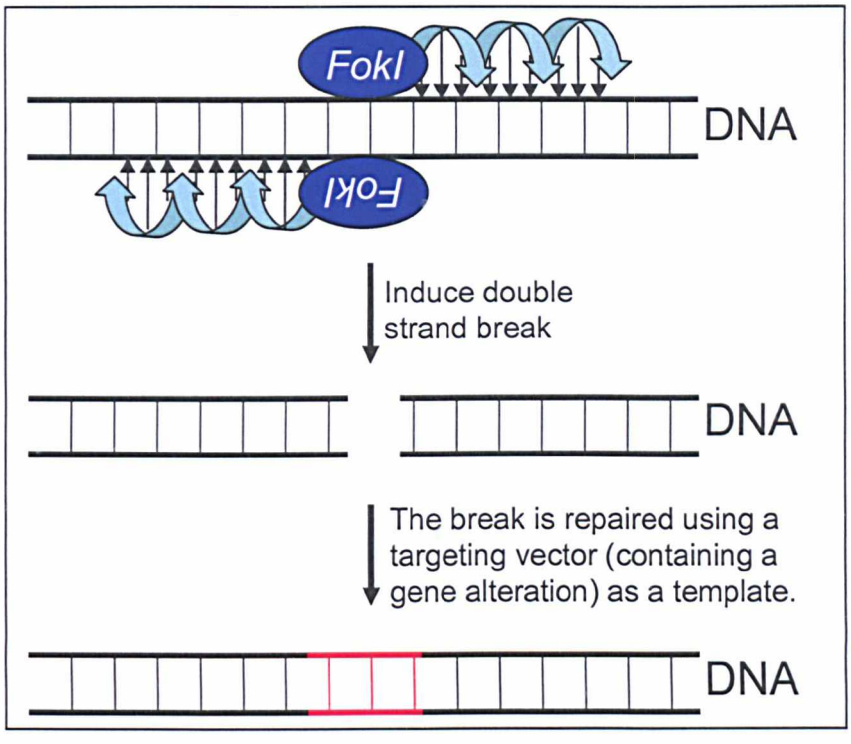



**Figure 1.7 Zinc finger endonucleases:** Illustration of a DNA bound zinc finger pair (top) whose FokI nuclease-activity induces a double strand break (middle). During a gene targeting experiment the DNA double strand break can be repaired using the targeting vector, containing the desired gene alteration, as a template.

## 1.1.4 Oligonucleotide-based approaches for specific gene silencing

Attempts to develop more efficient techniques other than gene targeting for improved loss-of-gene-function experiments have produced a number of oligonucleotide-based methods. These include anti-sense oligonucleotides (Zamecnik and Stephenson, 1978) and ribozymes (Cech *et al.*, 1981), which have routinely been used for over 10 years for the targeted degradation of specific RNAs. Although these methods work

effectively in simple systems (e.g. protozoan (Tetrahymena) [Chech and Brehm, 1991]; *E. coli* [Geller, 2005]), they are not quite as efficient in more complex mammalian organisms (Curcio *et al.*, 1997). Developments in the RNA interference (RNAi) field over the last few years have led to the preferred use of short interfering RNAs (siRNAs) for targeted gene silencing studies. This technology was first discovered in *C.elegans* when it was observed that injection of double stranded RNA, homologous to an endogenous gene, resulted in the post-transcriptional silencing of that gene (Figure 1.8A; Fire *et al.*, 1998). Traditional gene silencing by RNAi was achieved by two main steps: firstly dsRNA is recognised by the enzyme Dicer (RNase III nuclease) and is processed into small double-stranded molecules (called siRNA; 21- to 23-nucleotides) (Zhang *et al.*, 2002); secondly, the siRNA is bound by an RNA-Induced Silencing Complex (RISC) which has RNase activity and is responsible for the degradation of the target RNA (Figure 1.4A; Hammond *et al.*, 2000).

A limitation of traditional RNAi is that the introduction of dsRNA into mammalian cells often provoked a cytotoxic response (Hunter *et al.*, 1975). To overcome this, siRNAs (21-22nt) were directly used to produce specific gene silencing (Figure 1.8B; Elbashir *et al.*, 2001); however their application was further limited by their transient method of delivery. Consequently, the polymerase-III H1-RNA gene promoter was cloned into a mammalian expression vector to develop pSUPER; for the stable synthesis of siRNA-like transcripts (short hairpin RNAs [shRNAs]) (Brummelkamp *et al.*, 2002). As a direct result, conditional shRNA expression constructs are currently under much review, as they could restrict gene down-regulation to a particular time point, which would be highly attractive if the gene were essential for the cell's survival. One approach has involved tagging the polymerase-III U6 promoter to well characterised inducible response elements, which is normally combined with polymerase-II DNA promoters (Matsukura *et al.*, 2003). A limitation of this approach, however, was that a larger concentration of the inducer was required to completely suppress shRNA expression. More recently, polymerase-II directed synthesis of shRNA has been demonstrated, which should allow tissue specific siRNA expression, a feature not possible whilst expressing from a polymerase-III promoter (Xia *et al.*, 2002; Zhou *et al.*, 2005); consequently the future expression of shRNA from conventional inducible promoters (e.g. the Tet and Lac systems; discussed later see section 1.2.3) can be conceived. Finally, siRNA is an invaluable tool for gene

function studies, however its application is limited by the inability to completely abolish gene expression and to make specific gene alterations for mutagenic studies

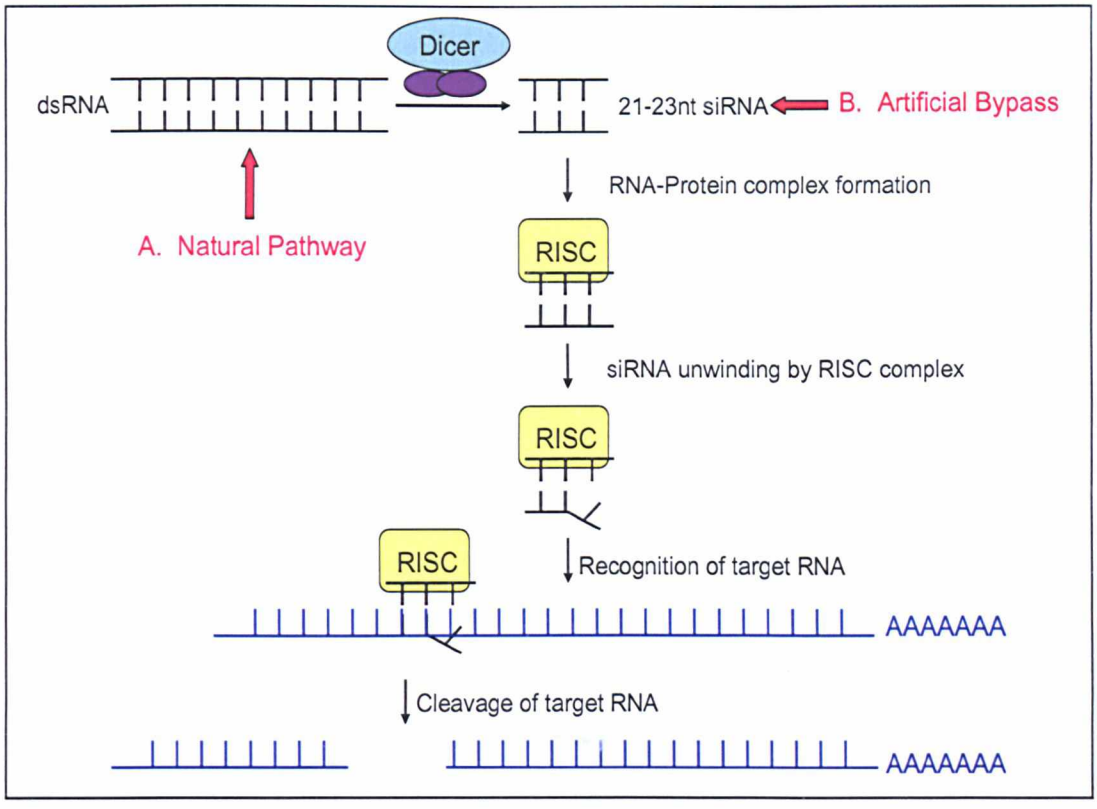



**Figure 1.8 RNA interference:** RNA interference pathway illustrating the natural pathway (**A.**) and the entry point for siRNA (**B.**).

## 1.1.5 Site-specific recombination

As discussed earlier, the directed introduction of specific mutations into the mammalian genome by gene targeting is a powerful strategy for gene function studies, although its application is greatly hindered by its low efficiency (Doetschman *et al.*, 1987; Thomas and Capecchi, 1987). The development of an additional technology for the site-specific integration of DNA into the mammalian genome has been greatly received and has significantly improved gene function studies, particularly for applications involving transgenic mice. Such technology utilises site-specific recombinases, which are enzymes that catalyse a recombination reaction between two particular sites. The most commonly used site-specific recombinase for genetic manipulation is the Cre-recombinase system which was first demonstrated in

mammalian cells almost 20 years ago (Sauer and Henderson, 1988). Cre recombinase is a 38kDa protein required by bacteriophage P1 for genome partitioning after replication (Sternberg and Hamilton, 1981) and it catalyses the site-specific recombination between two *loxP* sites. The *loxP* sequences are made up of an 8bp asymmetric spacer region flanked by two 13bp inverted repeats (Figure 1.10A) (Hoess *et al.*, 1982).

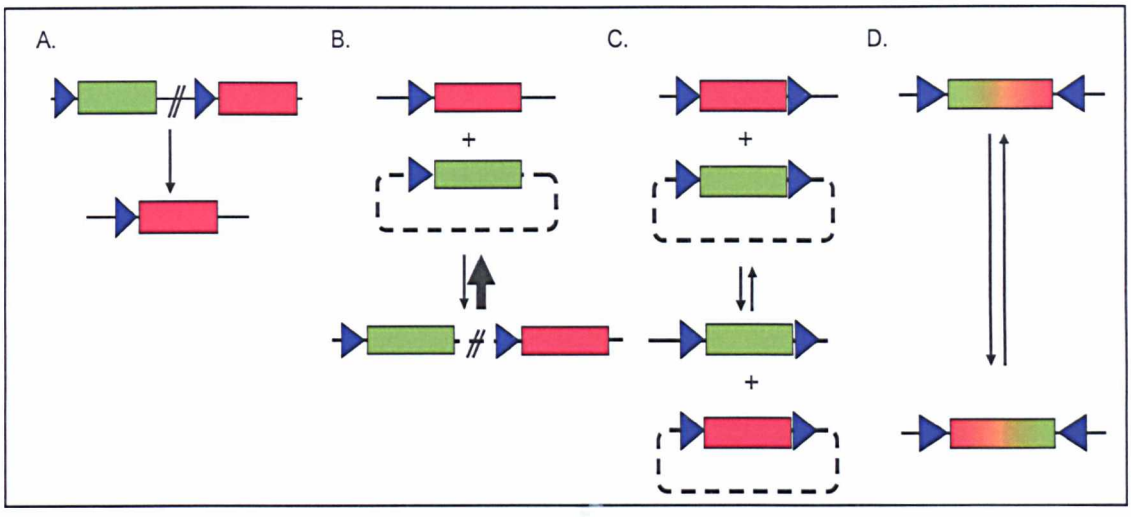



**Figure 1.9 Reactions catalysed by site-specific recombinases such as Cre or Flp: A.** Recombinase mediated excision removes the sequence flanked by recognition sites (represented by blue triangle). **B.** Recombinase-mediated insertion introduces vector DNA at a recognition site. **C.** Recombinase-mediated cassette exchange (RMCE) swaps one piece of DNA for another that are both flanked by recognition sites. **D.** Recombinase-mediated inversion reverses the DNA sequence flanked by inverted recognition sites.

More recently another site-specific recombinase, Flp, has been characterised from yeast, which recognises an FRT site (Kilby *et al.*, 1993; Dymecki, 1996). The wild-type FRT site is 48bp long and also contains an 8bp spacer, however unlike the *lox* site, it is flanked by three 13bp inverted repeats (Andrews *et al.*, 1985). A minimal FRT site (with only two 13pb inverted repeats), which is only 34bp in length, is effective at the excision reactions (Figure 1.9A) but not integration (Figure 1.9B; Senecoff *et al.*, 1985). In mammalian cells Flp is less active than Cre due to its reduced activity at 37°C. Depending on the combination of *lox*/FRT sites, Cre/Flp are capable of excising regions flanked by *lox*/FRT sites (Figure 1.9A), gene insertions or replacement of transfected DNA into a chromosomally located *lox*/FRT sites (Figure 1.9B and C, respectively) and inverting sequences flanked by inverted *lox*/FRT sites

(Figure 1.9D). The variety of applications these tools provide has been invaluable for gene knockout (loss-of-function) and activation (gain-of-function) experiments (Gu *et al.*, 1993; Fukushige and Sauer, 1992; Rodriguez *et al.*, 2000).

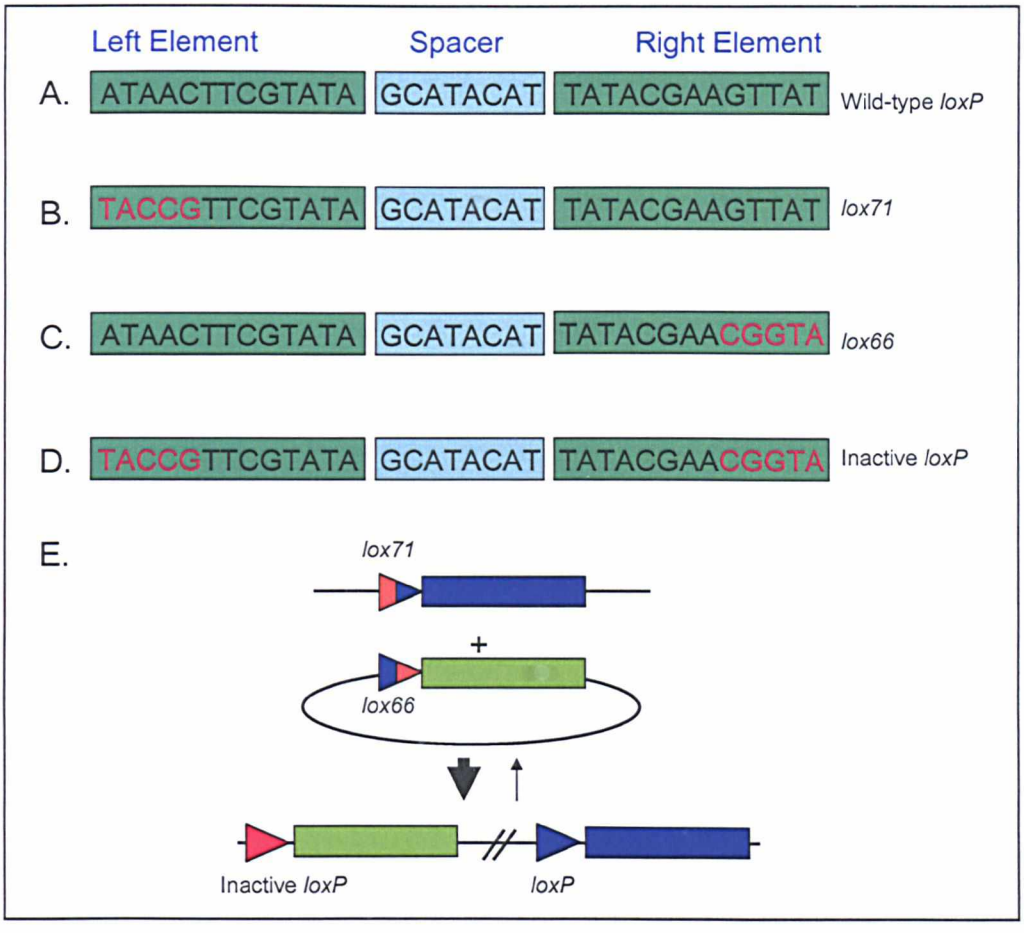



**Figure 1.10 The mutant *loxP* system: A.** Sequence of wild-type *loxP*. **B.** Sequence of mutant *lox71* site showing mutation in left element (in red). **C.** Sequence of mutant *lox66* site showing mutation in right element (in red). **D.** Sequence of double mutant inactive *loxP* site, the product of recombination between a *lox71* and a *lox66* site. **E.** Illustrates the generation of incompatible *loxP* sites after recombination between *lox71* and *lox66*. This inhibits excision reaction after insertion has taken place during transient Cre expression.

Wild-type *loxP* and FRT sites are inefficient at the integration reaction (Figure 1.9B) as they tend to re-excise the recombined product in excess transient Cre and Flp expression. The Cre-*loxP* system has improved on this limitation by developing mutant *loxP* sites, which have increased the frequency of Cre-mediated insertion or replacement by limiting the favoured subsequent excision reaction (Figure 1.9B and

C). The most commonly utilised mutant *lox* sites used for Cre-mediated integration are the *lox71* (Figure 1.10B) and *lox66* (Figure 1.10C) sites, which have 5bp mutations in the left and right 13bp arms, respectively. Recombination between these two sites generates a wild-type (Figure 1.10A) and double mutant *lox* (Figure 1.10D) site, which are thought not to recombine with each other (Figure 1.10E; Araki *et al.*, 1997). In addition, heterospecific *loxP* sites have been developed (*lox511*, *lox2272*, *lox5171*) to facilitate recombinase mediated cassette exchange (RMCE) (Figure 1.9C; Feng *et al.*, 1999). These mutant sites contain alterations to the 8pb spacer regions of their *loxP* sites and they can only recombine with heterospecific sites of the same type (Hoess *et al.*, 1986).

A further recombinase system has recently been developed for use in mammalian cells: the phage ΦC31 integrase (Smith and Thorpe, 2002; Groth *et al.*, 2000). ΦC31 integrase mediates the insertion of the ΦC31 phage DNA into target sites present in both the phage and host (*Streptomyces*) genomes. The sites of recombination between the phage and the host are called attP (**att**achment site in **p**hage genome) and attB (**att**achment site in **b**acterial genome), respectively (Groth and Calos, 2004). Unlike wild-type Cre and Flp systems, ΦC31 is a unidirectional recombinase as recombination between attP and attB generates attL and attR sites which are no longer substrates for integrase, in the absence of additional factors (Thorpe and Smith, 1998). The insertion frequency of wild-type ΦC31 integrase in mammalian cells is lower than seen for Cre, however, this technology is still in the early stages and further manipulations of ΦC31 are likely to increase its efficiency further (Sclimenti *et al.*, 2001).

Site-specific recombinases are now being deployed for the directed integration of a GOI for its overexpression (gain-of-function) under the transcriptional control of an endogenously located gene (Soukjarev *et al.*, 1999; Feng, 1999; Shmerling *et al.*, 2005). It is hoped that this method will eradicate the unpredictable patterns and reduced levels of expression observed by conventional gene overexpression methods (by random integration) caused by the choice of promoter, the position of integration (position effects) and epigenetic factors (Henikoff, 1998). Recently, Shmerling and colleagues (2005) showed that targeted integration of the Green Fluorescent Protein (GFP) gene downstream of the endogenous, continuously expressed, β-actin promoter

resulted in strong, ubiquitous GFP expression in all tissues of the developing mouse (Shmerling *et al.*, 2005). These findings led to Shmerling and colleagues developing a novel system that can utilise the β–actin promoter for generating high levels of expression of any gene of interest (presently in ES cells), using Cre-mediated cassette exchange (Figure 1.11). Downstream of the endogenous β–actin promoter the GFP gene is flanked by inverted *lox* sites (*lox511, loxP*). A simple cassette exchange vector is used to replace the GFP gene with the GOI and the integration event can be screened by analysing cells for loss of GFP expression (Figure 1.11).

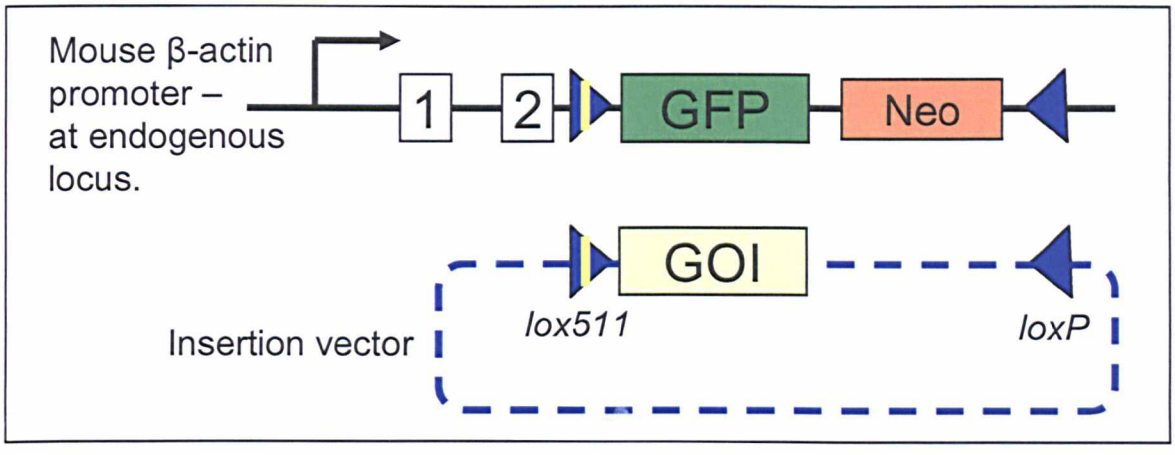



**Figure 1.11 A system for strong ubiquitous expression of a transgene:** Illustrates the strategy developed by Shmerling *et al.* (2005) for Cre-mediated cassette exchange replacing the GFP gene with a gene of interest (GOI), placing it under the control of the strong, ubiquitous endogenous β–actin promoter.

Gene function studies also frequently use site-specific recombinases for the generation of mutations in the genome (e.g. subtle mutations) and the subsequent removal of the drug resistance cassettes after selection for gene targeting events (Mansour *et al.*, 1990). An example which incorporates both of these features is described by Zhang and Lutz (2002) who targeted the CREB-binding protein (implicated in many intracellular signal transduction pathways) with a point mutation to investigate the effects on its activity. Firstly, a targeting step was carried out to incorporate the mutation (located within an inverted copy of exon 5) and selection in neomycin helped identify successfully integrated clones (Figure 1.12). A clone that had undergone the targeting event was still expressing the wild-type protein due to the pattern of splicing, which skips the drug selection cassette and the inverted mutated exon 5, as they are both located within the intron (Figure 1.12). Such clones can be

transformed to express the mutant form of the CREB-binding protein by expressing Cre-recombinase within the cells, which catalyses an inversion event and places the mutated exon 5 in a position so that it will be included in the transcript (Figure 1.12). The efficiency of this event was increased by using mutant *loxP* sites (*lox66* and *lox71*) to prevent subsequent repeated inversion reactions. If desired, the drug-resistance cassette could also be removed by an excision event using *flp*-recombinase, as the neomycin gene was flanked by FRT sites (Figure 1.12).

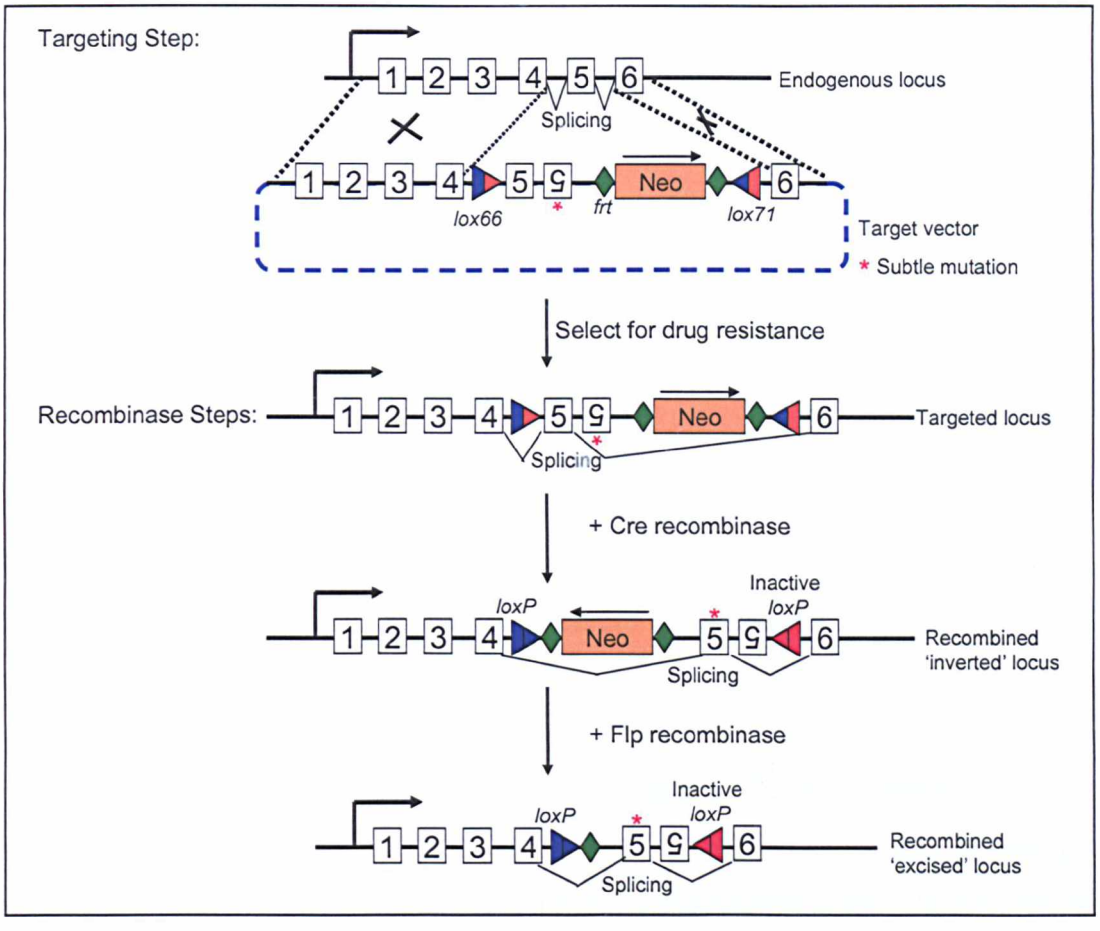



**Figure 1.12 Cre recombinase-mediated mutagenesis of the CREB-binding protein in ES cells:** Firstly, ES cells are targeting with a vector containing a point mutation in a second copy of exon 5 in the antisense orientation (top). Integration of the targeting event can be selected for in neomycin. Targeted clones express wild-type CREB-binding protein due to the pattern of splicing skipping the drug resistance cassette and the mutated exon 5 (see targeted locus). Treatment of the targeted clones with Cre recombinase results in an inversion event placing the mutated exon 5 in the sense orientation which is now included in the transcript after splicing (see recombined 'inverted' locus). If required, the neomycin resistance cassette can be excised from the genome by expression of flp recombinase (e.g. if it is suspected to be interfering with expression of the CREB-binding protein [see Recombined 'excised' locus]).

Limitations of the site-specific recombinase technology include the numerous steps involved in generating the required genetic alterations. In addition, the recombined sites remain in the genome whilst functional analyses are carried out; however these sites are often located in the non-coding introns and therefore do not usually interfere with expression of the GOI (for e.g. see Figure 1.12; Zhang and Lutz, 2002). In summary, site-specific recombinases not only improve gene overexpression studies by allowing the accurate targeting to a specific predetermined locus, they also enhance the scope for gene targeting experiments by facilitating the generation of subtle mutations in the mammalian genome. In addition, the timing of Cre expression can be controlled precisely, therefore the site-specific recombinase approach can also be utilized for conditional gene overexpression or mutagenesis studies (discussed later, see section 1.2.1).

### 1.1.6 Limitations of constitutive gene transfer approaches

Gain-of-function studies most often entail the random integration or transient expression of a cassette coding for a GOI and clones expressing the transgene are isolated after selection in the appropriate drug. The degree of overexpression is normally analysed by Western blot and calculated as a percentage above endogenous levels. This approach is limited by the unknown effects caused at the random integration site, which leads to variations in the amount of absolute expression levels. Subsequently, numerous clones (both overexpressing and wild-type) must be analysed in order to generate a representative picture of the phenotypic effect. In addition, the effects of proteins that reduce cell viability whilst up-regulated above endogenous levels cannot be analysed using constitutive gene overexpression methods, due to a difficulty in generating viable stable clones. For example, overexpression of the human *RAD52* gene, which is involved in DNA repair (discussed in Section 1.1.3), is not tolerated in HT1080 cells and its expression is lost from the cells after several passages (Yanez and Porter, 2002).

Loss-of-function studies, whether they involve gene targeting or RNAi, are also limited to those genes that are not essential for the cell's survival to generate stable

viable clones for analyses (e.g. Sonoda *et al.*, 1998); for example, genes that are involved in apoptosis, cell-cycle and development. It has been shown that *RAD50* (Luo *et al.*, 1999), *RAD51* (Lim and Hasty, 1996) *and Mre11* (Xiao and Weaver, 1997) deficient ES cells, which are involved in DNA repair (discussed in Section 1.1.3), result in lethal phenotypes. In mouse transgenics, if complete loss of a gene's product resulted in embryonic lethality, gene function at later stages of development cannot be analysed. Restrictions associated with uncontrolled gene regulation can be overcome by transient overexpression or silencing of the GOI, however, there is only a limited amount of time for phenotypic analysis before expression is lost (or regained), and the proportion of the cells that receive the vector is variable and always less than 100%.

## 1.2 Conditional gene expression for gene function studies

### 1.2.1 Recombinase-mediated conditional gene expression

As mentioned earlier, conditional gene activation can be achieved using the site-specific recombinase system. Typically, conditional gene activation involves separating the promoter and the start of the gene with a stop codon, flanked by two recombinase recognition sites (Figure 1.13A; Lakso *et al.*, 1992; Kolb, 2002). When ready to analyse the effects of the gene's expression the cells or mice are treated with the appropriate recombinase to excise the stop codon and activate the GOI. The advantage of this technique is that it restricts the expression of a toxic gene to a particular time point. Similarly, conditional inactivation of an essential GOI to generate a knockout phenotype can be controlled by flanking essential gene sequences (e.g. start codon) with the appropriate recombinase recognition sites and regulating the timing of recombinase expression (Figure 1.13B) (Sauer, 1998). Many examples of conditional gene manipulations involving site-specific recombinases are sited in the literature regarding mice (e.g. BRCA1 [Xu *et al.*, 1999] and BRCA2 [Ludwig *et al.*, 2001]; both genes are involved in DNA repair and are implicated in breast cancer); however by comparison, there are fewer examples involving cultured cells (e.g. *BRCA2* disruption in DT40 cells [Hatanaka *et al.*, 2005] and disruption of *RAD17*

ORF in colon epithelial cells [Wang *et al.*, 2003]; *RAD17* is required for the response to DNA damage and replication stress).

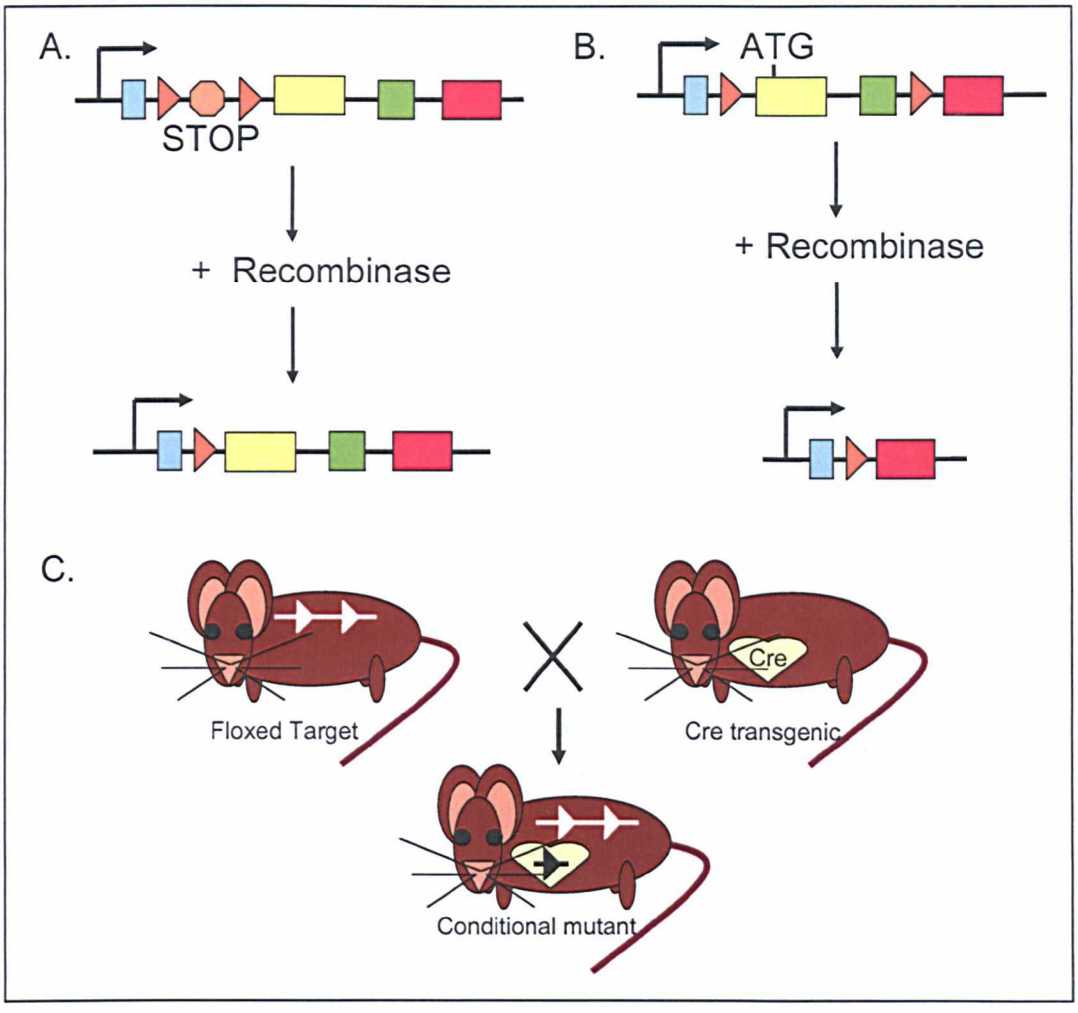



**Figure 1.13 Recombinase-mediated gene activation and inactivation: A.** Conditional activation of a GOI can be accomplished by the deletion of a floxed stop codon located within an early section of the GOI's ORF. **B.** Conditional gene inactivation can be achieved by the deletion of essential gene sequences (e.g. start codon) using flanking recombinase recognition sites. **C.** A method for tissue specific conditional recombination. A mouse carrying a loxP flanked gene (top left mouse) is bred with a Transgenic mice expressing Cre recombinase only in a particular tissue (linked to tissue specific promoters; top right mouse). In the resultant double Transgenic mouse (bottom mouse) excision of the *loxP*-flanked gene would have only taken place in the Cre-expressing tissue.

In mice, conditional site-specific recombination is extremely successful due to the ability to generate mice expressing Cre-recombinase typically from a ubiquitous or tissue specific promoters (by randomly integrated [Hennet *et al.*, 1995] or targeted

insertion into endogenous loci [Rickert *et al.*, 1997]); or alternatively inducible promoters (Gossen *et al.*, 1995; No *et al.*, 1996). Crossing the desirable Cre-expressing mouse with a mouse harbouring a *loxP*-flanked segment of target gene will result in the restricted modification of that locus at a particular time of development or tissue (Figure 1.13C). Alternatively, Cre can be delivered to somatic tissues via the infection of mice with adenoviral Cre expression vectors, although site-specific recombination in every desirable cell is not guaranteed when using this method (Rohlmann *et al.*, 1996; Wang *et al.*, 1996). A similar drawback is experienced when adopting conditional site-specific approaches in mammalian tissue culture cells. After the addition of Cre recombinase by transient transfection (the typical approach) only a proportion of the cells receive the plasmid and undergo the recombination event successfully (always less than 100%), therefore a selection or screening method is required for identifying the recombined clones. If the recombination event results in the expression of a toxic gene or the down-regulation of an essential gene then the ability to generate viable clones following the selection (or screening) process will be compromised. Attempts to overcome this limitation in tissue culture involve stable expression of Cre (Chiba *et al.*, 2000) and flp (Sawicki *et al.*, 1998) recombinases from inducible promoters which allow the control over the enzyme's expression by the addition or removal of an inducer. A limitation of this type of approach, however, is that low levels of recombinase expression can still remain when the inducible promoter is suppressed (leaky expression), and may lead to a background of unwanted recombination events.

### 1.2.2 Applications and advantages of an inducible promoter system

Both recombinase-mediated conditional gain- and loss-of-function techniques are unidirectional and the expression levels cannot easily be reversed or altered to alleviate any stresses subjected on the cells or animal model. It is therefore desirable to overcome the problems associated with permanent overexpression and silencing of a GOI from an inducible promoter to generate conditional gain- or loss- of function systems. Strict control over gene expression is a valuable tool for studying complex gene function processes in cultured cells and mammalian developmental systems.

Inducible gene regulation allows the expression of a gene in both a quantitative and temporal manner. Several regulatory mechanisms that allow the controlled expression of a GOI by taking advantage of regulatory elements found in *E.coli, Drosophila* and *S. cerevisiae* have been developed. The success of each of these systems has relied on their ability to achieve reproducible regulation, with low uninduced expression levels (tight regulation), in simple culture systems and higher eukaryotic organisms without any detrimental side effects (toxicity); this allows investigators to accurately predict and control where and when transgenes or gene mutations are expressed.

Conditional gene inactivation using site-specific recombination generally results in the complete elimination of that genes activity; usually due to the removal of its important regulatory elements from the genome (Figure 1.13B). The inability to achieve similar levels of gene inactivation (tight regulation) using inducible promoter systems has been one of the major limitations of adopting such a strategy. Tight inducible gene regulation is important when attempting to express a GOI only at a particular time point. This will facilitate the study of a GOI whose product is toxic; involved in cell arrest phenotypes (such as pro-apoptopic and anti-proliferative genes); or whose expression in small amounts results in a phenotype. Tight regulation is particularly important when attempting to generate a conditional knockout phenotype, as tight regulation from the inducible promoter upon its suppression is necessary to characterise the effects of loss-of-gene function fully (Carpenter and Porter, 2004).

A situation that would benefit from tight inducible gene expression would be the study of the effects of stable I-*SceI* expression on genomic stability in cultured cells. To achieve this, it is predicted that tight gene regulation, when uninduced, would be necessary to prevent uncontrolled cutting at pre-integrated (or cryptic?) I-*SceI* sites. In addition, our group have previously attempted to characterise the effects of human *RAD52* overexpression on homologous recombination and gene targeting (see section 1.2; Yanez and Porter, 2001). The constitutive overexpression approach adopted was limited by a toxicity of the RAD52 protein, even at low levels of up-regulation. It would therefore be advantageous to adopt a conditional approach for studying the function of the *RAD52* gene further, under more restrictive conditions.

## 1.2.3 Eukaryotic-based inducible gene expression systems

Inducible systems developed from eukaryotes are based on the heat shock, metallothionein, interferon-γ and hormone-dependent promoters. The applications of these systems, however, are limited to cell culture as they have had reduced success in mouse models due to pleiotropic effects from the inducers (inducer affecting expression at other endogenous regulatory units), lack of regulation and high basal expression (For review, see Gringrich and Roder, 1998). These systems are not usually applied for a GOI interest whose expression must be tightly regulated (e.g. an enzyme or toxic gene). A summary of the different types of inducible systems based on eukaryotic elements are illustrated in Table 1.2.

**Table 1.2 Eukaryotic-based inducible systems:** A summary of a few inducible systems derived from eukaryotes, highlighting advantages and disadvantages.

| Eukaryotic-based Inducible System | Pleiotropic Effects | Induction Ratio | Kinetics | Leakiness |
|---|---|---|---|---|
| **Heat shock** Kothary *et al.*, 1989 | Many | 4- to 10-fold | Fast 1h | low |
| **Heavy metal ions** Filmus *et al.*, 1992 | Many | 5- to 10-fold | Fast 16h | High |
| **Interferon** Kuhn *et al.*, 1995 | Many | 2- to 50-fold | Slow 3-4 days | Low |
| **FK506** Belshaw *et al.*, 1996 | Few | 1.5-fold | Fast 16h | Very low |
| **Steroid hormone** Kuo *et al.*, 1994 | Many | 200-fold | Fast 24h | High |
| **Estrogen receptor** Zhang *et al.*, 1996 | None | Not applicable | Slow 3-4 days | Low |
| **Ecdysone** No *et al.*, 1996 | None | $10^4$-fold | Fast 20h | Very low |

*The ecdysone system is the most promising eukaryotic derived inducible system*

The most promising eukaryotic-based inducible system to date is the ecdysone system and has successfully been applied to both mammalian culture and transgenic animal systems (No *et al.*, 1996). It is based on the insect moulting hormone ecdysone as an inducer. Ecdysone can rapidly stimulate metamorphosis in *Drosophila melanogaster* via the ecdysone receptor. The ecdysone receptor alone is inefficient at DNA binding; however once it fuses to the ultraspiracle (USP) protein it can form a heterodimer and bind to several nuclear receptors, driving expression of the appropriate genes. USP is the insect homologue of the vertebrate retinoid X receptor (RXR), which can substitute for USP in the mammalian ecdysone inducible system (Yao *et al.*, 1993).

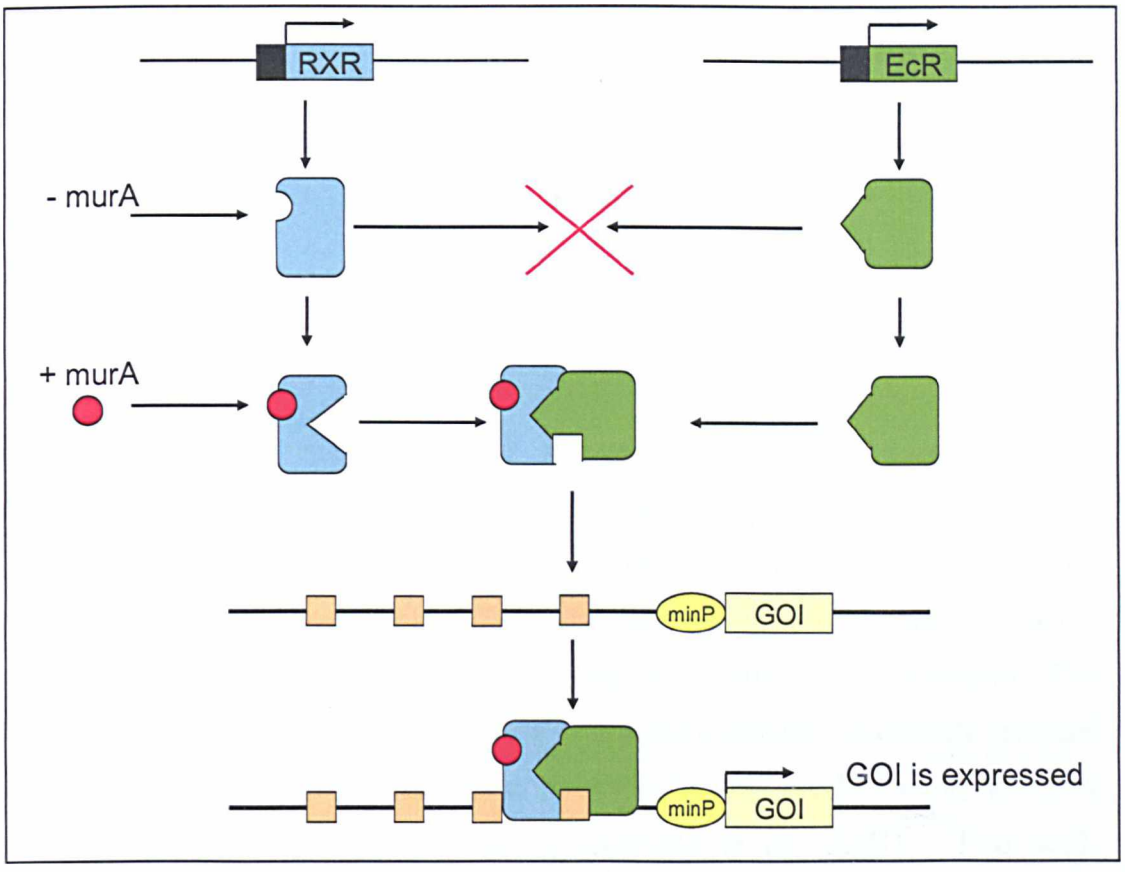



**Figure 1.14 The ecdysone inducible system:** The modified ecdysone receptor (EcR) is expressed along with the RXR protein and in the presence of ecdysone (or murA; red circle) they bind to form a heterodimer. This fusion protein is then able to activate expression of a GOI by binding to one of the four response elements (EcRE; yellow squares). Optimum induction is achieved by placing the four EcREs upstream of a minimal promoter (minP) to drive expression of the transgene.

The ecdysone receptor has been modified, by fusing it to the herpes simplex virus activation domain VP16, to allow it to specifically bind to an optimised ecdysone-responsive promoter when stably expressed in mammalian cells. The optimised promoter is made up of a minimal heat shock promoter (from *Drosophila*) linked to a series of ecdysone receptor response elements. When cells containing all the appropriate expression cassettes are treated with ecdysone (or muristerone A) the modified ecdysone receptor dimerises with RXR and drives expression of a GOI that is located downstream of the optimised ecdysone-responsive promoter (Figure 1.14). Fast kinetics of gene expression can be achieved and induction levels reaching 4 orders of magnitude (up to $10^4$-fold) have been demonstrated (No *et al.*, 1996). Crucially, the effects of prolonged exposure of mice to muristerone A have been examined and there appears to be no long term consequences. The major limitation of this inducible technique is generating clones containing all three transgenes (ecdysone receptor, RXR and the optimised ecdysone-responsive promoter linked to the GOI) integrated at genomic locations suitable for expressing in sufficient amounts for reproducible regulation (Yamomoto *et al.*, 2001).

### 1.2.4 Inducible systems based on the regulatory elements of prokaryotes

As discussed earlier, eukaryotic-based systems for inducible expression often demonstrate pleiotropic effects (e.g. interferon or heat shock; Table 1.2) or high basal levels of expression when uninduced leading to modest induction values (e.g. metal ions and steroid hormone; Table 1.2). Alternatively, inducible gene expression systems have also been developed using the regulatory circuits of prokaryotes. The advantage of using these types of systems is that they should demonstrate minimal pleiotropic effects, as prokaryotic transcriptional activators are less likely to affect gene expression in eukaryotic systems (Yamamoto *et al.*, 2001). Two well-characterised regulatory circuits based on the *lac* and *tet* operons of *E.coli* have successfully demonstrated their ability to regulate gene expression in a monospecific manner in both mammalian cells and transgenic animals and their features are summarised in Table 1.3.

**Table 1.3 Prokaryotic-based inducible systems:** A summary of two well-characterised inducible systems based on prokaryote regulatory circuits, highlighting their advantages and disadvantages.

| Prokaryotic-based Inducible System | Pleiotropic Effects | Induction Ratio | Kinetics | Leakiness |
|---|---|---|---|---|
| Lac<br>Baim *et al.*, 1991 | Cytotoxic | LacR-VP16:<br>1000-fold | LacR-VP16:<br>Slow 24-72h | High |
| Tet<br>Gossen *et al.*, 1995 | None | tTA $10^5$-fold<br>rtTA 1000-fold | Fast 12-24h | Very Low |

*The lac inducible system is base upon the lac operon of E. coli*

The lac-based system is adapted from the well-characterised *E. coli* lac operon. The *E.coli* lac genes are suppressed when the lac repressor (lacR) binds to the lac operator (lacO), which is located between the promoter and the regulated genes. In this way the lacR sterically inhibits entry of RNA polymerase for transcriptional initiation (Figure 1.15A, top). When the inducer molecule, β-galactosidase (or a synthetic analogue IPTG), is present it binds to the lacR preventing it from repressing expression at the lac operator (Figure 1.15A, bottom). The highly specific interactions of the lac operon circuit made it an obvious candidate for transcriptional regulation in mammalian cells.

Several attempts at generating the ideal lacR inducible system in mammalian cells have been documented (Hu and Davidson, 1997; Brown *et al.*, 1987). One approach involved fusing the lacR to the eukaryotic activation domain of the herpes simplex virus protein, VP16, converting the lacR from a repressor to a transcriptional activator (termed lac activator). The latter manipulation meant that IPTG did not have to be present continually to keep the transcriptional unit silent and it was hoped that gene activation would be achieved more rapidly. The regulated genes were placed downstream of a minimal promoter and a tandem array of numerous lacO sites (Labow *et al.*, 1990). This arrangement of the lacR machinery in mammalian cells generated a system that allowed the induction of a GOI by up to 1000-fold by addition the inducer IPTG (Baim *et al.*, 1991; Figure 1.15B).

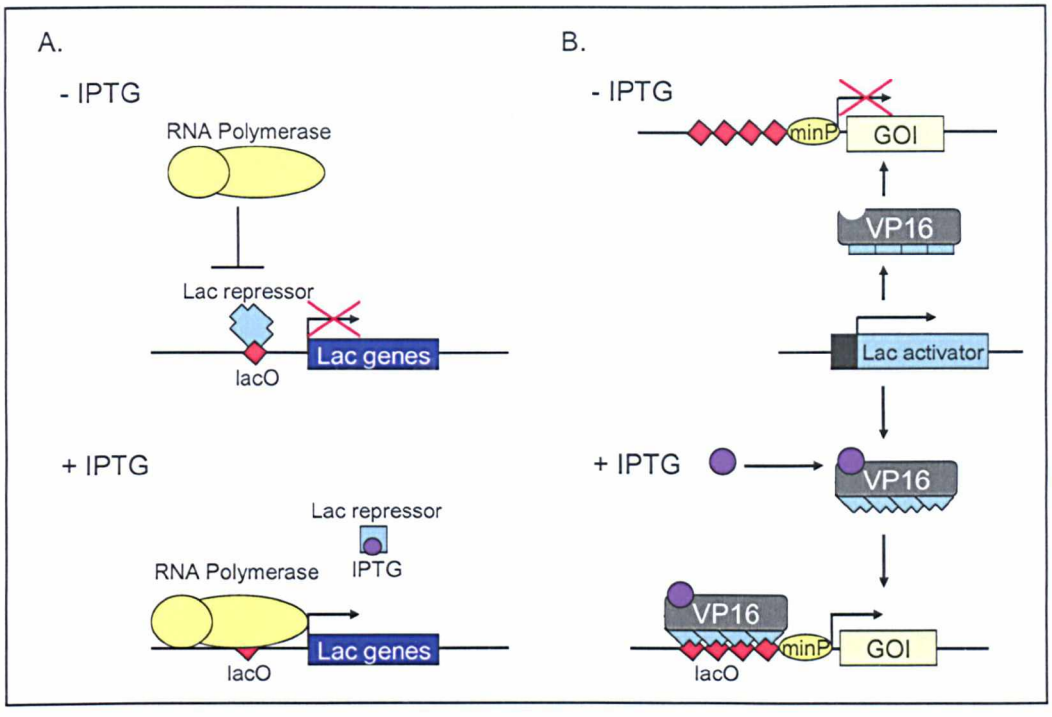



**Figure 1.15: The lac operon and inducible system: A.** A schematic representation of the lac operon system. The lacR binds to the lacO as a tetramer in the absence of IPTG (top) and inhibits expression of the lac genes by preventing RNA polymerase access. This process is reversed by the addition of IPTG (bottom). **B.** In the lac inducible system the lac activator cannot bind to the lacO in the absence of IPTG and under these conditions the GOI is suppressed (top). After the addition of IPTG the lac activator can bind to the lacO and transgene expression is resumed (bottom).

A limitation of this approach, however, is the temperature-sensitive binding of the lacR to the lacO and it means that the cell cultures have to be grown at elevated temperatures. This results in temperature-induced pleiotropic effects and therefore limits the applications of the lac system. Additional modifications have been made to improve the efficiency of the lac-based system and these include the fusion of a nuclear localisation signal (NLS) to the lacR, which resulted in a greater accumulation of the repressor in the nucleus and a heightened sensitivity to IPTG induction (Hu and Davidson, 1991). Although this approach did remove the temperature-sensitivity, it also resulted in higher basal levels of uninduced expression, which appeared resistant to IPTG regulation (Liu *et al.*, 1992). Furthermore, cytotoxic levels of IPTG were required to fully induce the GOI (Figge *et al.*, 1988). Early attempts to generate transgenic mice harbouring lac-inducible gene expression generated disappointing results due to heavy methylation of the bacterial sequences (Scrable and Stambrook,

1997). More recently, however, the lacR has been codon-optimised to facilitate expression in the mammalian cells (Cronin *et al.*, 2001).

*The well-characterised tetracycline system is based upon the tet operon of E. coli*

A second prokaryote-based inducible system, the tetracycline system, was developed by Gossen and Bujard (1992) and was modelled on the transposon-10 specific tetracycline-resistance (Tet) operon of *E. coli* (Figure 1.16A). Similarly to the lac operon, expression from the tetracycline operators (TetO) is suppressed by highly specific binding of the tetracycline repressor homodimer (TetR; Figure 1.16A, top; Chopra, 1985). Addition of the antibiotic tetracycline (or a derivative such as doxycycline) alleviates the suppression as it prevents the TetR from binding to the TetO sequences (Figure 1.16A, bottom; Hillen and Berens, 1994).

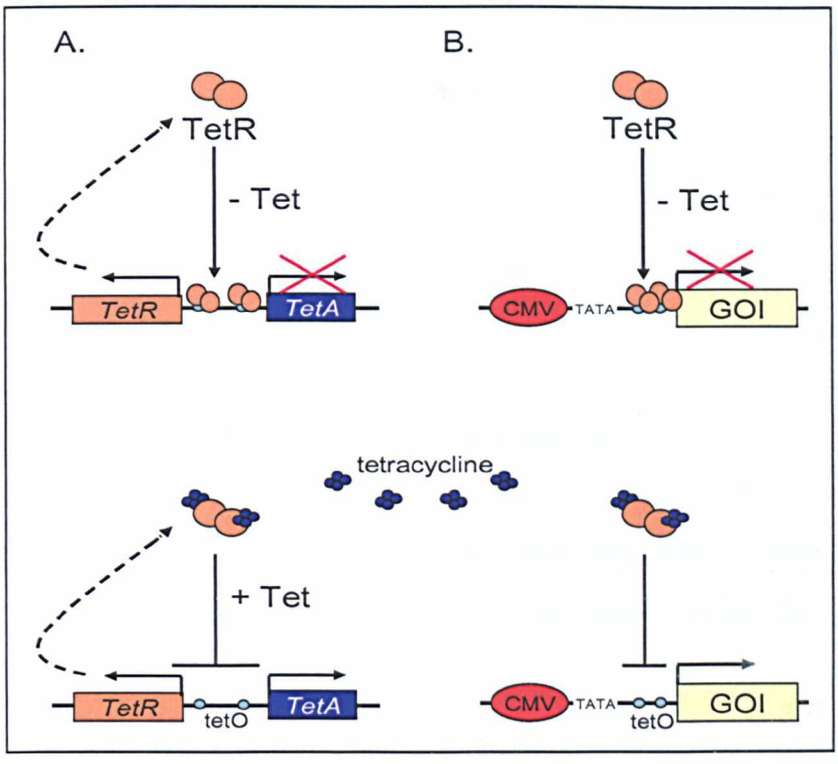



**Figure 1.16 The tetracycline system: A.** A schematic representation of the tetracycline operon in *E. coli*. The TetR is continuously expressed and in the absence of tetracycline it binds to the tetO sequences and prevents expression of the tetracycline resistance genes (tetA). When tetracycline is added the TetR cannot bind to the tetO sequences and the tetA genes are expressed. **B.** The T-REx system which uses the TetR to inhibit GOI expression in the absence of tetracycline. When tetracycline is present the TetR cannot bind to the two tetO sequences and the GOI is expressed.

Expressing the TetR in HeLa and rodent cells demonstrated TetO-linked GOI expression when tetracycline was added; however, an inadequate amount of TetR was produced to achieve tight regulation when tetracycline was removed (Gossen *et al*, 1993). Yao and colleagues (1998) appeared to have overcome this limitation by creating the T-REx system. This system deploys a complete CMV promoter followed by only two tetO elements that are bound in the absence of tetracycline by the original wild-type tetR (expressed from a strong CMV promoter) to inhibit expression of the downstream GOI (Figure 1.16B, top). To induce expression of the downstream gene, tetracycline is supplemented to the medium (Figure 1.16B, bottom). It is claimed that this system yields a higher level of induced expression than any other regulated mammalian expression system (Yao *et al.*, 1998), although recent literature demonstrates only a 50-fold induction value when regulating the transferin gene using the T-REx system (Jones *et al.*, 2005).

To adapt the tetracycline operon system further for highly inducible gene expression in mammalian cells the TetR was fused to the activating domain of the herpes simplex virus protein VP16, to create a transactivator (tTA) hybrid protein. The tTA can bind to a series of seven tandemly arranged TetO sequences linked to a minimal CMV promoter, termed a tetracycline response element (TRE), to drive expression of a downstream transgene. This arrangement is known as the Tet-OFF system (Figure 1.17A). Gene suppression is achieved by adding tetracycline as it prevents the tTA from binding to the TRE. Transient TRE-tagged luciferase expression has been shown to be tetracycline regulated in tTA expressing HeLa cells by up to five orders of magnitude (Gossen and Bujard, 1992). This could also be achieved by using very small concentrations of tetracycline and gene silencing was obtained in a dose dependant, rapid manner. Finally, unlike the lac-based system no temperature dependence was observed.

Although tetracycline is now thought to be non-toxic to eukaryotic cells at the low concentrations required for gene silencing, the effect of prolonged tetracycline exposure on transgenic animals and for gene therapy applications was previously unknown. It was also predicted that gene induction was being delayed due to residual tetracycline after its removal. To address this problem a second tetracycline system, the Tet-ON system (Figure 1.17B), was developed in which the TetR was altered so

that it required tetracycline for gene induction. This was achieved by randomly mutagenising the TetR and searching for a mutant with reverse DNA binding properties *in vivo*. The chosen mutant, termed the reverse tetracycline repressor (rTetR), contained alterations at four amino acids when compared to the wild-type TetR ($Glu^{71} \rightarrow Lys^{71}$, $Asp^{95} \rightarrow Asn^{95}$, $Leu^{101} \rightarrow Ser^{101}$ and $Gly^{102} \rightarrow Asp^{102}$ [Hillen and Berens, 1994]). These alterations did not affect tetracycline binding (Hinrichs *et al.*, 1994) and consequently the same changes were made to the tTA, generating a reverse tTA (rtTA; Gossen *et al.*, 1995). In this system tetracycline is required for the transcriptional activation of the GOI (Figure 1.17B).

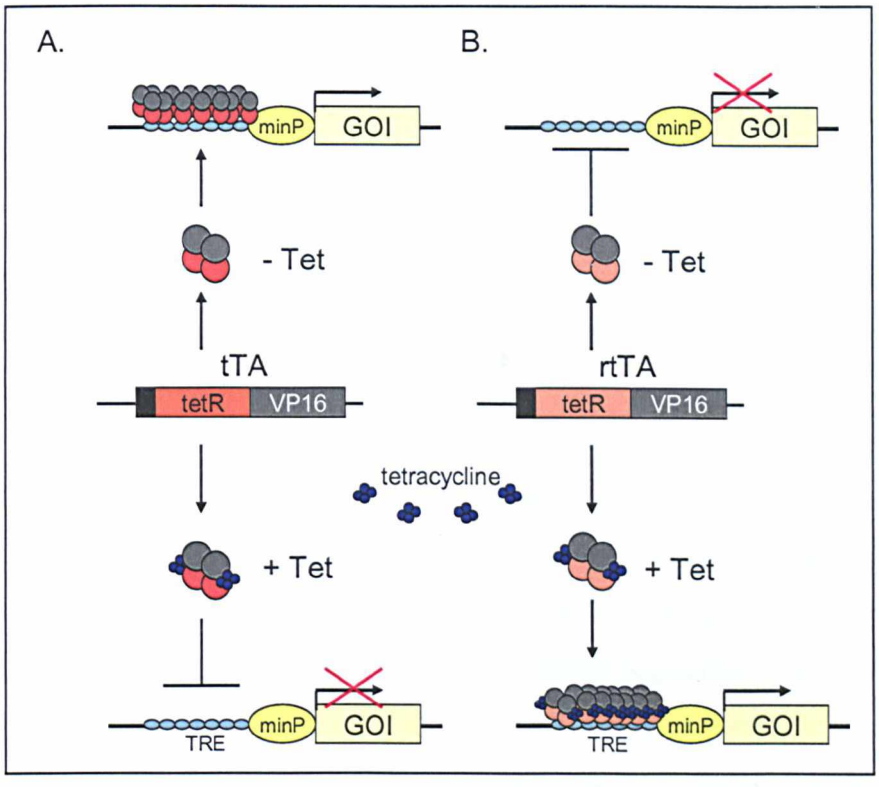



**Figure 1.17 The Tet-OFF and Tet-ON inducible systems:** **A.** The Tet-OFF system: The tTA (fusion of TetR and VP16 activation domain) can bind to the TRE (blue circles) in the in the absence of tetracycline to drive expression of the GOI. In the presence of tetracycline the tTA is inhibited from binding to the TRE and the GOI is down regulated. **B.** The Tet-ON System. In the absence of tetracycline the rtTA (reverse tTA) cannot bind to the TRE (blue circles) and the GOI is down-regulated. In the presence of tetracycline, however, the rtTA can bind to the TRE and GOI is expressed.

## 1.2.5 Choosing an appropriate inducible system for gene function studies

An ideal inducible system should facilitate elevated transgene expression which can rapidly convert to tight uninduced expression, and the reverse, at a particular time point (e.g. during development), in a dose dependant manner. The well-characterised tetracycline (Tet-OFF/ON) system is by far the most commonly used inducible system for generating regulated gene expression. This is largely due to the highly specific attraction between the TetR and the TetO and the TetR and the tetracycline molecules; and the vast amount of data available on the various tetracycline inducers (Gossen and Bujard, 1993). Compared to the lac system the tetracycline system has reduced basal expression levels (Liu *et al.*, 1992) and the inducer has a lower toxicity in mammalian cells (Figge *et al.*, 1988). The system also allows the expression over several orders of magnitude (up to $10^5$-fold) at concentrations that do not affect the rate of growth or morphology of cells in culture. In addition, the tetracycline regulatory system functions efficiently in transgenic mice due to the exceptional pharmacokinetic properties of tetracycline, allowing it to penetrate the placenta and the blood/brain barrier (Boothe, 1985).

Recently a comparison of the tetracycline and ecdysone system was carried out in the human embryonic retinoblastoma cell line 911 and it revealed that the Tet-ON system was shown to be the leakiest when uninduced, followed by Tet-OFF system. The ecdysone system not only demonstrated the tightest regulation but also the highest inducible expression (Meyer-Ficca *et al.*, 2004). The high uninduced basal expression levels observed for the Tet-ON and -OFF systems were thought to be due to the transient methods of analysis, as chromatinisation of the tetracycline promoter has been shown to be an important factor for tight regulation (Gossen and Bujard, 1992).

The ecdysone system may potentially be a more promising inducible system for the future, however, at present less is known about the effects of long term expression of the regulatory elements and exposure to the inducer molecules (muristerone A) on mammalian cells, transgenic mice and its suitability for therapeutic practices. It is therefore predicted that by optimising the tetracycline system for tight regulation it would produce an ultimate harmless approach for gene function studies, whose

advantages far outweigh that of any other inducible system, for use in vertebrate cells, mouse models and possibly gene therapy.

## 1.3 Approaches for optimising the tetracycline system further

### 1.3.1 Alternative methods to avoid laborious transfection procedures

Establishing the tetracycline system in cultured cells is a time-consuming two-step process; as it is firstly necessary to generate a cell line expressing the (r)tTA and then the TRE-tagged GOI can then be stably transfected. Apart from the interferon system (Kuhn et al., 1995; see Table 1.2), which requires only a single transfection step, this feature is typical of most inducible systems (e.g. Ecdysone system requires the transfection of 3 different constructs [EcR, RXR1 and GOI; see Figure 1.14]). Following the successful transfection of both constructs, cells must then be screened to find a clone demonstrating a desirable pattern of tetracycline regulated expression (Gossen *et al.*, 1992).

Creating stable (r)tTA expressing cell lines has proven difficult in the past as clones tended to lose their regulative capacity due to the toxicity of the protein when expressed at high levels (known as squelching [Gil and Ptashne, 1988; Gallia and Khalili, 1998]) and epigenetic silencing (Pikaart *et al.*, 1998). The (r)tTA toxicity is almost certainly a result of the VP16 activation domain (fused to the TetR) sequestering the host cells transcriptional machinery ('squelching') when expressed within the cells in large amounts (Barron *et al.*, 1997). Transcriptional silencing of the (r)tTA is likely to be caused by its non-mammalian genetic code (from *E. coli* [TetR] and Herpes Simplex Virus [VP16]); resulting in it being recognised by the host cell as foreign DNA for gene silencing by methylation. To limit these effects, an improved 'humanised' tTA and rTA have been created. They have been codon optimised, by eliminating potential splice sites and CG dinucleotides, for their enhanced stable expression in mammalian cells. In addition, the VP16 activating

domain is replaced with three synthetic F-domains to help reduce 'squelching' (Krestel *et al.*, 2004).

When generating mice with tetracycline inducible transgenes, insufficient production of the rtTA (or less often tTA for transgenic mice) in all tissues when expressed from a CMV promoter is often achieved (Chung *et al.*, 2002). An alternative strategy in ES cells, therefore, commonly involves gene targeting the rtTA into an endogenous locus (typically the ROSA26 locus) to allow its expression from the host cell's ubiquitous promoter (Wutz and Jaenisch, 2000). This approach would likely limit any 'squelching' effects caused by the VP16 activation domain by controlling the levels of rtTA expression; however, epigenetic silencing by methylation is not likely to be reduced, unless the codon optimised version was being targeted. Also, rtTA expression from the ROSA26 endogenous promoter was not ubiquitous, as the rtTA protein was not detected in the mouse mature neurone cells, therefore, applications involving these tissues are limited (Sonntag *et al.*, 2004).

Whilst establishing the tetracycline system in mammalian cells it is also recommended that the two constructs ((r)tTA and TRE-GOI) are transfected separately, so that they do not integrate into the same genomic site. This will avoid interference of TRE-regulation by the nearby CMV promoter, which is driving (r)tTA expression. Various strategies to overcome these lengthy procedures have been developed, such as self-regulatory systems in which the (r)tTA is also under the control of a bi-directional TRE (Figure 1.18 [Shockett *et al.*, 1995; Mohammadi and Hawkins, 1998; Strathdee and McLeod, 1999]). An advantage of this approach is that little tTA expression is produced during the suppressed state and therefore should reduce any transcriptional squelching. This does also suggest that 'leaky' expression from the TRE is required to re-induce GOI/tTA expression when tetracycline is removed from the medium (Figure 1.18A). This would not be acceptable for applications where tight regulation is a priority. In addition, this system only achieves moderate levels of induction (~200-fold in HeLa cells; Monammadi and Hawkins, 1998) and has also been reported to be cytotoxic (Gallia and Kjalili, 1998).

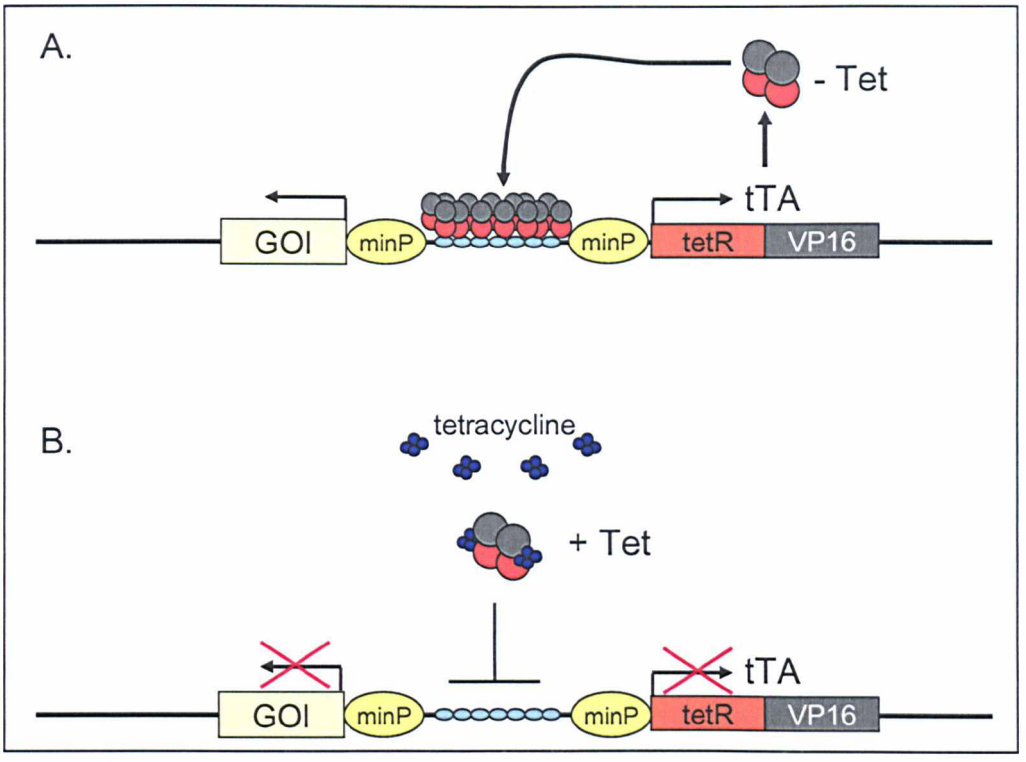



**Figure 1.18 Autoregulated tTA expression from a bi-directional TRE: A.** Binding of the TRE by tTA in the absence of tetracycline not only results in expression of the GOI but also more tTA (autoregulation). **B.** When tetracycline is added, the tTA cannot bind to the TRE and the GOI and tTA are down regulated. For key to symbols see Appendix I.

## 1.3.2 Optimising the tetracycline system for tight regulation

Despite such advances in tetraycline inducible technology their applications remain limited by the difficulty in isolating clones with tightly regulated transgenes. An example of this limitation was documented when the tetracycline system was used for a potential gene therapy approach to kill mouse cancer cells. The application of this treatment was limited due to leaky expression of the uninduced diphtheria toxin A transgene from the TRE, which resulted in the inhibition of protein synthesis, followed by cell death (Paulus *et al.*, 1997; Keyvani *et al.*, 1999).

Depending on the inducible transgene to be expressed or suppressed, the slightest dribble of expression from the TRE when uninduced may result in, for example, cell death (leaky expression of a toxic gene e.g. *RAD52,* diphtheria toxin), uncontrolled

cutting (leaky expression of an enzyme e.g. I-*SceI*), uncontrolled recombination (leaky expression of a recombinase e.g. Cre), or uncontrolled gene silencing (e.g. leaky shRNA expression). It is important under these circumstances that expression can be delayed until a specific time of observation or development (transgenic mouse). In addition, to create a knockout phenotype, complete transgene suppression is required, as leaky expression may complement for the loss of expression from the endogenous alleles (Carpenter and Porter, 2004).

Clonal variations in transgene induction and basal expression levels result from position effects at the site of random integration of the TRE-tagged GOI transgene. These position effects are caused by the integration of the inducible cassette nearby enhancer elements which have been reported to interfere with expression from the TRE (Damke *et al.,* 1995). Such effects vary between chromosomal locations and an 'ideal' integration site has not yet been established. The isolation of tightly regulated clones, therefore usually involves tediously screening numerous colonies for inducible transgene expression, usually by Western analysis. Various methods for reducing position effects have been investigated, including flanking the inducible cassette with insulators from the chicken β–globin HS4 region. It was hoped that this would guard against nearby enhancers; however moderate basal expression levels were still observed (Anastassiadis *et al.,* 2002; Izumi and Gilbert, 1999). This suggested that additional factors were leading to leaky uninduced expression and that there may be residual activity of the tTA at the TRE promoter. To investigate this, the tTA and the rtTA were fused to a steroid binding domain and were sequestered in an inactive state until the appropriate androgen was added. This technique was shown to greatly improve leaky expression levels for the Tet-ON system, but was less successful for the Tet-OFF system (Anastassiadis *et al.*, 2002; Qu *et al.,* 2002).

Another consideration when implementing a tetracycline inducible system is that the induction values and uninduced expression levels vary between cell types. The tetracycline system was originally developed in HeLa cells (Gossen and Bujard, 1992), using a *luciferase* reporter gene. They observed that luciferase could be induced in stable HeLa clones by up to $10^5$-fold and uninduced basal expression levels were barely detectable. Induction of reporter genes in other cell types, however, has been reported to be reduced when compared to HeLa cells (Ackland-

Berglund and Leib, 1994; Howe *et al.*, 1995). Previously, induction values generated for stable inducible luciferase in HT1080 cells were shown to be up to 40-fold (Sullivan et al., 2001) and 60-fold in mouse embryonic fibroblasts (Gould *et al.*, 2000).

The difference in inducibility of tetracycline-regulated gene expression in different cell lines appears to be due to a high basal uninduced expression level in those cells, resulting from additional factors which are capable of activating the TRE rather than a reduced induced expression. In fact, transient TRE driven luciferase was even observed for non-tTA expressing HT1080 cells (personal communication, Carpenter). Reports have also demonstrated the activation of the TRE by viral proteins (Herrlinger *et al.*, 2000), interferon activity (Rang and Will, 2000) and more recently GATA transcription factors (Gould and Chernajovsky, 2004). The latter recognition sites lie directly on the TetO repeat sequences and could potentially compete with the tTA for binding. Since GATA has altered expression in many different cells (especially cancer cells), it may help explain the varying degrees of gene regulation that are observed between cell types. In addition, examination of the CMV region of the TRE promoter by Gould and Chernajovsky (2004), using the MacDNASIS Pro V.3.6. database, revealed other potential transcription factor binding motifs, including AP2 (Williams and Tjian, 1991), LFA-1, IFN-γ, XRE-I and GCF (Figure 1.19).

Within the last few years a new TRE-containing vector has become commercially available, called pTRE-TIGHT (clontech), which claims to achieve minimal uninduced expression levels. This has been obtained by altering the sequences between the TetO repeats (Figure 1.19; in grey) so that they are no longer bound by additional factors (e.g. interferon) other than the (r)tTA. The recent report suggesting that the TRE is also recognised by GATA transcription factors was presumably not known at this time and they still remain in the tetO repeat sequences. Finally, additional technology for generating tight regulation by altering the properties of the TRE include the creation of a second-generation TRE (SG-TRE); which utilises a shortened CMV promoter along with eight tetO sequences (Mohommadi *et al.*, 2004). In HeLa cells, undetectable uninduced expression was obtained, which in turn facilitated an induction value approaching 100,000-fold (transiently).

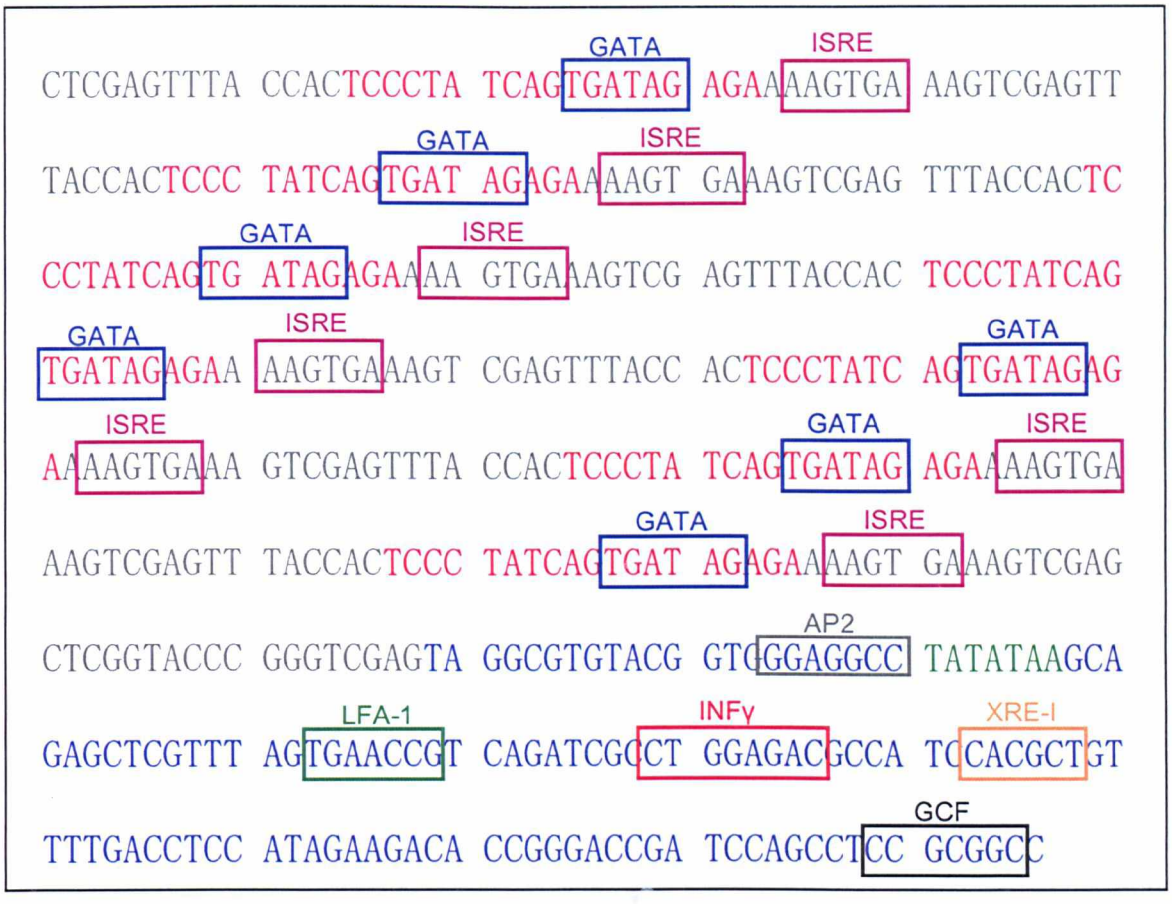



**Figure 1.19 Sequence analysis of the TRE:** The TRE is made up of seven TetO repeats (in red) and a CMV minimal promoter (in blue). Transcription factor binding motifs are illustrated by the boxed areas. ISRE - interferon response element.

### 1.3.3 Strategic approaches for an optimised tetracycline system

Various systems have been designed to overcome the limitations of variable expression levels observed between clones due to position effects. One such strategy is the commercially available Flp-In™ T-REx™ system (Invitrogen), where stable cell lines expressing the TetR and containing a pre-integrated FRT site facilitates the repeated integration of a transgene into the same locus to eradicate expression variability between clones (Figure 1.20). In addition, the integrated transgene is under the control of a T-REx promoter and can be inserted into the genome whilst down-regulated, an attractive feature when attempting to overexpress a gene whose product may compromise cell's survival (e.g. *RAD52* (Yanez and Porter, 2002) and *diptheria toxin* (Paulus *et al.*, 1997; Keyvani *et al.*, 1999)). A limitation of this strategy,

however, is that the site of repeated integration is not pre-screened for induction and low basal uninduced expression levels, prior to integration of the GOI. In addition, the presence of the nearby promoter elements driving the drug resistance cassette is likely to interfere with expression from the inducible promoter (Damke *et al.*, 1995).

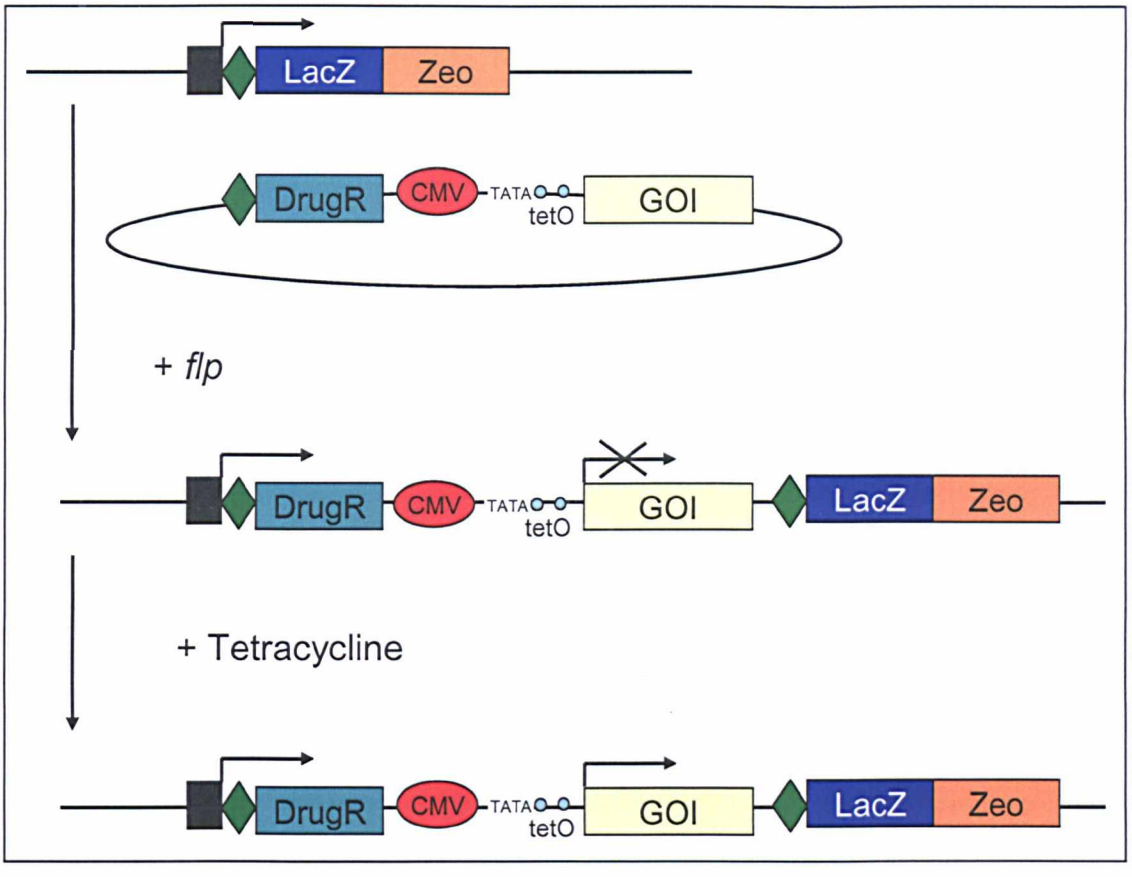



**Figure 1.20 The Flp-In™ T-REx™ system (Invitrogen):** The chief features of the Flp-In T-REx system for generating clones with minimal variations in GOI expression are illustrated. Firstly, a stable cell line expressing lacZ/Zeo and containing an FRT site (Green diamond) is generated. Secondly, Flp-mediated insertion of a construct containing a drug resistance cassette and a TRE-driven GOI is targeted into the FRT site within stably generated clone. After selection in the appropriate drug tetracycline is added to induce GOI expression. For key to symbols see Appendix I.

A 'select and invert' system has also been developed by Sullivan and colleagues (2001), which exploits a positively and negatively selectable marker gene to identify those rare tightly-regulated clones (Figure 1.21). The marker gene and a silent GOI, positioned in the reverse orientation, were flanked by *loxP* sites. After a tightly regulated clone has been isolated, expression of Cre recombinase catalyses an inversion event to place the GOI downstream of the tightly regulated promoter

(Figure 1.21). The Cre-mediated inversion event took place in approximately 16% of unselected clones, therefore, the GOI could potentially be repressed whilst inverted clones were being isolated. It may have been possible to increase this frequency further by using a combination of mutant *loxP* sites (e.g. *lox71* and *lox66*; see Figure 1.10) to reduce unwanted subsequent inversion events (Araki *et al.*, 2001). This strategy was largely limited, however, by the low stringency of the selection method failing to adequately isolate tightly-regulated clones. In addition, for every new GOI a further tightly regulated clone had to be identified, which could potentially lead to clonal variations in expression levels.

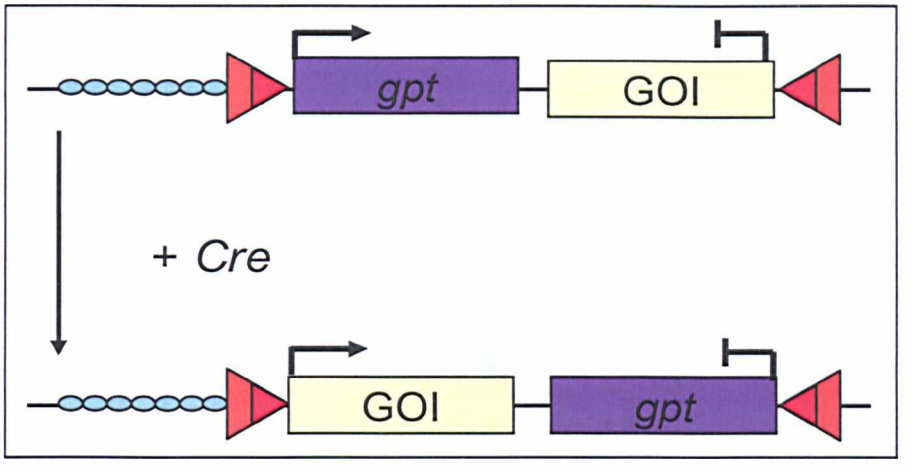



**Figure 1.21 The 'select and invert' system:** A positively and negatively selectable marker gene (*gpt*) is used to more simply screen for clones demonstrating tight regulation from the TRE (blue circles). The *gpt* gene is linked to a promoterless GOI in the reverse orientation and the two ORF are flanked by inverted *loxP* sites (red triangles). Clones in which the *gpt* gene is tightly regulated are selected and upon transient expression of Cre an inversion event places the GOI under the control of the TRE. For key to symbols see Appendix I.

Earlier this year, a targeted transgenesis strategy was developed which, to date, appears to be the most desirable system for generating clones with tightly regulated transgenes (Puttini *et al.*, 2005). Puttini and colleagues describe a system that firstly involves randomly integrating a TRE driven puromycine resistance cassette, linked by an internal ribosomal entry site (IRES) to the SEAP gene, into the genome of RCCD1 epithelial cells (Figure 1.22 (top); Blot-Chabaud *et al.*, 1996). The IRES facilitates the expression of both the puromycin and SEAP cassettes from the TRE and the SEAP gene enables the tightness and inducible expression from the TRE to be measured using a simple assay. Downstream of the SEAP gene is an I-*SceI*

recognition site and the 3'end of the hygromycin gene (non-functional; Figure 1.22). After clones with tight regulation have been identified (using expression of the SEAP gene as an indicator) I-*SceI* mediated gene targeting is carried out, by transfecting a construct containing a TRE-linked GOI (with 0.9kb of homology to the target locus) and the promoter and 5' end of the hygromycin gene (containing 0.7kb of homology to the target locus) (Figure 1.22). The gene targeting event places the GOI downstream of the tightly regulated TRE and corrects the two non-functional hygromycin resistance cassettes, to generate a method of selection for desirable clones (Figure 1.22).

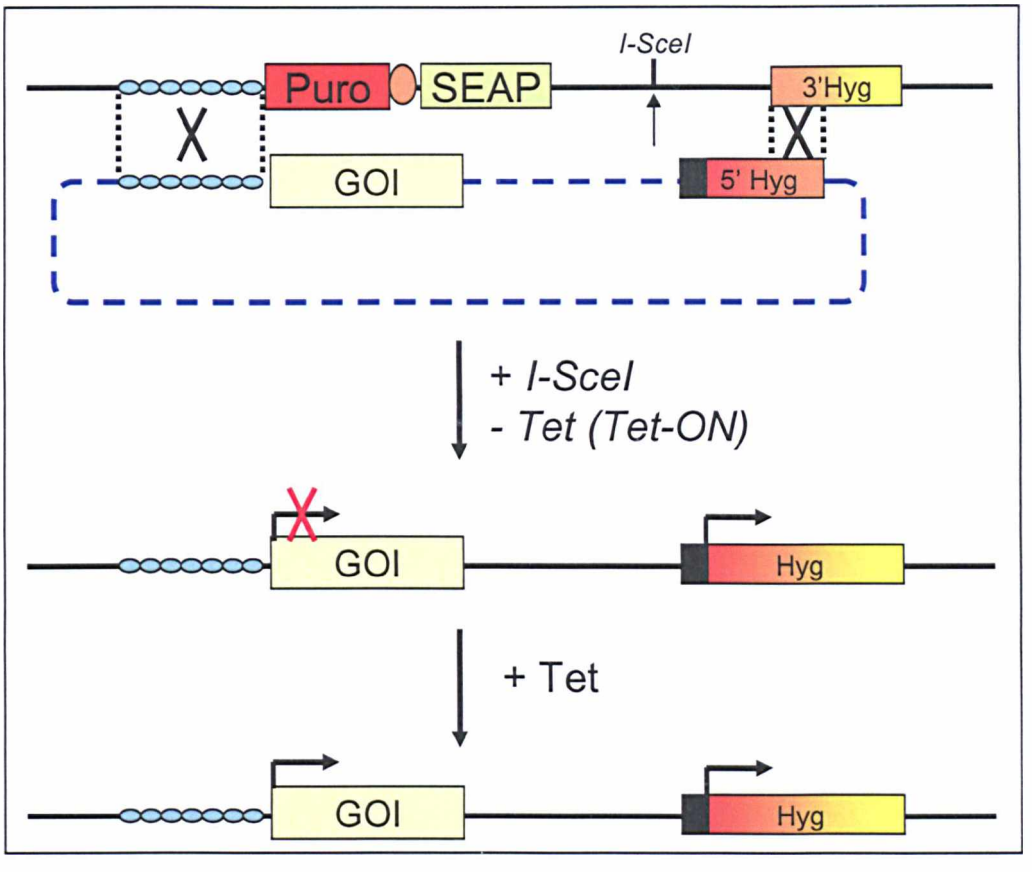



**Figure 1.22 The targeted transgenesis strategy:** Outline of the strategy used for targeted, double strand break-mediated transgenesis (Puttini *et al.*, 2005). Puromycin-resistant clones expressing tightly regulated SEAP are selected and used for the gene targeting event. The targeting construct is co-transfected with an I-*SceI* expressing vector which encourages a gene targeting event which places the GOI downstream of the tightly regulated promoter. This event can be carried out in the absence of tetracycline (Tet-ON system, therefore the GOI is down-regulated) and the TRE can be induced when its effects can be observed. Internal ribosomal entry site (IRES) – orange oval. For key to other symbols see Appendix I.

This approach not only allows the pre-selection of tightly regulated clones prior to the targeting event, it also facilitates the selection of targeted clones whilst the GOI is continuously suppressed (in the absence of Tet using the Tet-ON system); which is advantageous for examining the effects of toxic genes at a particular time point. In fact, Puttini and colleagues successfully generated clones expressing tightly regulated *4.1N* gene, whose product is involved in interfering with cell proliferation (Ye *et al.*, 1999). Previous attempts to generate cells expressing this particular gene had failed (Puttini *et al.*, 2005). Despite these advantages, this system is limited by a low targeting efficiency and unpredictable TRE expression after the targeting event; most likely due to the insertion of the hygromycin genes enhancer elements effecting TRE regulation. In addition, variations in gene expression within a clone cannot be examined using the SEAP screening method, therefore a more sensitive approach for isolating uniformly expressing, tightly-regulated clones would be more desirable.

## 1.4 Project Goals

### 1.4.1 Project Aims

The aims of this thesis were to:

1. Design an optimised strategy for the easy generation of clones expressing tight, tetracycline regulated transgenes.

2. Utilise the developed system to examine the effects of a given GOI on genomic stability and homologous recombination.

### 1.4.2 Addressing aim one: Developing an optimised tetracycline system

The primary goal of this thesis was to develop an improved system to optimise the use of the tetracycline system in mammalian cells. The main criteria for an ideal tetracycline expression system are:

1. The easy isolation and characterisation of clones supporting tight TRE-driven expression.
2. The ability to repeatedly integrate various GOIs into the predetermined tightly-regulated genomic loci to eliminate clonal variations in GOI expression.
3. The GOI should be inserted under restrictive conditions so that analysis of the early effects of its expression can be delayed until a convenient time after integration.

To achieve the first condition the gene coding for the green fluorescent protein (GFP) was tagged to the TRE and used as a reporter to monitor the degree of regulation in cell clones, by flow cytometry. Flow cytometry can measure the fluorescence emitted

from individual cells within a clone (or a pool) and it was predicted that this would provide a more sensitive technique of analysing variations in protein expression levels, when compared to the traditional method of Western blot. After a clone stably expressing the TRE-GFP cassette from a genomic location that facilitates tight regulation was identified, it was intended that the GFP gene would then be swapped with a GOI. To carry out the gene replacement step, Cre-mediated insertion was planned using the two mutant *loxP* sites, *lox71* and *lox66* (Figure 1.10). These sites were positioned directly upstream of the inducible GFP gene in the stably integrated 'target' locus and upstream of a promoterless GOI within an 'insertion' construct (Figure 1.23). Various designs of this 'screen and insert' approach were tested and the endeavour to incorporate all of the above criteria into a simple system is described in this thesis.

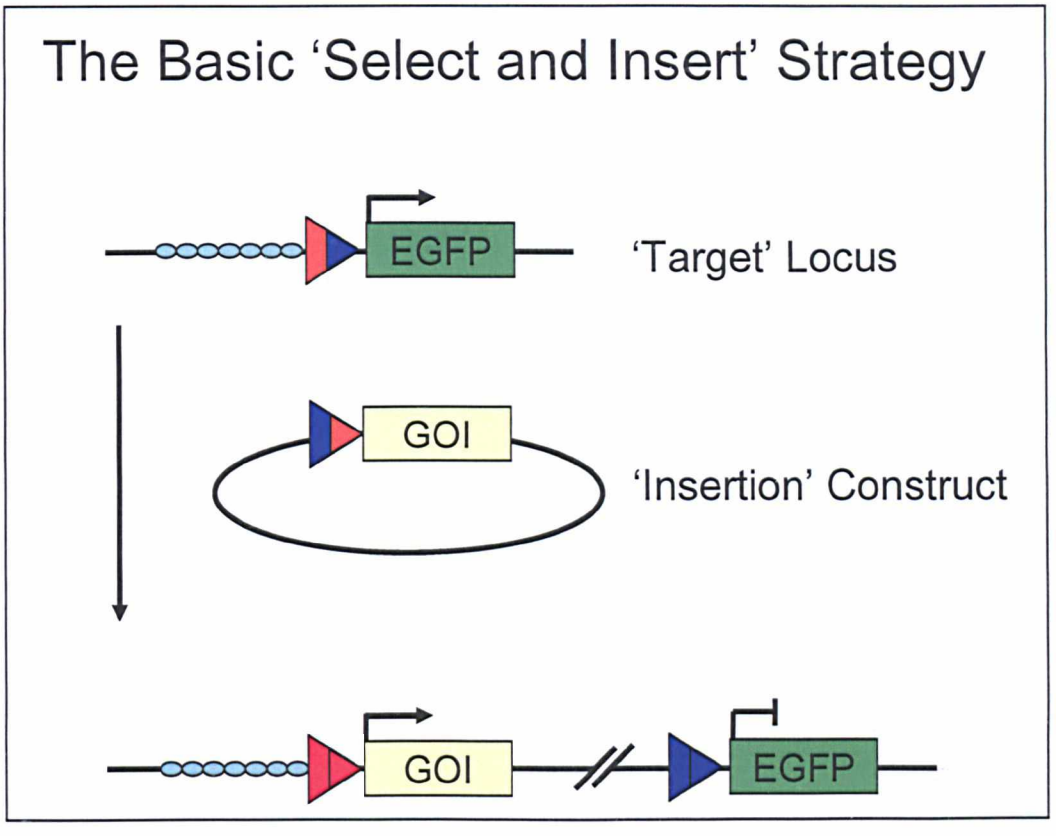



**Figure 1.23 The 'screen and insert' system:** Illustrates the basic 'screen and insert' approach. The GFP gene was used as a marker to identify clones demonstrating tight regulation using FACS analysis. Once tightly regulated clones were identified the GOI was inserted downstream of the tightly regulated TRE (blue circles) using Cre-mediated recombination. The mutant *loxP* sites, *lox71* ('target' locus) and *lox* 66 ('insertion' construct) were used to optimise the insertion frequency. For key to symbols see Appendix I.

### 1.4.3 Addressing aim two: Utilising the 'screen and insert' approach

After successful development of a cell line containing tightly regulated tetracycline expression using the 'screen and insert' approach, it was then deployed to study the effects of various GOIs that may affect genomic stability and gene targeting frequencies (research interests of the Gene Targeting Group). Candidate genes that were predicted to require tight inducible gene regulation for adequate functional analysis included toxic genes (e.g. *RAD52;* Yanez and Porter, 2002), essential genes (e.g. *RAD51*; Yanez and Porter, 1999a; *RAD50*; Luo *et al.*, 1999; *MreII*; Xiao and Weaver, 1997), enzymes (e.g. *I-SceI, Cre*; Schonig *et al.*, 2002) and shRNA (Matthess *et al.*, 2005). Due to time constraints the results are limited to conditional gene overexpression studies, rather than conditional gene silencing, by describing the successful generation of clones containing tightly regulated, never switched on before, I-*SceI* and *RAD52* genes.

# Chapter Two
# Materials and Methods

## 2.1 Preparation of nucleic acid

### 2.1.1 Isolation of plasmid DNA

#### 2.1.1.1 *Transformation of E.coli and culture conditions*

*E.coli* XL1-Blue competent cells (Stratagene, #200249) were used for standard cloning procedures and were transformed with plasmid DNA as per the manufacturers' instructions. Bacterial cultures were grown at 37°C with LB broth/agar, supplemented with ampicillin (50μg/ml, Sigma). When selecting for bacteria containing a zeocin expression construct they were grown in low salt LB broth/agar. All media and glassware were sterilised by autoclaving and the ampicillin was filter sterilised using a 0.2μm syringe filter (Sartorius).

#### 2.1.1.2 *Plasmid preparation*

Qiagen mini (#27196) and HiSpeed maxiprep (#12663) kits were used to prepare plasmid DNA from bacterial cultures by following the manufacturers' instructions.

### 2.1.2 Isolation of genomic DNA

A confluent plate of mammalian adherent cells were washed with PBS and covered with lysis buffer (0.5ml/cm$^2$) supplemented with 150 μg/ml of proteinase K (Roche). The culture dish was rocked at 37°C for at least 4 hours. An equal volume of isopropanol was added to the lysate and the culture dish was rocked at room temperature until the DNA precipitated. The DNA was retrieved using the end of a plastic pipette tip and after drying off residual isoproponol, was then resuspended in sterile TE.

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

## 2.1.3 Oligonucleotide synthesis

Gene Jockey (Biosoft) was used to design approximately 20 base pair primer sequences (see Table 2.1) by selecting for those with a 45-55% CG contents. Oligonucleotides generating synthetic *loxP* and FRT sites (Table 2.2) were based on published results by Araki *et al.* (1997) and Luetke *et al.* (1997), respectively. Oligonucleotides were purchased from sigma-genosys. The lyophilised nucleotides were resuspended in sterile water to a concentration of 1µg/µl and stored at -20°C.

**Table 2.1 PCR and sequencing primers**: The oligonucleotides used for PCR and sequencing.

| PrimerName | Primer Sequence (5'-3') | Used For: |
|---|---|---|
| Forward | | |
| TetOF | TTTAGTGAACCGTCAGATCGCC | PCR and Sequencing |
| TIGHT1 | ACGAGGCCCTTTCGTCTTCA | PCR and Sequencing |
| INSF | GGATCTGCGATCTAAGTAAG | Sequencing |
| LOXGPTF | AAGTGCAGGTGCCAGAACAT | Sequencing |
| Reverse | | |
| LucR | CCCCTTTTTGGAAACGAACAC | PCR and Sequencing |
| LOXGPTR | AACATGTCCCAGGTGACGAT | Sequencing |
| NeoR | AGAAGGCGATAGAAGGCGATGC | PCR |
| HygR1 | CGCTCGTCTGGCTAAGAT | PCR |
| BSgptR | CCACGGCTTACGGCAATAATGC | PCR |
| RevTIGHT1 | TCCTCTAGAGATATCGTCGAC | PCR |
| ISCE1R | ACTCGAACTGCATACAGTAG | PCR |
| RAD52-2R | CCAAAGATAAAGCCTTGGAC | PCR |

**Table 2.2 Oligonucleotide Pair Sequences**: The oligonucleotides used to generate *loxP*, FRT & MCS.

| Oligo | Sequence (5'-3') |
|---|---|
| lox71F | TACCGTTCGTATAGCATACATTATACGAAGTTAT |
| lox71R | CTAGATAACTTCGTATAATGTATCGTATACGAACGGTAGC |
| lox71RBF | AATTCTACCGTTCGTATAGCATACATTATACGAAGTTATACTAGTG |
| lox71RBR | GATCCATCAGTATAACTTCGTATAATGTATGCTATACGAACGGTAG |
| lox66A | TCGAGAATTCATAACTTCGTATAGCATACATTATACGAACGGTAG |
| lox66B | AGCTCTACCGTTCGTATAATGTATCGTATACGAAGTTATGAATTC |
| lox66RBS | CTAGATAACTTCGTATAGCATACATTATACGAACGGTAGAAT |
| lox66RBA | TCGAATTCTACCGTTCGTATAATGTATGCTATACGAAGTTAT |
| FRTT1 | AATTAGATCTGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC |
| FRTT2 | AATTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCAGATCT |
| FRTI1 | AGCTGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCGAATT |
| FRTI2 | CATGAATTCGAAGTTCCTATACTTTCTAGAGAATAGGAACTTC |
| MCSinsF | CGCTAGCAGCTGGTCCGCGGACTAGTCCCGGGAAGCTTCTCGAGAGGCCTCATATGCATGCCATGGCCGG |
| INSmcsRcor | CCATGGCATGCATATGAGGCCTCTCGAGAAGCTTCCCGGGACTAGTCCGCGGACCAGCTGCTAG |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

### 2.1.4 Quantification of nucleic acid

DNA (5µl) was diluted in sterile $H_2O$ (950µl) in a clean cuvette and a spectrophotometer (Eppendorf, BioPhotometer) was used to measure the Optical Density (OD) at $260_{nm}$. The concentration of DNA was calculated using the following formula:

$$[DNA]\mu g/\mu l = OD260_{nm} X \{dilution\ factor\ (1/200=X200)\} X\ 0.05\mu g/\mu l.$$

Alternatively, DNA was quantified using ethidium bromide-stained agarose gels to compare known quantities of DNA under ultraviolet (UV) light (see 2.2.3).

## 2.2 DNA manipulation

### 2.2.1 Restriction digestion of plasmid DNA

Plasmid DNA was digested with New England Biolabs (NEB) restriction enzymes for at least 1.5 hours at 37°C (unless otherwise recommended in the NEB catalogue). Digestions were carried out using 1 Unit (U) of enzyme per µg of plasmid DNA in 1X BSA and 1X the appropriate buffer (NEB).

### 2.2.2 Restriction digestion of genomic DNA

Typically, genomic DNA from 100,000 cells was digested with 40U of restriction enzyme (see 2.2.1) in a total volume of 500µl. The digest was supplemented with spermidine (1mM, Sigma) and was generally incubated overnight. The following day an additional 20U of restriction enzyme was added plus 50µg/ml of RNAse (Boehringer Mannheim) and incubated for a further 2 hours.

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

## 2.2.3 Gel electrophoresis

Agarose (BioGene LTD) gels (ranging from 0.5-2%) were prepared using 1XTAE buffer and 0.1μg/μl of ethidium bromide (Invitrogen). The gels were placed inside the electrophoresis tanks (GIBCO Horizon 11.14 system) and were filled to the appropriate level with 1XTAE. Prior to loading, DNA samples were combined with 10 X loading buffer. Electrophoresis was carried out in a constant field of 100V and the DNA was visualised by transillumination with UV light at 312nm. The DNA ladder used to predict fragment sizes was the 1KB ladder from Invitrogen. It was made to a concentration of 0.1μg/μl using loading buffer (1X) and $H_2O$ and 10μl (1μg) were typically loaded onto the gel.

## 2.2.4 DNA clean-up

### 2.2.4.1 *Phenol/Chloroform treatment and precipitation of DNA*

To clean up the DNA, the DNA solution was made up to an appropriate volume with $H_2O$ (at least 100μl). This volume was then doubled by adding half phenol and half chloroform (at 1:1 ratio). The mixture was mixed well by vortex then centrifuged for 2 minutes at 13,201gs (Sigma Laborentifugen 3K10E centrifuge). The DNA-containing aqueous layer was removed and placed into a fresh tube.

To precipitate the DNA, twice the sample volume of 100% ethanol and 1/10 of the sample volume of 3M sodium acetate (pH 5.2) was added. This sample was mixed well and placed at -20°C for at least 30 minutes. The mixture was then spun at 13,201gs for 30 minutes. The supernatant was removed and the pellet was washed in 70% ethanol. The mixture was centrifuged again and the supernatant was removed. The DNA pellet was allowed to air-dry before resuspending in sterile TE.

#### 2.2.4.2 *Purification of DNA fragments from agarose gels*

DNA bands, visualised under UV light, were removed from agarose gels using a scalpel blade. The DNA was purified from the agarose using the QIAquick gel extraction kit (Qiagen).

### 2.2.5 Preparation of DNA fragments for cloning procedures

To dephosphorylate the 5' terminal groups of plasmid DNA, Calf Intestinal Phosphatase (CIP, New England Biolabs) was used, as described in the manufacturers' instruction. To fill-in the 5' overhang of plasmid DNA to form blunt ends a restriction digest was supplemented with dNTPs (0.1mM, Pharmacia), the DNA polymerase Klenow (1U/μg DNA ends) and the enzyme buffer (New England Biolabs) and left for 15 minutes at room temperature as per the manufacturers' instructions. DNA fragments for cloning were generated by restriction digest, often preceded by one of the above treatments, followed by gel purification. Ligation of DNA ends was carried out using T4 DNA ligase (New England Biolabs) in a 20μl reaction volume and incubated at 16°C overnight.

### 2.2.6 Radiolabelling DNA

After plasmid digestion and gel extraction, 25ng of short DNA fragments were labelled with 25μCi of $\alpha^{32}$P dCTP (Amersham) using the Random Primed DNA Labelling Kit (Roche) as per the manufacturers instructions. The 25μl reaction was incubated for at least one hour at 37°C then the volume was increased to 100μl using sterile TE. The unincorporated label was removed by passing the mixture through a MicroSpin S-300 HR column (Amersham Biosciences). The probe volume was then made up to 1ml by adding 50μl of salmon testes DNA (10mg/ml, Sigma) and 850μl of water. Before the labelled probe was added to the hybridisation buffer/membrane it was firstly incubated at 95°C for 10 minutes, followed by 15 minutes on ice.

The 1Kb ladder (Invitrogen) was end-labelled with $\alpha^{32}$P dCTP in 1 x nick translation buffer, ATG mix (0.123mM, Pharmacia) and Klenow (5U, Promega) in a final volume of 20µl and incubated at room temperature for 1.5 hours. The mix was increased to 100µl using TE and the unincorporated label was removed by passing the mixture through a MicroSpin S-300 HR column (Amersham Biosciences). The ladder is then made up to 1ml in TE and 1 x loading buffer and stored at -20°C until required.

### 2.2.7 Oligonucleotide phosphorylation and hybridisation

Oligonucleotide pairs (see Table 2.2) were independently diluted in sterile $H_2O$ to a concentration of 0.1µg/µl. 4µl of each oligonucleotide were combined and heated to 80°C for 5min, followed by 5min on ice, then phosphorylated using T4 polynucleotide kinase (PNK, NEB) as described in the manufacturers' instructions. Annealing took place after heat inactivating the PNK at 65°C for 20min and leaving to cool at room temperature.

## 2.3 Vector cloning

### 2.3.1 Generating transactivator (tTA) expressing constructs

The vector pRK5-itTA, containing an 'improved' or humanised transactivator (htTA) gene, was kindly donated by Rolf Sprengel (Krestel *et al.*, 2004). This vector was cut with *Eco*RI and *Bam*HI releasing the htTA open reading frame (ORF) and was subsequently cloned into the *Eco*RI/*Bcl*I sites of pZeoSV (Invitrogen) to create pZeoSVhtTA.

## 2.3.2 Generating the 'target' constructs

For scale plasmid maps of each of the six 'target' vectors created see Appendix II.

### 2.3.2.1 *Generating 'target' constructs containing the original TRE.*

'Target' Construct One:

pUHD10-3 (www.zmbh.uni-heidelberg.de/bujard/homepage.html) was donated by H. Bujard (Gossen and Bujard, 1992). pTRElox71 was created by cloning the annealed lox71F/R oligonucleotide pair (Table 2.2) containing *Sac*II/*Xba*I compatible ends into the same sites in pUHD10-3. The EGFP ORF was isolated as an *Xba*I fragment from pEGFP (Clontech) and inserted downstream of the *lox71* site into the *Xba*I site of pTRElox71, generating pTRElox71EGFP (Appendix II).

'Target' Construct Two:

The ORF of d2EGFP (from pd2EGFP, Clontech) was extracted as a *Not*I/*Age*I fragment and inserted into the same sites of pTRElox71 (see 2.3.2.1) to generate pTRElox71d2EGFP (Appendix II).

### 2.3.2.2 *Generating 'target' constructs using pTRE-TIGHT (Clontech)*

'Target' Construct Three:

pTRE-TIGHT (Clontech) was digested with *Eco*RI and *Bam*HI and the lox71RBF/R (Table 2.2) oligonucleotide pairs were cloned into these sites, creating pTlox71. pTlox71 and pd2EGFP (Clontech) were digested with *Not*I and *Bam*HI and the d2EGFP ORF was inserted downstream of the *lox71* site to produce pTlox71dGFP (Appendix II).

'Target' Construct Four:

pTlox71dGFP was cut with *Afl*III and then the overhangs were end-filled. A puromycin cassette (including promoter and polyA) was cut from pBL-Puro/R (for construction information see Yun *et al.*, 2004) with *Eco*RI and *Xho*I and

was also end-filled. The puromycin cassette was cloned downstream of the d2EGFP ORF (in the opposite orientation) to generate pTlox71dGFPpuro (Appendix II).

'Target' Construct Five:

pTlox71dGFPiresHyg (Appendix II) was generated by cloning the *Nhe*I/*Bsp*HI IRESHyg fragment from pIRESHyg3 (Clontech) into the *Xba*I/*Afl*III site of pTlox71dGFP.

'Target' Construct Six:

An FRT site was added to the 'Target' construct by cloning the FRTT1/2 oligonucleotide pair (Table 2.2) into the *Hind*III/*Nco*I site of pTlox71dGFP to create pTFRTlox71dGFP (Appendix II).

## 2.3.3 Generating the 'insertion' constructs

For scale plasmid maps of each of the 'insertion' vectors created see Appendix III.

### 2.3.3.1 *Generating plox66Luc, plox66Hyg and plox66Neo*

plox66Luc:

Oligonuclotide pairs, *lox66*A/B (Table 2.2), were cloned into the *Xho*I/*Hind*III site of pGL3Basic (Clontech) to created plox66Luc (Appendix III).

plox66Hyg:

plox66Luc was cut with *Nco*I and pIRESHyg3 (Clontech) was cut with *Hind*III and both fragments were end-filled. Following this, they were digested with *Xba*I and ligated together to place the hygromycin ORF downstream of the *lox66* site, creating plox66Hyg (Appendix III).

plox66Neo;

plox66gptLuc (see 2.3.3.3) and pSV2neo (Southern *et al.*, 1982) were both cut with *Bgl*II and *Bam*HI. The neomycin ORF was then cloned downstream of the *lox*66 site to create plox66Neo (Appendix III).

### 2.3.3.2 *Generating plox66LucIresHyg*

plox66LucIresHyg:

pIRESHyg3 (Clontech) was cut with *Xho*I/*Nhe*I and the IRES-Hyg fragment was cloned into the *Xba*I/*Sal*I site of plox66Luc to generate plox66LucIresHyg (Appendix III).

### 2.3.3.3 *Generating pINSluc, pINSmcs and pINSneoMCS*

pINSluc:

The *Bam*HI and *Bgl*II *gpt* ORF and polyA fragment from pBSgpt (Itzhaki and Porter, 1991) was cloned into the *Bgl*II site of pGL3-Basic (Clontech), to create pGL3gptLuc. Annealed *lox*66(RBS/A) oligonucleotide pairs (Table 2.2) were cloned into the *Nhe*I/*Xho*I sites of pGL3gptLuc to generate plox66gptLuc. Oligonucleotide pairs FRTI1/2 (Table 2.2) were annealed and cloned into the *Hind*III/*Nco*I site of plox66gptLuc create pINSluc (Appendix III).

pINSmcs:

pINSluc was cut with *Bst*BI and *Fse*I to remove the luciferase ORF and the annealed multiple cloning site olignucleotide pairs (MCSinsF/INSmcsRcor) (Table 2.2) were then cloned into this site to generate pINSmcs (Appendix III).

pINSneoMCS:

pINSmcs and plox66neo were cut with *Bgl*II and *Pfm*I and the *gpt* ORF was replaced with the neomycin ORF to create pINSneoMCS (Appendix III).

#### 2.3.3.4 *Generating pINSneoISCE1and pINSneoRAD52*

pINSneoISCEI and pINSneoRAD52:

pINSneoMCS was cut with *Eco*RI, end-filled then cut with *Sal*I. pFB580 (contains human RAD52 cDNA cloned into pUC18; donated by F. Benson and S. West, ICRF; see Yanez and Porter, 2002) and pCMV-I-SceI (gift from M. Jasin; see Rouet *et al.*, 1994) were also cut with *Eco*RI, end-filled then cut with *Sal*I to release the *RAD52* and *I-SceI* ORF, respectively. *RAD52* and *I-SceI* ORF were then cloned into pINSneoMCS to generate pINSneoRAD52 and pINSneoISCEI, respectively (Appendix III).

### 2.3.4 Generating pTIGHTgpt and pTIGHT-ISCE1

pBSgpt (Itzhaki and Porter, 1991) was cut with *Bgl*II and *Bam*HI and the *gpt* ORF was cloned downstream of the TRE into the *Bam*HI site of pTRE-TIGHT to create pTIGHTgpt. pTRE-TIGHT (Clontech) and pCMV-I-SceI were both cut with *Sal*I and *Eco*RI and the *I-SceI* ORF was cloned downstream of the TRE to create pTIGHT-ISCEI.

### 2.3.5 Sequencing of plasmid DNA and PCR products

Plasmid DNA and PCR products were sequenced using 18-22bp primers (for sequences see Table 2.1) designed to span the region of interest. The sequencing reaction consisted of a 3.2pMol primer concentration plus 100ng/kb of plasmid DNA or 90ng/kb of PCR product. A sequencing service was used, provided by The Genomics Core Laboratory (MRC), who use ABI3730xl and ABI 3130xl sequencers (Perkin Elmer) which provide high-throughput sequencing results. The resulting traces were retrieved using FETCH (Softworks) software and analysed using ABI-PRISM (Perkin Elmer) and Gene Jockey (Biosoft) software.

## 2.4 Polymerase Chain Reaction (PCR) analysis

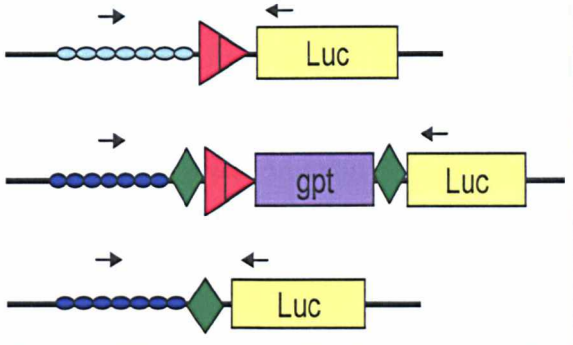
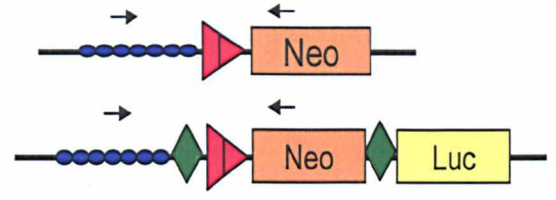
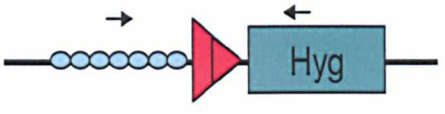
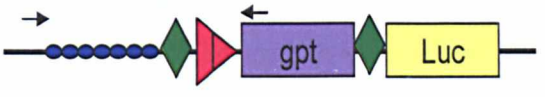
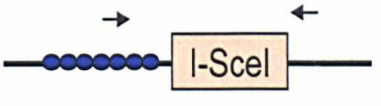
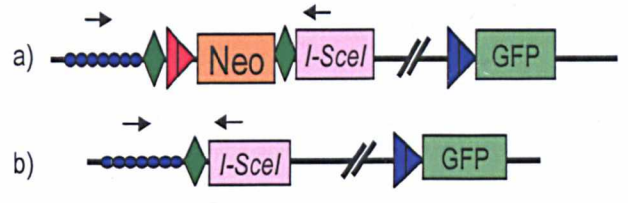
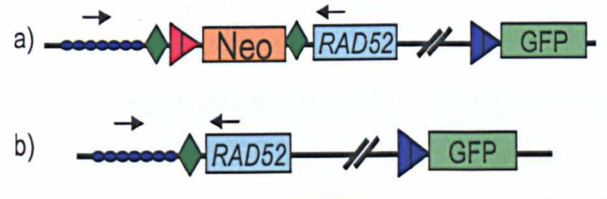
### 2.4.1 General PCR conditions

PCR was carried out using a Gene E thermal cycler (Teche). Reactions were made up in a total volume of 25µl using dNTPs (0.25µM, Pharmacia), primers (100ng each, Sigma-Genosys), reaction buffer (1X, Qiagen), Taq polymerase (1.25U. Qiagen) and plasmid DNA template (~10pg). A typical PCR reaction consisted of 95°C for 10 minutes, followed by 30 cycles of 95°C (1min), between 55-65°C (1min) and 72°C (1min/kb), finishing off with 72°C for 10 min. For individual primer pair (Table 2.1) conditions and the PCR product expected see Table 2.3.

### 2.4.2 PCR on cell pellets

PCR was also performed on cell pellets. 100-10,000 tissue culture cells suspended in PBS and spun in a sterile eppendorf tube, the PBS was then removed and the cell pellets were snap-freezed in dry-ice. Alternatively, whilst picking individual colonies that required a diagnostic PCR executed on them, 100µl (PBS/cells) was extracted when scraping the colony and half the cells were place into a clean eppendorf (for PCR analysis) and half into an individual tissue culture well for propagation. Again the tubes were spun down and cells were snap-frozen on dry-ice. The pellets could then be stored at -80°C until the PCR reaction was performed. Cell pellets were resuspended in a 25µl mixture containing pronase (0.6µg/µl, Roche), PCR reaction buffer (1X, Qiagen) and sterile water. The reaction was placed at 50°C for one hour, 95°C for 10 minutes then onto ice. The 25µl pronase mixture was made up to 50µl using the same PCR conditions described above (2.4.1).

**Table 2.3 PCR conditions:** PCR primer pairs (for sequence see Table 2.1) used in this study along with their combined annealing temperature, their expected PCR product size and a cartoon illustrating the regions amplified (not to scale).

| Primer Pair | Anneal. Temp. | Product Size | Regions Amplified |
|---|---|---|---|
| TetOF<br>LucR | 60°C | a) 549bp<br>b) 2.359kb<br>c) 472bp |  |
| TetOF<br>NeoR | 63°C | a) 902bp<br>b) 944bp |  |
| TetOF<br>HygR1 | 62°C | a) 592bp |  |
| TIGHT1<br>BSgptR | 60°C | a) 629bp |  |
| TetOF<br>RevTIGHT1 | 66°C | a) 932bp |  |
| TetOF<br>ISCER | 56.7°C | a) 2.494kb<br>b) 427bp |  |
| TetOF<br>RAD52-2R | 56.7°C | a) 2.596kb<br>b) 529bp |  |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

## 2.5 Southern blot analysis

### 2.5.1 Genomic DNA separation and transfer

Digested genomic DNA (See 2.2.2) was separated on a 0.7% agarose gel by running overnight at 25V. Radiolabelled (see 2.2.5) and unradiolabelled (see 2.2.3) 1kb DNA ladder (Invitrogen) was also run alongside the samples to illustrate the separation. The following day the gel was exposed to UV light for 2 minutes and the excess agarose was cut away. The gel was gently agitated in denaturing solution for 2 x 15 minutes, rinsed in sterile water, followed by 2 x 30 minute washes in neutralising solution. Southern transfer was carried out overnight as described by Sambrook (1989), after which the DNA was crosslinked to the membrane (Gene Screen Plus, Perkin Elmers Life Sciences) using the autocross link settings on a UV crosslinker (UV Stratalinker 1800, Stratagene).

### 2.5.2 Membrane hybridisation and washes

The membrane was incubated at 68°C in hybridisation buffer (church mix) for at least two hours. During this time the probe was radiolabelled (see 2.2.5) and after the denaturation step and 10 minutes on ice, the labelled probe was added to the hybridisation bottle and incubated overnight, again at 68°C. The following day the blot was firstly washed in 2xSSC + 0.5xSDS for 2 x 15 minutes at 65°C. Finally the membrane was washed in 0.2xSSC + 0.5xSDS for 2 x 15 minutes or until reduced radioactivity was observed when the membrane was monitored using a Geiger counter.

### 2.5.3 Developing Southern blot

The membrane was wrapped neatly in Saran wrap and exposed to a phosphoimager cassette. The resulting images were analysed using ImageQuant (Version 5.2,

Molecular Dynamics) software and from there they were exported to Photoshop software (Adobe) for further manipulation.

## 2.6 Protein analysis

### 2.6.1 Protein sample preparation

Equal numbers of cells from 80% confluent 6-well plates were spun in PBS-A to pellet the cells. The cell pellets were resuspended in PBS-A (5µl/$10^5$ cells) plus an equal volume of 2xsample buffer. The samples were boiled at 95°C for 5 minutes then spun at 15,493gs for 10 minutes. The supernatant was transferred into a fresh tube and then stored at -80°C until required.

### 2.6.2 SDS-PAGE and transfer

Equal amounts of each protein sample (3.5µl) were made up to 25µl with 1 x sample buffer. They were boiled at 95°C for 5 minutes and then spun briefly. The samples were run, alongside a protein marker (GIBCO), using a vertical electrophoresis cell (Model V15.17, GIBCO) on a 4% SDS stacking gel and a 12% SDS separating gel, for 3.5 hours at 200V in 1 x running buffer as per the manufacturers' instructions. Proteins were transferred onto Immobilon-P membrane (Millipore) using a transfer cell (Bio-Rad) at 100V for 1.5 hours in ice cold transfer buffer as per the manufacturers' instructions. The membrane was left in blocking solution overnight at 4°C.

### 2.6.3 Immunodetection

After removal from the blocking solution the membrane was placed into a sealable plastic pocket and incubated with the primary antibody (Table 2.3), which was diluted into 8ml of blocking solution. This was carried out at room temperature whilst agitating gently for 2 hours. Following this, the membrane was washed for 3 x 10 minutes in 150ml of PBS-T. The membrane was subsequently incubated with the secondary antibody (conjugated to horseradish peroxidase; Table 2.3) and washed as described for the primary antibody. Using the ECL detection system (Amersham) and

Hyperfilm ECL (Amersham), immunodetection of signal was carried out as per the manufacturers' instructions. Membranes were stored at 4°C overnight and then reprobed as described above using antibodies designed to examine the protein loading (e.g. Anti-actin).

**Table 2.4 Primary and Secondary Antibodies:** Listed are the primary and secondary antibodies used for Western analysis throughout this study.

| Primary Antibodies | | | | |
|---|---|---|---|---|
| Target | Company Purchased | Catalogue Number | Dilution Used | Host Species |
| HA Tag | Stratagene | 3F10 | 1/500 | Rat |
| GFP | Clontech | 632380 | 1/1000 | Mouse |
| Actin | Sigma | A-2066 | 1/1000 | Rabbit |
| **Secondary Antibodies** | | | | |
| Target | Company Purchased | Catalogue Number | Dilution Used | Host Species |
| Rat | Sigma | A-9037 | 1/1000 | Goat |
| Rabbit | Sigma | P-0448 | 1/1000 | Goat |
| Mouse | Sigma | P-0447 | 1/1000 | Goat |

## 2.7 Luciferase assay

### 2.7.1 Protein sample preparation

A known number of tissue culture cells were plated and grown to ~80% confluency in a 6-well plate, washed with PBS-A and then treated with 500µl of room temperature 1 x passive lysis buffer (Promega). The plate was gently agitated for at least 20 minutes until a white precipitate was formed. The cell lysate was transferred to a sterile

eppendorf and spun at 13,201gs for 2 minutes to pellet the cell debris. The supernatant was transferred to a fresh tube and stored at -80°C until required.

### 2.7.2 Measuring the luciferase activity

To measure the luciferase activity, 20µl of the cell lysate was added to 100µl of room temperature luciferase assay reagent (Promega) in an appropriate luminometer tube. A reading, in relative light units, was immediately measured using a Biorbit 1253 luminometer connected to an IBM PC (on one occasion a TD-20/20 [Turner Designs] luminometer was used). Ten values of luciferase activity were taken at 2 second intervals and the average was recorded.

## 2.8 Mammalian cell culture

### 2.8.1 Cell types and culture conditions

Two types of cell lines were used in this study; the human fibrosarcoma cell line, HT1080 (Rasheed *et al.*, 1974) and the human telomerase immortalised retinal epithelial cells, hTERT-RPE1 (Clontech). Both cell types required the same culture and growth conditions. The transactivator expressing cells, HT2 cells, were G418-resistant HT1080 cells which were previously made by a colleague in the laboratory (see Sullivan *et al.*, 2001 for methods).

Cells were grown in a humidified incubator (NAPCO-5410) at 37°C with 5% $CO_2$ and sterile manipulation of all cells was carried out using a FASTER BHA 48 flow hood. The cells were cultured in 1 X Dulbecco's modified eagle medium (GIBCO), L-Glutamine (2mM, GIBCO), penicillin and streptomycin (100U/ml and 100µg/ml respectively, GIBCO), non-essential amino-acids (0.4mM, GIBCO) and sodium pyruvate (1mM, GIBCO).

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

When required the media was also supplemented with one or more of the following drugs: G418 (400μg/ml, GIBCO), hygromycin B (100μg/ml, Sigma), Zeocin (200μg/ml, Invitrogen), puromycin (0.4μg/ml, Sigma), mycophenolic acid (10μg/ml, Sigma), xanthine (100μg/ml, Sigma), HAT supplement (1X, GIBCO). In addition, the cells were also treated at various other times with the following drugs: HT supplement (1X, GIBCO), tetracycline (both 1μg/ml, Sigma) and 5-aza 2'-deoxycytidine (1μM, Sigma).

### 2.8.2 Cell growth and passage

Cells were plated and left until they had reached 80% confluency. At this point the medium was removed and the cells were washed in PBS-A. Room temperature trypsin/EDTA (GIBCO) was added (enough to cover all the cells) and left for five minutes until the cells detached form the bottom of the plate. The cells were then resuspended in fresh medium and a fraction (usually $^{1}/_{10}$) was replated into a new culture plate.

### 2.8.3 Freezing and thawing

Trypsinised cells were resuspended in medium and centrifuged at 700g at 4°C for 5 minutes. The medium was then removed and the cell pellets were resuspended in 1ml freezing medium/1million cells and 500μl aliquots were placed into cryovials (NUNC). Cryovials were stored at -80°C in polystyrene containers and then transferred to liquid $N_2$ for long term storage.

To thaw the cells cryovials were removed from cold storage and placed in a 37°C water bath until the freezing medium had thawed. The cells were then placed into a universal tube containing fresh medium and the tube was then spun at 700g for 5 minutes. Following this, the medium was removed and the cells were resuspended in fresh medium and plated into a sterile culture dish and stored in the 37°C tissue culture incubator.

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

## 2.8.4 Stable transfection of tissue culture cells by electroporation

### 2.8.4.1 *Preparation of linear plasmid for transfection*

Prior to stable transfection, constructs were linearised using the appropriate restriction endonuclease (see 2.1.1; Table 2.5). In some cases (for the 'Target' constructs) the plasmid was removed of most of its vector sequences. After digestion the DNA was purified by phenol/chloroform extraction followed by ethanol precipitation (see 2.2.4.1). Removal of the final 70% ethanol was carried out in a sterile flow hood and the DNA was resuspended in sterile TE at 1-2μg/μl. A sample of linear DNA was analysed using a 0.7% agarose gel (see 2.2.3) to check and confirm fragment concentration.

### 2.8.4.2 *Electroporation*

Both HT1080 and hTERT-RPE1 cells were grown to 80% confluency were trypsinised and resuspended in cold PBS-A. An aliquot was taken for counting the number of cells. The cells were pelleted and resuspended in cold PBS-A at 10 million cells per ml and then incubated on ice for 5 minutes. Eight million cells (800μl) were placed into an electroporation cuvette (0.4cm gap, Bio-Rad) with linearised plasmid DNA (Table 2.5; generally, for co-transfections: 20μg of 'target' construct and 1μg of puromycin resistance construct was used. For single-transfections: 10μg of construct was transfected). The cells were subjected to one electrical pulse (400V, 250μF, Bio-Rad Gene Pulser with capacitance extender) and then left on ice for a further 5 minutes. The electroporated cells were then transferred to a vial containing warm medium and dispensed amongst numerous plates (usually 0.8 million cells/15cm plate).

48 hours after the transfection the appropriate drug selection was added (Table 2.5) to the medium (see 2.8.1) and selection was carried out for a further 14-18 days. When the colonies where visible and big enough, they were transferred to the well of a 24-well plate using a micropipette tip (Cloning cylinders (SciQuip) were used for

transferring hTERT-RPE1 cells). After expanding into a 6-well plate the selected colonies were then analysed further for various traits (e.g. PCR (see 2.4.2), Southern analysis (see 2.5), GFP expression analysis (see 2.9) or Protein analysis (see 2.6)).

**Table 2.5 Linearising Plasmid DNA for Stable Transfection:** The restrictions enzymes used to linearise the vectors prior to stable transfection into HT1080 (all constructs listed were transfected) and hTERT-RPE1 (only those marked with * were transfected) cells. In addition, the amount of DNA and drug selection used are listed.

| Construct Type | Vector Name | Enzyme(s) to linearise prior to transfection | Co-transfection with *SpeI* linearised pBL-Puro/R (1μg) | Amount of DNA (μg) | Drug resistance (Usually 14-18 days) |
|---|---|---|---|---|---|
| 'Target' | pTRElox71EGFP | *Pvu*I | Yes | 20 | Puro |
| | pTRElox71dGFP | *Bgl*I & *Ssp*I | Yes | 20 | Puro |
| | pTlox71dGFP | *Bgl*I & *Ssp*I | Yes | 20 | Puro |
| | pTlox71dGFPpuro | *Apa*LI | No | 10 | Puro |
| | pTlox71dGFPiresHyg | *Bgl*I & *Pme*I | No | 10 | Hyg |
| | pTFRTlox71dGFP* | *Bgl*I & *Ssp*I | Yes | 20 | Puro |
| tTA | pZeoSVitTA* | *Bgl*II | No | 10 | Zeo |
| I-*SceI* Site | pDR-Neo | *Bsp*HI | No | 10 | Hyg |
| I-*SceI* | pTIGHT-ISceI* | *Xmn*I | Yes | 20 | Puro |

## 2.8.5 Transient transfection by lipofection

Lipofection was routinely carried out for the transient expression of a gene from a construct. Unless otherwise stated, 250,000 HT1080 cells or 100,000 hTERT-RPE1 cells were plated into a 6-well plate in 2ml of antibiotic free medium 24 hours prior to the transfection. A total of 4μg (or 6μg for hTERT-RPE1 cells) of circular plasmid was used (For co-transfections: 2μg of 'insertion' construct and 2μg of Cre-recombinase expressing plasmid (pMC-Cre, donated by H. Gu) was used for HT1080

cells. Alternatively for hTERT-RPE1 cells, 3μg of 'insertion' construct and 3μg of pMC-Cre were used). The DNA was mixed with 250μl of Optimum (GIBCO) and in a second tube 250μl of Optimum was mixed with 10μl of lipofectamine 2000 reagent (Invitrogen). Both volumes were incubated for 5 minutes, then combined and incubated for a further 20 minutes at room temperature. After the incubation, the mixture was added to the medium of the cells plated the day before. The transfection was left for 4 hours at 37°C at 5% $CO_2$, after which the medium was replaced with fresh medium. The following day, the transfected cells were divided between 5 X 9cm plates and a further 24 hours later the appropriate drug was added and screening commenced as described previously. Alternatively, if no selection was being applied the cells were plated at very low density (50-500 cells/9cm plate) and screened in the appropriate way (GFP expression analysis (2.9), PCR (2.4.2)).

**Table 2.6 Transiently transfected constructs:** List of the plasmids used which were transiently transfected into HT1080 and hTERT-RPE1 cells and the conditions used.

| Construct Type | Vector Name | Amount to transfect (μg) HT1080 | Amount to transfect (μg) hTERT | Co-transfection with pMC-Cre (2μg HT1080; 3μg hTERT) | Drug resistance (Usually 14-18 days) |
|---|---|---|---|---|---|
| 'Insertion' | plox66Luc | 2 | - | Yes | None |
| | plox66Hyg | 2 | - | Yes | Hyg |
| | Plox66Neo | - | 3 | Yes | G418 |
| | plox66LuciresHyg | 2 | - | Yes | Hyg |
| | pINS | 2 | 3 | Yes | gpt |
| | pINSmcs | 2 | - | Yes | gpt |
| | pINSneoMCS | 2 | 3 | Yes | G418 |
| | pINSneoISCEI | 2 | 3 | Yes | G418 |
| | pINSneoRAD52 | 2 | 3 | Yes | G418 |
| **Flp expression** | pCAGGS-FLPe | 4 | 6 | No | None |
| **TRE driven Luc** | pTIGHT-Luc | 4 | 6 | No | None |
| **TRE driven GFP** | pTFRTlox71dGFP | 1-15 | 1-15 | No | None |
| **TRE driven I-SceI** | pTIGHT-ISceI | 4 | - | No | G418 |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

### 2.8.6 Staining colonies

To score colonies, cells were stained with crystal violet stain. Colony plates were firstly washed in PBS, then treated for 10 minutes with 100% ethanol and then soaked for 30 minutes in crystal violet stain. After the incubation time the plates were rinsed in water and left to air dry.

## 2.9 GFP expression analysis

### 2.9.1 UV Microscopy

Colonies were first analysed visually for gain or loss of GFP expression using an Axiovert S100TV microscope (Zeiss). Cells were analysed in 6-well/9cm and 15cm plates in PBS-A to reduce any autofluorescence caused by the media. Fluorescent light was emitted from an HBO100 halogen lamp with an Atto Arc controller (Zeiss), controlled by a lambda 10-2 filterwheel (Sutter Instruments Co.). The focus was controlled by an H308 focus driver (Prior).

### 2.9.2 Fluorescence Activated Cell Sorting (FACS)

#### 2.9.2.1 *Preparing cells for FACS analysis*

Cells containing tetracycline regulated GFP (approximately 100,000 cells) were plated into +/- tetracycline-containing medium 48 or 72 hours prior to FACS analysis, in 6-well plates. On other, occasions an 80% confluent plate of HT1080 cells and a 60% confluent hTERT-RPE1 cells was used. To prepare the cells for FACS analysis they were firstly washed in PBS-A, trypsinised and then resuspended in 500µl (per 0.5 million cells) of cold PBS-A before storing on ice.

### 2.9.2.2 *FACS settings*

Prior to flow cytomtery, the samples were mixed, by pipetting up and down. On a high flow rate setting, 40,000 cells were counted per sample, using an Argon ion laser turned to 488nm (fluorescence channel FL-1, with fluorescence channel FL-3 as a negative control) on a Becton Dickinson FACScan machine. On almost all occasions gating the healthy population, using the forward scatter (FSC) and side scatter (SSC) plots, did not result in any change to the overall FACS profile, therefore no gating was used. Channel gains were set so that the peak of non-fluorescing wild-type cells (or autofluoresence) was set at 5-10 on a log scale. The typical FACS settings used, are shown in Table 2.7. Acquisition and storage, along with the analysis of data was carried out using CellQuest software (Becton Dickinson). In this study plots are displayed as FL-1 (log scale) versus the number of counts. See Appendix IV for instructions on how to interpret a FACS plot.

**Table 2.7 GFP FACS Settings:** Settings used on FACScan for analysing GFP fluorescence in HT1080 and hTERT cells.

| Channel | Voltage | AmpGain | Mode |
|---|---|---|---|
| FSC | E-1 | 3.02 | Lin |
| SSC | 181 | 2.18 | Log |
| FL-1 | 500 | 1.00 | Log |
| FL-3 | 421 | 1.00 | Log |

## 2.10 Propidium Iodide (PI) nuclei staining

### 2.10.1 Preparation of stained nuclei

To check the ploidy of selected clones, propidium iodide nuclei staining was carried out. Approximately 40,000 cells were placed into a snap-cap tube and centrifuged for 1500gs for 4 minutes. The cell pellets were resuspended in 900µl of nuclei prep

Solution One supplemented with 25µl of propidium iodide (1mg/ml, Sigma) and 1µl of RNase (10mg/ml, Boehringer Manheim). The tubes were mixed gently and left at room temperature, in the dark for 30 minutes. After the incubation period, 900µl of nuclei prep Solution Two plus 40µl of propidium iodide (1mg/ml, Sigma) were added to the tubes, mixed gently and then placed at 4°C overnight.

## 2.10.2 FACS analysis of stained nuclei

Treated cells were passed through a mesh using a pipette to get rid of cell debris and 1ml was placed into a clean FACS tube and then placed onto ice. 20,000 stained nuclei were then counted using the FACScan as described above (see 2.9.2.2). This time, however, the machine was set for a low flow rate and the laser was turned to the FL-2 channel. Channel gains were set so that the diploid peak (for diploid cells) laid around 200 units on the FL-2 height scale; hence the tetraploid peak would be at 400 units. Typical settings are shown in Table 2.8. Again, acquisition and storage, along with the analysis of data was carried out using CellQuest software (Becton Dickinson). In this study plots are displayed as FL-2-height (linear scale) versus the number of counts.

**Table 2.8 FACS settings used for PI nuclei analysis:** Settings used on FACScan for analysing PI nuclei in HT1080 and hTERT-RPE1 cells.

| Channel | Voltage | AmpGain | Mode |
|---|---|---|---|
| FSC | E-1 | 8.02 | Lin |
| SSC | 270 | 4.42 | Lin |
| FL-1 | 150 | 1.00 | Lin |
| FL-2 | 415 | 4.00 | Lin |
| FL-3 | 288 | - | Log |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

## 2.11 Appendix of solutions

All chemicals used were from BDH, unless otherwise stated.

### 2.11.1 Standard solutions

| | |
|---|---|
| 50 X TAE | per litre: 242g tris base, 57.1ml glacial acetic acid, 100ml 0.5M EDTA (pH8). |
| T.E. | 10mM tris/HCl (pH7.4), 1mM EDTA (pH8). |
| 20 X SSC | 3M NaCl, 0.3M trisodium citrate (pH7). |
| Nick translation buffer | 50mM tris/HCl (pH7.2), 10mM $MgSO_4$, 0.1M dithiothreitol. |
| DNA gel loading buffer | 0.25% bromophenol blue (Sigma), 0.25% (v/v) xylene cyanol FF (Sigma), 15% Ficoll type 400 (Pharmacia). |

### 2.11.2 Bacterial culture and transformation reagents

| | |
|---|---|
| LB broth (GIBCO) | Bacto-tryptone $10gL^{-1}$ (Merck), bacto-yeast extract @ $5gL^{-1}$, NaCl @ $10gL^{-1}$. |
| Low Salt LB Broth | Bacto-tryptone $10gL^{-1}$ (Merck), bacto-yeast extract @ $5gL^{-1}$, NaCl @ $5gL^{-1}$. |
| LB Agar | 1.5% (w/v) bacto-agar (Merck) in LB broth. |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

| | |
|---|---|
| Low Salt LB Agar | 1.5% (w/v) bacto-agar (Merck) in low salt LB broth |
| 10 X Ligation Buffer | 0.5M tris-HCl (pH7.5), 100mM $MgCl_2$, 100mM DTT. |

## 2.11.3 Cell Culture

| | |
|---|---|
| Freezing Medium | 10% (v/v) DMSO (Sigma), 50% (v/v) FCS, 40% complete medium. |
| HT1080 and hTERT-RPE1 Complete Medium | per 500ml 1 X DMEM (Sigma): 10% (v/v) FCS, 2ml 5mg/ml Pen/strep (GIBCO), 5ml 100mM Na Pyruvate (GIBCO), 5ml 200mM L-Glutamine (GIBCO), 20ml (100X) non-essential amino acids (GIBCO). |
| PBS-A | per litre: 10g NaCl, 0.25g KCl. 1.44g $Na_2HPO_4$, 0.25g $KH_2PO_4$. |
| Cell Lysis Buffer | 10mM NaCl, 10mM tris (pH8), 10mM EDTA (pH8), 0.5% (v/v) SDS. |
| Crystal Violet stain | Enough crystal violet is added to 500ml $dH_2O$ to create a deep purple solution. |
| Nuclei Prep Solution One | per 20ml: 40µl 5M NaCl, 20mg Na citrate, 120µl 10% NP40. |
| Nuclei Prep Solution Two | per 20ml: 300mg citric acid, 1712mg sucrose. |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

## 2.11.4 Southern Analysis

| | |
|---|---|
| Denaturation Solution | 0.5M NaOH, 1.5M NaCl. |
| Neutralising Solution | 1.5M NaCl, 0.5M tris-HCl. |
| Church Mix | per litre: 13.8g $Na_2HPO_4$, 143g $NaH_2PO_4$, 10g BSA, 70g SDS, in $ddH_2O$. |

## 2.11.5 Western analysis

| | |
|---|---|
| Blocking Solution | per 500ml PBS: 1g BSA (Sigma). |
| 2 X Sample Buffer | 1M tris HCl (pH6.8), 20% (v/v) glycerol, 20% (v/v) SDS, 10% (v/v) β-mercapthoethanol (Sigma). |
| 5 X Running Buffer | per litre: 15g tris base, 72g glycine, 5g SDS. |
| Stacking Gel | 0.125M tris-HCl (pH6.8), 0.1% (w/v) SCS, 4% (v/v) acrylamide/bis (Scotlab), 0.5% (v/v) ammonium persulphate, 0.05% (v/v) TEMED (Sigma). |
| Separating Gel | 0.375M tris-HCl (pH8.8), 0.1% (w/v) SCS, 15% SDS, 15% (v/v) acrylamide/bis (Scotlab), 0.5% (v/v) |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

| | |
|---|---|
| | ammonium persulphate, 0.1% (v/v) TEMED (Sigma). |
| Transfer Buffer | 25mM tris base, 192mM glycine, 20% (v/v) methanol, pH8.3. |
| Bromophenol Blue | 1g bromophenol blue (Sigma) dissolved in $dH_2O$. |
| PBS-T | 0.1% (v/v) tween-20 (Sigma) dissolved in PBS-A. |

N.B. The recipes for routinely used solutions are listed in the Appendix of solutions (section 2.11).

# Chapter Three

# Generating Tetracycline Inducible GFP Cell-Lines

## 3.1 Introduction

A basic outline of the proposed strategy is illustrated in Figure 3.1 and for simplicity, at this stage, can be thought of as a two-step process. The first step involves transfecting transactivator (tTA) expressing HT1080 cells with a 'target' construct in which the GFP gene is linked to a tetracycline response element (TRE). This results in GFP expression only in the absence of tetracycline when the tTA is able to bind to the TRE. Conversely, in the presence of tetracycline the tTA is unable to bind to the TRE and the GFP gene was down-regulated (Figure 3.2). The second step involves the insertion of a gene of interest (GOI) downstream of the tightly regulated promoter using site-specific recombination (see section 1.1.4) and results in the GOI being under the control of the TRE in a similar manner to the GFP ORF (Figure 3.1).

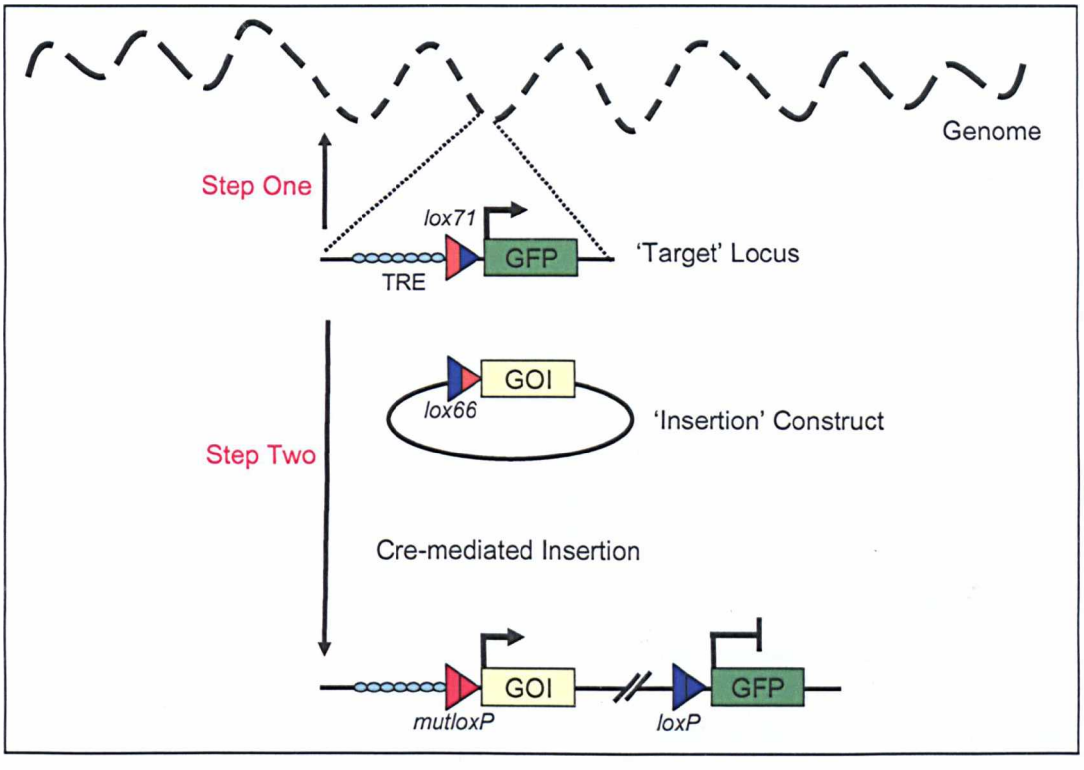



**Figure 3.1 The Basic outline of a 'Select and Insert' Strategy:** A cartoon illustrating the basic 'Select and Insert' Strategy. The first step involves the generation of cells expressing tight, inducible GFP from an integrated 'target' construct. The second step involves the insertion of a promoterless GOI by *Cre*-mediated recombination. For a key to symbols see Appendix I. Not to scale.

This chapter describes the experiments carried out to establish step one of the system, the generation of clones expressing tightly regulated GFP (Figure 3.1). Cells stably expressing the 'target' construct were isolated and FACS analysis was used to identify those clones expressing tightly regulated GFP (i.e. those that demonstrate high amounts of GFP fluorescence when induced (-Tet) and minimal amounts of basal GFP expression when uninduced (+Tet)) (Figure 3.2). Ideally, uninduced levels of GFP expression would be no higher than the background fluorescence detected in GFP-negative cells.

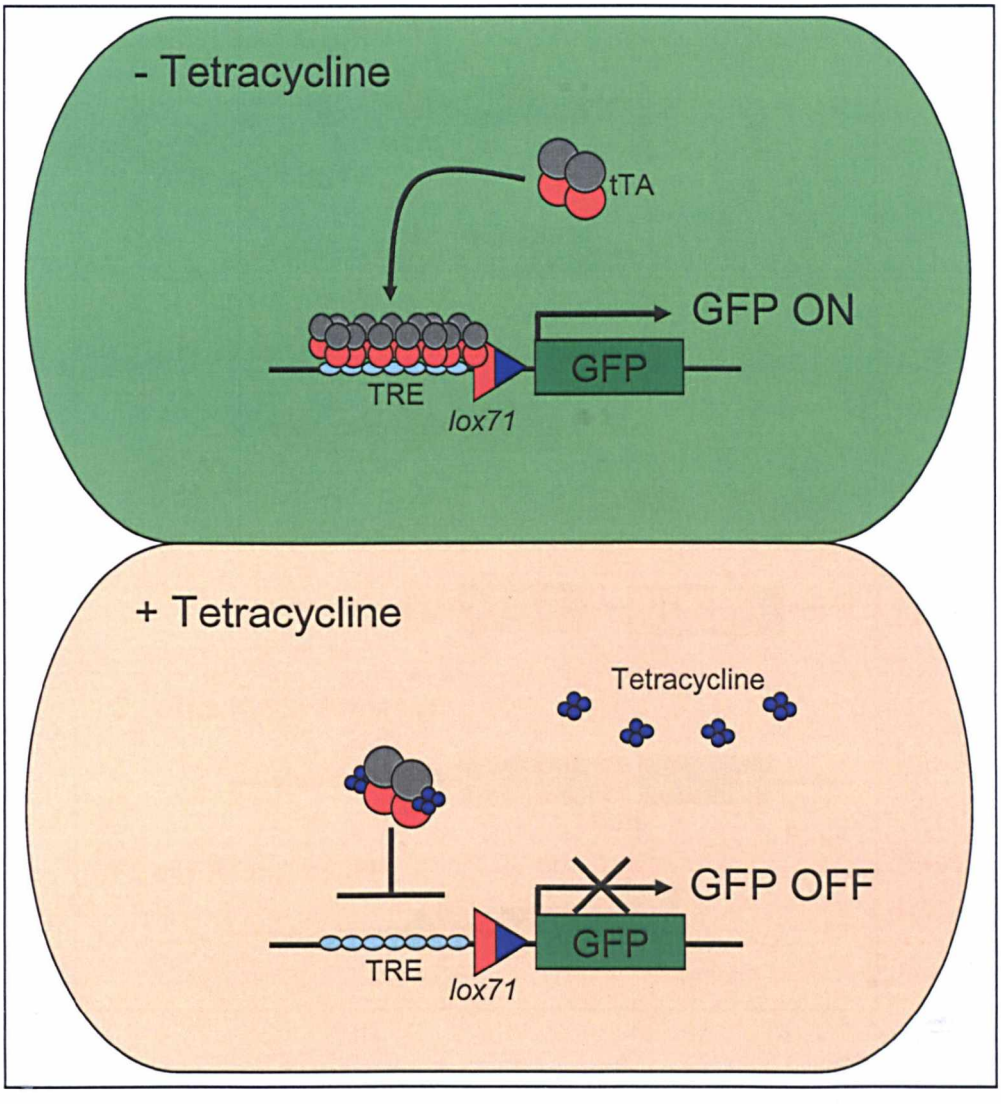



**Figure 3.2 The Tetracycline System:** Illustrates how expression of the GFP ORF was regulated by tetracycline. In the absence of tetracycline the transactivator was able to bind to the TRE and the GFP was expressed. However, in the presence of tetracycline the transactivator was unable to bind to the TRE and the GFP was down-regulated. For a key to symbols used see Appendix I. Not to scale.

As expected, the position at which the 'target' construct integrates into the genome affects the degree to which tight regulation was achieved, such position effects led to numerous clones being screened. Two types of TRE were tested for their ability to tightly regulate gene expression and the tTA was later replaced with a humanised form to prevent its down-regulation by methylation, which was observed in HT1080 cells. Various different forms of the 'target' vector were tested and these are summarised in Figure 3.3. For a scale plasmid map of each of the 'target' constructs tested see Appendix II.

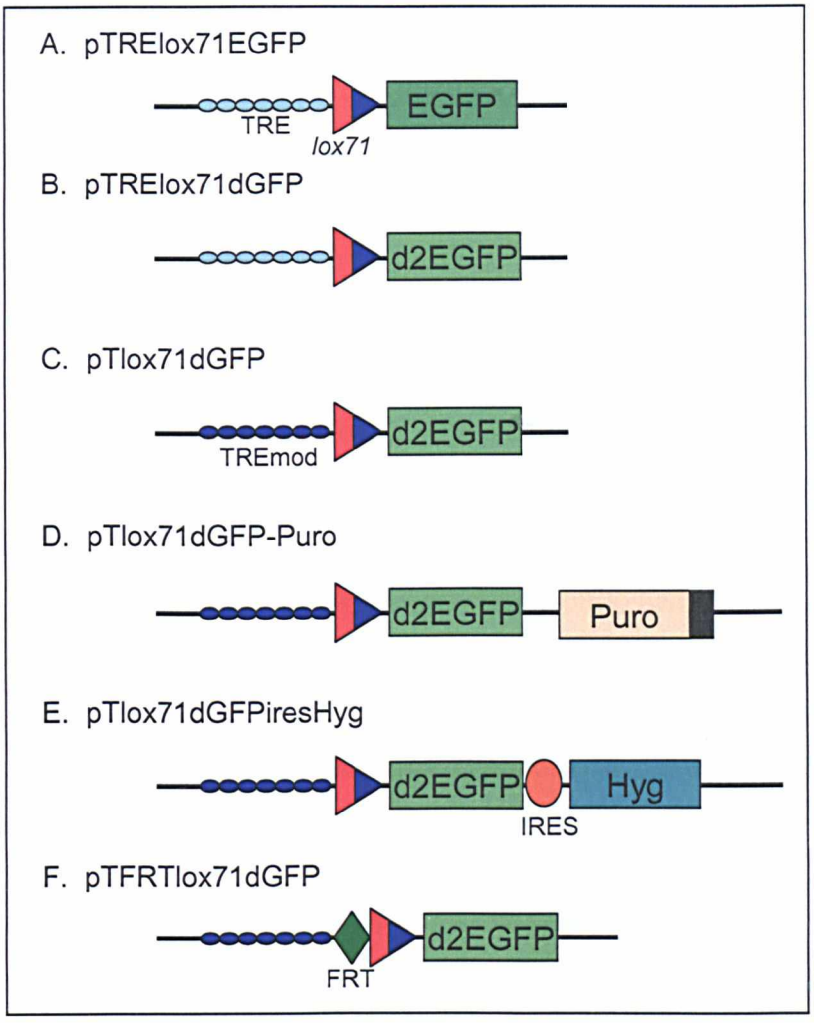



**Figure 3.3 The 'target' constructs created and tested:** Six 'target' constructs were tested altogether. Illustrations of **A.** pTRElox71EGFP; **B.** pTRElox71dGFP; **C.** pTlox71dGFP; **D.** pTlox71dGFP-Puro; **E.** pTlox71dGFPiresHyg (IRES – internal ribosomal entry site); and **F.** pTFRTlox71dGFP are shown. Not to scale. For key to symbols used see Appendix I. For scale plasmid maps of each of the 'target' constructs see Appendix II.

# 3.2 Results

## 3.2.1 Inducible GFP expression using HT2 (tTA expressing) cells

*Co-transfection was used to generate clones with stably integrated 'target' construct*

All 'target' constructs were generated as described in the materials and methods (see 2.3.2). Co-transfection was used to avoid the need to generate a 'target' construct that included a drug-resistance cassette (this does not apply to 'target' vectors illustrated in Figure 1.3 D and E), as enhancer activity in such cassettes were thought to interfere with expression from the TRE (Damke *et al.*, 1995). To optimise the co-transfection procedure 'target' vector pTRElox71EGFP was used (Figure 3.3A) and unless otherwise stated, tTA expressing HT1080 cells (HT2 cells; Sullivan *et al.*, 2001) were co-transfected with linear 'target' plasmid and the puromycin-resistance construct (pBL-Pur/R; Yun *et al.*, 2004; Figure 3.4). The optimal DNA concentrations were determined by carrying out a series of transfections using various amounts of each plasmid DNA, maintaining a 20-fold weight excess of 'target' plasmid to selection construct throughout (Table 3.1). Although 100% of colonies were GFP-positive when 40μg of the 'target' construct was used, only two colonies were actually generated. The greatest number of puromycin-resistant colonies was obtained when 20μg of 'target' construct was used and of these 54% of the colonies were GFP-positive. All further co-transfections were therefore carried out with 20μg of 'target' construct and 1μg of puromycin-resistance construct.

**Table 3.1 Co-transfection efficiencies with various amounts of plasmid DNA:** Each transfection was electroportated as described in section 2.8.4.2 using the various concentrations of DNA listed below. After 14 days selection in puromycin, colonies were scored for GFP expression.

| ['Target'] Construct (μg) pTRElox71EGFP | [pBL-PurR] Construct (μg) | Number of PuroR Colonies | Number of GFP +ve Colonies | % of GFP +ve Colonies |
|---|---|---|---|---|
| 10 | 0.5 | 10 | 0 | 0% |
| 20 | 1 | 65 | 35 | 54% |
| 40 | 2 | 2 | 2 | 100% |
| 100 | 5 | 1 | 0 | 0% |

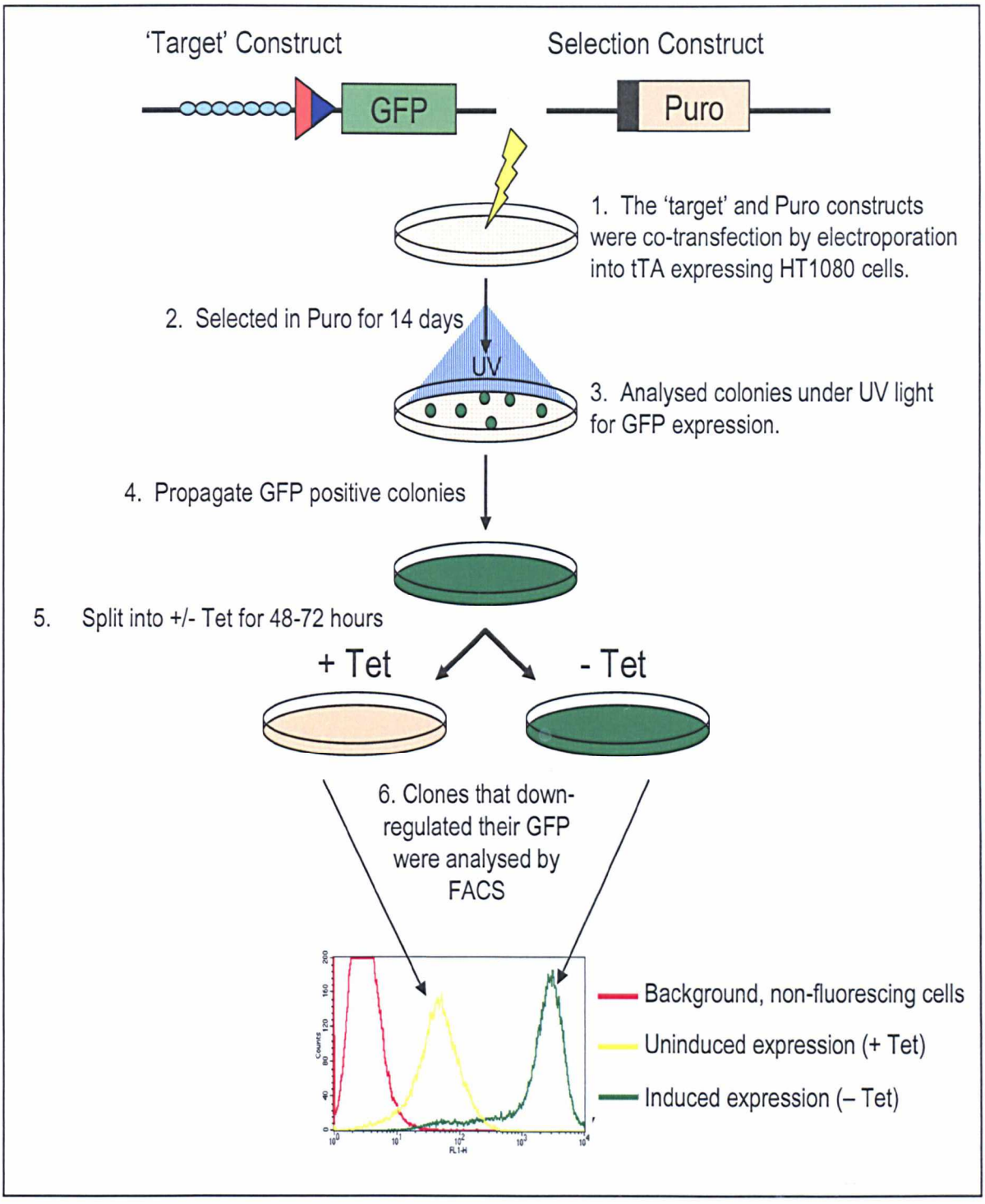



**Figure 3.4 Scheme for the generation of clones with tightly regulated GFP:** HT2 cells were co-transfected with a 'target' construct and a puromycin-resistance plasmid (pBL-Pur/R). Cells were selected in puromycin for 14 days and GFP-positive colonies were identified using a UV microscope. GFP-positive colonies were placed in medium with or without tetracycline and analysed by FACS. Clones with tight, inducible and uniform GFP expression were selected. A representative FACS profile is shown at the bottom of the figure. For instructions on how to interpret a FACS plot also see Appendix IV.

Following the electroporation step, typical methods for assessing the efficiency of regulation of each new 'target' construct involved 14 days selection in puromycin; followed by the analysis of colonies using a UV microscope for GFP fluorescence (Figure 3.4). GFP-positive colonies were picked and grown in the presence or absence of tetracycline for 48-72 hours and then analysed by FACS (Figure 3.4). As outlined below, the majority of GFP-expressing clones produced down-regulate their GFP expression upon the addition of tetracycline to the medium; but the degree of down-regulation depends upon the type of 'target' construct used (Figure 3.3) and the particular clone analysed (i.e. the position at which the 'target' construct integrates into the genome). Those clones that demonstrate uniform expression profiles and negligible uninduced expression (+Tet) were chosen for further manipulations. Finally, tetracycline was used in this study, rather than the commonly used analogue doxycycline, as it was found by ourselves and others to be more easily and reliably removed from the cells to re-induce expression from the TRE (Rennel and Gerwins, 2002).

*Tight regulation was not achieved using 'target' construct based on the original TRE*

The first experiments involved the 'target' construct pTRElox71EGFP (Figure 3.3A). The EGFP ORF was cloned downstream of the original TRE, developed by Gossen and Bujard (1992), and a *lox71* oligonucleotide was inserted in between (For cloning steps see 2.3.2.1). After co-transfection and selection (Figure 3.4), numerous (100s) GFP-positive clones were analysed of which over 61% down-regulated their GFP expression upon the addition of tetracycline to the medium. It soon became clear, however, that such clones showed very slow GFP-depletion kinetics and that most expressed high amounts of GFP, making it difficult to overlay the induced and background profiles on the same plot (Figure 3.5A).

Clone 6 was generated using pTRElox71EGFP. It demonstrated an induced GFP expression profile ~1000-fold greater than background (Figure 3.5B). The uninduced expression level, however, was 10-fold greater than background (Figure 3.5B). The GFP-induction value (Non-corrected; see Appendix IV) for this clone was approximately 200-fold. To achieve this amount of uninduced expression, Clone 6 cells were exposed to tetracycline for 144 hours (6 days), this was most likely due to

the long (16 hour) half-life of the EGFP protein (Figure 3.5C). It was therefore decided to replace EGFP with a destabilised form (d2EGFP), which has a half life of 2 hours. A clone produced in later experiments, containing d2EGFP, reached uninduced expression levels at 48 hours after exposure to tetracycline (Figure 3.5D).

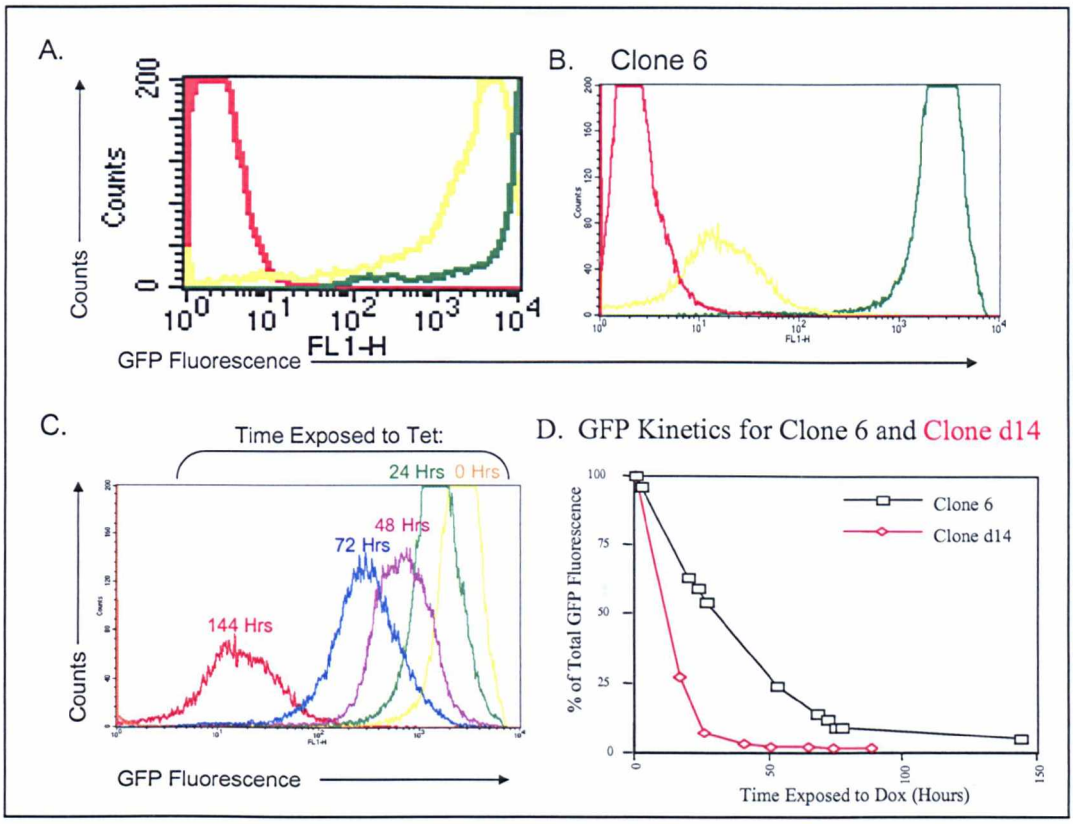



**Figure 3.5 Analysis of Clone 6 GFP expression profiles:** **A.** A typical FACS profile for a clone expressing inducible EGFP. The uninduced profile was only exposed to tetracycline for 48 hours. **B.** Illustrates the FACS profile for Clone 6 showing induced (in green), uninduced (in yellow) and background, nonfluorescing cells (HT2 in red). The GFP induction value was approximately 200-fold and inunduced expression was 10-fold greater than that of background (after a 144 hour exposure to tetracycline). **C.** Demonstrates the turning-off of GFP in response to tetracycline treatment for various time points. It shows that the amounts of GFP fluorescence reduces progressively each day after treatment and takes up to six days to reach maximum uninduced expression levels. **D.** A graph comparing the GFP switch-off kinetics of Clone 6 and a d2EGFP containing clone, Clone d14. The d2EGFP clone reaches uninduced expression levels by 48 hours after the addition of tetracycline.

A second 'target' construct (pTRElox71dGFP) was generated in which EGFP was replaced by d2EGFP (Figure 3.3B; for methods see section 2.3.2.1). Over 200 GFP-positive puromycin-resistant colonies transfected with pTRElox71dGFP were screened but none was found to turn-off its GFP expression completely in the

presence of tetracycline (Figure 3.6). The lowest uninduced expression levels of GFP were approximately 5-fold higher than the background signal from untransfected HT2 cells (for examples see Clones d14 and d46, Figure 3.6). As expected, induction ratios varied between clones. Induction values (See appendix IV) for clones with the lowest uninduced expression ranged from ~60-fold (Clone d3) to ~250-fold (clone d14) (Figure 3.6).

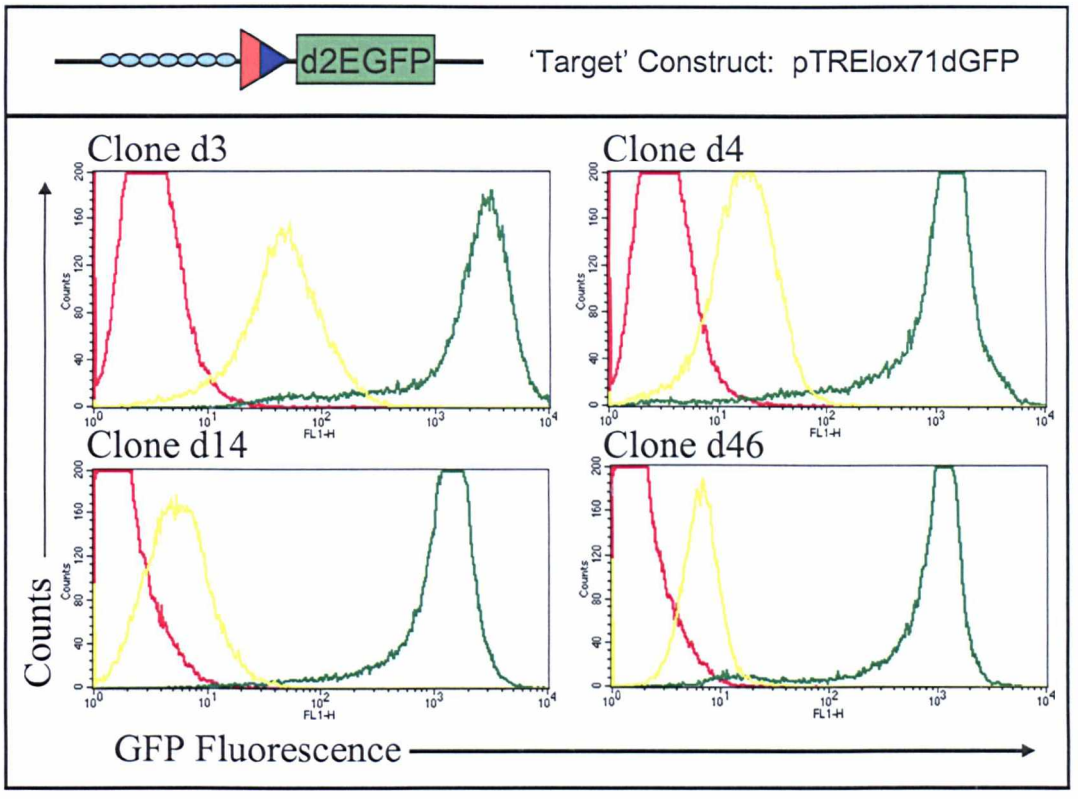



**Figure 3.6 FACS profiles showing four favourite clones containing pTRElox71dGFP:** The figure represents the FACS profiles of the four most desirable clones generated using the 'target' construct pTRElox71dGFP (top; not to scale). All clones demonstrate high, uniform induced expression (green), however, they also have uninduced (yellow) expression levels at least 5-fold greater than that of background (non-fluorescing HT2 cells, red). For instruction on how the FACS profiles are read see Appendix IV.

*Reduced uninduced GFP expression was achieved using a modified TRE*

Based on the above results it was decided to replace the original TRE with a newly available modified version, $TRE_{mod}$ (from pTRE-TIGHT, Clontech). Reports (Clontech themselves) suggest that the $TRE_{mod}$ can generate clones with no uninduced

levels of GOI expression. The $TRE_{mod}$ contains the same seven direct repeats of the 19bp tetracycline operator sequences (tetO) as for the original TRE; however the sequence located between the tetOs has been altered (Figure 3.7). The modifications were made in such a way so that the $TRE_{mod}$ was less likely to be recognised by additional factors that may be present within the cell (e.g. Interferon [Rang and Will, 2000]). The recognition sites for GATA transcription factors, however, still remain in the TREmod sequence (See section 1.3.2 for more details [Gould and Chernajovsky, 2004]). In addition, the $TRE_{mod}$ was upstream of a truncated CMV minimal promoter, again lacking any enhancer sequences that make up the complete CMV promoter (Figure 3.7).

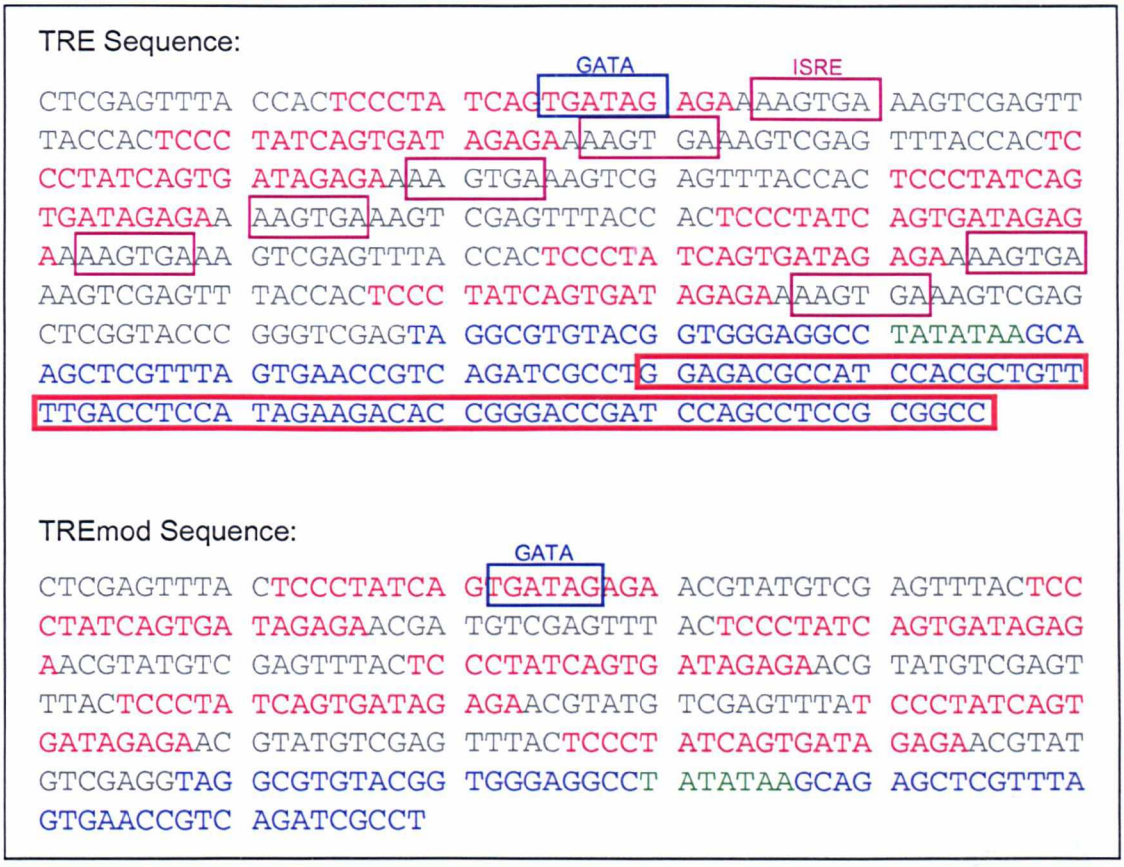



**Figure 3.7 A sequence comparison of TRE and TREmod:** Nucleotide sequences differences of the original TRE, from pUHD10-3 (Gossen and Bujard, 1993), and TREmod, from pTRE-TIGHT (Clontech), are compared. The seven tetO repeat sequences are shown in red; the CMV minimal promoter is shown in blue and the location of the TATA box is in green. Alterations include a truncated CMV minimal promoter (see red box) and modifications to the sequence situated between the seven tetO repeats (in grey). Such alterations have removed the sites which are thought to be recognised by interferon (ISRE).

A third 'target' construct (pTlox71dGFP, Figure 3.3C) was therefore generated. This construct is equivalent to pTRElox71dGFP, but it contains the $TRE_{mod}$ instead of the original TRE. pTlox71dGFP was co-transfected into HT2 cells as described above (see 3.2.1.1). In total, 63 puromycin-resistant colonies were analysed by UV microscopy and almost 100% of these were GFP-positive. Of these, 8 (12.7%) had very low uninduced levels of expression: less than 2-fold higher than the background signal of untransfected HT2 cells (e.g. Clones T18, T22, T24, Figure 3.8). One clone (Clone T21, Figure 3.8) had uninduced expression levels virtually indistinguishable from that of background. Non-corrected induction (see Appendix IV) values for such clones ranged from ~200-fold (Clone T24) to ~700-fold (Clone T18) (Figure 3.8). When compared to observations from the previous section (see 3.2.1.2), these results confirm that $TRE_{mod}$ was superior to TRE in its ability to prevent uninduced transcription.

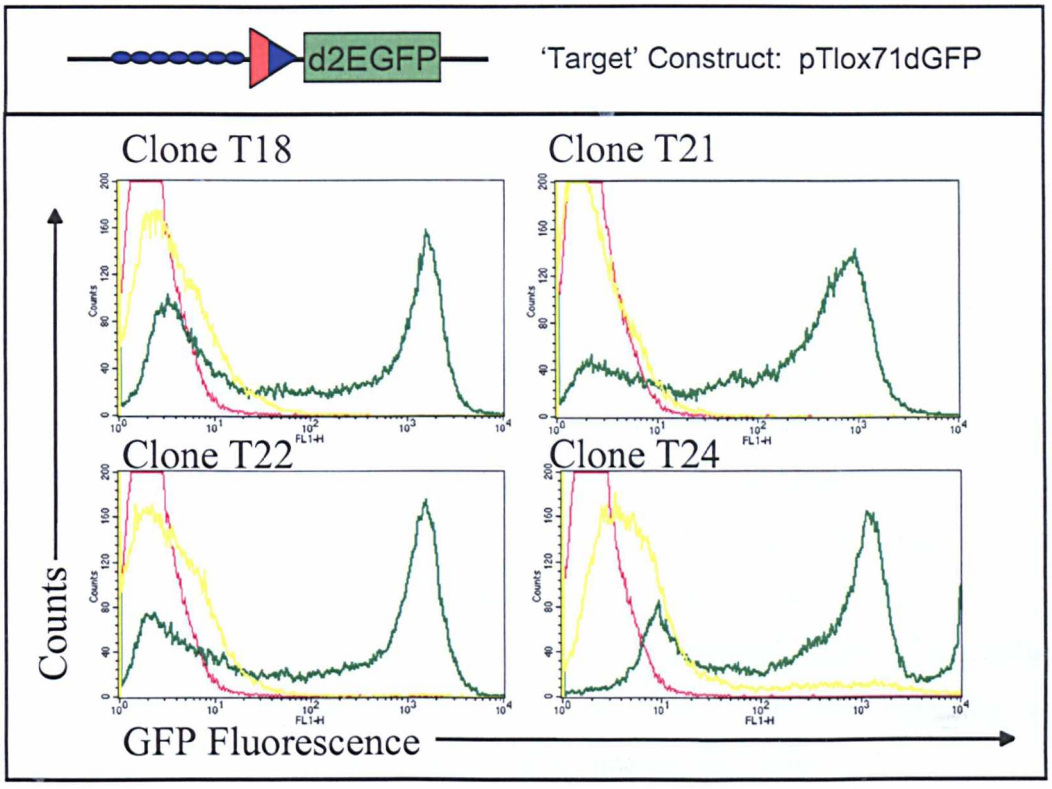



**Figure 3.8 GFP expression profiles of four pTlox71dGFP-transfected clones:** FACS profiles are shown for the four most desirable clones generated using the 'target' construct pTlox71dGFP (top; not to scale). Almost 13% of clones analysed had uninduced (yellow) expression levels less than 2-fold greater than background signal (red) (Clone T18) and almost 2% had uninduced (yellow) expression levels virtually identical to that of background signal (red) (Clone T21). The induced (green) profiles, however, were mosaic in appearance. For instructions on reading the FACS plots see Appendix IV.

Despite the improvement in uninduced GFP expression levels, a new problem was evident from the FACS profiles (Figure 3.8). Within clones, a mosaic pattern of GFP expression was observed in the induced state. It appeared that a significant proportion of cells within a given clone were down-regulating their GFP expression, even when tetracycline was not present. To eliminate the possibility that such mosaicism was simply a result of impure colonies, clones were plated at low density and subcloned. The resulting colonies showed poorer FACS profiles than the parental clones (data not shown); however this was most likely a consequence of prolonged passaging. This suggested that mosaicism was a reflection of progressive GFP gene silencing during clonal expansion which could possibly result from DNA methylation, histone deacetylation, or even gene deletion. It was previously suggested by Damke and colleagues (1995) that increased cell confluency interfered with expression from the TRE. To test this, GFP expression was measured as a function of cell density in several affected clones, but no correlation between density and mosaicism was observed (data not shown).

*Mosaicism can be reduced by selection for a linked drug resistance cassette*

To investigate whether the GFP cassette was being lost or downregulated, a fourth 'target' construct, pTlox71dGFP-Puro (Figure 3.3D), was generated. A puromycin cassette was cloned into pTlox71dGFP in the reverse orientation to the GFP ORF as described in the materials and methods (see 2.3.2.2). It was hoped that this construct would select for those clones still containing and expressing from the 'target' construct. It was also anticipated, however, that the puromycin cassette may impair the ability to achieve tight regulation, due to the potential action of it's enhancer elements on the minimal promoter of the TRE. On this occasion, linear pTlox71dGFP-Puro (10μg) was transfected alone into HT2 cells.

All 28 GFP-positive clones analysed by FACS for inducible GFP expression demonstrated non-uniform expression profiles for both the induced and uninduced states. FACS profiles for some of the clones obtained containing this 'target' construct are shown in Figure 3.9. These results appear to confirm the expected

detrimental effects of nearby enhancer elements on expression from the TRE, effects were seen for both the induced and uninduced expression states (Damke *et al.*, 1995).

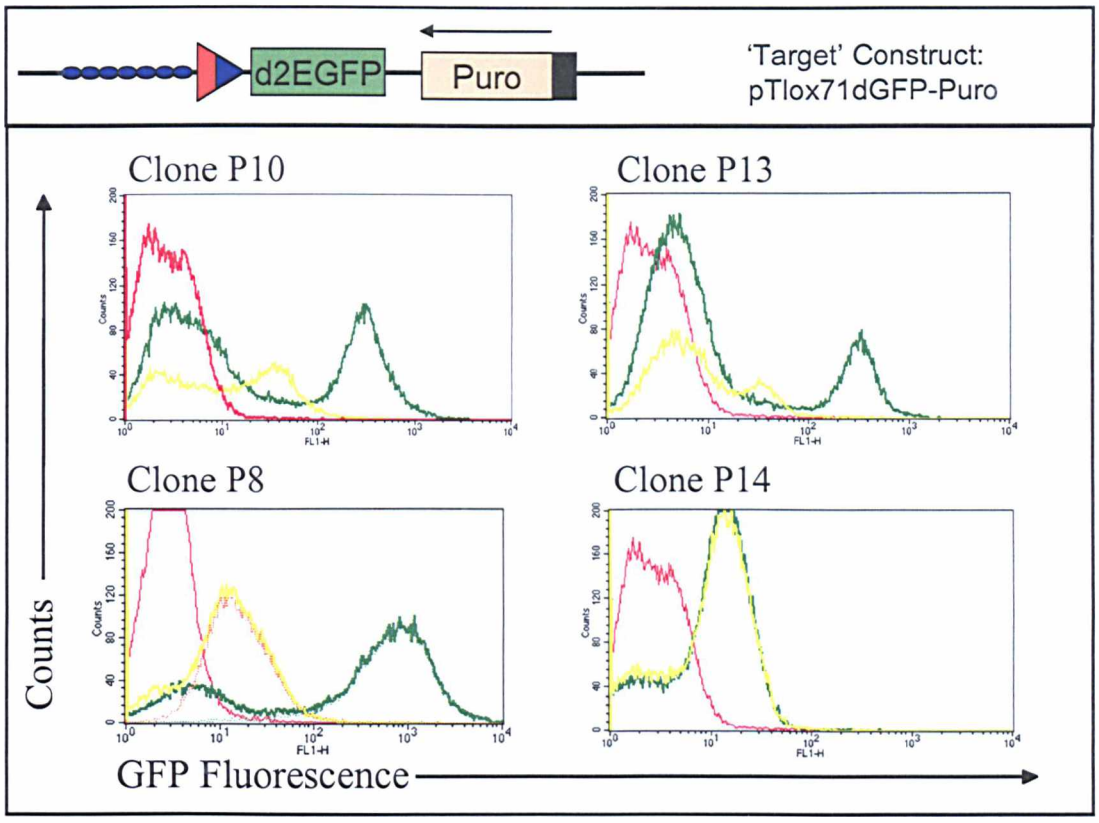



**Figure 3.9 GFP expression profiles of four pTlox71dGFPpuro-transfected clones:.** The presence of the puromycin cassette greatly affected GFP expression for both the induced (green) and uninduced (yellow) states. Illustration of pTlox71dGFPpuro (top) is shown. For instructions on how to read FACS plots see Appendix IV.

To overcome this problem a fifth 'target' construct (pTlox71dGFPiresHyg (Figure 3.3E)) was generated in which an IRES-hygromycin cassette was placed downstream of the GFP gene in pTlox71dGFP (See section 2.3.2.2). Hygromycin-resistant clones generated with this construct are expected to become hygromycin-sensitive if they down-regulate their GFP expression cassette. Selection in hygromycin should therefore improve the FACS profile by killing those cells in which induced expression levels are poor. This 'target' construct was linearised (see Table 2.5 for conditions) and electroporated (10µg) into HT2 cells that were then selected in hygromycin for 14 days.

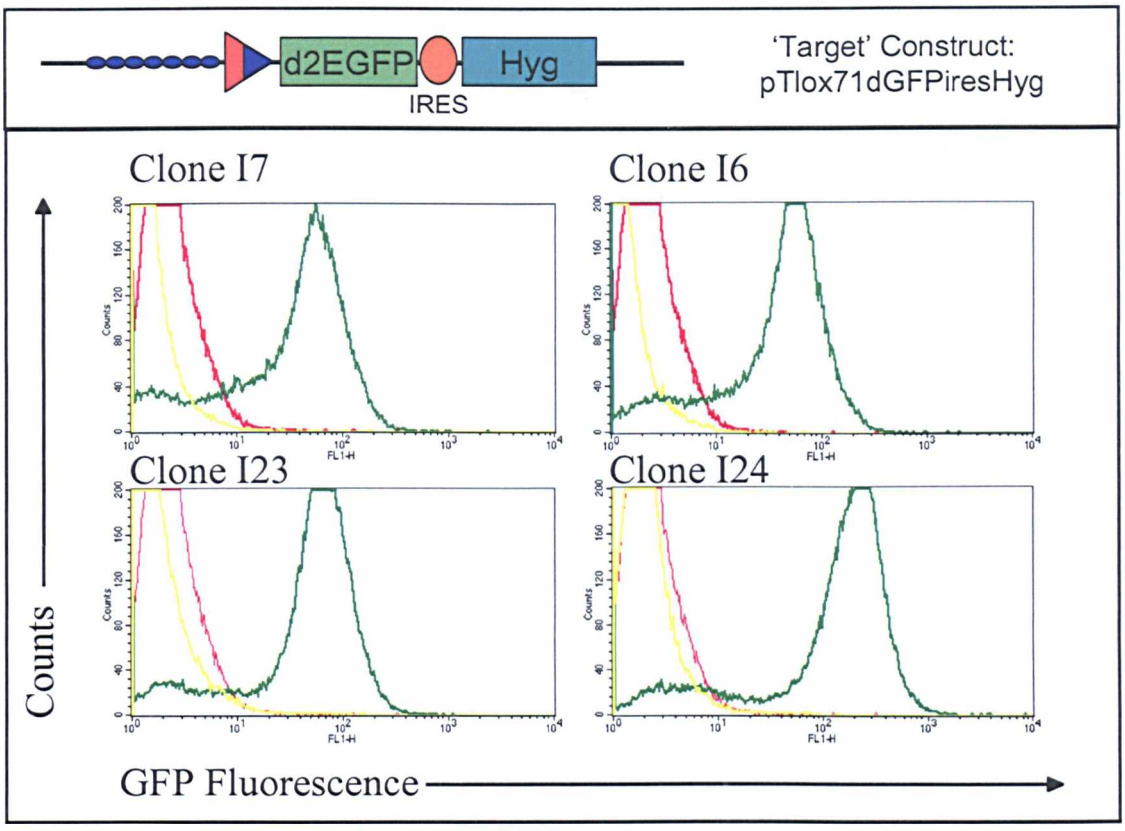



**Figure 3.10 GFP expression profiles of four pTlox71dGFPiresHyg-transfected clones:** This figure shows FACS profiles for four clones generated using the 'target' construct pTlox71dGFPiresHyg. 93% of the clones analysed had uninduced expression levels (yellow) as low as that of background fluorescence (red). However, induced expression (green) was reduced when compared to previous 'target' construct containing clones. See Appendix IV for instructions on how to interpret FACS plots.

Out of 30 colonies analysed, 28 (93%) were found to have uninduced expression levels at least as low as the background signal of untransfected HT2 cells (Figure 3.10). In fact, in some cases the uninduced expression levels were lower than background. The reasons for this were unknown. Unexpectedly, the maximum fluorescence (Figure 3.10) observed was ~10-fold lower than that seen with the previously tested 'target' constructs (Figures 3.6 and 3.8). As a result, the induction values for clones with the lowest background were lower than that for pTlox71dGFP, ranging from ~60-fold (Clone I7) to ~100-fold (Clone I24) (Figure 3.10). These results suggested that the IRES may impair the amount of induced GFP expression, possibly by reducing the efficiency of translation of the GFP ORF. The regulation of the GFP gene formed a representation of what to expect for the inserted gene of

interest, however, it was unclear what effects removing the IRES after the insertion of the GOI would have on $TRE_{mod}$ expression. A further disadvantage of using this 'target' construct for future experiments was therefore highlighted.

As anticipated the FACS analysis revealed a much-improved profile for the induced state (Figure 3.10). For clones such as I6, the removal of hygromycin from the medium for one week resulted in reversion to a mosaic pattern (Figure 3.11A), similar to that observed in previous experiments (Figure 3.8), confirming that hygromycin selection was responsible for the removal of low GFP-expressing cells in the induced state (-Tet). These results indicated that silencing of either the $TRE_{mod}$-d2EGFP or the tTA gene was responsible for down-regulation of GFP expression.

*The mosaic GFP expression profiles were a result of DNA methylation of the tTA*

CpG methylation is a mechanism used by mammalian cells to silence regions of the genome. The d2EGFP gene itself is unlikely to be silenced by methylation as its ORF has already been codon-optimised for expression in mammalian cells. The tetracycline system, however, is based upon DNA sequence originating from *E.coli* and it is therefore likely that any CpG rich regions were recognised as 'foreign' DNA and methylated (Yoder *et al.*, 1997). To determine whether the GFP gene silencing involved DNA methylation, two clones expressing GFP from either pTlox71dGFP (Clone T15) or pTlox71dGFPiresHyg (Clone I6, Figure 3.10) were grown in the presence or absence of 5-aza 2'deoxycytidine (5Azc), a methylation inhibitor, for 24 hours. After a further 24 hours FACS analysis was carried out (Figure 3.11B).

The results revealed that 5Azc treatment increased the proportion of GFP-positive cells (M2, Figure 3.11B) and decreased the number of GFP-negative cells (M1, Figure 3.11B), suggesting that methylation was the mechanism of GFP silencing in clones such as T15 and I6. To determine which sequence (the tTA or $TRE_{mod}$) was being silenced, a plasmid (pTIGHT-Luc, Clontech) containing the luciferase ORF downstream of a $TRE_{mod}$, was transiently transfected into cells (Clone T15 and I6) which had either been treated or untreated with 5Azc for 24 hours. A luciferase assays was carried out 24 hours after the removal of 5Azc and corrected for cell

number (Figure 3.11C). In this experiment, only tTA-encoding DNA (i.e. not $TRE_{mod}$) was exposed to 5Azc (assuming all 5Azc was efficiently removed prior to transfection). Thus any effect of 5Azc on luciferase expression would most likely reflect demethylation of tTA rather than $TRE_{mod}$.

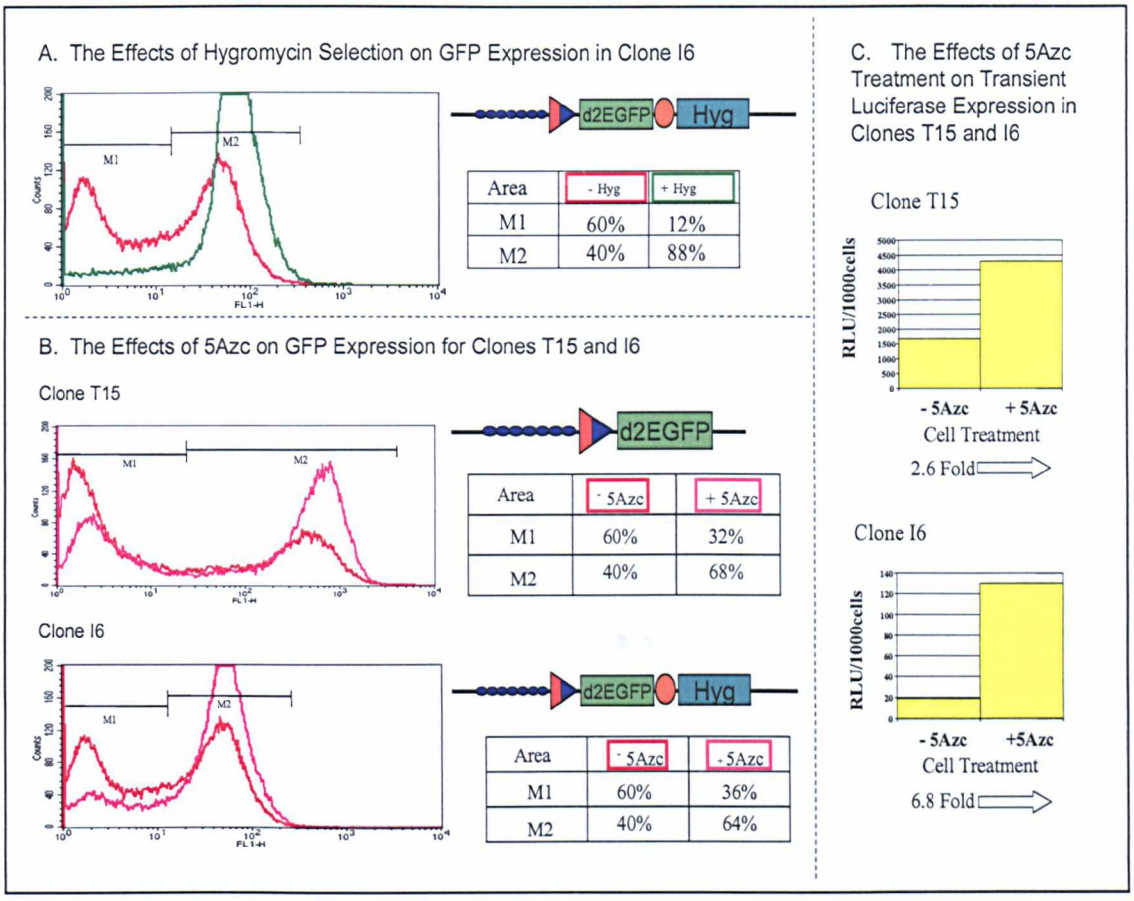



**Figure 3.11 The effects of hygromycin selection and 5-Azc on expression from the TRE: A.** Shows the effects of hygromycin expression on Clone I6. There was a marked increase in non-fluorescing cells (M1) when the selection pressure was not present. **B.** Shows the effects of the methylation inhibitor, 5-Azc, on GFP expression. There was an increase in GFP expressing cells (M2) when methylation was inhibited, as shown for Clones T15 and I6. **C.** Illustrates the effects of 5-Azc treatment on transient luciferase expression. An increase in luciferase expression was observed after treatment with 5-Azc for Clones T15 and I6, suggesting that methylation of the tTA was taking place.

The results clearly showed that there was an increase in the amount of luciferase expression in the 5Azc treated cells (Figure 3.11C). It therefore appears that the tTA cassette is progressively silenced by methylation during expansion of HT2 cells and their derivatives, causing increasing numbers of cells within a 'target' clone to spontaneously down-regulate their GFP expression. This effect would have been hard

to detect by Western analysis, the technique commonly used to observe gene induction, as cell to cell variations in expression within a clone cannot be detected by this method.

### 3.2.2 Inducible GFP in cells with an 'improved' transactivator (itTA)

*Successful generation of HT1080 clones expressing the improved transactivator*

An 'improved' tTA (itTA) gene (Krestel *et al.*, 2004) was obtained from Rolf Sprengel (Max-Planck-Institute for Medical Research, Germany). The codon-optimised tTA sequence was generated by eliminating potential splice sites and GC dinucleotide sequences using a PCR-based method. This technique, described previously by Stemmer *et al* (1995), uses PCR to synthesise a long synthetic DNA sequence from a large number of oligonucleotides. This method relies on DNA polymerase to generate progressively longer fragments of DNA, instead of using the conventional DNA ligase method. In addition, the VP16 activation domain was replaced with three synthetic activation domains, called F-domains (PADALDDFDLDML), to help reduce 'squelching' (discussed in section 1.3.1; Barron *et al.*, 1997). It was hoped that this improvement would limit the amount of transactivator gene silencing described in the previous section and therefore improve expression of GFP from the modified TRE.

The itTA ORF was cloned next to an SV40 promoter/enhancer in a zeocin expression vector (pZeoSV, Invitrogen) to produce pZeoSVhtTA (see section 2.3.1 for materials and methods). This vector (10μg) was transfected into 8 million wild-type HT1080 cells by electroporation (for conditions see Table 2.5). After selection in zeocin, 24 colonies were picked and assayed for their ability to regulate TRE-containing transcripts. This was done by transient transfection, using Lipofectamine 2000, with pTIGHTLuc (Clontech) in either the presence or the absence of tetracycline. Luciferase assays were carried out (see section 2.7 for methods) and after correcting for cell numbers, an induction value (non-corrected) was generated for each clone by dividing the luciferase activity in the absence of tetracycline by the activity in the presence of tetracycline. Clone Rht14 (shown in red, Figure 3.12) was chosen for

further experiments as it demonstrated a high amount of induction and low uninduced expression. Clone Rht1 was disregarded as it was later found to be tetraploid (data not shown).

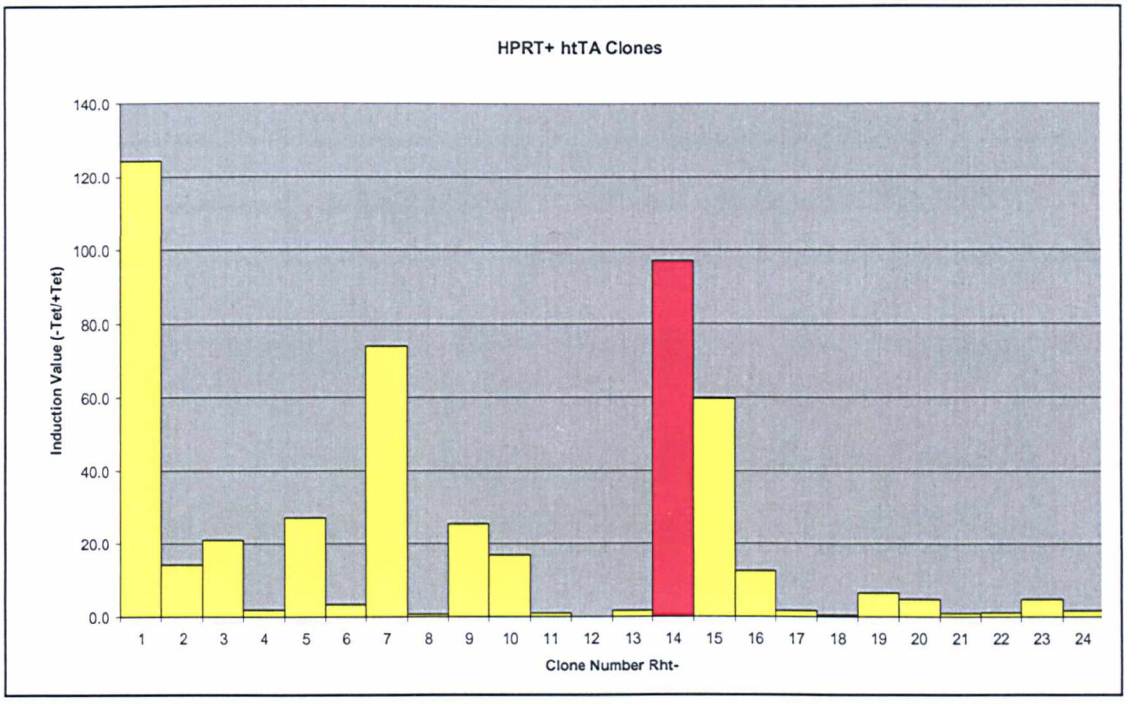



**Figure 3.12 Testing itTA expressing clones for their ability to regulate TRE driven luciferase gene expression:** 24 itTA expressing clones were selected in zeocin and once established were plated in the presence or the absence of tetracycline and their ability to regulate TRE-driven luciferase (from pTIGHT-Luc) was measured. The induction value of each tTA expressing clone was calculated. Clone Rht14 was selected (red) for further experiments as it demonstrated low uninduced expression levels and relatively high induction of luciferase.

*The improved transactivator supports tight regulation and minimal mosaicism*

pTFRTlox71dGFP (Figure 3.3F) was generated as described in the materials and methods (see 2.3.2). This 'target' construct was similar to pTlox71dGFP (Figure 3.3C), however it also contained an FRT site between the TREmod and the *lox71* site. The FRT site was included as part of a revised 'insertion' strategy that is explained later in this study (Chapter 5). This final 'target' construct was co-transfected with the puromycin-resistance construct (pBL-Pur/R), as described in 3.2.1, however on this occasion the transfection was carried out using the newly developed Rht14 cells. 32 puromycin-resistant GFP-positive clones were analysed by FACS. Of these, 7 (21.9%) had uninduced expression levels <2-fold greater than background

fluorescence (e.g. Clones Rht14-3 and 20, Figure 3.13). Furthermore, 2 clones (6.25%) were detected with uninduced GFP expression indistinguishable from background fluorescence (e.g. Clones Rht14-10 and 19, Figure 3.13). This frequency (6.25%) was higher than previous experiments (1.7%) using a similar construct in HT2 cells (Figure 3.8).

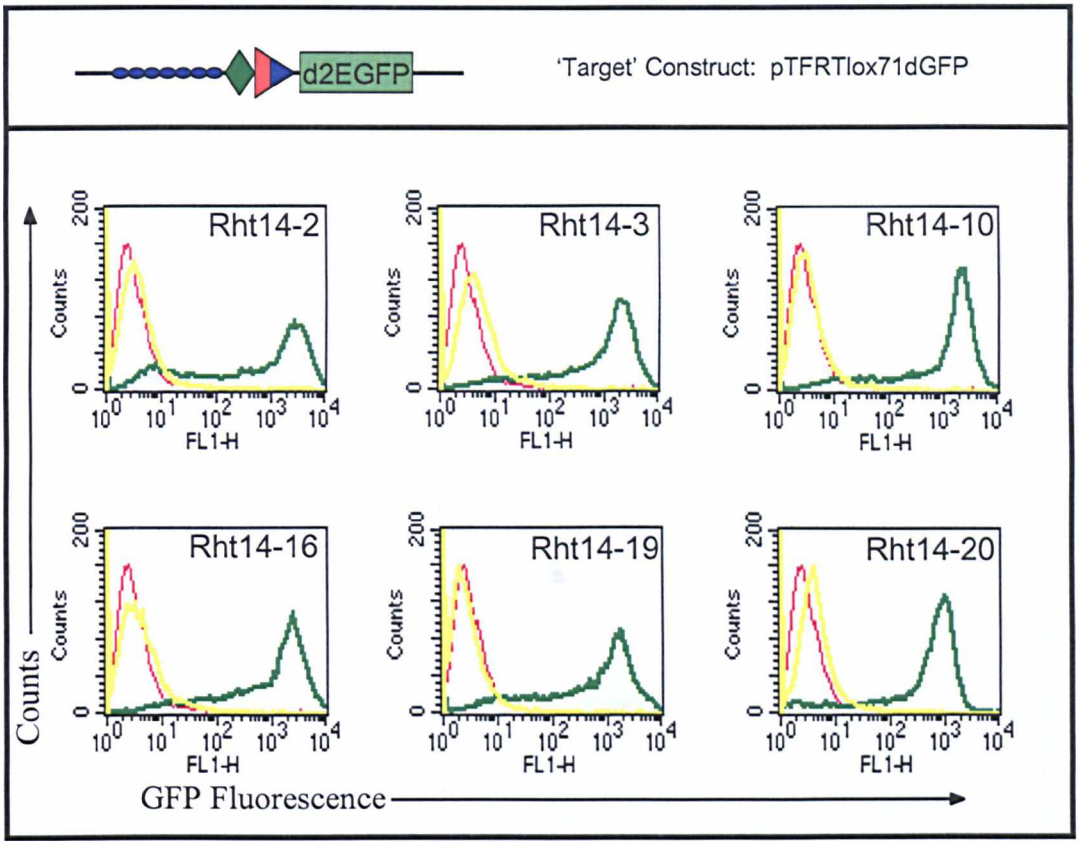



**Figure 3.13 FACS profiles from six favourite Rht14-derived clones containing pTFRTlox71dGFP:** Almost 30% of pTFRTlox71dGFP (top; not to scale) -derived Rht14 clones analysed had uninduced (yellow) expression similar to that of background (red) and 6% of those (Clones Rht 14-10 and 19) were identical to that of non-fluorescing cells (red). Of the desirable clones, the non-corrected induction values ranged from ~250-fold (Rht14-20) to ~1000 fold (Rht14-10) and the induced (green) expression profiles had improved. For instruction of how to interpret FACS plots see Appendix IV.

Induction values (see Appendix IV for an explanation) for clones with the lowest uninduced expression ranged from ~250-fold (Clone Rht14-20, Figure 13.12) to ~1000-fold (Clone Rht14-10, Figure 3.13), these were similar to values for equivalent clones expressing non-humanised tTA (ranged from~200-700-fold, Figure 3.8). FACS analysis also revealed an improved induced profile with reduced mosaicism,

which was consistent during several weeks in culture (Figure 3.14). This result was in agreement with previous experiments, which suggested that methylation of the tTA sequence was the cause of GFP gene silencing in clones such as I6 and T15 (Figure 3.11). Slight variations in GFP expression profiles were observed but were most likely caused by differences in cell confluency prior to analysis, as suggested previously (Damke *et al.,* 1995). When the cells were deliberately allowed to approach confluency before analysis, GFP expression profiles showed clear mosaicism (Figure 3.14). Subsequently, FACS analyses were therefore carried out on cells that were no more than 80% confluent (i.e. still dividing normally).

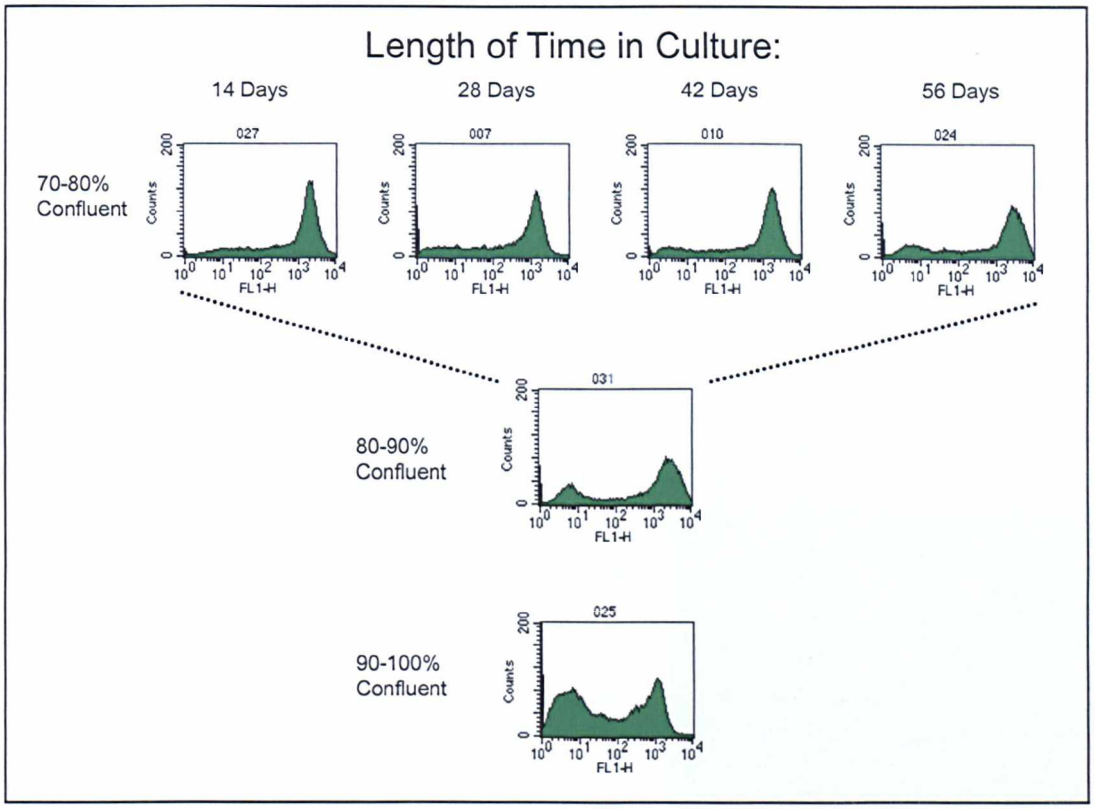



**Figure 3.14 The effects of time and confluency on the GFP expression profile of Clone Rht14-10:** The GFP expression profile of Rht14-10 appeared consistent for at least 8 weeks in culture. Increased confluency of Rht14-10 resulted in a disruption to the uniform pattern of GFP. All FACS analyses were carried out in the absence of tetracycline.

*The desirable clones were analysed for ploidy and copy number*

The ploidies of 7 well-regulated clones were checked by propidium iodide staining and flow cytometry, as described in section 2.10.1. Of the 7 clones analysed, one appeared to be tetrapoloid (Rht14-21, data not shown) and was rejected. A tetraploid

set of chromosomes would result in more than one copy of the 'target' construct being present in the genome of that particular clone and would likely result in unpredictable expression and regulation of any inserted GOI.

For a similar reason, clones with a single copy number of the 'target' construct were also required. This would ensure that the FACS profile of GFP expression would represent the fold-induction and tight regulation expected for the inserted GOI, following *Cre*-mediated recombination. Therefore, before continuing, tightly regulated clones were screened by Southern blot for 'target' site copy number (Figure 3.15A). Genomic DNA was extracted and digested with the appropriate enzyme, separated, blotted and probed with a portion of the GFP gene (see section 2.5 for methods). The results (Figure 3.15B) demonstrate that Clones Rht14-10, 16, 19 and 20 all contained one copy of the GFP cassette. Two of these clones (Rht14-10 and Rht14-19) were used for further experiments (Figure 3.15C).

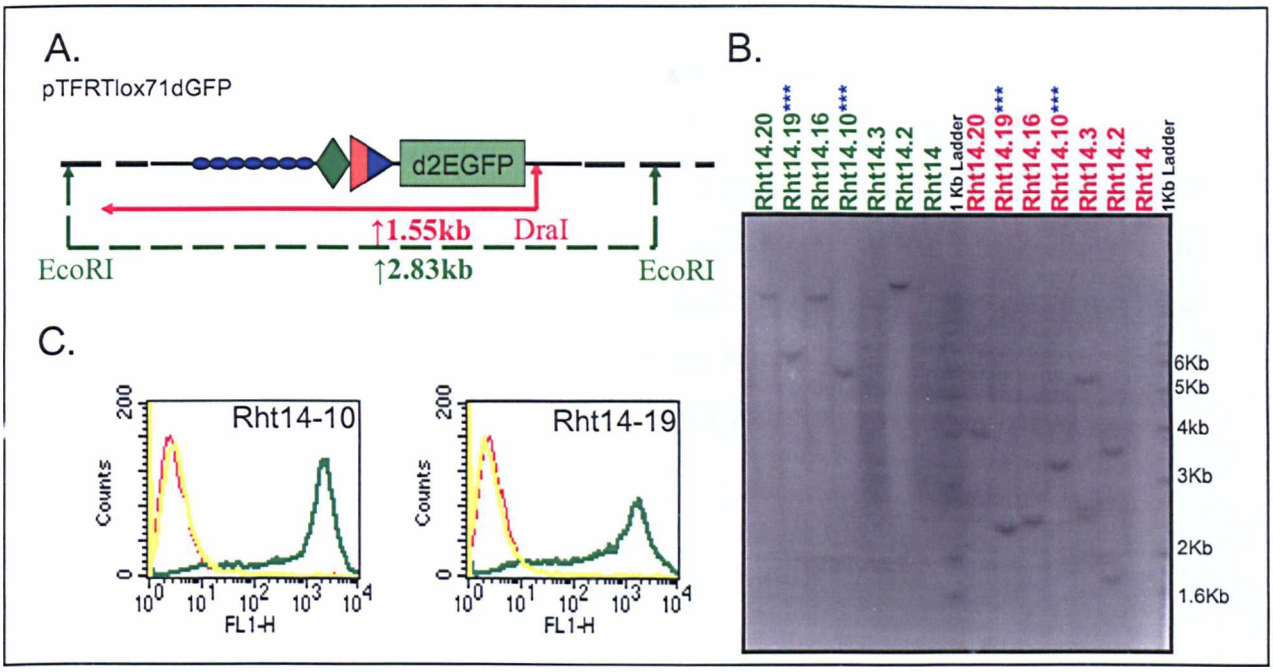



**Figure 3.15 Southern analyses to determine 'target' copy number in Rht14-derived clones: A.** Illustrates the restriction patterns expected at the 'target' locus for Rht-GFP-derived clones. Genomic DNA was digested with *Dra*I (red) and *Eco*RI (green) which cut once inside the 'target' construct or not at all, respectively. The minimum sizes expected when probed with a portion of the GFP gene are shown. **B.** A single copy number was represented by a single band for all digests carried out. Clones Rht14-10, -16, -19 and -20 all appeared to have a single copy number of the 'target' construct, however clones Rht14-2 and -3 did not. **C.** Clones Rht14-10 and -19 are used for further experiments, highlighted using *** in **B.**

### 3.2.3 The kinetics of GFP induction in clones Rht14-10 and -19

To measure the kinetics of GFP induction, clones Rht14-10 and -19 were analysed by FACS at different time points after the removal or addition of tetracycline. The mean fluorescence of all cells at each time point was generated using CellQuest software (Becton Dickinson). The mean GFP fluorescence when completely induced was set at 100% and all the other samples were calculated as a percentage of the induced mean GFP fluorescence which was then plotted against time (For Rht14-10 see Figure 3.16C). For GFP switch-off, it took approximately 90 hours to completely turn-off GFP expression (Figure 3.16A). However, 24 hours after the addition of tetracycline Clone Rht14-10 was only expressing 7% of the total GFP fluorescence and 1% by 48 hours (Figure 3.16A).

The kinetics of an inserted GOI would be affected by the stability of the protein generated, as a stable protein will take longer to degrade after gene switch-off. Alternatively, unstable proteins (such as d2EGFP) will degrade quickly and would therefore demonstrate faster kinetics. It took 48 hours to completely turn on the GFP gene (Figure 3.16B). The kinetics of Rht14-10 (Figure 3.16) and Rht14-19 (data not shown) were relatively slow (even for a quickly degrading protein) and suggest that the tetracycline system is limited to applications where quick on-off expression is not required (e.g. not suitable for expression during a particular point in the cell-cycle).

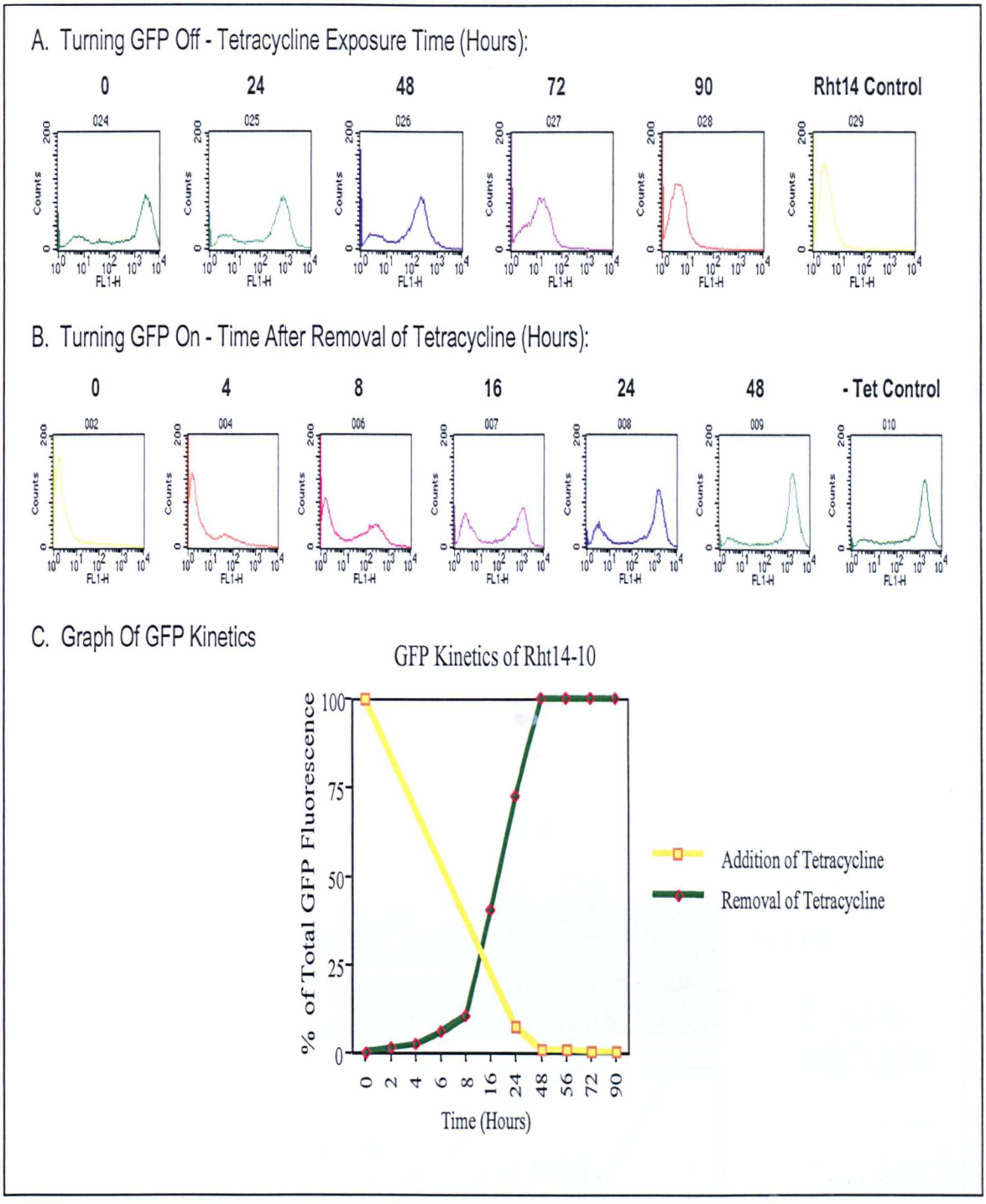



**Figure 3.16 GFP kinetics of Rht14-10: A.** Shows the kinetics of GFP switch-off for clone Rht14-10 after various lengths of time exposed to tetracycline. **B.** Illustrates the kinetics of GFP switch-on for clone Rht14-10 at various time intervals after removal of tetracycline. **C.** Here the GFP kinetics are illustrated on a graph as a percentage of mean induced GFP fluorescence. It appeared to take up to 90 hours to completely turn off d2EGFP and 48 hours to completely re-establish GFP.

### 3.2.4 Comparing the sensitivity of FACS with traditional methods

Whilst working with Clones Rht14-10/19 a number of advantages of using this FACS based method over traditional methods for analysing inducible gene expression (i.e. Western analyses) became obvious and are discussed below.

*Low GFP-expressing clones were detected by FACS but not by Western blot*

Tight, inducible gene expression is usually demonstrated by Western blot. However, the sensitivity of the Western determines the accuracy. Figure 3.17 shows a Western analysis for GFP expression on protein samples taken from cells with the displayed FACS profiles (for methods see section 2.6). The FACS profile represents cells that have been exposed to tetracycline for various time intervals. Western analysis suggested that the GFP protein disappears after 24 hours of exposure to tetracycline, however FACS analysis clearly illustrates that there is still GFP present in these cells which is approximately 25-fold greater than that of non-fluorescing cells.

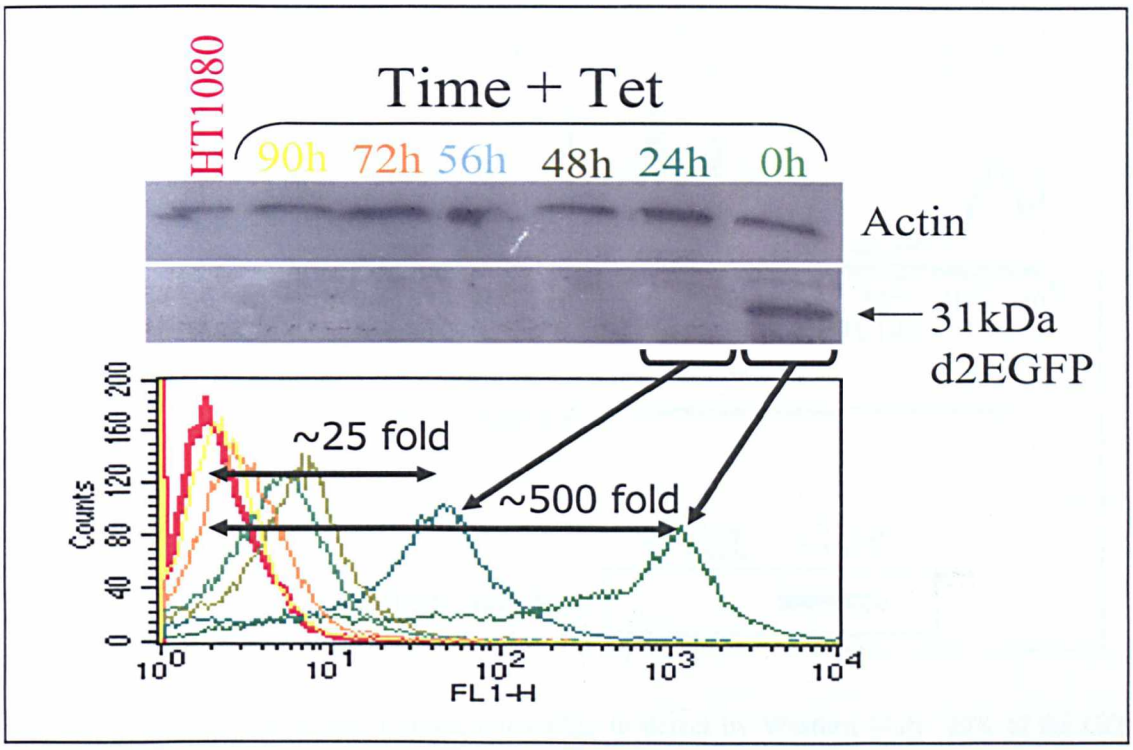



**Figure 3.17 FACS is more sensitive than Western analysis:** Rht14-19 cells were exposed to tetracycline for various time intervals. The cells were then analysed by both FACS and Western blot for GFP expression. GFP was not detected after 24 hours by Western, however, GFP expression was still 25-fold greater than that of background (red) on the FACS profile.

*Uniform induced expression profiles were not observed for all GFP-positive clones*

Approximately 20% of the clones analysed displayed non-uniform expression profiles in the induced state when analysed by FACS, despite having tight regulation (Figure 3.18). It was not possible to pick up these variations in GFP expression within a clone by Western analysis.

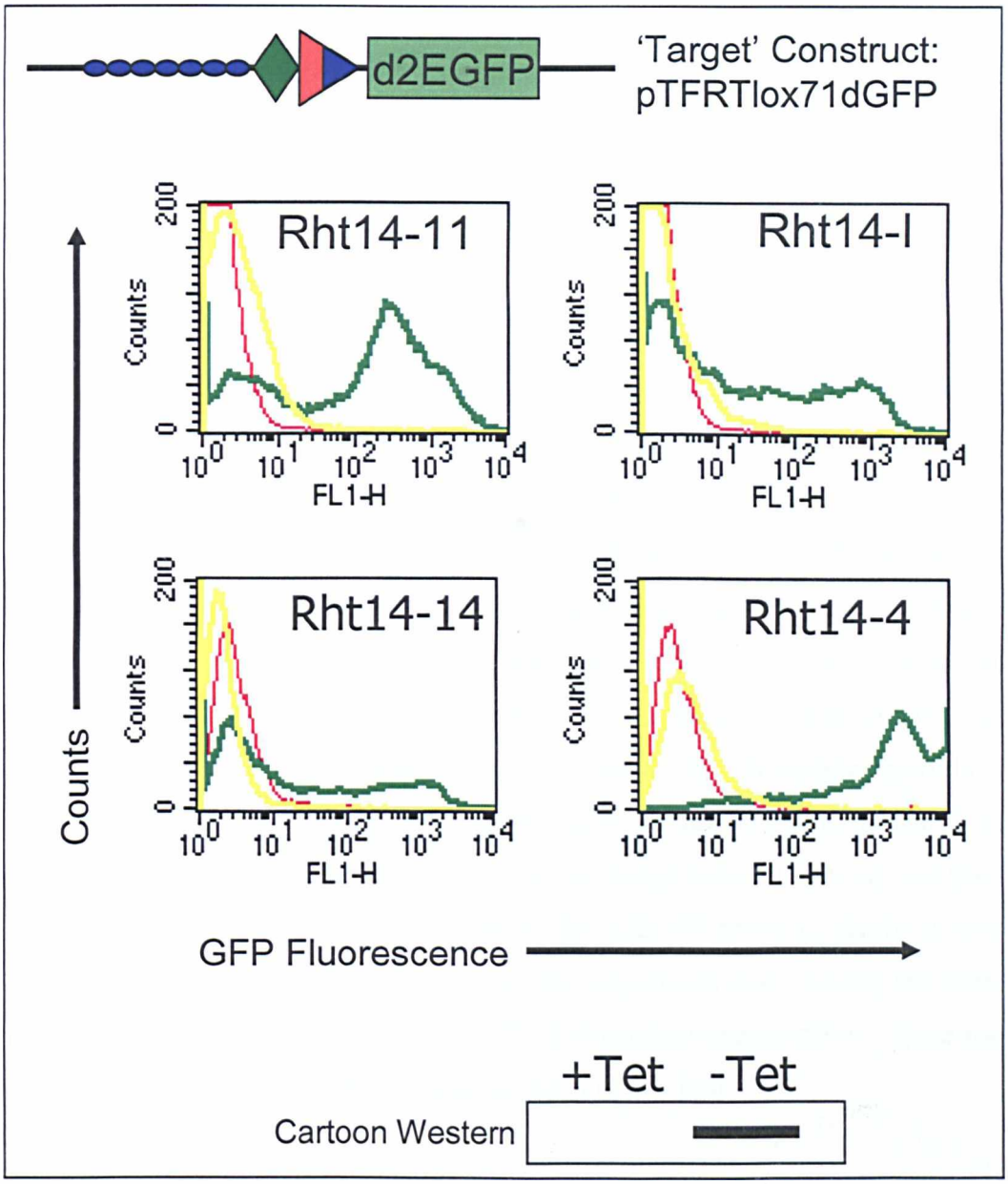



**Figure 3.18 Variations within a clone, impossible to detect by Western blot:** 20% of the GFP expressing clones analysed, generated using the 'target' construct pTFRTlox71dGFP (top; not to scale) in Rht14 cells, had non-uniform expression profiles in the induced stated despite having low uninduced expression. These variations within a clone would not be detected by Western blot and would simply be represented as an on and off signal (see cartoon Western; bottom).

*Analysis of dose response revealed a non-uniform pattern of switch-Off*

It had previously been suggested that the amount expression from the TRE could be controlled by the addition of different concentrations of tetracycline, calculated from a dose response curve (Damke *et al.*, 1995). A dose response curve was normally generated by quantifying the protein expression seen on a Western blot, generated from cells exposed to increasing concentrations of tetracycline and plotting the results. To generate a dose response curve for Rht14-10/19 FACS plots were generated for different concentrations of tetracycline and the mean fluorescence from each plot was determined using CellQuest (Becton Dickinson) software and the % of the total GFP fluorescence was calculated.

This experiment revealed that GFP expression was not silenced incrementally with progressive increases in tetracycline concentrations, but rapidly silenced completely by demonstrating a fast exchange between the on and the off states (Figure 3.19A). For example, when attempting to express GFP at 50% of the total fluorescence the dose response curve suggests using between 1-2ng/ml of tetracycline. However, the FACS profiles at 1-2μg/ml indicate that not many cells within the clone actually expressed at 50%. Rather, two main populations of cells exist, those expressing 100% induced levels and those expressing ~0% uninduced levels. A similar analysis was carried out for a clone expressing the original EGFP 'target' construct in Rht14 cells (Figure 3.19B) to determine whether the rapid exchange between the on and the off states was dependant on the unstable nature of the d2EGFP protein. Again, a similar result was obtained for EGFP and supported the hypothesis that varying the dose of tetracycline did not incrementally regulate TRE-dependant transcription. Once more, these findings would not have been observed by Western blot.

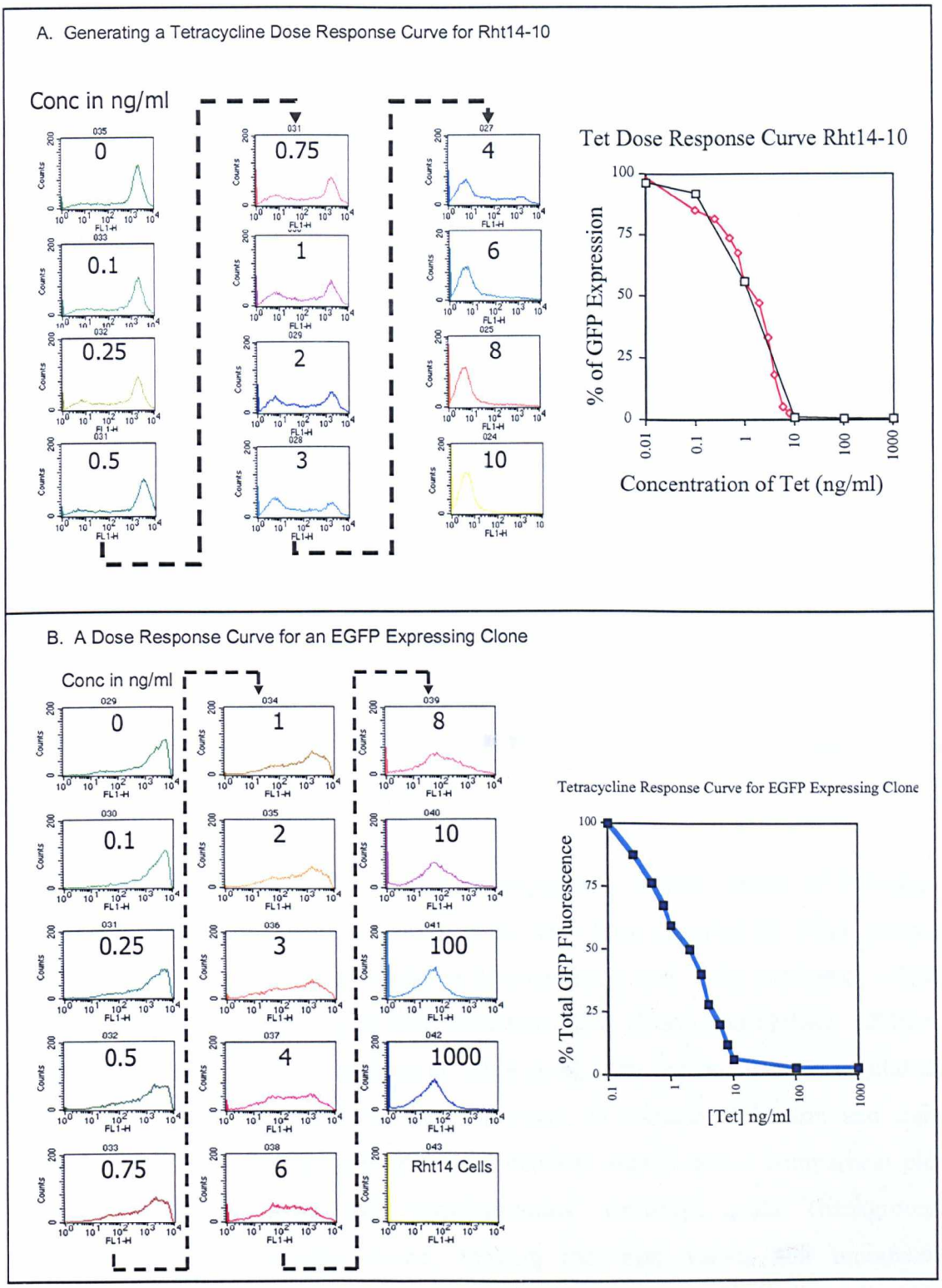



**Figure 3.19 Tetracycline dose response curves for Rht14-10 and an EGFP containing clone:** This figure illustrates the concentrations of tetracycline required to regulate two different forms of GFP (**A.** d2EGFP and **B.** EGFP). It revealed that there was not the uniform pattern of switch-off expected (Damke *et al.*, 1995) and that there was a rapid exchange between the on and the off states.

## 3.3 Discussion

### 3.3.1 The TREmod can achieve tight, highly inducible GFP regulation

The first two 'target' constructs tested, pTREIox71EGFP and pTREIox71dGFP (Figure 3.3A and B), were made using the original TRE and failed to generate any sufficiently regulated clones in HT2 host cells; the lowest uninduced levels of expression observed were ~5-fold greater than background (Figure 3.6). When the TRE was replaced with a modified version (TREmod), to generate pTlox71dGFP (Figure 3.3C), uninduced expression levels identical to background fluorescence were generated at a frequency of ~1.7% in HT2 transfectants (Figure 3.8) and 6% in Rht14-derived clones (Figure 3.13). The modifications made to the TRE were therefore crucial in generating clones with tightly regulated GFP. The induction (non-corrected) values achieved for the tightly regulated clones when using TREmod for expression, ranged from 250- to 700-fold (T24 and T21, respectively; Figure 3.8) for HT2-derived clones and 500– to 1000-fold (Rht14-19 and Rht14-10, respectively; Figure 3.13) for Rht14-derived clones.

Using similar TRE-tagged GFP cassette arrangements, a wide variety of induction values and basal uninduced expression levels have been reported by other groups; examples include the generation of clones demonstrating very leaky expression which in turn leads to a modest 10-fold GFP induction value (Izumi and Gilbert, 1999) to clones exhibiting a 2000-fold induction value along with apparently tight regulation (Qu *et al.*, 2004). The methods used, however, to measure induction and tight regulation have not always been consistent between studies and a comparison plot between uninduced cells and non-fluorescing wild-type cells (background fluorescence) is not always shown, making induction values and uninduced expression levels between studies difficult to compare. The system described in this thesis involves a more detailed comparison of uninduced TRE-expression compared to background expression levels and has therefore led to it being one of the more sensitive strategies described for successfully identifying tightly regulated clones.

The improvement in tetracycline regulation observed when using the TREmod has recently been supported by Tomac and colleagues (2004) who carried out a direct comparison of three different types of TREs (including the original TRE (Gossen and Bujard, 1992), TREmod (Clontech), and tk-TRE (Backman *et al.*, 2004)) in their ability to tightly regulate the luciferase ORF using the Tet-ON system in HEK293 cells. They demonstrated that the TREmod could achieve the highest overall induction (50-fold), with the lowest uninduced expression levels, when compared to the activities of the other two promoters. The second-generation tetracycline-regulated promoter is an alternative version of the TRE which has recently been developed and it is reported to produce the tightest, most inducible transgene regulation (Mohammadi *et al.*, 2004). The TRE is made up of eight tetO sequences (instead of seven) and a shortened CMV minimal promoter (like TREmod; Figure 3.7) and these modifications can facilitate tight transgene regulation and extremely high inducibility ($10^5$; transiently) in HeLa cells (Mohammadi *et al.*, 2004). This induction value is 500- to 1000-fold higher than the regulation efficiency in HeLa cells using the original TRE. This promoter may therefore be desirable for expressing a GOI that requires tight regulation, along with a high level of induced expression.

### 3.3.2 GFP is a suitable reporter for analysing expression from the TRE

Using the reporter gene, GFP, as a marker for analysing for TRE-regulation, was extremely successful. Being able to generate tightly regulated clones (using the TREmod) at a frequency of 6% may have suggested that the need to generate a system such as the 'select and insert' strategy, which is being designed and tested in this study, is no longer necessary. However, comparison of a FACS analysis with a Western blot revealed that uninduced GFP expression levels as high as 25-fold greater than background were not detected by immunoblot analysis (Figure 3.17). This discrepancy obviously depends on the particular antibodies used for the Western blot; however, it does indicate a potential limitation when using Western analysis as proof of tight regulation.

GFP expression analysis not only provided an indicator for tight-regulation by identifying levels of expression that could not be detected by more traditional methods, but it also highlights problem clones that are not uniformly expressing the transgene. 20% of the clones analysed demonstrated mosaicism, i.e. variations in GFP expression between cells within a clone (Figure 3.18). A study which assessed GFP expression from a cassette integrated at random genomic sites at various passage numbers concluded that only rarely does a transgene integrate at a site that permits long-term expression in every cell of a clone; typically the transgene is lost over a period of 60-250 generations from a clonal population (Migliaccio *et al.*, 2000). By using FACS analysis, clones demonstrating variations in GFP expression could easily be distinguished from uniformly expressing clones; however deviations of expression within a clone would not be picked up by Western blot and would ultimately go undetected.

In addition, tTA silencing by methylation would not have been so easily identified by Western analysis and again may have otherwise gone unnoticed. These types of limitations of monitoring expression from the TRE would not be so important if expression of the linked transgene was essential for the cells survival, as all down-regulated clones would be lost from the population (Carpenter and Porter, 2004). However, when it is important that all cells within a clone are expressing at the same amount, these limitations need to be resolved.

### 3.3.3 The GFP profiles improved after a method of selection was added

To reduce the mosaicism by killing-off those cells that down-regulated their GFP expression a selection was added to the 'target' construct. Initially, a puromycin cassette was included (Figure 3.3D), however this cassette resulted in undesirable GFP regulation for both the induced and uninduced profiles. This confirmed previous reports that suggested that inclusion of enhancer elements near or around a TRE impairs tetracycline regulated gene-expression (Damke *et al.*, 1995). A second selectable 'target' construct was therefore generated which included an IRES-hygromycin cassette downstream of the GFP gene (pTlox71dGFPiresHyg, Figure

3.3D). Although selection in hygromycin was successful in reducing the mosaicism (Figure 3.10) and uninduced levels of GFP expression were extremely low, induced levels of GFP expression (and therefore induction values) were lower than that seen for the previous 'target' constructs tested (Figure 3.10).

From these experiments it could be concluded that the induction of GFP was impaired when proceeded by an IRES-hygromycin cassette. The reasons for this effect were unclear; however it was possible that high over-expression of hygromycin reduced cell viability or that expressing a gene as part of a bicistronic message reduced the total amount of expression that could be achieved. One of the attractions of the planned 'screen and insert' strategy is that the pattern of expression observed for GFP will be adopted by the GOI following its insertion. If the presence of an IRES is effecting the amount of expression from the TRE then it is unsuitable for use in this system.

### 3.3.4 The 'improved' tTA achieves long-term stable expression.

Evidence that the mosaicism was caused by methylation of the tTA came initially from experiments using the methylation inhibitor 5Azc (Figure 3.11). To reduce the problem of mosaicism and provide further evidence that is was caused by down-regulation of the tTA, an 'improved' transactivator (itTA; Krestel *et al.*, 2004) was used and cells expressing itTA were successfully generated. Clone Rht14 was chosen for further experiments as it demonstrated transient inducible luciferase and relatively low transient uninduced luciferase expression. The induced FACS profiles were much improved in the Rht-derived clones and TRE expression was consistent through numerous weeks in culture (Figure 3.14). Cell confluency appeared to be the main factor affecting the GFP expression profile (Figure 3.14).

Most documented humanised transactivators involve the Tet-ON version (rtTA) (Urlinger et al., 2000; Anastassiadis *et al.*, 2002), most likely due to the need for a codon optimised system for work involving transgenic mice. An alternative improved rtTA, called rtTA2S-M2, contains 5 amino acid substitutions embedded in a synthetic

codon optimised sequence for increased sensitivity to tetracycline; and a minimal VP16 activation domain to reduce 'squelching' (Urlinger *et al.*, 2000). Similarly, using GFP as a reporter of TRE expression, Hillen and colleagues (2000) showed that these improvements have helped towards reducing uninduced expression and mosaicism, however only in a low percentage of clones (Urlinger *et al.*, 2000). When the TRE-GFP cassette is flanked by two chicken β-globin insulators, however, a reduced number of clones with high basal and mosaic expression were obtained. Finally, for both cases the pattern of GFP expression was stable over several passages (34 passages - 7 months in culture; Qu *et al.*, 2004).

An alternative method for maintaining tTA expression in mammalian cells has been tested which involves linking its expression to a drug resistance cassette using an IRES sequence (Izumi and Gilbert, 1999). This arrangement means that any cell which down-regulates or rejects its tTA cassette will automatically be removed from the population by the selective pressure. Using an IRES-hyg linked GFP as the reporter transgene, this approach was shown to greatly improve the mosaic pattern of induced TRE expression, when compared to an IRES-less approach (Izumi and Gilbert, 1999). This investigation was limited, however, by failing to examine the effects of prolonged passage on the GFP expression profiles in these clones.

### 3.3.5 The 'desirable' clones generated using an itTA and a TREmod.

In summary, cells expressing tight, inducible, stable GFP were successfully generated (3.13) using the itTA ORF and the TREmod. Those clones which maintained a diploid genome (86%) and contained a single copy of the 'target' construct (Rht14-10, -16, -19 and -20) became the 'desirable' clones (Figure 3.15). Rht14-10 and Rht14-19 were primarily used in further experiments. Rht14-10 has an induction value of nearly 1000-fold and uninduced expression as low as that of background. Rht14-19 has an induction value of 500-fold and also has uninduced expression as low as that of background (Figure 3.15). The kinetics for both clones were analysed and it revealed that it took almost 90 hours for the GFP gene to completely turn off and 48 hours for the GFP gene to turn completely on (For Rht14-10 see Figure 3.16). These results

suggested that applications for the tetracycline system alone must be limited to those which do not require rapid switch on/off. In other words, unless modified, this system will not be suitable for expressing a GOI at a particular point in the cell-cycle or during a particular experiment (e.g. during a transfection).

### 3.3.6 Further work

In light of the above findings it would be desirable to generate a bank of inducible clones, each tightly regulated but with varying induction values, where the amount of induced expression could be chosen to suit the GOI. Presently, there does not appear to be much variation in inducible expression (500-1000-fold) for tightly regulated clones, however, this may be due to the initial visual method of selection for GFP expression using a UV microscope, which may favour those clones that are expressing highly. It may be easier to generate a full range of clones with varying induction values by cell-sorting pools of transfected cells for various levels of GFP expression, prior to selection in puromycin.

In addition, it would be interesting to study the regions of the genome which facilitate tight regulation, in clones such as Rht14-10 and -19, as it would allow the identification of a few rare genomic regions that permits long-term expression and reduced leakiness. This could be achieved by a process known as inverse PCR (Ochman *et al.,* 1988). Inverse PCR is a method for rapidly amplifying sequences that flank a region of known sequence (in this case the TRE-GFP cassette). This method utilises the polymerase chain reaction (PCR), but it has primers orientated in the reverse direction of the usual orientation. The template for the reverse primers is a restriction fragment that has been ligated upon itself to form a circle. Sequencing of the resulting PCR products and a database search for the 'unfamiliar' sequence will reveal where in the genome the construct has integrated.

# Chapter Four

# Testing a Simple 'Screen and Insert' Strategy

## 4.1 Introduction

In the previous chapter tetracycline-regulated *loxP*-tagged GFP expressing clones were generated. This chapter describes how these clones were used to test and establish a simple 'screen and insert' strategy. Three 'target' clones were used and their characteristics are summarised in Figure 4.1. The choice of 'target' clone depended on what was available at the time of the experiment. Early experiments used clones 6 and d46 (Figure 4.1A and B, respectively; both HT2-derived clones), neither of which were very tightly regulated or characterised fully; however, they were suitable for initial strategy testing by examining the efficiency of Cre-mediated insertion in HT1080 cells. Clones Rht14-19 (Figure 4.1C) and Rht14-10 were used in later experiments for developing two 'screen and insert' approaches. The first 'screen and insert' strategy, System One, will be described in the second half of this chapter and the results using clone Rh14-19 (Figure 4.1C) will be discussed.

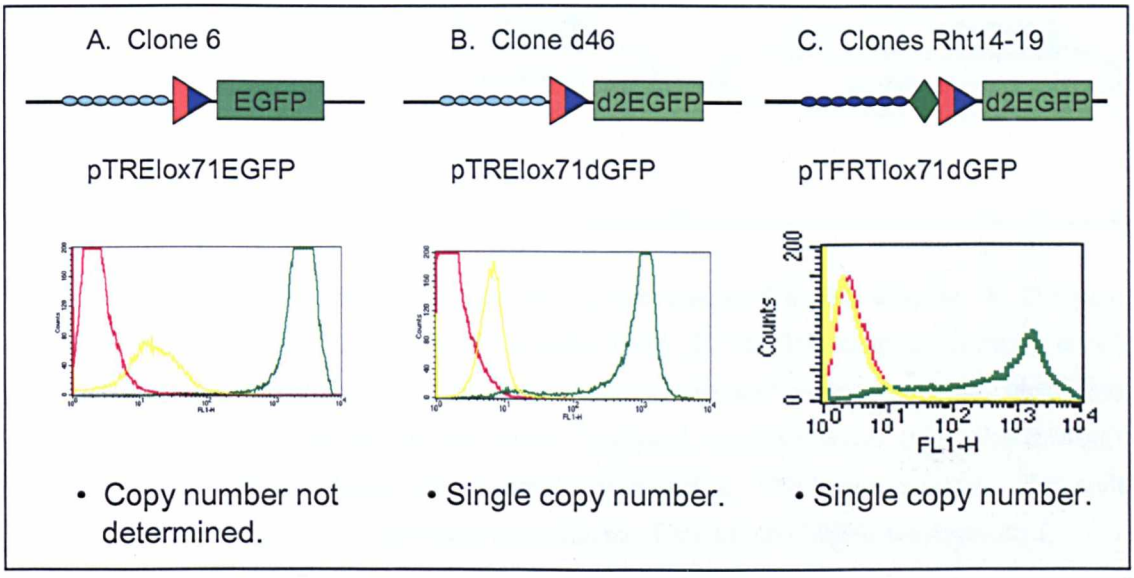



**Figure 4.1 The 'target' clones used to test the insertion step:** **A.** Shows the 'target' construct used to generate clone 6 (pTREIox71EGFP) and its GFP expression profile. **B.** Shows the 'target' construct used to generate clone d46 (pTREIox71dGFP) and its GFP expression profile. **C.** Shows the 'target' construct used to generate clones Rht14-19 (pTFRTIox71dGFP) and its GFP expression profiles. DNA, not to scale. For instruction on how to interpret the FACS plots see Appendix IV.

The system under development principally involves Cre-mediated insertion of a GOI downstream of a previously characterised tightly regulated promoter in a 'target' clone (Figure 4.2A). It is expected that the regulation of the GOI in the resulting clone will mimic the GFP expression in the parental. The mutant *loxP* system (*lox66/lox71,* see section 1.1.4) was used to reduce unwanted excision events after insertion during transient *Cre*-recombinase expression.

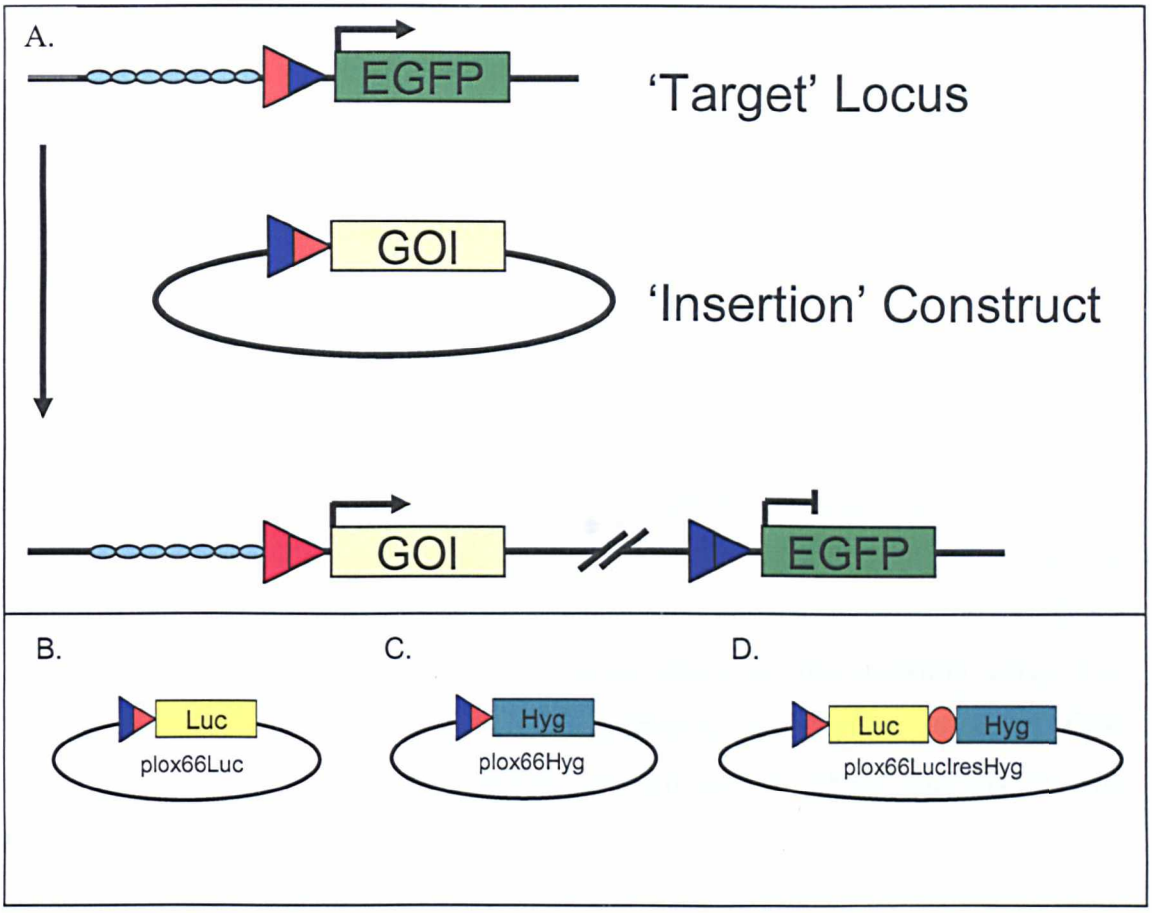



**Figure 4.2 The three 'insertion' constructs tested and discussed in this chapter: A.** The basic 'screen and insert' strategy under investigation in this study. **B.** The first 'insertion' construct tested, plox66Luc (non-selectable). **C.** Illustrates the second 'insertion' construct created, plox66Hyg (selectable). **D.** The first 'screen and insert' 'insertion' construct tested (plox66LucIresHyg), incorporating the GOI (luciferase) and a method of selection. DNA, not to scale. For scale representations of the 'insertion' vectors see Appendix III. For a key to symbols see Appendix I.

The general criteria for designing an 'insertion' construct included a mutant *lox66* site cloned upstream, and a polyadenylation signal (polyA) cloned downstream of a promoterless GOI. A promoterless GOI was important as it reduced the chance of isolating random integrants after the insertion event. Similarly, selectable markers

together with their promoter/enhancer elements were not included in the 'insertion' constructs due to their possible effects on expression from the TRE after insertion (see Figure 3.8). For testing purposes the luciferase ORF was used as the model GOI.

Three 'insertion' constructs will be discussed in this chapter. The first two (plox66Luc and plox66Hyg; Figure 4.2B and C; Appendix III) were simple test constructs to examine the efficiency and viability of *Cre*-mediated insertion in HT1080 cells. Initially it was hoped that the frequency of *Cre*-mediated recombination would be high enough so that clones which had undergone the insertion event could be identified by fluorescence microscopy, by analysing for those that had lost GFP expression. Preliminary results revealed, however, that a method of selection was required.

The third 'insertion' construct (plox66LuciresHyg, Figure 4.2D; Appendix III) was designed in accordance to the preliminary findings to generate the first 'screen and insert' strategy (System One) and it takes into account the criteria listed above. It contains an internal ribosomal entry site (IRES) to express both the GOI and the hygromycin-resistance ORF from the TRE after insertion without the need for additional promoter/enhancer elements. After selection, the insertion event was further characterised by loss of GFP expression, tetracycline regulated GOI expression, PCR for the recombination junction and Southern analysis for the expected DNA rearrangements.

## 4.2 Results

### 4.2.1 Analysing pools of plox66Luc transfected cells

*Optimising the transfection efficiency of HT1080 cells using Lipofectamine 2000*

Lipofectamine 2000 was chosen to co-transfect the 'insertion' and *Cre*-expressing constructs as it had previously been shown to achieve very high transfection efficiencies for transient expression (Vasquez *et al.*, 2001). To determine the amount

of DNA required to attain optimal transfection efficiencies using Lipofectamine 2000 in HT1080 cells, pTRElox71dGFP (Figure 4.1B) was used. Various amounts of the construct were transfected, as described in the materials and methods, into Rht14 cells (itTA expressing HT1080 cells, discussed in section 3.2.2) and analysed for GFP expression by FACS after 48 hours. As suggested in the manufacturers' instructions, optimal tranfection efficiency was achieved when using approximately 3-4µg of DNA, with at least 90% of cells expressing transient GFP (Figure 4.3).

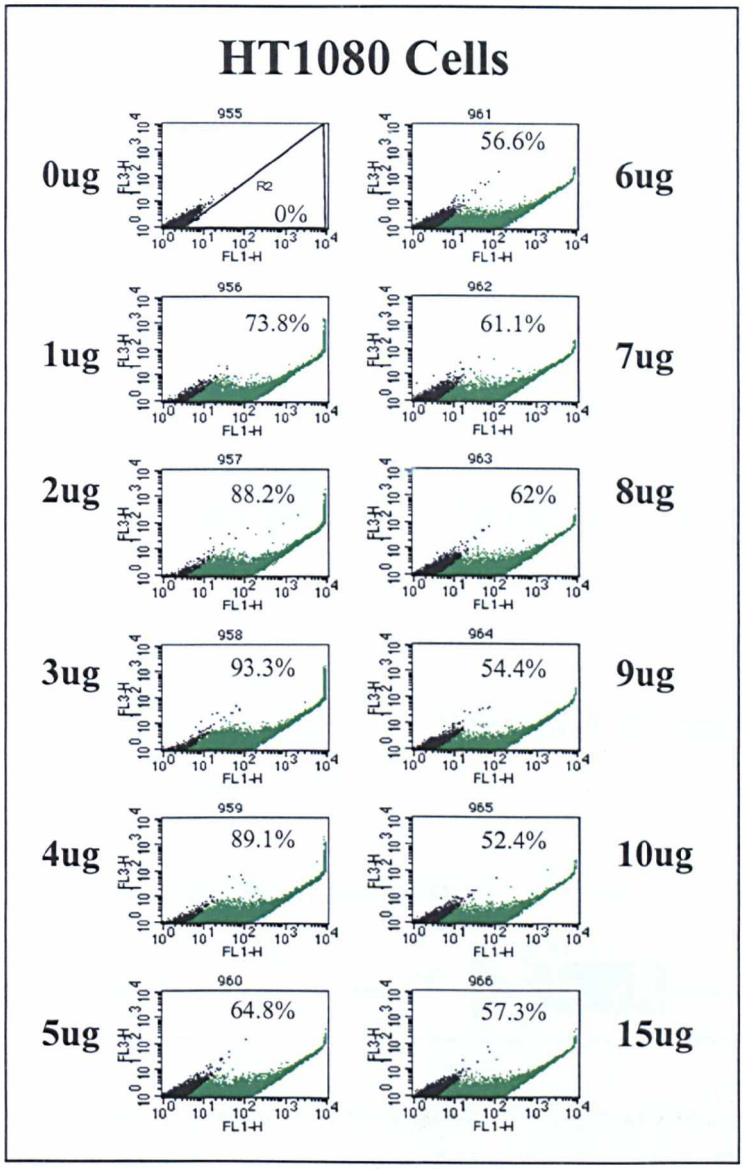



**Figure 4.3 Optimising the transfection efficiency using Lipofectamine 2000:** The indicated amounts of pTlox71dGFP were transfected into Rht14 cells with Lipofectamine 2000 (see section 2.8.5).

*Testing for Cre-mediated insertion without a means of selection*

The first 'insertion' construct tested (plox66Luc; Figure 4.4) contained a promoterless luciferase gene preceded by a *lox66* site. A luciferase assay is extremely sensitive; therefore even a low efficiency of insertion events should be detectable by this method. In addition the insertion of plox66Luc is not selectable but it may be possible to identify targeted clones by analysing for loss of GFP expression. Experiments using this construct were carried out prior to the generation of the most desirable 'target' clones Rht14-10 and -19, therefore one of the earlier clones (clone 6) was used to test the efficiency of the insertion event. Clone 6 expressed tetracycline regulated EGFP (from the original TRE) in HT2 cells, with a GFP induction value of ~200-fold and uninduced expression ~10-fold above background (Figure 4.1A). The 'target' copy number of clone 6 was not determined; however for the purpose of detecting whether the desired insertion event was possible, a single copy number was not essential.

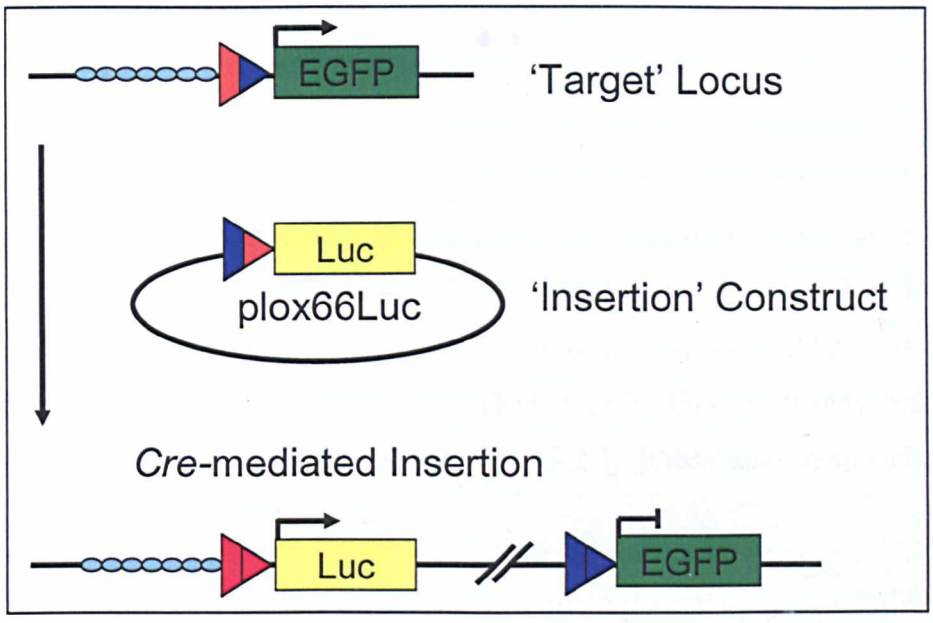



**Figure 4.4 Targeted insertion of plox66Luc in clone 6 cells:** The expected *Cre*-mediated insertion of plox66Luc downstream of the TRE in clone 6 cells is illustrated (not to scale). Pools of transfected cells were analysed for inducible luciferase expression and loss of GFP expression by FACS analysis. Colonies formed from transfected cells plated at low density were also analysed for loss of GFP expression using a UV microscope. Insertion results in tetracycline regulated luciferase expression and loss-of-GFP exression. For key to symbols see Appendix I.

Two million clone 6 cells were co-transfected (by lipofection; for methods see section 2.8.5) with a *Cre*-expressing plasmid (pMC-Cre) and plox66Luc and replated in either the presence or the absence of tetracycline. For each experiment a control was carried out in which pMC-Cre was omitted. To determine whether the insertion event had taken place (Figure 4.4) pools of transfected cells were subjected to luciferase assays, PCR and FACS analysis.

Luciferase assays were carried out 48 hours following transfection and the values generated were corrected for cell number. The results clearly revealed tetracycline responsive luciferase expression in pools of transfected cells (Figure 4.5A), showing evidence that the expected insertion event had taken place. The induction values (non-corrected: luciferase activity –Tet/+Tet) observed for luciferase activity in the pools when tetracycline was removed from the medium was 193-fold (50029/259; Figure 4.5A), this was similar to the fold-induction of GFP observed for clone 6. In addition, the uninduced luciferase activity was ~8-fold (259/33; Figure 4.5A) greater than that of untransfected and control cells, similar to the levels of uninduced GFP expression in clone 6. It therefore appears that GFP expression in clone 6 was a good indicator of GOI expression levels following *Cre*-mediated integration. The very slight increase in luciferase activity between the wild-type and the control samples was most likely a result of random integration of the 'insertion' vector into sites with a nearby enhancer. Over time, pools of transfected cells were found to lose their luciferase activity (Figure 4.5B). As clone 6 was derived from HT2 cells and not those expressing humanised transactivator (Rht14 cells) this result may reflect tTA silencing by methylation, as discussed in section 3.2.1. It may also be silencing of the newly integrated luciferase gene itself.

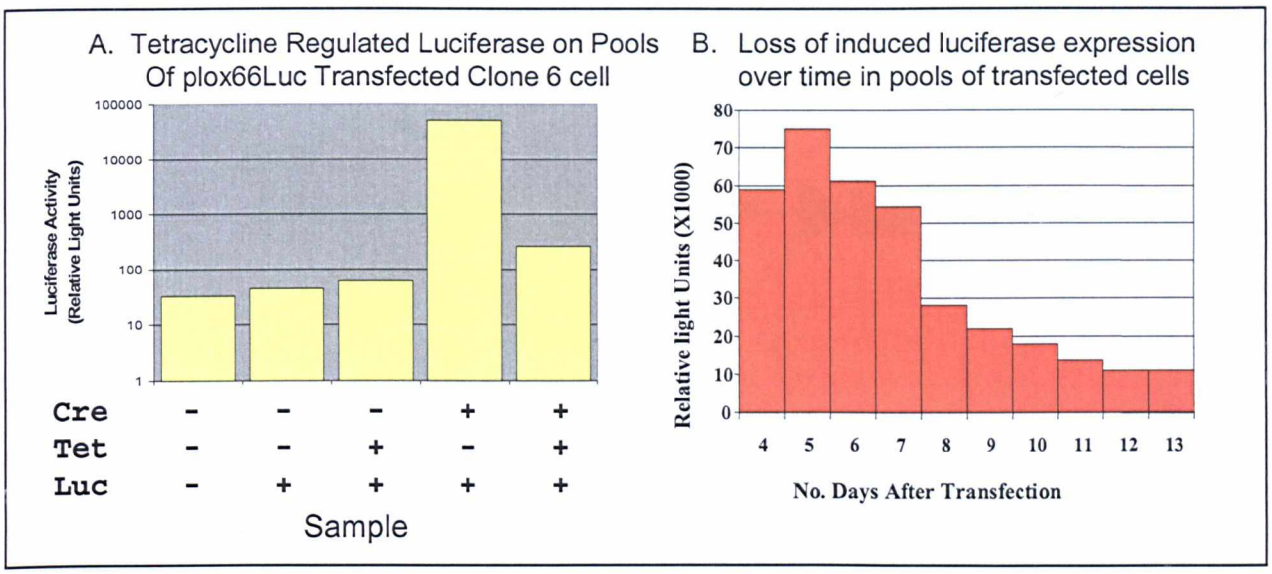



**Figure 4.5 Analysis of luciferase activity in pools of plox66Luc transfected clone 6 cells: A.** Illustrates tetracycline induction of the luciferase ORF after insertion downstream of the TRE in clone 6 cells. Untransfected clone 6 cells, as well as the control transfection (-pMC-Cre for both +/- Tet), were also assayed as described in the materials and methods (see section 2.8.5 and 2.7). **B.** Illustrates the loss of luciferase expression over time.

The luciferase assays strongly suggested that the desired integration event had occurred; PCR was used to provide direct molecular evidence for the targeted insertion. PCR conditions for amplifying the recombination junction (Figure 4.6A) were optimised with the primers TetOF and LucR (see Table 2.1 for sequences) and pTIGHT-Luc (Clontech) as the DNA template. pTIGHT-Luc is a vector containing the TREmod followed by the luciferase ORF and was obtained when pTRE-TIGHT (Clontech) was purchased. PCR analysis of pooled transfected cells (from a separate transfection experiment to that shown in Figure 4.5A) revealed that the amplified product (of 473bp) was only observed for cells transfected with both pMC-Cre and plox66Luc (for samples containing both 10,000 and 50,000 cells per pellet) and not for cells from the control experiment (-pMC-Cre; Figure 4.6B). These findings provide molecular evidence that the expected insertion event was taking place. Furthermore, the results suggest that there was at least one recombined cell per 10,000 cells transfected. In principle a more accurate figure for the frequency of

recombination might be determined by PCR analyses on smaller numbers of transfected cells, but this was not pursued.

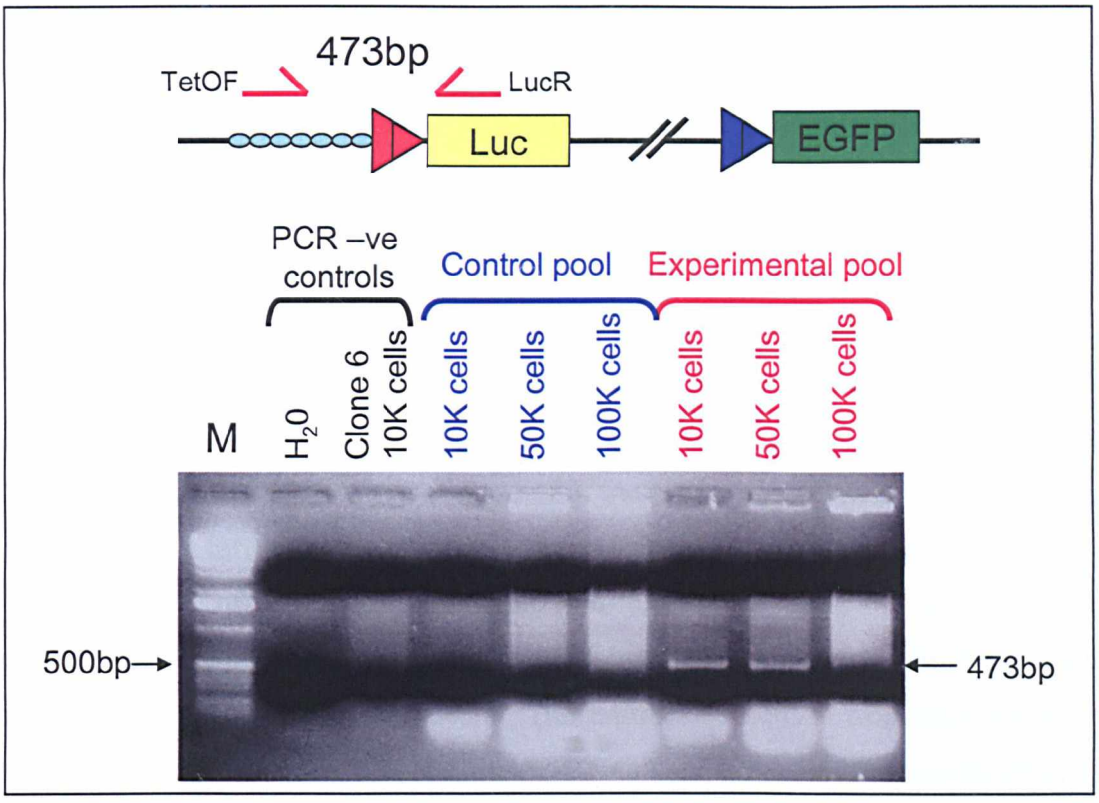



**Figure 4.6 Detecting *Cre*-mediated insertion of plox66Luc by PCR: A.** The expected genome structure (not to scale) after *Cre*-mediated insertion is illustrated and the approximate positions of the PCR primers TetOF and LucR are indicated. **B.** Ethidium bromide-stained agarose gel of PCR products from plox66Luc transfected clone 6 cells (with and without pMC-Cre). For methods see section 2.4. M = 1kb ladder (Invitrogen).

## 4.2.2 The recombination frequency was higher than previously reported

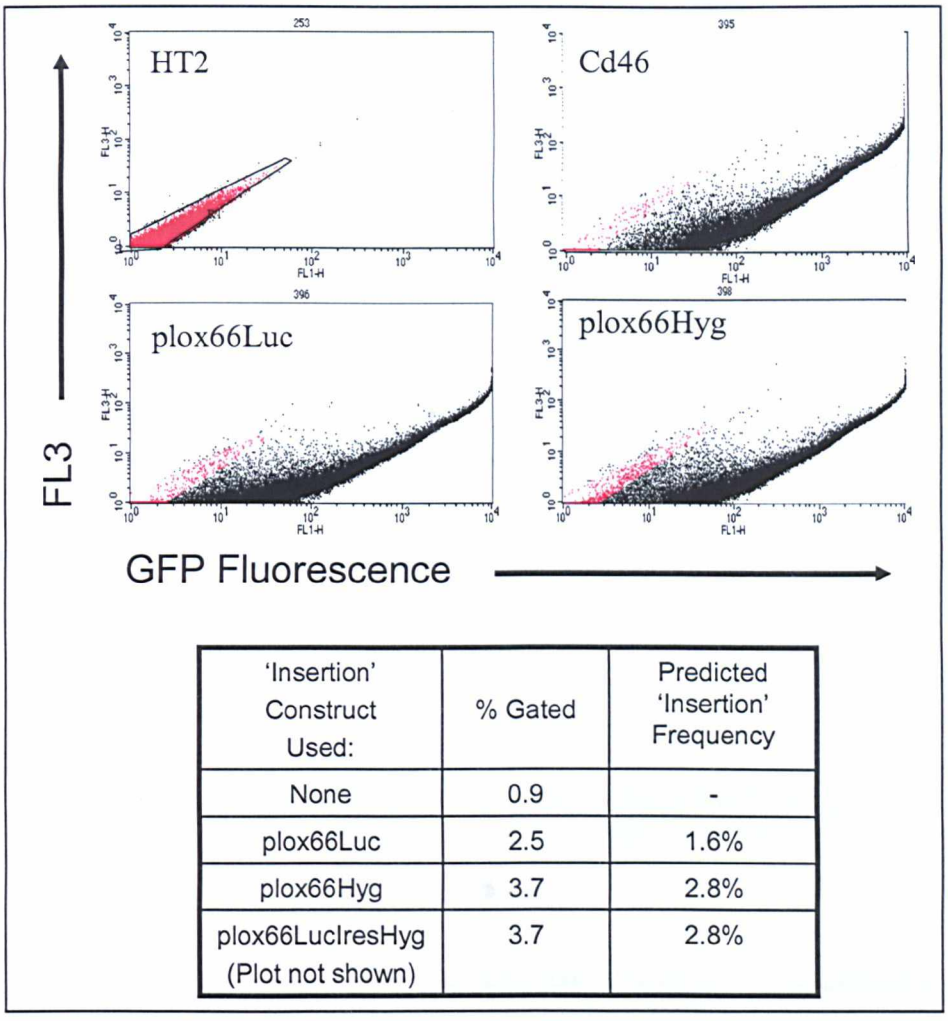
*A high insertion frequency was predicted from analysis on pools of transfected cells*

To estimate the recombination frequency, the luciferase activity in pools of transfected clone 6 cells were compared to the average luciferase expression from clones of cells thought to be expressing tetracycline-regulated luciferase at a similar level (Clones IRESHyg1-10, see Figure 4.12). The average luciferase activity from 500,000 IRESHyg1-10 cells was 10,461,300 RLU, whereas 500,000 Clone 6 cells,

co-transfected with pMC-Cre and plox66Luc, was 12,507 RLU (50092/4 [2 million pooled transfected clone 6 cells had a luciferase activity of 50092 RLU, Figure 4.5A]). Therefore, the frequency of *Cre*-mediated insertion could be estimated as 0.12% (12,507/10,461,300 X 100), or approximately 1 recombination event per 1000 cells transfected.

This value was not consistent with data generated by FACS, which examined the proportion of non-fluorescing cells four days after the insertion transfection was carried out in clone d46 cells (Figure 4.1B). Clone d46 was known to have only a single copy of the 'target' construct (data not shown) and therefore loss of GFP expression after an insertion event should take place. By gating on the non-fluorescing profile of HT2 cells (Figure 4.7), the proportion of cells within this range was calculated for the parental clone d46 (background) and pooled transfected cells. The proportion of cells lying within the non-fluorescing range was calculated as a percentage. After subtracting the background, a value was generated for the predicted recombination frequency (Figure 4.7). The estimated insertion frequency when cells were transfected with plox66Luc (Figure 4.2.A) was 1.6%. Both frequencies (1.6% and 0.12% [predicted from the luciferase assay]) were much greater than previous reports by Araki and colleagues (1997) who observed an absolute recombination frequency of approximately 0.002% when using a similar system.

If the predicted insertion frequency was as high as that suggested above (by FACS) it was assumed that clones which had undergone the desired event could be identified by examining for loss of GFP expression under a UV microscope. After co-transfecting clones d46 (Figure 4.1B) with pMC-Cre and plox66Luc (Figure 4.4), cells were plated at low density and left for colonies to form (10 days). The resulting colonies were then analysed under the UV microscope for loss of GFP expression. After examining many (1000+) clones, only GFP-positive colonies were observed. This therefore suggested that the recombination frequency was not as high as that predicted by FACS and that a selection for the insertion event was required.




| 'Insertion' Construct Used: | % Gated | Predicted 'Insertion' Frequency |
|---|---|---|
| None | 0.9 | - |
| plox66Luc | 2.5 | 1.6% |
| plox66Hyg | 3.7 | 2.8% |
| plox66LucIresHyg (Plot not shown) | 3.7 | 2.8% |

**Figure 4.7 Predicting the recombination frequency by FACS:** 500,000 Clone d46 cells were co-transfected with either no DNA (none), plox66Luc, plox66Hyg or plox66LucIresHyg (Figure 4.2B, C and D, respectively) and pMC-Cre. Four days after transfection the cells were analysed by FACS for loss of GFP expression. By gating on the non-fluorescing profile of HT2 cells (in red), all cells that lay within this range were calculated and the number of non-fluorescing cells observed in the parental Clone d46 (background levels) was subtracted.

*The absolute recombination frequency was generated using plox66Hyg*

To determine the frequency of *Cre*-mediated insertion, a second 'insertion' construct was generated, called plox66Hyg (Figure 4.8). Colonies that had undergone the insertion event after co-transfection with plox66Hyg and pMC-Cre into clone d46 (Figure 4.1B) were identified by tetracycline regulated hygromycin resistance and loss of GFP fluorescence. In addition, the expected genome rearrangements were

demonstrated by Southern blot. Again, a control transfection was carried out where pMC-Cre was omitted.

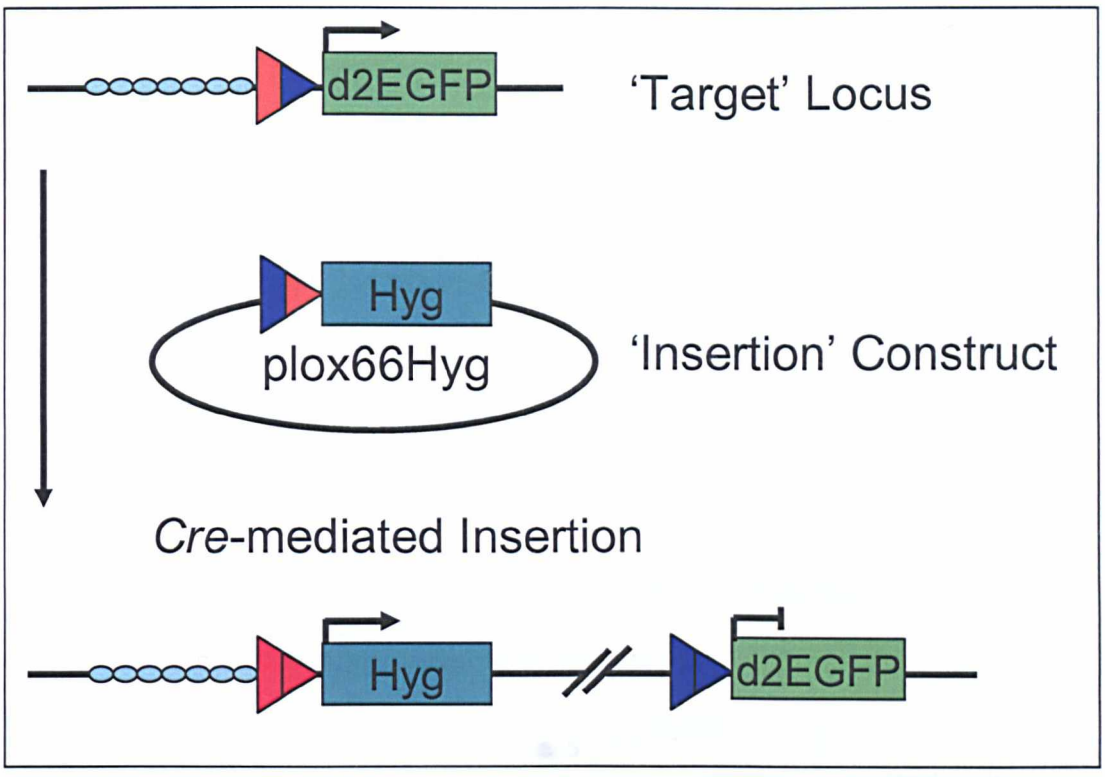



**Figure 4.8 Targeted insertion of plox66Hyg in clone d46:** The expected *Cre*-mediated insertion of plox66Hyg downstream of the modified TRE in clone d46 is illustrated. Not to scale. For a key to symbols see Appendix I.

61 hygromycin-resistant colonies were generated per 500,000 clone d46 cells transfected. Of these, 57 (93%) no longer expressed GFP when analysed by fluorescence microscopy (For FACS analysis of three such clones, Hyg1-3, see Figure 4.9A) and became hygromycin-sensitive upon the addition of tetracycline to the medium. The control experiment (-pMC-Cre) yielded no hygromycin-resistant colonies. The recombination frequency obtained using plox66Hyg (93%), was much higher than that reported by Araki and colleagues (1997) who observed that 2-16% of clones generated after selection had undergone the desired insertion event. They were not, however, using a promoter-trap style approach as being described here which is expected to have generated the increased insertion frequencies seen in this study (Figure 4.8).

Confirmation that the desired insertion event had taken place was demonstrated for one of the hygromycin resistant clones by Southern blot (Hyg 1; Figure 4.9A). Genomic clone d46 and Hyg1 DNA was digested with *Dra*I and probed with a portion of the TRE (*Stu*I/*Xho*I fragment) as described in the materials and methods (Section 2.5). *Dra*I cuts at a single known site in the 'target' locus, the size of the *Dra*I fragment from the parental clone d46 (Figure 4.9C) can be used to determine the position of the genomic *Dra*I site upstream of the TRE (Figure 4.9B). From this, the size of the band after insertion of plox66Hyg can then be calculated (Figure 4.9B) and confirmed by Southern analysis on Hyg1 DNA (Figure 4.9C).

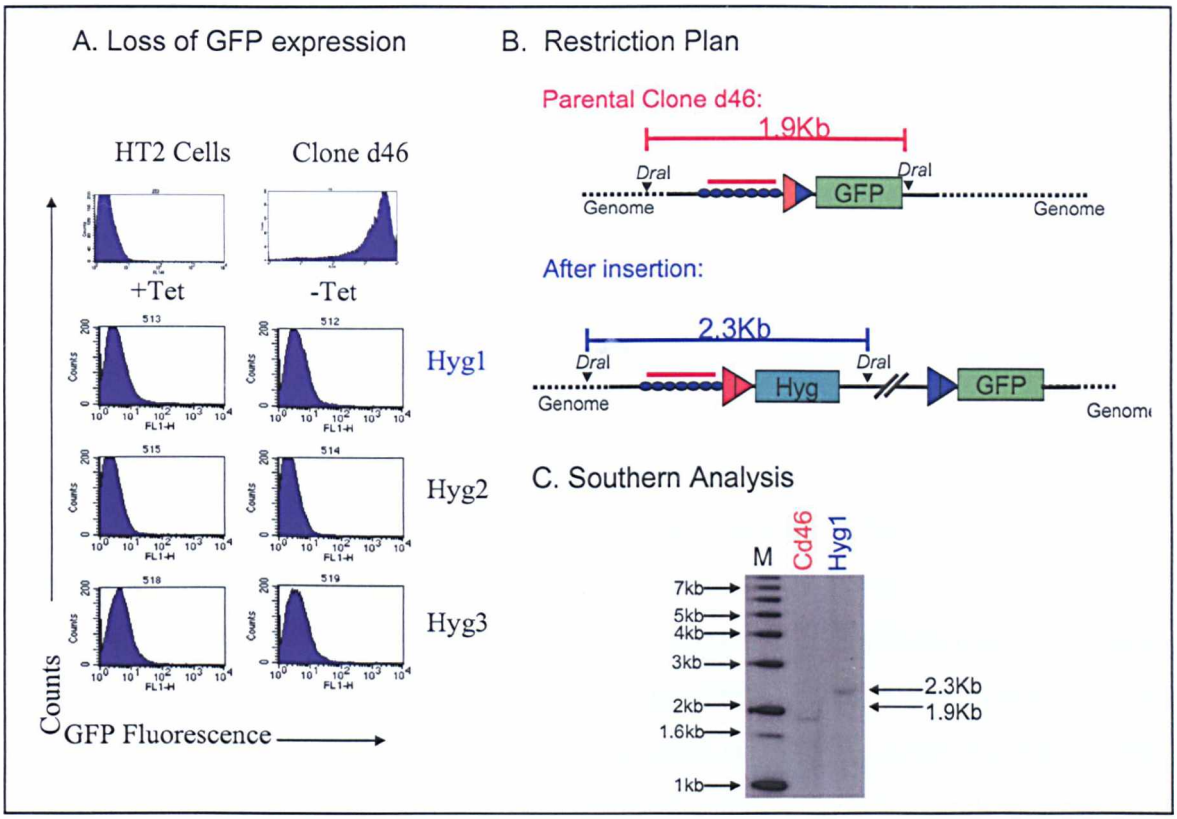



**Figure 4.9 FACS and Southern analysis to confirm the insertion event:** **A.** The hygromycin-resistant clones generated after co-transfecting plox66Hyg and pMC-Cre into clone d46 were analysed by FACS for loss of GFP expression. 57 out of 60 drug resistant clones had lost their GFP expression, results for only three clones are shown in this figure (Hyg1, Hyg2 and Hyg3). **B.** Map of 'target' locus in clone d46 (top; not to scale) and after Cre-mediated insertion of plox66Hyg (bottom; not to scale). DraI sites are indicated. The probe used for Southern analysis is indicated by the red bar. **C.** Genomic DNA from one of the three clones shown in A. (Hyg1) was digested with *Dra*I and probed with a (*Stu*I/*Xho*I) TRE fragment by Southern blot.

Since 57 clones were identified from 500,000 cells transfected to have lost their GFP expression, this suggested an absolute insertion frequency of ~0.01% (57/500,000 X 100). Again, this value was less than the value of 0.12% and 1.6% estimated by luciferase assay and FACS analysis (Figure 4.7), respectively. A value of 0.01% is nevertheless greater than the absolute recombination frequency of 0.002% reported by Araki and colleagues (1997). In this study similar results were obtained when using different 'target' clones, however, slight variations in the colony numbers were generated (ranging from 0.009-0.064%) suggested that the frequency of insertion varied from clone to clone. This variation in insertion frequency most likely reflects the accessibility of the integrated 'target' construct in each of the different clones (data not shown).

### 4.2.3 Testing a IRES containing 'Screen and Insert' strategy

*Designing a new 'insertion' construct – combining the GOI and selection cassette*

In view of the absolute Cre-mediated insertion frequency being 0.01% and considering that most GOIs will not be selectable it was necessary to devise a way in which a non-selectable GOI that had undergone the desired integration event could be isolated. With this in mind, a third 'insertion' construct (plox66LucIresHyg, Figure 4.10) was created which cloned an IRES-hygromycin cassette downstream of the luciferase ORF in plox66Luc (for methods see section 2.3.3.2). After insertion of this construct downstream of the TRE, expression of both the luciferase and hygromycin ORF would be controlled by tetracycline; in other words, selection in hygromycin can be used to identify clones in which a non-selectable GOI (luciferase) has undergone Cre-mediated insertion. The IRES avoided using promoter/enhancer elements that may have interfered with expression from the TRE after insertion. Also, by using a promoterless insertion construct, the background of hygromycin-resistant colonies resulting from random integration would be reduced. Colonies that underwent the insertion event (after transfection with plox66LucIresHyg and pMC-Cre [2μg of each by Lipofectamine 2000]) were expected to lose their GFP expression and have tetracycline inducible hygromycin-resistance and luciferase expression (Figure 4.10).

This is the first of two 'screen and insert' strategies described in this thesis, that are applicable to any GOI and is therefore referred to as System One.

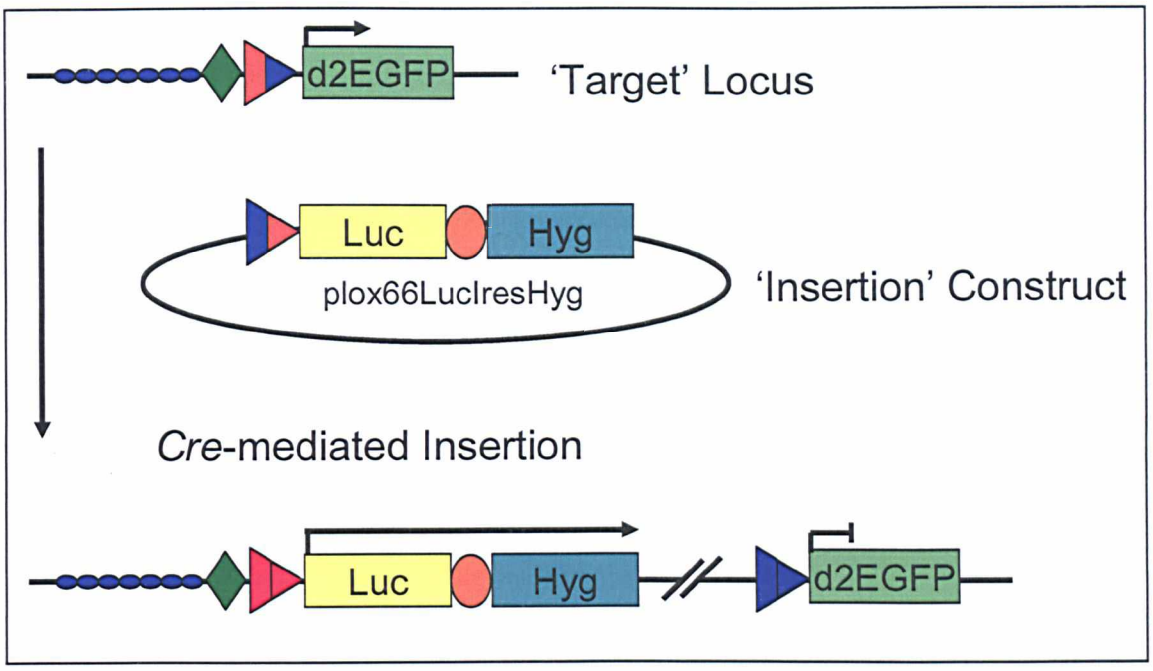



**Figure 4.10 'Screen and Insert' Strategy One:** The strategy is illustrated for a model GOI (luciferase) in an 'insertion' construct (plox66LucIresHyg) that undergoes Cre-mediated insertion into the 'target' locus of clone Rht14-19 (DNA - not to scale). Correct insertion generates clones that express tetracycline inducible luciferase and hygromycin, but no GFP expression. The desired clones can be selected for in hygromycin. For key to symbols see Appendix I.

*Successful isolation of hygromycin-resistant clones varied between 'target' clones*

Various 'target' clones were used for testing this construct (including clone d46, d14 and Rht14-19). Prior to the generation of Clone Rht14-19, a lot of time was spent testing the new IRES-based 'insertion' construct on 'target' clone d46, as the results using plox66Hyg were so encouraging. Generating healthy, hygromycin-resistant colonies from clone d46, using plox66LucIresHyg, was difficult to achieve and their ability to survive picking was much reduced. From three different experiments only six colonies were successfully picked and propagated with only one of those demonstrating the expected phenotypes (Table 4.1). An explanation as to why healthy hygromycin-resistant colonies were not generated still remains unclear. The

possibility that insertion of plox66LucIresHyg was hindered by the genomic location of the integrated 'target' construct being inaccessible, could be eliminated as the same 'target' clone (d46) was used in the previous experiments when using plox66Hyg (Table 4.1). Similarly, the 'target' locus of clone d46 was now known to have a functional *lox71* site as it successfully recombined with plox66Hyg (Table 4.1).

**Table 4.1 Results of insertion experiments using 'target' clone d46:** The colony numbers generated when 500,000 clone d46 cells were transfected by lipofection with the 'insertion' constructs, plox66Hyg and plox66LucIresHyg and clone Rht14-19 was transfected with plox66LucIresHyg. Exp: 'insertion' construct + pMC-Cre. Cont: 'insertion' construct – pMC-Cre.

| Clone | 'Insertion' vectors | Exp. No. | No. of HygR clones | No. Surv. Picking | Tet Regul. Luc. | Tet Regul. Hyg. | Lost GFP Exp. | No of recombined clones |
|---|---|---|---|---|---|---|---|---|
| d46 | plox66Hyg | Exp1 | 61 | 61 | - | 57 | 57 | 57 |
| | | Cont1 | 0 | - | - | - | - | - |
| d46 | plox66LucIresHyg | Exp1 | 3 | 0 | 0 | 0 | 0 | 0 |
| | | Cont1 | 0 | - | - | - | - | - |
| | | Exp2 | 6 | 3 | 1 | 1 | 1 | 1 |
| | | Cont2 | 0 | - | - | - | - | - |
| | | Exp3 | 40 | 3 | 0 | 0 | 0 | 0 |
| | | Cont3 | 0 | - | - | - | - | - |
| Rht14-19 | plox66LucIresHyg | Exp1 | 53 | 48/48 | 10/10 | 10/10 | 47/48 | 10/10 |
| | | Cont1 | 0 | - | - | - | - | - |

Concurrent experiments involved the generation of the most desirable clones, Rht14-10 and -19 (discussed in section 3.2.2). The 'target' construct used (pTFRTlox71dGFP) to generate clones Rht14-10 and -19 contained an FRT site between the TREmod and the *lox71* site (Figure 4.10; top). The FRT site was added for use in a modified 'screen and insert' strategy, the results of which will be discussed in chapter five. Clones Rht14-10 and -19 were generated shortly after the above problems with clone d46 were experienced.

A decision was made to test the IRES-based 'screen and insert' strategy on clone Rht14-19. The inducible GFP-expression demonstrated for Rht14-19 was approximately 500-fold and the uninduced GFP expression levels were identical to that of background (non-fluorescing) cells. Clone Rht14-19 also contained a single

copy of the 'target' construct, a feature particularly important for predicting the expression pattern for the inserted GOI and for the preliminary identification of clones that had undergone the insertion event, by analysing for loss of GFP expression.

After transfecting 500,000 Rht14-19 cells with plox66LucIresHyg and pMC-Cre, 53 hygromycin-resistant colonies were generated (Table 4.1). 48 clones were picked and plated into 24-wells with or without tetracycline supplementing hygromycin-containing medium. This experiment revealed that each of the clones were sensitive to hygromycin when tetracycline was present. The 48 clones were analysed under a UV microscope for loss of GFP expression and all except one clone, which appeared to contain a few GFP-positive cells, were no longer expressing GFP. In addition, no hygromycin-resistant clones were generated for the control experiments (-pMC-Cre; Table 4.1). These results therefore suggest that almost all the hygromycin-resistant clones generated had undergone the insertion event and an absolute 'insertion' frequency of 0.01% was calculated. This was a similar result to what had previously been obtained for Cre-mediated integration of plox66Hyg in clone d46 (Table 4.2).

**Table 4.2 The absolute recombination frequencies of clones d46 and Rht14-19:** The absolute insertion frequencies of clone d46 when transfected with plox66Hyg and plox66LucIresHyg and clone Rht14-19 when transfected with plox66LucIresHyg are shown.

| Clone | 'Insertion' Vector | % GFP –ve HygR clones | Absolute recombination frequency |
|---|---|---|---|
| d46 | plox66Hyg | 93% | 0.01% |
| d46 | plox66LucIresHyg | Less than 16% | Less than 0.0002% |
| Rht14-19 | plox66LucIresHyg | 98% | 0.01% |

To gain further evidence for the correct Cre-mediated insertion of plox66LucIresHyg in clone Rht14-19, ten GFP-negative clones (IRESHyg1-10) were chosen and analysed further by FACS, for inducible luciferase expression and for the expected genome rearrangements. FACS analysis verified that the ten clones were no longer expressing GFP (Figure 4.11).

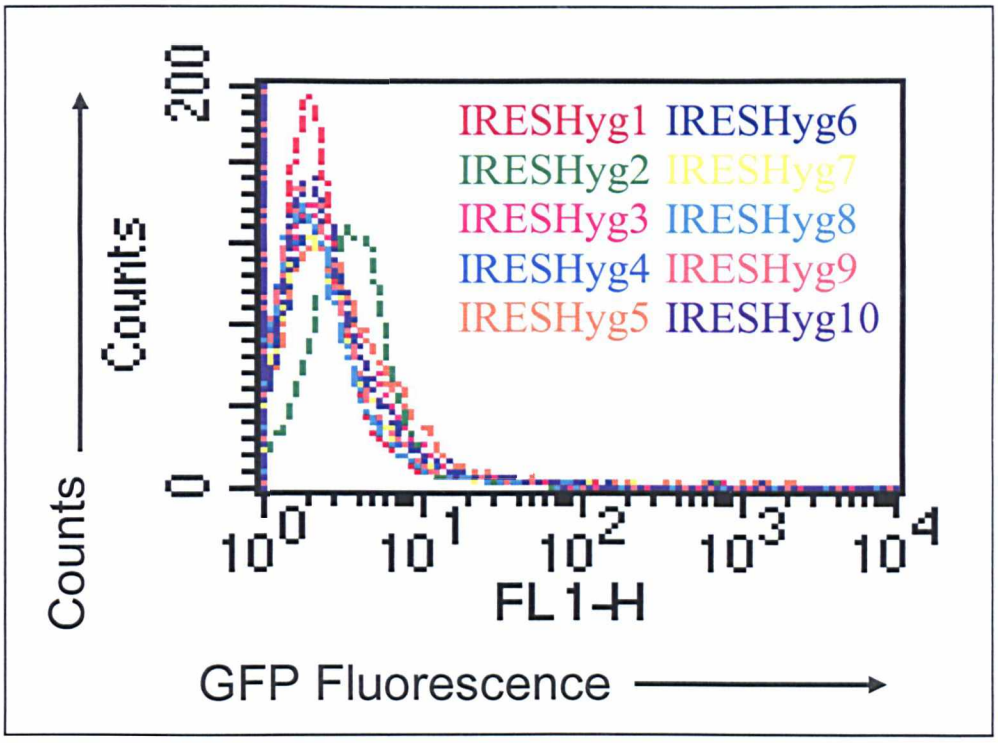



**Figure 4.11 Loss of GFP expression in ten of the hygromycin-resistant clones:** Illustrates that the ten hygromycin-resistant clones (IREShyg1-10) generated after transfection of clone Rht14-19 with plox66LucIresHyg and pMC-Cre were no longer expressing GFP. The ten FACS profiles are overlayed.

*Tetracycline-inducible luciferase was detected in the ten hygromycin-resistant clones*

The ten clones chosen (IRESHyg1-10) were split equally into medium with or without tetracycline. They were incubated for 48 hours and then analysed for inducible luciferase activity (See section 2.7). Each of the clones demonstrated tetracycline regulated luciferase expression (Figure 4.12). The mean induction value for inducible luciferase activity is 243.5-fold. This is very similar to the 244.4-fold mean inducible GFP expression calculated for the parental clone, Rht14-19, using CellQuest software (Becton Dickinson). The uninduced expression levels of luciferase activity, however, were 340-1700-fold greater than that of the control (Rht14-19 cells without luciferase expression) and the uninduced luciferase expression for the ten clones analysed did not appear to mimic the tight regulation seen for the parental clone Rht14-19. These values must reflect the highly-sensitive nature of the luciferase assay which has been shown to detect as little as $10^{-20}$ moles of luciferase and yields linear results over eight orders of magnitude (www.promega.com; Wood, 1991). FACS analysis, however,

requires at least $10^{-10}$ moles of EGFP for its detection, with a linear range over four orders of magnitude (see www.bdbiosciences.com); this therefore suggests that utilising flow cytometry for detecting tight uninduced expression levels is not the most sensitive method, as even smaller amounts of expression can be detected using luciferase as the reporter gene.

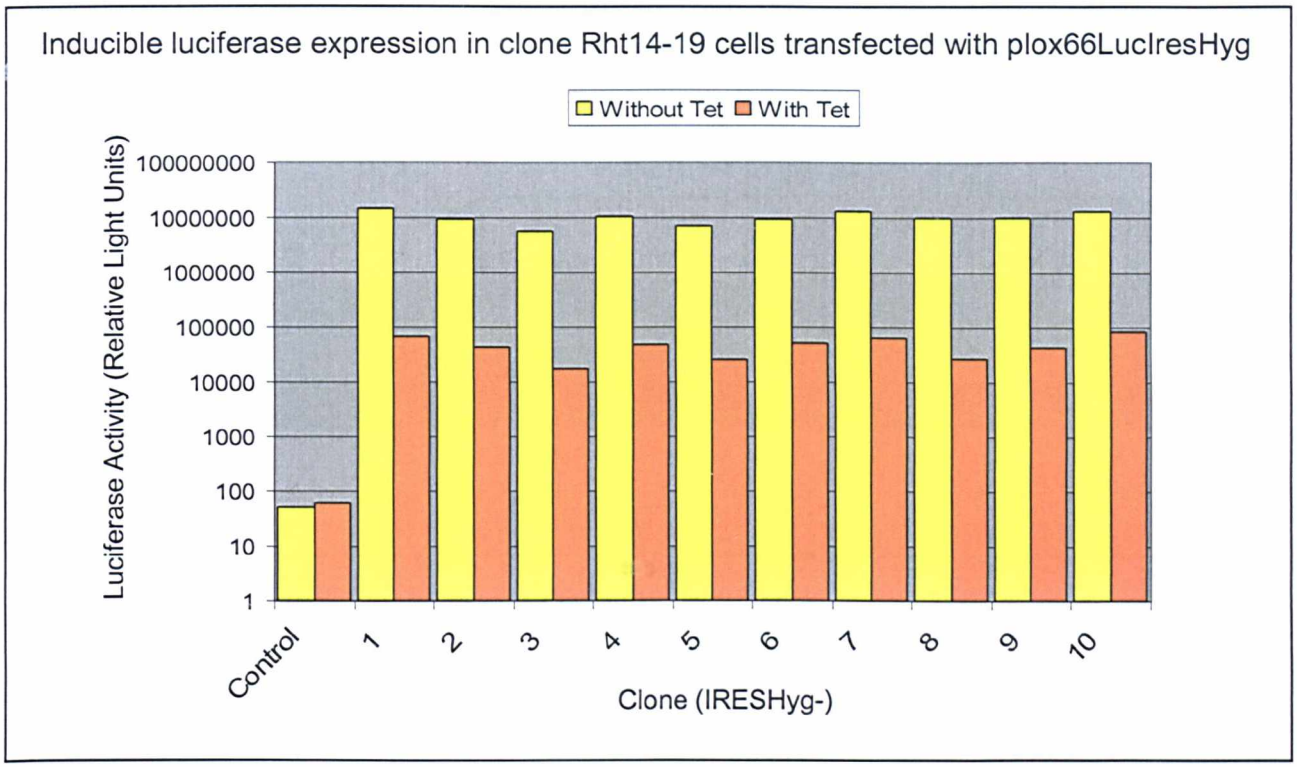



**Figure 4.12 Analysis of inducible luciferase expression:** Shows that each of the ten clones had tetracycline inducible luciferase expression (for methods see section 2.7). The yellow bars are clones grown in the absence of tetracycline and the orange bars are the clones grown in the presence of tetracycline for 48 hours. A log (y-axis) scale was used so that the uninduced values (orange) could be seen.

*The expected genome rearrangements were demonstrated by PCR and Southern blot*

Molecular evidence that the desired insertion event had taken place in the ten clones (IRESHyg1-10) was demonstrated by PCR amplification of the recombination junction and Southern blot analysis for the expected genome rearrangements. Cell pellets were treated with pronase (for methods see section 2.4) and the TetOF and LucR primers (for sequence see Table 2.1) were used to amplify the recombination

junction (Figure 4.13A). All ten clones were positive for this region, as confirmed by the production of a 517bp band (Figure 4.13B).

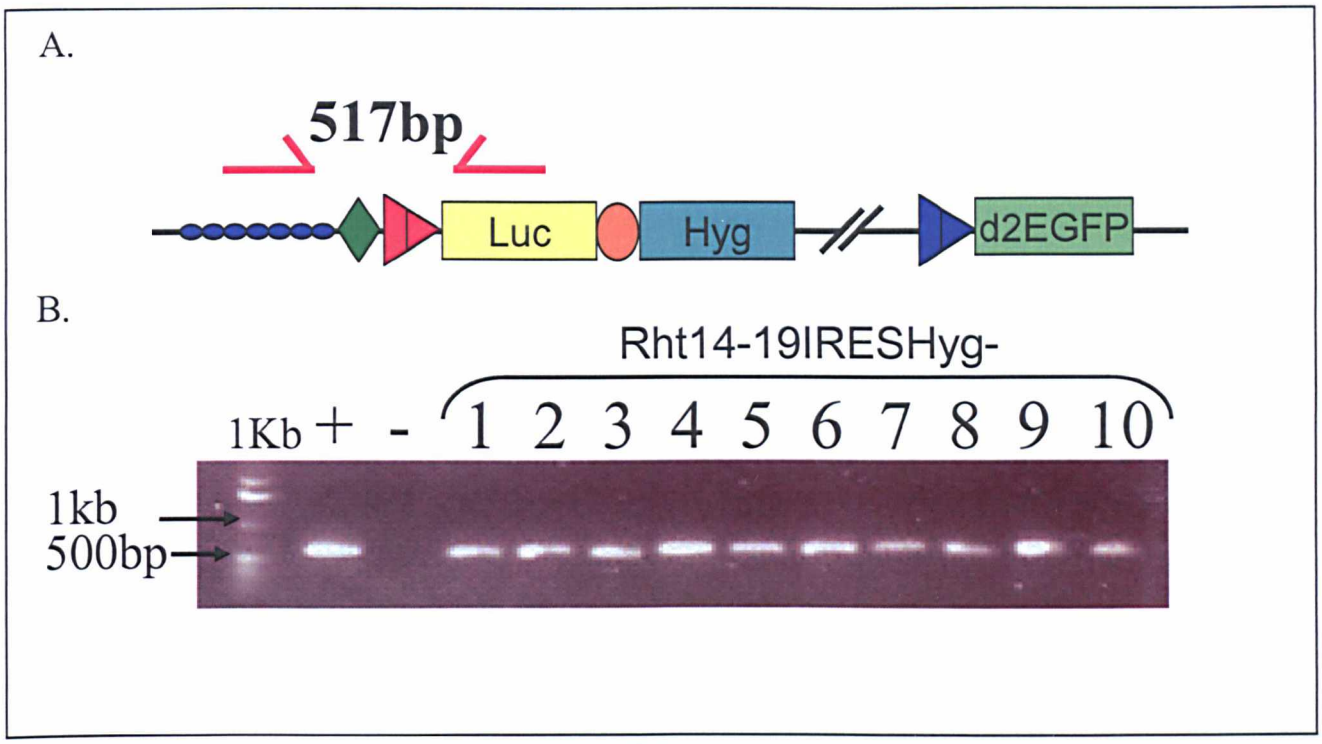



**Figure 4.13 PCR amplification of the recombination junction in hygromycin-resistant Rht14-19 cells transfected with plox66LucIresHyg: A.** A cartoon showing the recombination junction which was amplified by PCR to demonstrate that the insertion event had taken place in the ten hygromycin-resistant clones analysed. The approximate position of the primers are shown (TetOF/LucR) and the expected PCR product was 517bp (not to scale). **B.** An ethidium bromide-stained gel of PCR products from ten hygromycin-resistant, plox66LucIresHyg transfected Rht-14-19 clones.

Finally, genomic DNA from each of the ten clones under analysis was digested with *Dra*I and probed with a portion of the GFP gene (*Nco*I) (Figure 4.14A) (for methods see Section 2.5). Confirmation that the ten hygromycin-resistant clones had all undergone the insertion event was generated by Southern blot with the insertion event being represented by the conversion of a 2.4kb *Dra*I GFP fragment to a 2.1kb fragment (Figure 4.14B).

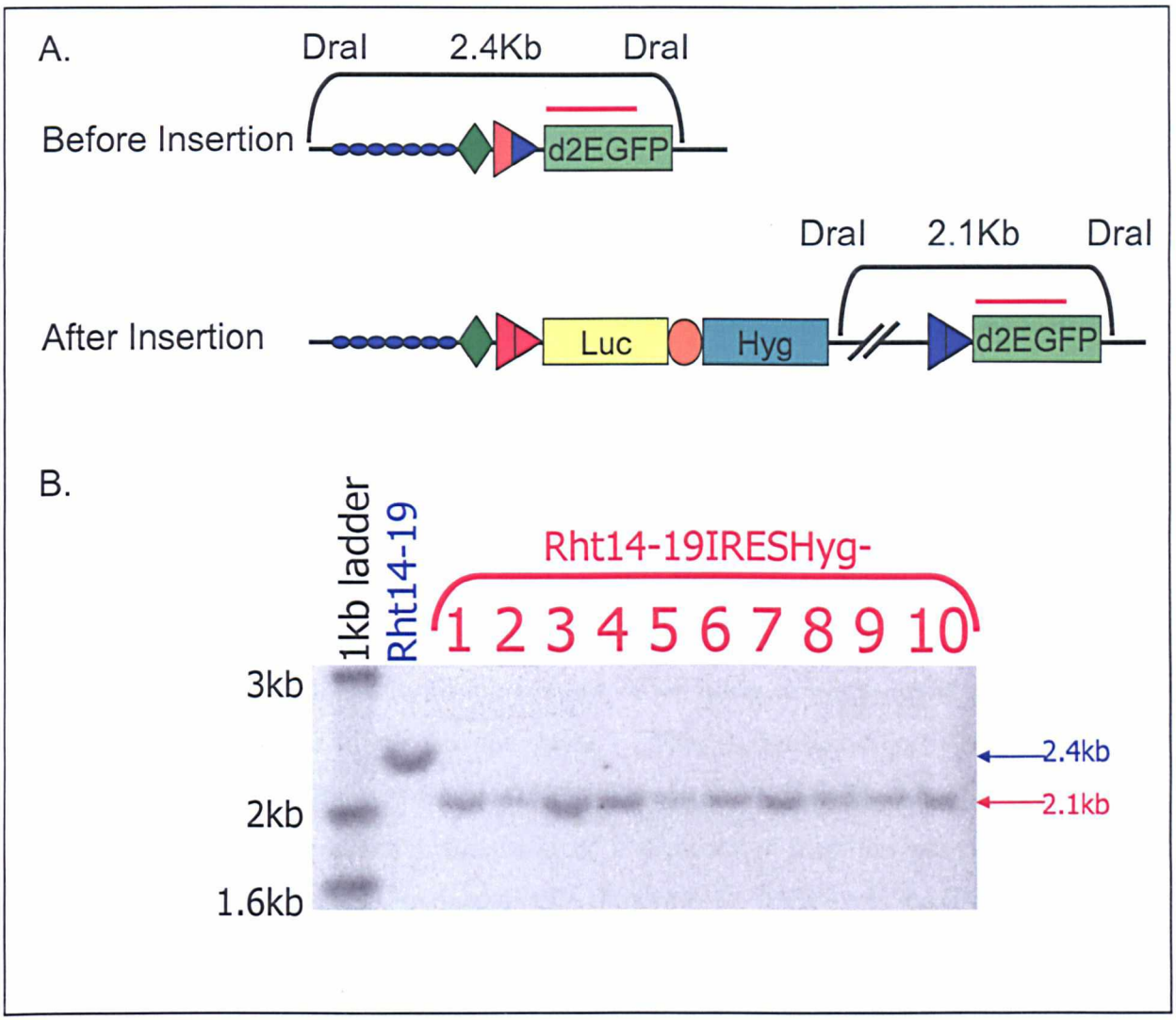



**Figure 4.14 Southern analyses of the ten hygromycin-resistant clones: A.** The expected restriction pattern when cut with *Dra*I and probed for GFP, before and after the insertion event (not to scale). **B.** The results of the Southern analysis on the ten hygromycin-resistant clones revealing that that expected rearrangements had taken place.

## 4.3 Discussion

### 4.3.1 A selection was required for the isolation of recombined clones

Evidence that the desired insertion event was taking place came initially from experiments using a promoterless luciferase construct (plox66Luc, Figure 4.2B). The results revealed that, after co-transfection of plox66Luc with pMC-Cre into clone 6 cells (Figure 4.1A), luciferase activity was regulated in a similar manner to the GFP expression seen in the parental 'target' clone (Figure 4.5A). The results therefore suggested that it was possible to generate predictable, reproducible expression of the 'inserted' gene. The amount of induced activity, 48 hours post transfection, was 0.12% of the induced activity of the average of ten other clones thought to be stably expressing luciferase at a similar level. This suggested that *Cre*-mediated recombination may have occurred in 0.12% of the transfected cells. A higher estimate for the absolute recombination frequency of *Cre*-mediated insertion was obtained by FACS analysis on pooled transfected cells (Figure 4.7). This experiment determined that 1.6% of transfected cells were no longer expressing GFP and these cells were therefore thought to have undergone the 'insertion' event.

After screening many colonies (>1000) under a UV microscope for loss of GFP expression it soon became apparent that using this method for identifying those clones that had undergone the insertion event was very inefficient and unproductive. This suggests that the approximated recombination frequencies (especially the FACS generated value (1.6%)) are largely overestimated. It is possible, however, that some clones tended to lose their 'insertion' construct during subsequent cell passages (perhaps due to lingering transient Cre-recombinase expression); although, using the mutant *loxP* (*lox71* and *lox66*) system is supposed to limit this frequency (Araki *et al.*, 1997).

It was subsequently shown that the absolute *Cre*-mediated recombination frequency for clone d46 (Figure 4.1B) when using plox66Hyg (Figure 4.2C) was 0.01%. In other words, 1 recombined colony would be present per 10,000 colonies examined;

verifying that the frequency was too low for their identification by visual screening methods. A more practical frequency would be around 1%, where only 100 clones would have to be screened to detect a desirable clone. It could have been possible to cell-sort the pooled transfected cells for loss of GFP expression. The products of the sort could subsequently be plated out at low density and screened for inducible luciferase. This was however never carried out due to the possibility of contamination during the sort and the decision was taken to test a potentially improved 'screen and insert' approach (Chapter 5).

### 4.3.2 A simple 'screen and insert' strategy was successfully established

*Successful isolation of hygromycin-resistant clones appeared to be clone dependant*

From previous results it was concluded that the *Cre*-mediated insertion event was successfully taking place, using plox66Hyg (Figure 4.2C), at a frequency high enough to generate stable clones using a selection. It was predicted that once a suitable method of selection was added to the 'insertion' construct, which allowed the expression of both the drug resistance ORF and the GOI, isolating clones that had undergone the 'insertion' event should be possible. The inclusion of promoter elements on the 'insertion' construct was avoided by cloning an IRES-hygromycin cassette downstream of the luciferase gene in plox66Luc (Figure 4.2B), to generate plox66LucIresHyg (Figure 4.2D).

Early experiments using clone d46 as the 'target' clone yielded poor and disappointing results and there was no obvious explanation as to why. The possibility that the 'target' locus was inaccessible or that the *lox71* site was inactive were quickly eliminated due to the encouraging results observed when using plox66Hyg using the same 'target' clone (see section 4.2.2). Also, it was later revealed that after insertion of plox66LucIresHyg into Rht14-19 cells, both the *lox66* and the IRES in the 'insertion' construct were functioning correctly to generate hygromycin-resistant colonies. It therefore appeared as though the problems observed were a combination of both the IRES and clone d46. It could be argued, however, that clone Rht14-19

(Figure 4.1C) was expressing almost twice as much GFP as clone d46 (Figure 4.1B) and perhaps after the insertion event not enough hygromycin was being produced in the latter clone to generate healthy drug-resistant colonies. Consequently, the 'target' clone Rht14-19 was predominantly used to test for the insertion event and the insertion frequency generated was identical to that achieved for plox66Hyg using clone d46.

*The construct design facilitated limited isolation of random integrants after selection*

For both insertion constructs (plox66Hyg [in clone d46] and plox66LucIresHyg [in clone Rht14-19]), the frequency of hygromycin-resistant clones that appeared to have undergone the insertion event was at least 93%. This value was much greater than previous reports by Araki and colleagues (1997) who observed that 2-16% of drug-resistant clones had undergone the desired recombination event. In this study, the high recombination frequency was due to the 'promoter-trap' feature of the 'insertion' construct. In the study by Araki *et al.* (1997), the selectable marker gene was expressed from its own promoter and as a result random integrants were presumably more predominant amongst the drug-resistant clones. The absolute recombination frequency must therefore be used to compare the results of the two studies and it was found that when using plox66Hyg (in clone d46, 0.01%) and plox66LucIresHyg (in clone Rht14-19, 0.01%) their frequencies were 5-fold greater than the 0.002% reported by Araki and colleagues (1997).

*Hygromycin-resistant Rht14-19-derived clones expressed inducible luciferase activity*

Ten hygromycin-resistant clones (IRESHyg1-10) were selected for FACS analysis, inducible luciferase activity, PCR and Southern blot analysis for the expected genome rearrangement. Each of the ten clones successfully demonstrated the above attributes; however the luciferase experiment yielded some disappointing results. It revealed that the uninduced luciferase expression in each of the ten clones was much higher than background levels (340-1400-fold greater). The half-life of luciferase is approximately 3 hours in mammalian cells (www.promega.com); this is similar to the d2EGFP protein. The kinetics of luciferase depletion following tetracycline addition

should therefore be comparable to that for d2EGFP depletion in the parental 'target' clone and it is consequently unlikely that the luciferase levels observed after 48 hours exposure to tetracycline are caused by incomplete protein degradation.

It is also unlikely that expression from the TRE is affected by insertion of the luciferase ORF itself. It is possible that the IRES in the 'insertion' construct is responsible for residual luciferase expression when uninduced. The decision to include an IRES in the 'insertion' construct was made before the results involving the 'target' construct, pTlox71dGFPiresHyg (Figure 3.10), had emerged. Nonetheless, results from the previous chapter suggested that the IRES is more likely to suppress expression of an upstream cistron rather than promote it. Conflicting reports claim that expression of either the first (Mizuguchi et al., 2000) or the second cistron (Houdebine and Attal, 1999) is favoured in a bicistronic message. It therefore appears as though the affects of an IRES on expression is relatively unclear.

It is credible that an alteration in TRE regulation after the insertion of the luciferase ORF had not actually taken place. The variation in regulation may simply be due to the cumulative and sensitive nature of the luciferase assay, when compared to FACS analysis. In other words, FACS analysis provides a representation of the GFP fluorescence from each individual cell within a clone and the mean GFP fluorescence can be estimated from the plot. The luciferase activity, however, is an additive measure of the total luciferase protein present within a cell lysate sample and has a broader scope for error as it relies on accurate cell number counts. To quantitate the luciferase assay further, the concentration of protein in the cell lysate used for the assay could be calculated, to normalise the results. Attempts were made to achieve this; however, the concentration of protein in the cell lyates was too low to be detected accurately. Further work is now required optimise this procedure.

In addition, the luminometer used (TD-20/20; Turner Designs) to measure the activity in pools of plox66Luc-transfected clone 6 cells, which incidentally demonstrated reproducible luciferase expression when compared to the parental 'target' clone 6 (Figure 4.5A), was not available for any further assays carried out in this study. Instead, an alternative luminometer (Biorbit 1253) was used, which consistently generated elevated luciferase readings. Finally, whilst attempting to duplicate the

luciferase results, it was also noticed that there were quite obvious variations in luciferase induction values generated for clones IRESHygl-10, each time the assays were repeated. This therefore highlights a further possible susceptibility for error when carrying out the luciferase assay. Again, these claims are merely speculation and further work is required to resolve these problems. Due to difficulties of comparing GFP and luciferase expression and the unreliable nature of the results, the luciferase assay was subsequently used as a method for simply detecting inducible luciferase expression and the values obtained were not investigated further.

### 4.3.3 Limitations of the first 'screen and insert' strategy - System One

The results generated so far have demonstrated that the basic principles of the 'screen and insert' strategy are working effectively. The tested IRES-based System One is limited, however by the need to express the GOI whilst selection for the desired insertion event (in hygromycin) is taking place. It would be desirable if the system allowed the selection for the insertion event whilst the GOI is completely suppressed. This feature would be particularly important whilst attempting to overexpress a GOI that compromises the cells survival and it would also limit GOI's expression to a particular time point (for example, during development or whilst being observed). As a result, an improved 'screen and insert' strategy, System Two, was developed which allows the selection of the insertion event without expressing the GOI. The newly proposed strategy, along with the results, is discussed in Chapter Five.

# Chapter Five

# Testing an Improved 'Screen and Insert' Strategy

## 5.1 Introduction

In the previous chapter a simple 'screen and insert' strategy (System One) was tested. It utilised an IRES-based technique to successfully select for clones that had undergone the insertion event. A limitation of this strategy, however, is that the GOI had to be continually expressed whilst the insertion event is selected for in hygromycin. This aspect is not attractive for the analysis of a GOI whose expression may compromise cell survival or whose phenotype is only evident upon initial expression. This chapter describes a second 'screen and insert' strategy, System Two, designed to overcome these restrictions. System Two was developed to facilitate the isolation of clones with suppressed, tightly regulated GOI, without using any promoter/enhancer elements or an IRES-based selection cassette. To achieve this an additional-step following Cre-mediated insertion is adopted; which involves using flp-recombinase to delete a stretch of DNA and to link the GOI to the TRE. It was hoped that by taking advantage of the reported high frequency of *flp*-mediated excision (Schaft *et al.*, 2001) that the desirable clones could be isolated under selection-free conditions.

The insertion step of the improved 'screen and insert' approach is achieved by *Cre*-mediated recombination (Figure 5.1), similar to that already described in Chapter Four. Clones that undergo this step are isolated by selection for the expression of the *E.coli* guanine phosphoribosyltransferase (*gpt*) gene. *Gpt* can be used as a positive and negative selectable marker gene in wild-type mammalian cells, where *gpt* expression is selected for in mycophenolic acid (MPA) and xanthine; and selection against *gpt* expression is carried out in 6-thioxanthine (6-TX) (Lupton *et al.*, 1991). The rationale behind *gpt* selection came from its capability to rescue the cell's purine nucleotide synthesis, when the *de novo* pathway was blocked (e.g. in MPA), by creating a salvage pathway (Figure 5.2). This is due to *gpt*'s ability to use xanthine

for purine nucleotide synthesis, a role which the mammalian *gpt* gene does very poorly (Mulligan and Berg, 1981).

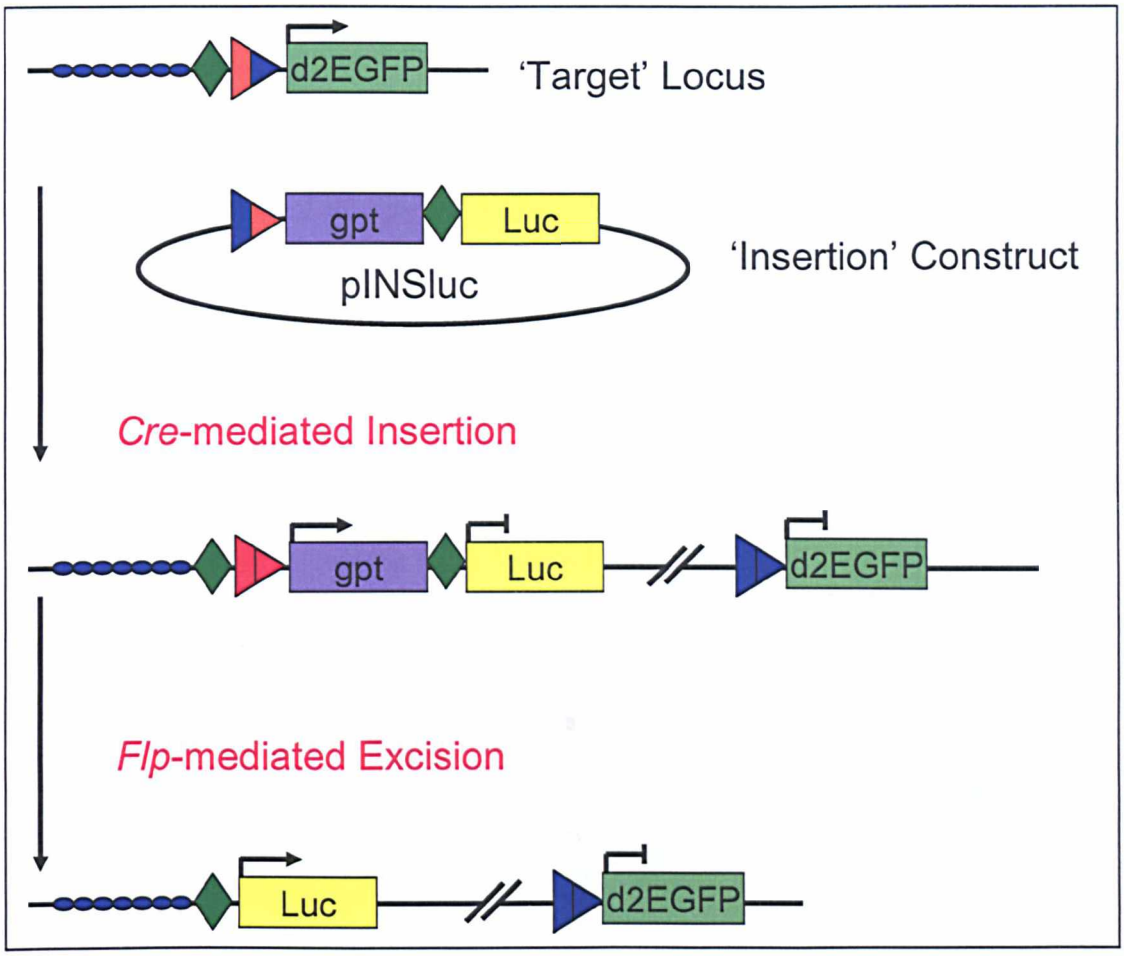



**Figure 5.1 Outline of the improved 'screen and insert' strategy:** Illustrates the improved 'screen and insert' strategy discussed in this chapter (DNA – not to scale). 'Target' clones which undergo the insertion step are selected for *gpt* expression. In the selected clones, the GOI is placed downstream of the silent TRE using *flp*-mediated recombination.

The deletion step of the improved 'screen and insert' strategy involves the expression of *flp* recombinase in the isolated *gpt*-positive clones to catalyse an excision reaction between the two FRT sites that flank the *gpt* ORF (Figure 5.1). This reaction places the GOI downstream of the TRE by removing the *gpt* ORF. It has previously been shown that *flp*-mediated excision can be achieved in 6% of transfected ES cells, using an improved *flp*-expression vector (pCAGGS-FLPe, Cambion). Molecular evolution was used to improve the thermostability of *flp* recombinase for improved performance

at 37°C (generating *Flpe*); as the wild-type enzyme has an optimum working temperature of 25-30°C, which is not desirable when working with mice and tissue culture cells (Buchholz *et al.*, 1998). In addition, *flpe* was cloned into an expression vector (pCAGGS (Araki *et al.*, 1997)) and was shown to achieved up to 10-fold improved recombination over other previously tested expression vectors (Schaft *et al.*, 2001). It was therefore hoped that the frequency of *flp*-mediated excision, using pCAGGS-FLPe, would be high enough to allow a simple PCR screen of individual colonies without the need for selection. Alternatively, when it is not necessary for the GOI to remain down-regulated, the new 'screen and insert' system also allows selection for the excision event by selecting for loss of *gpt*-expression (in 6-TX).

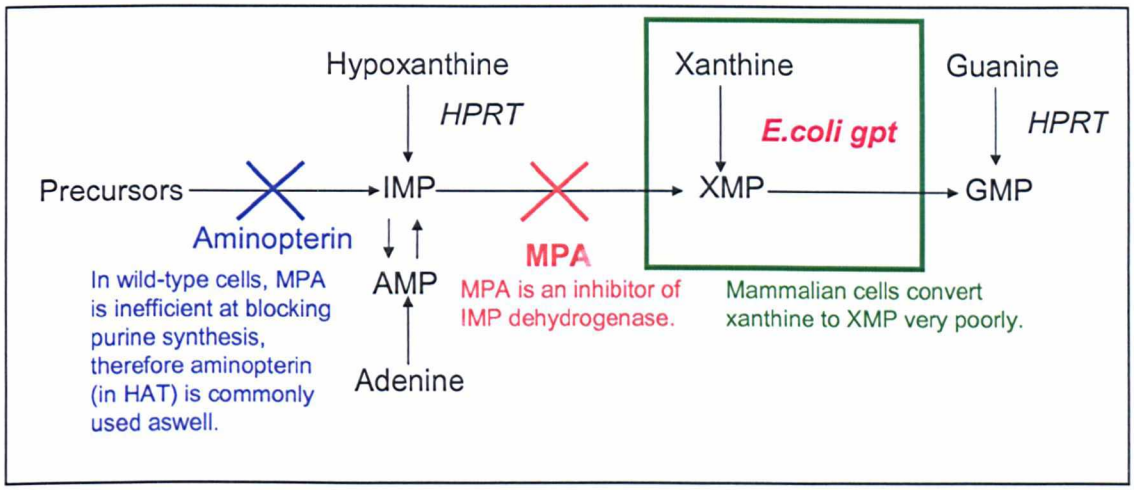



**Figure 5.2 Outline of the rational behind *gpt* selecton:** Illustrates the purine synthesis pathway in mammalian cells and the sites of inhibition by aminopterin (blue) and MPA (red). MPA specifically blocks the conversion of IMP to XMP by IMP dehydrogenase. During such a block, cells expressing the *E.coli gpt* can be rescued as they can convert supplemented xanthine to XMP, generating a salvage pathway (green).

## 5.2 Results

### 5.2.1 Testing the Cre-mediated insertion step in the modified System

*Design of a new 'insertion' construct in the improved 'select and insert' strategy*

A modified 'insertion' construct, pINSluc (Figure 5.3), was generated as described in the materials and methods (See 2.3.3.3). It contained the promoterless *gpt* and luciferase ORF (model GOI), with an FRT site cloned in between. Upstream of the gptFRTluc cassette is a mutant *lox66* site (Figure 5.3). After insertion of pINSluc downstream of the TRE, expression of the *gpt* gene can be selected for in MPA and xanthine. If the expected insertion event takes place, then clones isolated in MPA and xanthine should become sensitive to such selection upon the addition of tetracycline to the medium.

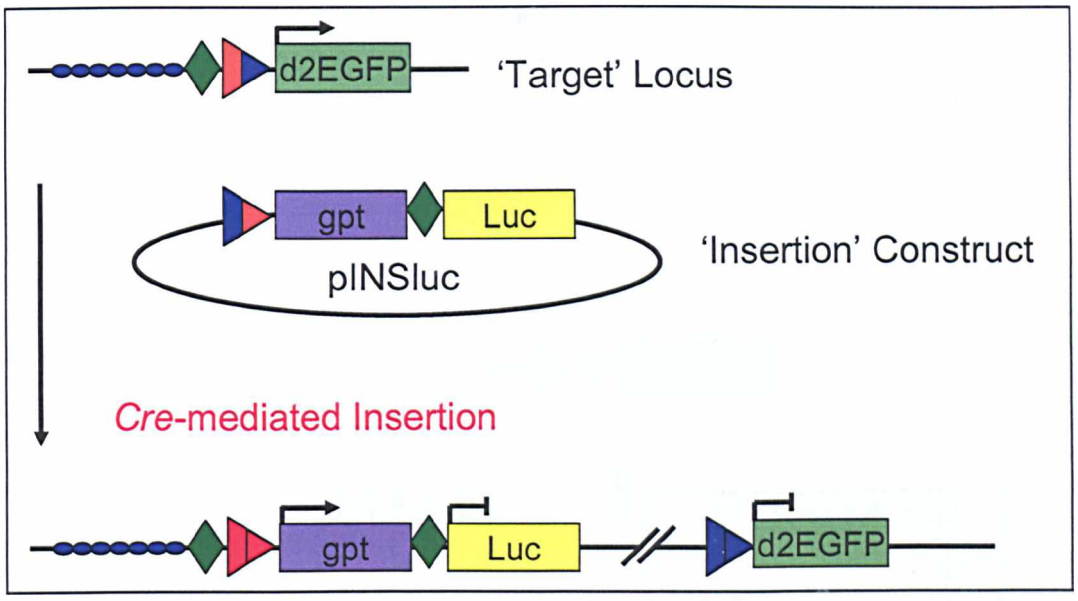



**Figure 5.3 Cre-mediated insertion in the improved 'screen and insert' strategy:** 'Target' clones are co-transfected with pINSluc and the *Cre*-expressing construct (pMC-Cre) (not to scale). Following Cre-mediated insertion the gpt gene is adjacent to the TRE so that in the absence of tetracycline, insertion events can be selected for in MPA and xanthine. Note that the GOI (luciferase) is not expressed at this stage. For key to symbols see Appendix I.

At this stage no luciferase expression is expected, as transcription from the TRE should terminate at the *gpt* polyA. In addition, because it is promoterless, pINSluc should generate very few *gpt*-expressing colonies resulting from its random integration, as was seen for plox66LucIresHyg (see section 4.2.3). Colonies that undergo Cre-dependant insertion are therefore expected to lose their GFP expression and have tetracycline inducible *gpt* expression.

*Clones Rht14-10 and Rht14-19 were used to test the new and improved strategy*

The 'target' clones used to test this new strategy were clones Rht14-10 and Rht14-19 (Figure 5.4), which are described in Chapter Three. Both clones were created using the FRT containing 'target' construct (pTFRTlox71dGFP) shown in Figure 5.4. Clones Rht14-10 and Rht14-19 both have uninduced expression levels identical to that of background and have maximum induction levels approximately 1000- and 500-fold greater than uninduced levels, respectively. Each clone also contains a single copy of the 'target' construct. Two clones were chosen to test the improved 'screen and insert' strategy, to detect any clone to clone variations in recombination frequency; as seen in Chapter Four when the 'insertion' construct, plox66LucIresHyg, was transfected into both clones d46 and Rht14-19 and different results were generated (See 4.2.3).

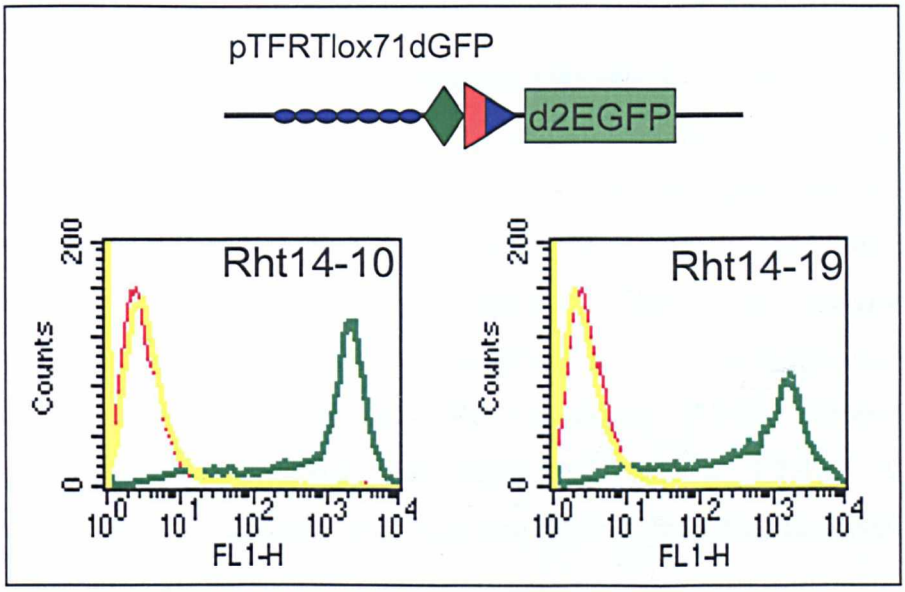



**Figure 5.4 Target clones Rht14-10 and -19:** The 'target' construct (pTFRTlox71dGFP) used (not to scale) to create clones Rht14-10 and -19 (top) and their GFP-expression profiles (bottom) are shown. For instructions on how to interpret FACS plots see Appendix IV.

*Optimised selection for gpt-expressing clones in wild-type cells required HAT medium*

Initially, the removal of non-*gpt* expressing colonies in MPA and xanthine was difficult and often a high background meant that isolating individual *gpt*-positive clones was difficult. A literature search revealed, however, that many groups have reported a difficulty in achieving clean forward selection for *gpt* expression in wild-type (i.e. HPRT+) human cells using MPA and xanthine alone (Lupton *et al.,* 1991; Trott *et al.,* 1995). Nevertheless, earlier papers suggested that MPA was sufficient to prevent extensive growth of some cells for colony generation, however in other cell lines the block in purine synthesis appeared to be delayed and initially incomplete. Mulligan and Berg (1981) consequently proposed using aminopterin to help block the *de novo* synthesis of all purines (Figure 5.2) and hypoxanthine to provide for AMP formation (Figure 5.2) and to inhibit mammalian *gpt* activity; both heightening the requirement of xanthine for GMP formation (Figure 5.2). On this basis, MPA and xanthine was supplemented with HAT (Hypoxanthine, Aminopterin, Thymidine) for the selection of *gpt*-positive clones.

*The insertion frequency of pINSluc is lower than what had been previously achieved*

After transfecting 500,000 Rht14-10 or Rht14-19 cells with pINSluc and pMC-Cre, 11 and 32 *gpt*-expressing colonies were generated, respectively. These colonies were analysed under the UV microscope and the proportion of those clones that were no longer expressing GFP were calculated (Table 5.1). In the experiment involving clone Rht14-10, 82% of the colonies generated were GFP-negative; for Clone Rht14-19, 69% of colonies generated were GFP-negative. These results suggested an absolute insertion frequency of pINSluc of 0.0018% (Rht14-10) and 0.0044% (Rht14-19) (Table 5.1), somewhat lower than the frequencies (0.01%) observed for plox66Hyg and plox66LucIresHyg (see Table 4.2, section 4.2.3). Control experiments were carried out where pMC-Cre was omitted from the transfection and no colonies were formed (Table 5.1).

**Table 5.1 Recombination efficiency of pINSluc:** Shows the number of *gpt*-expressing colonies obtained and the proportion of those which were no longer expressing GFP (cont – control).

| Clone | No. *gpt* +ve Col. | No. GFP -ve | % GFP -ve | Absolute Frequency |
|---|---|---|---|---|
| **Rht14-10** | 11 | 9 | 82% | 0.0018% |
| **Rht14-10 cont** | 0 | 0 | - | - |
| **Rht14-19** | 32 | 22 | 69% | 0.0044% |
| **Rht14-19 cont** | 0 | 0 | - | - |

For each of the 'target' clones, nine of the GFP-negative colonies were analysed further (Rht14-10IN1-9 and Rht14-19IN1-9). FACS analysis confirmed that these clones were no longer expressing GFP (Figure 5.5). The picked clones were plated in MPA and xanthine (+HAT), either with or without tetracycline in the medium. The growth of those clones exposed to tetracycline was impaired, however, they did not die and detach from the plate as had been seen for the hygromycin selection (section 4.2.3).

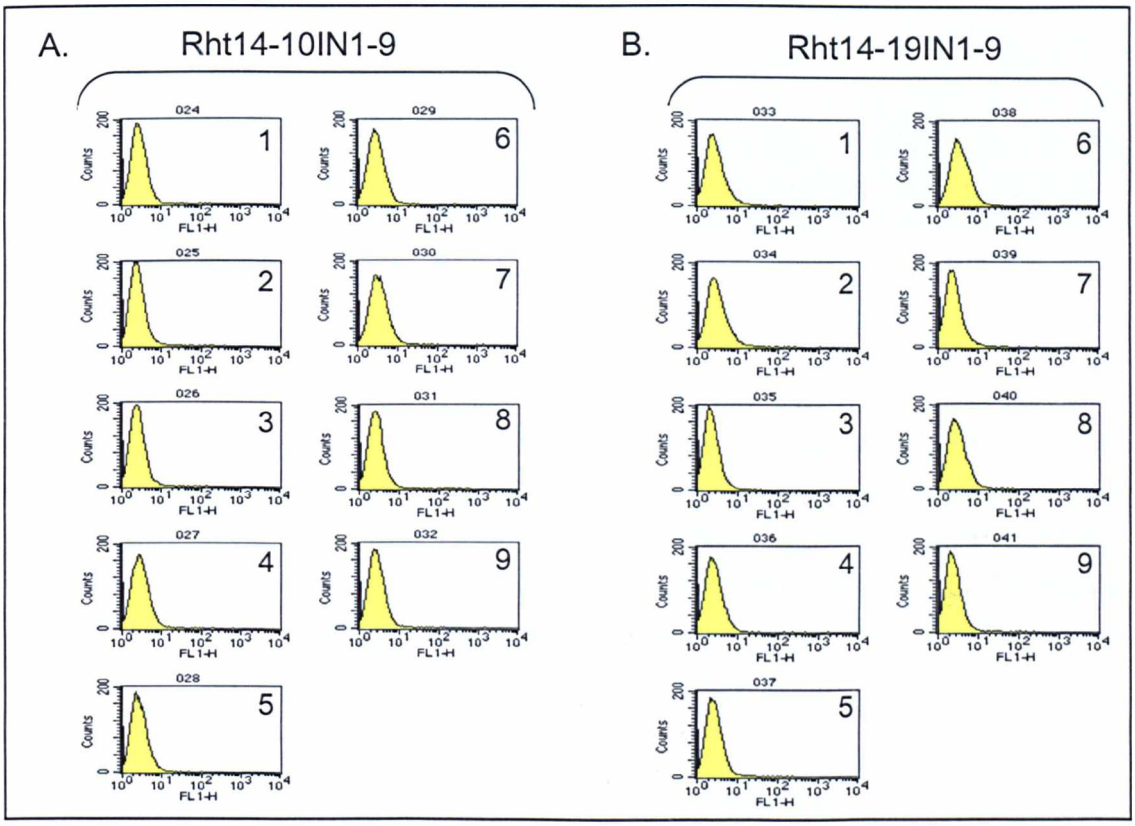



**Figure 5.5 FACS analyses on the selected *gpt*-expressing clones:** The FACS plots confirm that the selected clones were no longer expressing GFP for both the Rht14-10 experiment (**A.**) and the Rht14-19 experiment (**B.**).

*The recombination junction was successfully amplified by PCR*

Confirmation that the desired insertion event had taken place in the nine GFP-negative clones for each experiment was demonstrated by PCR amplification. Cell pellets were treated with pronase (See 2.4) and the primers TIGHT1 and BSgptR (for sequence see Table 2.1) were used to amplify the recombination junction (Figure 5.6A). All clones were positive for this region, as confirmed by the production of a 629bp band (Figure 5.6B). Further evidence for the desired insertion event was also demonstrated by Southern analysis (see Figure 5.13).

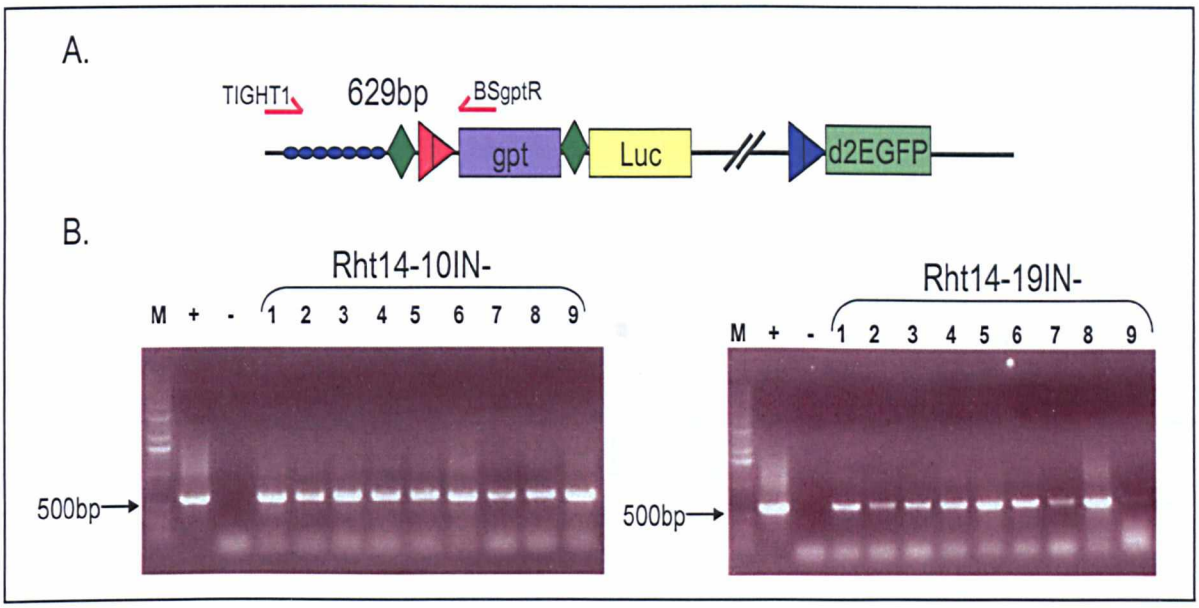



**Figure 5.6 PCR amplification of the recombination junction:** **A.** Map (not to scale) of integrated pINSluc showing the region amplified by PCR to help confirm that Cre-mediated insertion had taken place. **B.** Ethidium bromide-stained gel used to analyse PCR products generated by primers TIGHT1 and BSgptR acting on cell pellets of the indicated clones (for methods see section 2.1.3). The negative control (-) was Rht14-10 or Rht14-19 cell pellets and the positive control (+) was 10ρg of pTIGHTgpt DNA (see 2.3.4). M = 1kb Ladder (Invitrogen).

## 5.2.2 Testing the excision step of the new 'screen and insert' strategy

*Preparation of gpt-expressing cells for transfection with flp-recombinase*

Successful isolation of clones that had undergone the insertion step of the improved 'screen and insert' strategy led to subsequent testing of the excision step (Figure 5.7). Clones Rht14-10IN5 and Rht14-19IN5 were both selected to test the second step and were transferred to HT medium for 5 days prior to the removal of MPA and xanthine and the addition of tetracycline to the medium. Tetracycline was required at this stage to down-regulate expression from the TRE before the GOI (luciferase) became linked to the TRE. After at least 72 hours in tetracycline, *flp*-mediated excision was carried out (Figure 5.7). 500,000 Rht14-10IN5 and Rht14-19IN5 cells were both transfected with 4µg of pCAGGS-flpe using Lipofectamine 2000 (for methods see section 2.8.5). 24 hours after the transfection they were plated out at low density (50 cells, 100 cells, 250 cells and 500 cells per 15cm plate) for colony formation in tetracycline supplemented medium.

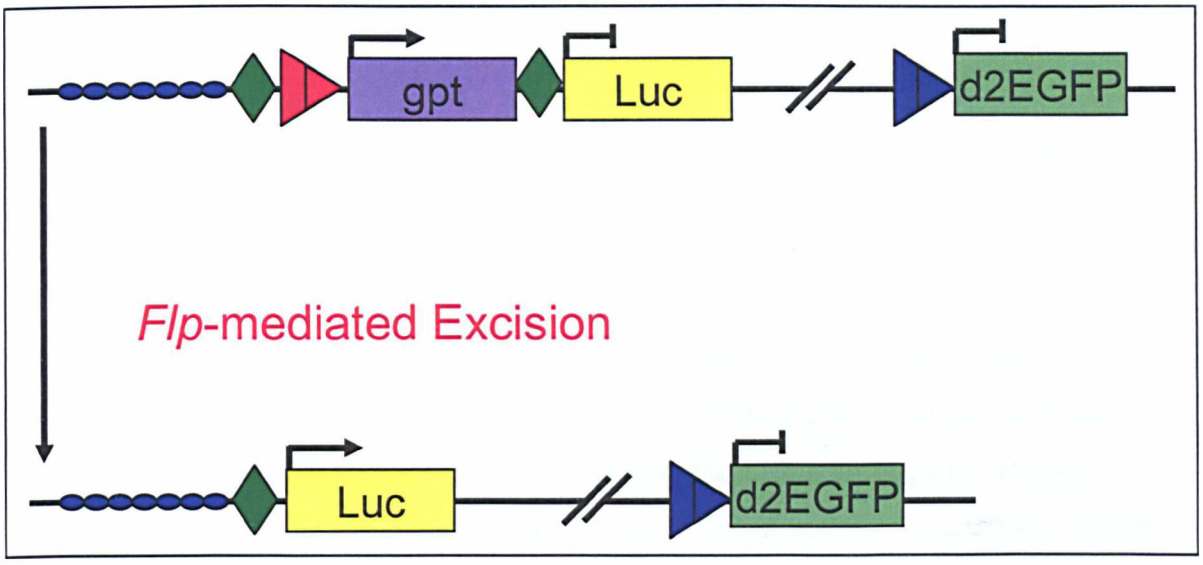



**Figure 5.7 The excision step of the improved 'screen and insert' approach:** The excision step of the strategy is illustrated (not to scale). When gpt-expressing clones derived from the insertion step are transfected with a *flp*-expressing construct (pCAGGS-FLPe) an excision event between the two FRT sites (green diamonds) will take place. This reaction is carried out in the presence of tetracycline so as not to express the GOI (luciferase) whilst isolating the desirable clones. For key to symbols see Appendix I.

*PCR analysis on pools and colonies demonstrated the expected excision event*

After 10 days in tetracycline supplemented medium the colonies were big enough to be picked and analysed by PCR. To allow for a possible low frequency of *flp*-mediated excision, pools of six clones were analysed together for the excision event. Once pools with the excised PCR product were identified, individual clones from the appropriate pools were analysed further. The size change expected when the excision event took place, using TetOF and LucR primers (for sequence see Table 2.1), is illustrated in Figure 5.8.

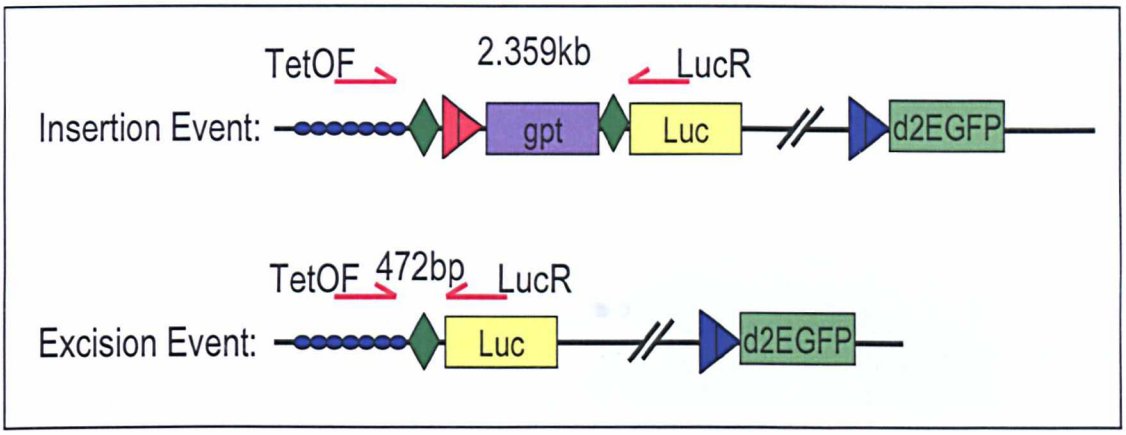



**Figure 5.8 The PCR screen used for identifying clones that had undergone the excision event:** A cartoon illustrating how PCR analysis was used to screen for clones that had undergone the excision event (for methods see section 2.4). .

For clone Rht14-10IN5, 16 pools of 6 clones (i.e. 96 clones in total) were analysed. Nine of the pools (pools 2, 5, 6, 7, 8, 9, 10, 11 and 14) had a PCR product of 472bp as expected for excision (Figure 5.9; top), therefore suggesting an excision frequency of at least 9.4% (9/96 X100). The individuals from pools 5 and 9 were analysed further by PCR and revealed that 7 out of the 12 (Clones Rht14-10IN5flp2, 3, 6, 7, 8, 9 and 11) were positive for the excision product (Figure 5.9; bottom). This therefore meant that a *flp*-mediated excision frequency of at least 14.6% (14/96 X100) could be estimated.

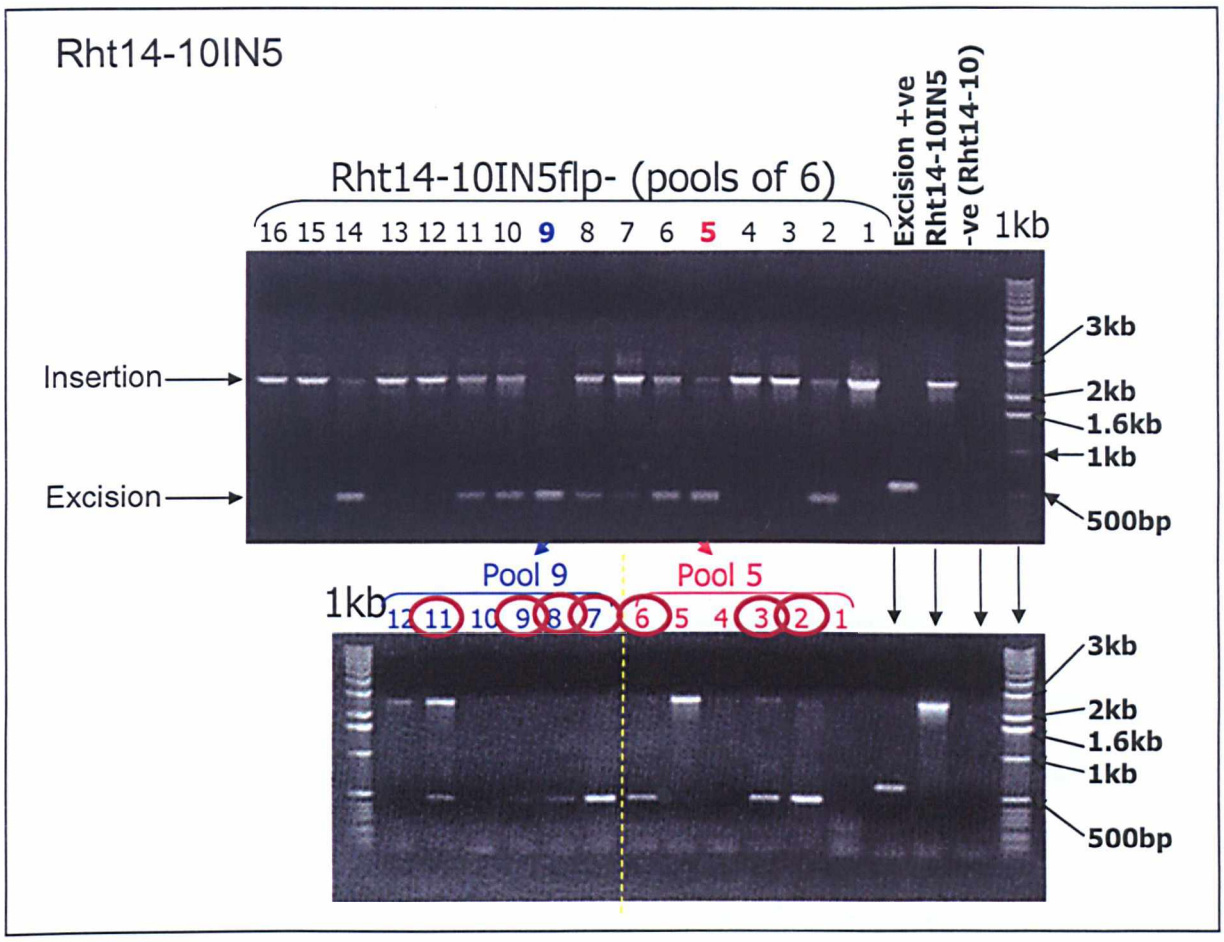



**Figure 5.9 PCR analysis for excision in clones generated from Rht14-10IN5:** Ethidium bromide-stained agarose gel of PCR products generated from 16 pools of six clones isolated after Rht14-10IN5 cells were transfected with flp-expressing plasmid (top). (Bottom) Ethidium bromide-stained agarose gel of PCR products generated from the individuals of pools 5 and 9. The PCR-negative control was generated from Rht14-10 cell pellets. The positive controls used were 10ρg of pTIGHTLuc construct (band expected: 549bp) and Rht14-10IN5 cell pellets (2.359kb) representing the approximate band sizes expected for the excised and un-excised products, respectively.

For clone Rht14-19IN5, 12 pools of 6 clones (i.e. 72 clones in total) were analysed. Two of the pools (pools 9 and 12) had the 472bp PCR product (Figure 5.10; top), therefore suggesting an excision frequency of at least 2.8% (2/72 X 100). The individual clones from pools 9 and 12 were analysed further by PCR and revealed that only one clone from each pool (Clone Rht14-19IN5flp26 and 32) generated the excision product (Figure 5.10; bottom). This therefore meant that the absolute *flp*-

mediated excision frequency could be calculated as 2.7%. This frequency was lower than that (14.6%) seen when using clone Rht14-10IN5. The two experiments were carried out using similar methods, however the tetracycline was removed from the medium during the lipofection step when using clone Rht14-10IN5 but not when using Rht14-19IN5. The presence of tetracycline in the medium may have affected the transfection efficiency when using Lipofectamine 2000.

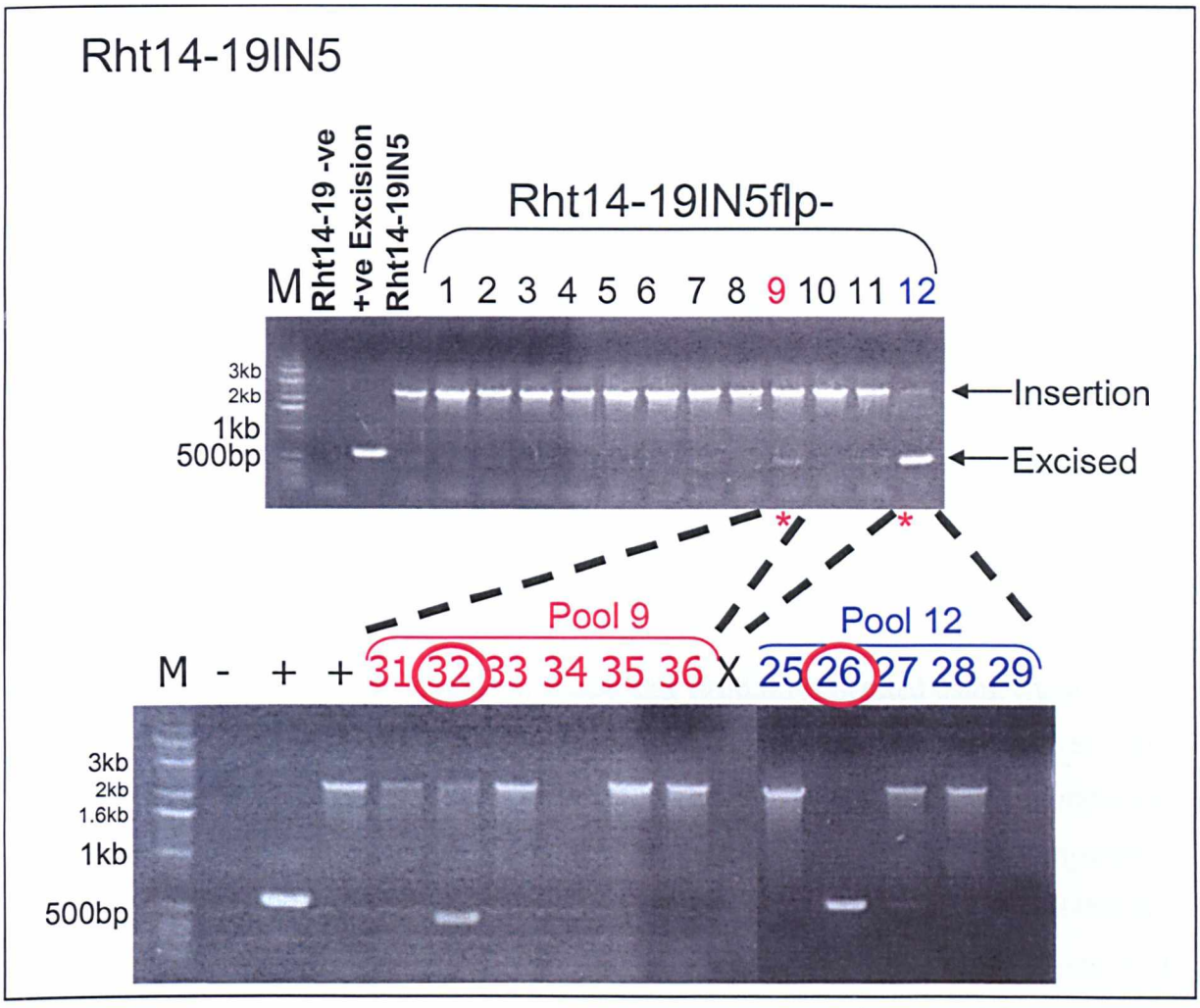



**Figure 5.10 PCR analysis for excision in clones generated from Rht14-19IN5:** Ethidium bromide-stained agarose gel of PCR products generated from 12 pools of six clones isolated after Rht14-19IN5 cells were transfected with flp-expressing plasmid (top). (Bottom) Ethidium bromide-stained agarose gel of PCR products generated from the individuals of pools 9 and 12. The PCR negative control was generated from Rht14-19 cell pellets. The positive controls used were 10pg of pTIGHTLuc construct (549bp) and Rht14-19IN5 cell pellets (2.359kb) representing the approximate band sizes expected for the excised and un-excised products, respectively.

*Tetracycline-inducible luciferase was detected in all of the identified excised clones*

All the identified excised clones were split equally into medium supplemented with or without tetracycline. They were incubated for 48 hours and then analysed for inducible luciferase activity (for methods see 2.7). Each of the seven clones from the Rht14-10IN5 (10flp2, 10flp3, 10flp6, 10flp7, 10flp8, 10flp9, 10flp11) experiment demonstrated similar tetracycline inducible luciferase expression to each other (Figure 5.11A). The induction values (= luciferase activity –Tet/+Tet), for the seven clones, were calculated and the mean induction was found to be 6214-fold (Figure 5.11A). This value is ~6-fold higher than the maximum induction value seen for GFP expression in the parental clone Rht14-10 (~1000-fold; Figure 5.4). The average uninduced levels of luciferase expression in the seven tested clones is 36.9-fold greater than the control experiment (Rht14-10 cells; Figure 5.11A); this is unlike the uninduced GFP-expression observed in the parental clone Rht14-10, which exhibits almost identical expression to that of background (Figure 5.4). Confirmation that luciferase expression was restricted until the excision event had taken place was also demonstrated (Figure 5.11B).

Similarly, the two excised clones (19flp26 and 19flp32) generated using clone Rht14-19IN5 demonstrated inducible luciferase activity (Figure 5.11C); however, their induction values were completely different to one another. A much higher induction value was obtained for 19flp26 (16063-fold above uninduced levels) compared to 19flp32 (905-fold above uninduced levels); generating an average of 8484-fold (Figure 5.11C). Maximum parental GFP expression was ~500-fold (Figure 5.4); therefore 19flp32 appeared to be more representative of this value than 19flp26. Uninduced levels of expression were relatively low when compared to previous experiments with 19flp26 being 1.6-fold greater than background and 19flp32 being identical to background expression levels, where background is clone Rht14-19 cells. Again, 19flp32 demonstrated a similar pattern to that seen for the parental GFP-expressing clone (Figure 5.4).

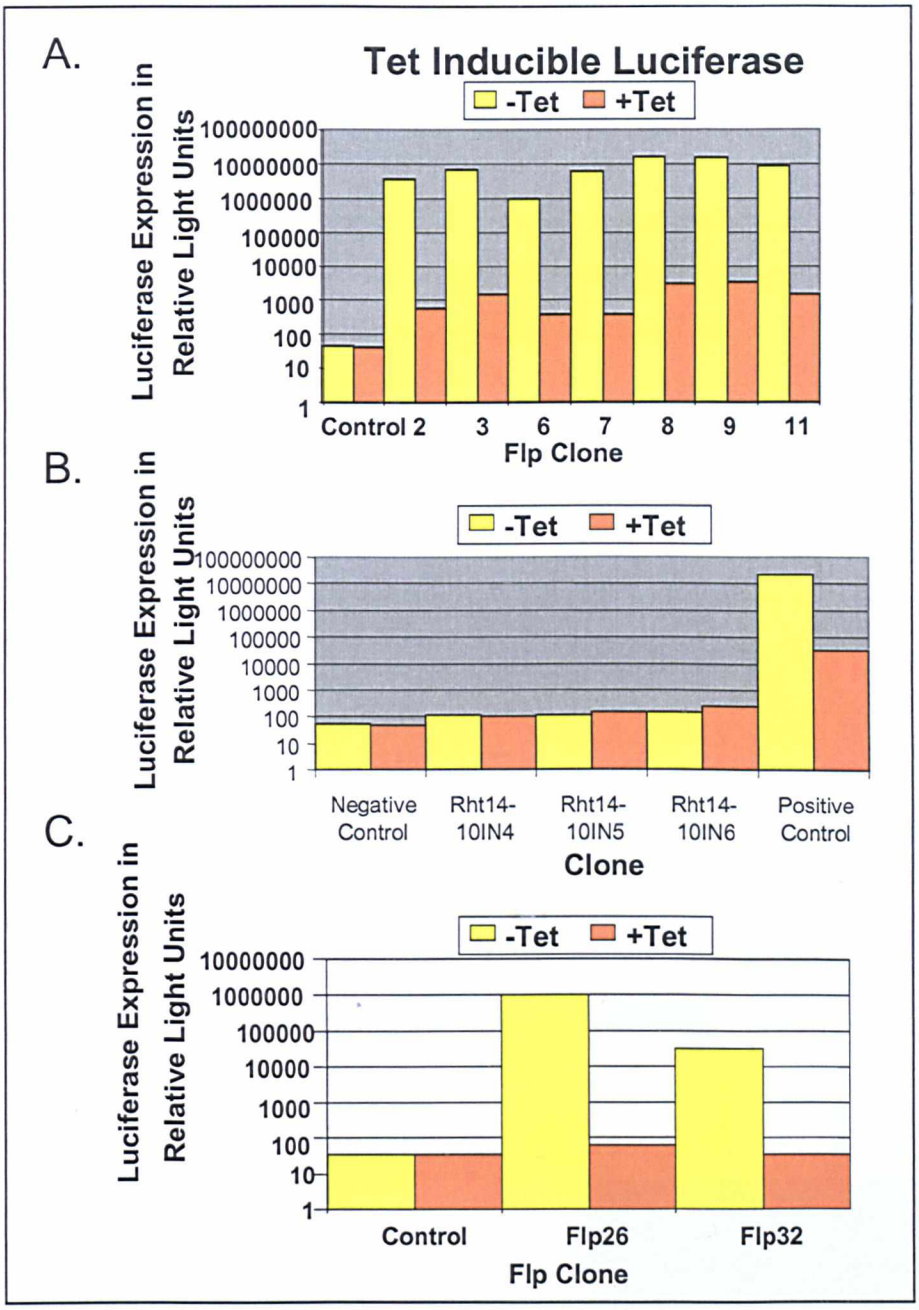



**Figure 5.11 Luciferase activity in excised clones generated from Rht14-10IN5 and Rht14-19IN5:** **A.** Tetracycline inducible luciferase expression in each of the 7 excised clones generated from Rht14-10flp is illustrated. The control experiment involved analysing non-luciferase Rht14-10 cells. **B.** The results of a luciferase assay illustrating that there is no luciferase expression prior to the excision event in clone Rht14-10IN4-6. The negative control used was non-luciferase Rht14-10 cells and the positive control used was a tetracycline-inducible luciferase expressing cell-line which has been discussed earlier (clone IRESHyg1, see Section 4.2.3). **C.** Tetracycline inducible luciferase expression in each of the two excised clones identified from the Rht14-19flp experiment is illustrated. The control experiment involved analysing non-luciferase Rht14-19 cells. For **A.** and **B.** and **C** all cells were plated in +/- Tet for 48 hours, after which a luciferase assay was carried out as described in section 2.7. The mean induction value (-Tet/+Tet) is shown in blue and the mean uninduced expression over background levels (the control) is shown in red).

*The expected genome rearrangements were demonstrated by Southern blot analysis*

Finally, genomic DNA from some of the clones generated from the experiments involving Rht14-10 and Rht14-19 were digested with *Bgl*II (Figure 5.12) and *Hind*III (Figure 5.13), respectively. Southern analysis was carried out on the digested DNA and probed with an *Nco*I fragment of the GFP gene (for methods see 2.5). Confirmation that both the insertion and excision event had taken place when using Clone Rht14-10 was demonstrated by the appropriate change in band sizes (Figure 5.12).

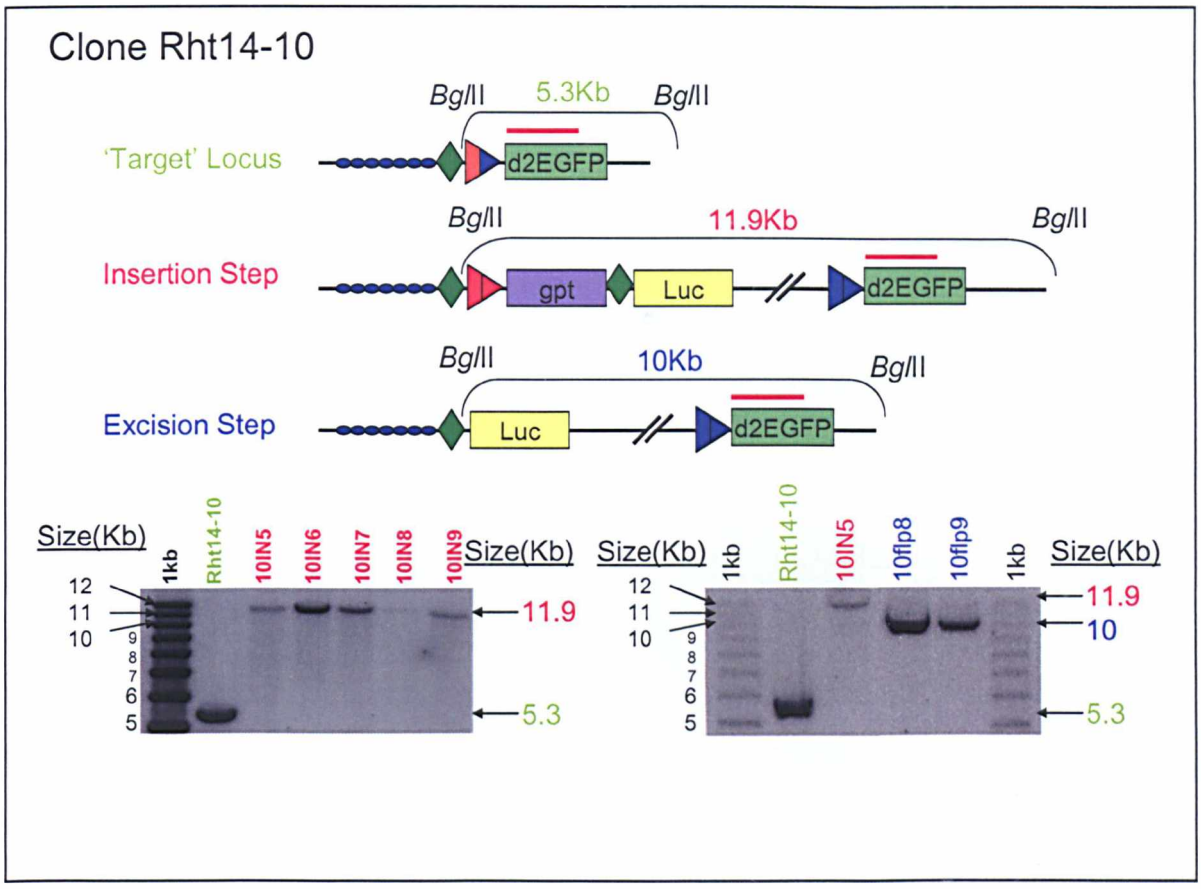



**Figure 5.12 Southern analyses demonstrating both the insertion and excision events in Rht14-10-dervived clones:** Top: The expected ( not to scale) fragments sizes when probed with the GFP (*Nco*I) gene for the 'target' clone Rht14-10, the insertion event (Clone Rht14-10IN5-9) and the excision event (10flp8 and 10flp9), is illustrated. Bottom: Genomic DNA was cut with *Bgl*II and Southern analysis (for methods see section 2.5) revealed that the expected genomic rearrangements had taken place.

For Clone Rht14-19, however, an explanation for the large differences in luciferase activity between 19flp26 and 19flp32 became apparent (Figure 5.13). From the Southern analysis it appeared as though 19flp26 had undergone the excision event and 19flp32 had not. It also revealed that 19flp32 was demonstrating the band size expected for the insertion event. It was therefore concluded that clone 19flp32 may have been contaminated with a few cells that had undergone the excision. These cells were picked up during the more sensitive PCR step and luciferase assay; however, there were not a high enough proportion of these cells to be detected by Southern blot.

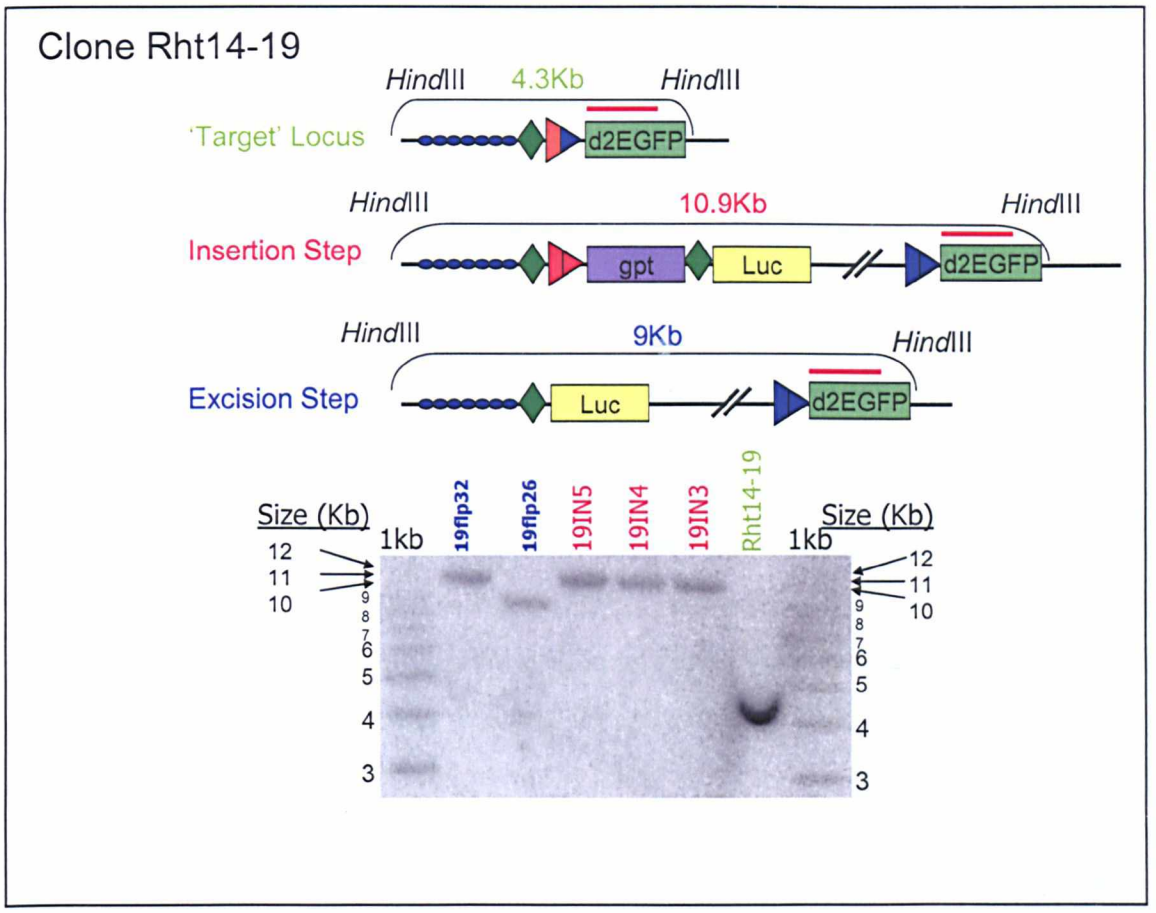



**Figure 5.13 Southern analysis demonstrating both the insertion and excision event in Rht14-19-derived clones:** Top: The expected ( not to scale) fragments sizes when probed with the GFP (*Nco*I) gene for the 'target' clone Rht14-19, the insertion event (Clone Rht14-19IN3-5) and the excision event (19flp26 and 19flp32), is illustrated. Bottom: Genomic DNA was cut with *Hind*III and Southern analysis (for methods see section 2.5) revealed that the expected genomic rearrangements had taken place for clone 19flp26 but not 19flp32.

### 5.2.3 Adapting the 'insertion' construct for easy cloning of any GOI

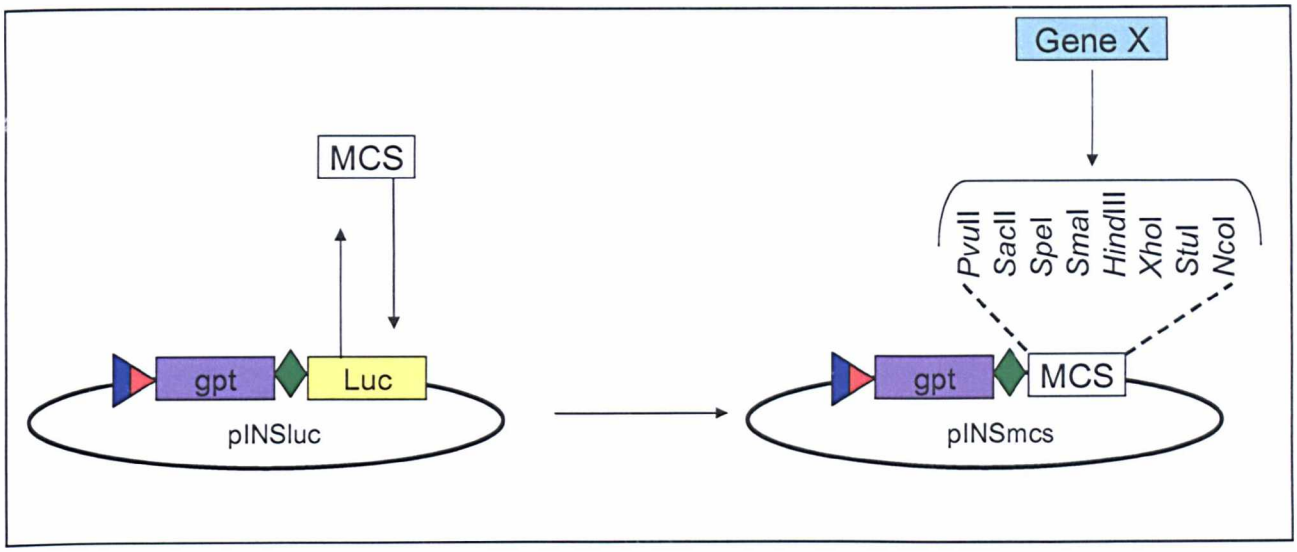



**Figure 5.14: Plans to add a MCS to the 'insertion' construct:** The luciferase ORF was removed and replaced with an oligonucliotide coding for unique sites (MCS) which would enable the easy addition of any GOI into the 'insertion' construct. The addition of the MCS generated pINSmcs (not to scale). For scale representations see Appendix IV. For key to symbols see Appendix I.

Further alterations to the pINSluc 'insertion' construct were carried out following the successful demonstration of the improved 'screen and insert' strategy. The changes involved replacing the luciferase ORF with an oligonucleotide coding for a multiple cloning site (MCS), generating pINSmcs, to allow the easy cloning of any future GOI (Figure 5.14). The MCS was designed by culminating all the restriction sites that were not present in pINSluc and care was taken so as not to introduce any start sites that may appear upstream of an integrated GOI. The cloning steps for inserting the MCS oligonucleotide are described in the materials and methods (see 2.3.3.3; for a scale vector representation of pINSmcs see Appendix III) and proof that successful insertion of the MCS had taken place was demonstrated by restriction digestion (data not shown) and sequence analysis (using the INSF sequencing primer, for sequence see Table 2.1) (Figure 5.15). Once this was achieved the system was ready to be used with any selected GOI.

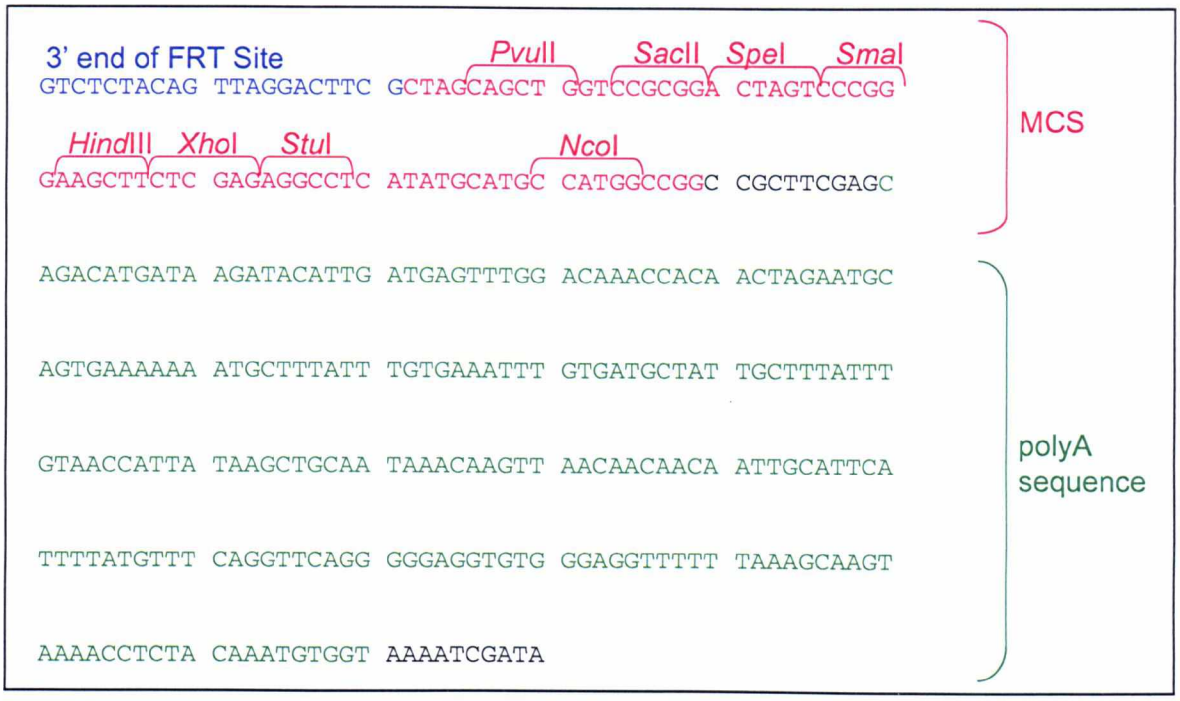



**Figure 5.15 The sequencing results of pINSmcs:** Shows the results of sequencing pINSmcs using the INSF primer (Table 2.1), which is located slightly upstream of the FRT site. The 3'end of the FRT site is shown in blue. The MCS is shown in red with a few of the selected unique cloning sites which the oligonucleotide provided. The green represents the polyA tail for the accurate termination of the GOI during transcription.

## 5.3 Discussion

### 5.3.1 Assessment of the new and improved 'screen and insert' system

The new and improved 'screen and insert' system was designed to enable the generation of clones with tightly regulated GOI, without expressing from the TRE. This was achieved using a strategy which firstly integrated the 'insertion' construct into the tightly regulated locus using a selection for *gpt* expression to combat the low *Cre*-mediated insertion frequency. The next step of the strategy deployed the reported high frequencies of *flp*-mediated excision to place the GOI downstream of the suppressed TRE using selection free conditions and colony analysis by a simple PCR screen. The 'insertion' construct was further redesigned to include a MCS for easy cloning of any future GOI.

### 5.3.2 Successful isolation of recombined clones using *gpt* expression

*A lower absolute insertion frequency was detected when compared to previous results*

After optimising the forward selection for *gpt* expression in HT1080 cells by using HAT supplement, colonies which had undergone the insertion event were successfully isolated. The number of colonies generated differed by almost 3-fold between clone Rht14-10 (11 colonies) and clone Rht14-19 (32 colonies). These differences may simply be due to the accessibility of the 'target' locus in the genome of each of these clones. The number of colonies generated when using clone Rht14-19, however, was almost half that seen for previous experiments (Table 5.2). When 500,000 Rht14-19 cells were co-transfected with pMC-Cre and plox66lucIresHyg (see Chapter Four) or pINSluc, 53 and 32 clones were generated, respectively. In addition, only 69% of gpt-expressing clones were no longer expressing GFP, compared to 98% when using the IRES-base 'insertion' construct. The absolute recombination frequency had also dropped from 0.01% (for plox66lucIresHyg) to 0.0044% (for pINSluc), a value more consistent with the previously reported 0.002% (Araki *et al.*, 1997).

**Table 5.2 A comparison of transfection efficiencies in clone Rht14-19:** A table illustrating a reduction in the number of colonies obtained and a drop in the absolute recombination frequency, when pINSluc was used in clone Rht14-19.

| 'Insertion' Construct | No. Cells Transfected | No. Colonies generated | No. GFP -ve colonies | % of GFP negative cells | Absolute recombination frequency |
|---|---|---|---|---|---|
| plox66LucIresHyg | 500, 000 | 53 | 47/48 | 98% | 0.01% |
| ploxLucIresHyg control | 500,000 | 0 | 0 | - | - |
| pINSluc | 500, 000 | 32 | 22 | 69% | 0.0044% |
| pINSluc control | 500,000 | 0 | 0 | - | - |

The reason why there was a decrease in colony numbers and an increase in what appeared to be random integration when using pINSluc, compared to the previously tested 'insertion' constructs, is unknown. It is possible that the *gpt*-selection is not as efficient as others (hygromycin) at generating healthy, thriving colonies. Additionally, supplementing the medium with hypoxanthine (in HAT) to diminish the activity of mammalian *gpt* (to help reduce background levels) may also have affected the activity of the *E.coli gpt* and hence restricted colony formation (Mulligan and Berg, 1980). It had already been reported that clean forward selection for *gpt*-expression was difficult to achieve (Trott *et al.*, 1995). It is therefore possible that the high number of GFP-positive clones observed after selection were not random integrants, but merely a result of a high background during selection. If this were the case then it was surprising to find no colonies forming on the control plates (-pMC-Cre; Table 5.2). Perhaps the background was generated on the experimental plates by *gpt*-negative cells thriving whilst situated nearby *gpt*-positive clones and if this were the case it is expect that they would not have survived being picked into individual wells containing MPA, xanthine and HAT. This hypothesis was never investigated as all GFP-positive clones were always discarded.

*The clone characteristics of an insertion event were successfully demonstrated*

Nine GFP-negative clones, generated from Rht14-10 and Rht14-19, were analysed further for the expected characteristics of an insertion event. All clones successfully amplified the recombination junction and Southern blot analysis helped prove they had all undergone the correct genomic rearrangements.

### 5.3.3 Successful isolation of clones containing tightly regulated GOI

*A high frequency of flp-mediated excision was detected for Rht14-10 but not Rht14-19*

A PCR strategy was designed for analysing the non-selected tetracycline exposed colonies, generated after transient expression of *flp*-recombinase in clones Rht14-10IN5 and Rht14-19IN5. The frequency of the excision event in HT1080 cells was unknown and it is expected to vary between 'target' clones, due to the changing position and accessibility of the locus within the genome. Pools of six clones were analysed together and positive pools were isolated and the individuals were examined further. The results revealed that clone Rht14-10IN5 demonstrated a *flp*-mediated excision PCR product in at least 14.6% of clones tested. This is a higher frequency than reports by Schaft and colleagues, who demonstrated a 6% *flp*-mediated excision frequency in ES cells when using the same flp-expression vector (Schaft et al., 2001). A lower frequency of 2.8%, however, was detected when using Rht14-19IN5. In fact, this frequency later reduced to 1.4% when it was revealed that one of the clones (19flp32) generated a positive PCR band due to contaminating cells. The low frequency of excision seen for Rht14-19 was unlikely to be caused by an inaccessible 'target' locus due to the higher frequencies of *Cre*-mediated insertion observed earlier.

Both experiments were carried out using similar methods; however for clone Rht14-10IN5 the medium was replaced with fresh (-tetracycline) medium during the 4 hour transfection incubation step. The Lipfectamine 2000 protocol does suggest removing all antibiotics from the medium prior to the transfection. It was therefore possible that the tetracycline present in the medium may have affected the transfection efficiency for clone Rht14-19IN5. In fact, when these conditions (those adopted for clone Rht14-19IN5) were repeated using clone Rht14-10IN5 no excised colonies were found, even after analysing 48 individual clones (data not shown).

*The characteristics expected after the excision event were present in almost all clones*

Inducible luciferase was detected in all of the excised clones analysed (including 19flp32). Again, the amount of luciferase activity and induction varied from clone to clone and the average fold-induction was 6-times greater than that of the parental GFP-expressing clone for experiments involving Rht14-10IN5 and up to 30 times greater when using Rht14-19IN5. Again, as discussed in Chapter Four (see section 4.3.2), the luciferase techniques still require further optimisation to facilitate the generation of more reproducible results. If the assays still prove unreliable, then other quantitative methods for measuring gene expression could be adopted, such as Western analysis and RT-PCR.

Finally, Southern analysis confirmed that the correct genomic rearrangements had taken place in the clones generated from clone Rht14-10IN5. It also revealed that the majority of cells of clone 19flp32 had not undergone the excision event. Amplification of the excised PCR product would be favoured over the unexcised PCR product due to its smaller size. The positive PCR band identifying 19flp32 was therefore likely to be contaminating excised cells. Consequently, confirmation of the excision by Southern analysis after identification by PCR is recommended prior to further experiments.

### 5.3.4 Future applications of the developed 'screen and insert' system

The 'insertion' construct was redesigned to incorporate a MCS to allow easy cloning of any subsequent GOI. Consequently, the developed 'screen and insert' strategy (System Two) can now be used to generate HT1080 cells expressing tightly regulated, never expressed before, GOI. The system should suite any GOI, however it would be particularly desirable for those whose overexpression or down-regulation may compromise the cells survival. For example, uncontrolled overexpression of the *RAD52* gene, which is involved in homologous recombination during DNA repair, has been reported to reduce cell viability and is lost during extensive passages (Yanez and Porter, 2002). The system would enable the overexpression of the GOI in healthy cells at a particular time point whilst the effects can be observed. Afterwards, the GOI expression levels can also be returned to endogenous levels to alleviate any resulting stresses or to analyse the cells ability to recover.

Additionally, this 'screen and insert' approach could also be applied for the controlled expression of a recombinase (such as Cre recombinase) or a rare cutting endonuclease (such as *I-SceI*) to achieve regulated recombination and cutting in every cell, respectively. It would be predicted that tight regulation of such genes would be important for preventing a background of uncontrolled reactions during leaky uninduced TRE expression. Conversely, conditional loss-of-function of a reported essential gene (e.g. *RAD51* (Sonoda *et al.*, 1998), *RAD50* (Luo *et al.*, 1999); both involved in DNA repair by homologous recombination) could also be achieved by avoiding lethality after gene silencing. Expression from the inducible cassette could be used to rescue the cells following targeting of the endogenous genes, either by gene targeting or interference. The timing of GOI down-regulation to produce a knockout phenotype could be controlled to a particular time point (e.g. during development or time of observation).

# Chapter Six

# Establishing the System in Non-Transformed Cells

## 6.1 Introduction

In the previous chapters (Four and Five) a 'screen and insert' strategy was established and tested in the human fibrosarcoma cell line, HT1080. This transformed cell line is commonly used as they are immortal, adherent and very easy to maintain in culture. Despite this, the use of transformed cells for biological research has been criticised due to their gross phenotypic and morphological characteristics. These include loss of contact inhibition, growth in low serum and a tendency to lose or gain whole or partial chromosomes, which can lead to genetic instability (polyploidy) (Bodnar *et al.*, 1998). Due to their nature they also exhibit abnormal gene expression profiles when compared to 'normal' cells and there are concerns that this may confuse the study of many biological mechanisms (Jiang *et al.*, 1999). It would therefore be more advantageous to set up the 'screen and insert' system in a cell type whose genomic stability and expression profile represents that observed in a typical wild-type cell, as much as possible.

A decision was made to set up the system in the telomerase immortalised retinal epithelial cell line, hTERT-RPE1 (Clontech). This cell line was created from human retinal pigment epithelial (RPE) cells that were immortalised by stably expressing human telomerase reverse transcriptase (hTERT). Telomerase is an enzyme actively expressed in germline cells to maintain the length of their telomeres and subsequently prolonging their life-span. Telomeres are specific DNA repeat sequences that stabilise chromosomal ends (Blackburn *et al.*, 1991). Telomerase is not expressed in somatic cells (e.g. RPE cells), consequently telomeres are degraded after subsequent divisions and when a critical length is reached they enter a non-dividing state called senescence (Harley *et al.*, 1990). It has been shown that telomerase expression in normal cells does not appear to induce changes associated with a malignant

transformed phenotype, therefore, they provide a good model cell-type to establish the 'screen and insert' system further (Jiang, *et al.*, 1999).

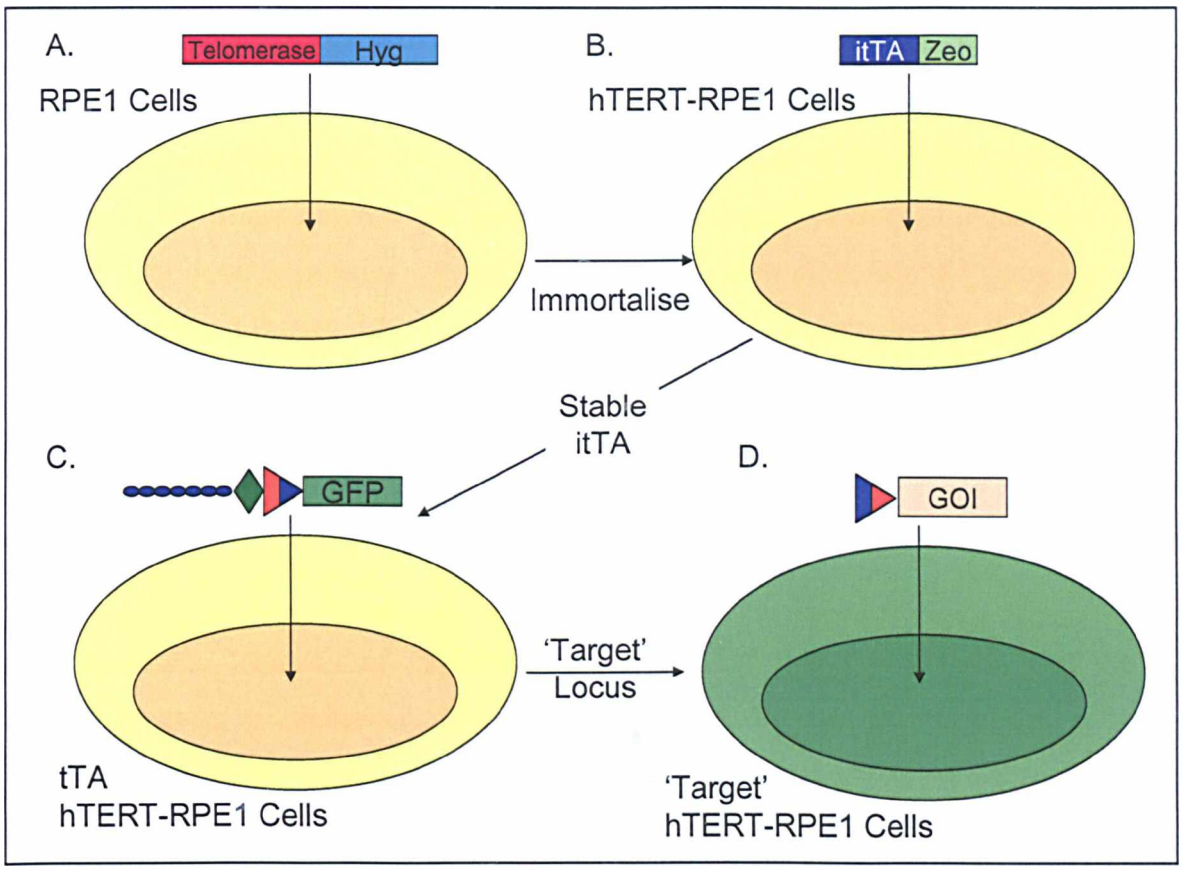



**Figure 6.1 The steps involved in establishing the 'screen and insert' system in hTERT-RPE1 cells:** **A.** Retinal Pigment Epithelial (RPE) cells were immortalised by stable integration and expression of the telomerase ORF (Clontech). **B.** hTERT-RPE1 cells will be transfected with the 'improved' humanised transactivator (itTA) so that the tetracycline system can be established in these cells. **C.** The itTA expressing hTERT-RPE1 cells will subsequently be transfected with the 'target' construct, containing TRE driven GFP. **D.** Finally, the hTERT-RPE1 cells expressing the itTA and tightly regulated 'target' construct will be used to test the *Cre*-mediated insertion event. For key to symbols see Appendix I.

The retinal epithelial cells were obtained from clontech and were already stably expressing the telomerase gene (Figure 6.1A). The first task was therefore to generate 'improved' tTA expressing hTERT-RPE1 cells (Figure 6.1B). As these cells had not already been used by the group the optimal transfection and selection conditions were determined. When this was achieved the selected itTA expressing clone was used to generate cells with tightly regulated GFP expression (Figure 6.1C). The FRT

containing 'target' construct (pTFRTlox71dGFP, Figure 6.1C) was used so that all 'insertion' constructs (Appendix III) could be tested on these cells.

Attempts to establish the tetracycline system in hTERT-RPE1 cells has not been documented a great deal, especially reports regarding the tightness of gene expression that can be achieved. It is already know that the amount of uninduced expression from the TRE differs between cell-types, resulting in various degrees of induction and control; low basal uninduced expression levels have been demonstrated using HeLa cells (Gossen and Bujard, 1995), and now HT1080 cells (Carpenter and Porter, 2004). A failure to generate clones with tight, reproducible, inducible gene-expression could be a possible limitation when using a new cell-type; however, it is hoped that the use of the improved TRE sequences (TREmod; see section 3.2.1, Figure 3.7) will overcome this.

The successful isolation of inducible GFP-expressing clones will mean that the 'insertion' event can subsequently be tested (Figure 6.1D). An additional limitation of using these cells may be that they have a reduced capacity to be successfully transfected. A reduced transient transfection efficiency could limit the successful isolation of recombined clones following *Cre*- and *flp*-expression, particularly during the latter non-selectable stage. A hygromycin-resistance cassette was included on the vector used to express the telomerase ORF in the hTERT-RPE1 cells (Figure 6.1A). This implicated that the 'insertion' constructs, plox66hyg and plox66LucIresHyg (Appendix III), could not be used in experiments involving the telomerase immortalised cells. A new simple 'insertion' construct that contained a neomycin ORF was created for examining the efficiency of the *Cre*-mediated insertion step in hTERT-RPE1 cells. In addition, to test the *flp*-mediated excision event, pINSLuc was used.

## 6.2 Results

### 6.2.1 Optimising the transfection efficiencies of hTERT-RPE1 cells

*Optimising the transfection efficiency of hTERT-RPE1 cells using Lipofectamine 2000*

For all steps of the developed 'screen and insert' strategy it was important that plasmid DNA could efficiently be transfected into the nucleus of the cells. It was not known how easy hTERT-RPE1 cells were to transfect and what method should be adopted. The steps of the system that required the highest efficiency of transfection were those that involved transient expression of the recombinases, in particular the non-selectable *flp*-mediated excision step. Previously, Lipofectamine 2000 had been used in HT1080 cells for transient expression; therefore this method was tested on the new cell-type. To determine the amount of DNA required to attain optimal transfection efficiencies in hTERT-RPE1 cells using Lipofectamine 2000, pTRElox71dGFP (Figure 4.1B) was used. Various amounts of the construct were transfected (for methods see section 2.8.5) into hTERT-RPE1 cells and analysed for GFP expression by FACS after 48 hours. The largest proportion of GFP expressing cells was observed when using 7μg of DNA. Unfortunately, the proportion of GFP expressing cells was only 33% (Figure 6.2A), suggesting that these cells were harder to transfect than HT1080 cells.

The hTERT-RPE1 cells were also plated out at various densities and transfected with 7μg of plasmid DNA, to account for the transfection efficiency being affected by cell-cell contact (i.e. not actively dividing due to contact inhibition) and cell surface area exposure to the lipofection media. The optimal transfection efficiency was achieved when the cells were 40% confluent with 35% of cells expressing transient GFP (Figure 6.2B). An increased survival after lipofection at a reduced confluency was also observed when using hTERT-RPE1 cells compared to HT1080 cells. A similar transfection efficiency (30%) was achieved for these cells by a colleague when optimising the transfection efficiency using FuGene. If more time were available, a full screen of different transfection methods for optimised transient expression in

hTERT-RPE1 cells would have been be carried out and the importance of this will be revealed in the subsequent chapter.

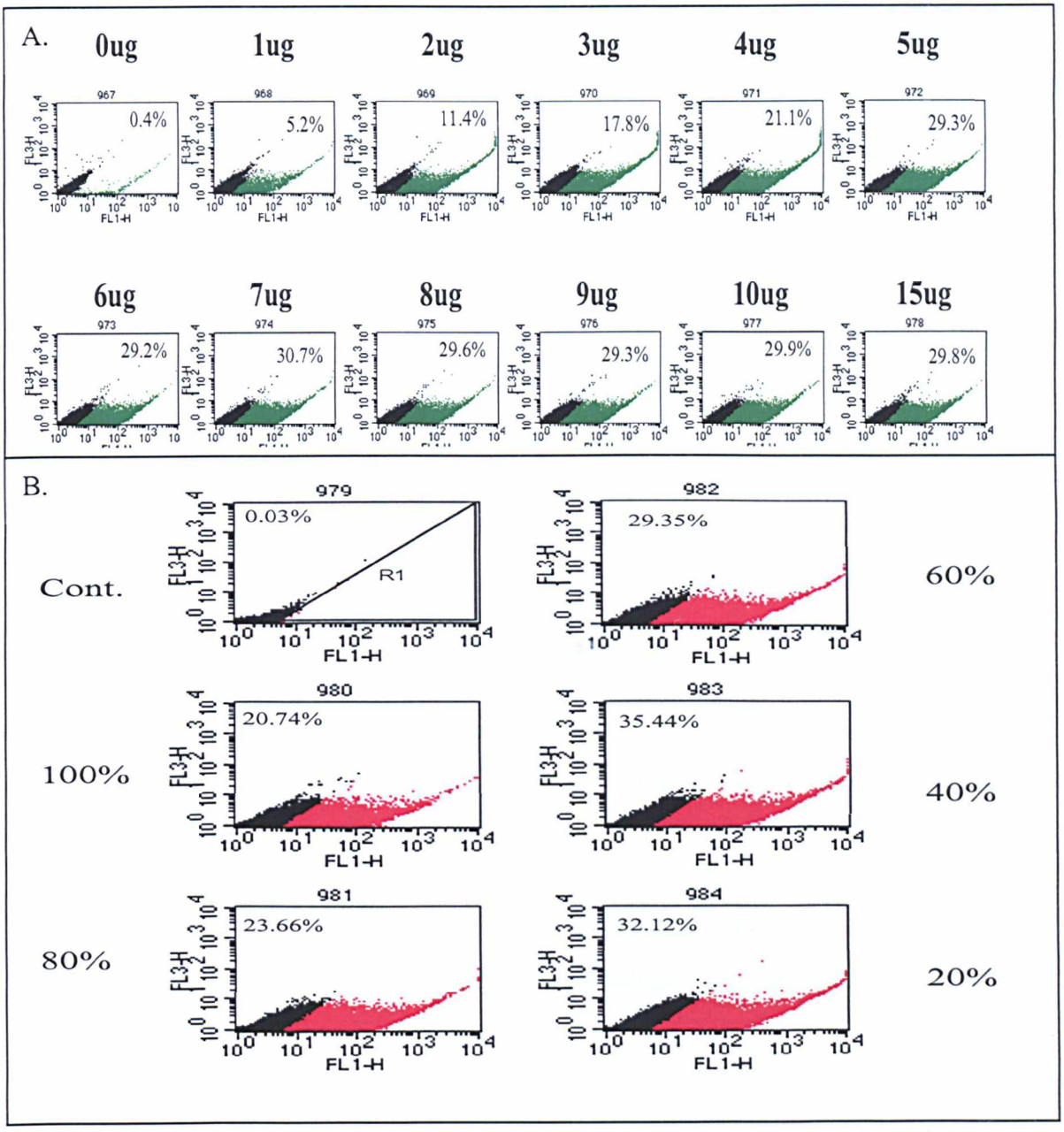



**Figure 6.2 Optimising the transfection efficiency of hTERT-RPE1 cells using Lipofectamine 2000: A.** Various amounts (shown) of pTlox71dGFP was used to transfect hTERT-RPE1 cells using Lipofectamine 2000 (see 2.8.5). **B.** hTERT-RPE1 cells were plated at various densities and transfected with 7µg pTlox71dGFP using Lipofectamine 2000 (see 2.8.5).

*Successful generation of zeocin-resistant hTERT-RPE1 colonies using electroporation*

As a new cell type was being established and the appropriate drug concentrations for selection had not previously been determined, hTERT-RPE1 cells were placed at various concentrations of zeocin, puromycin and neomycin to determine the optimum selection conditions. A concentration of 200μg/ml was determined for zeocin and 400μg/ml of neomycin, these were the same concentrations required for selection in HT1080 cells. For puromycin, however, 10 times the concentration used for HT1080 cells was required (4μg/ml) to completely kill hTERT-RPE1 cells.

Electroporation of HT1080 cells was a standard procedure carried out within the laboratory and the conditions had already been optimised previous to the start of this study. The conditions required for hTERT-RPE1 cells, however, were unknown; therefore as a starting point, the same procedure used for HT1080 cells was tested on the new cell type (for methods see section 2.8.4). The generation of an ample number of colonies (50+) for analysis, was all that was required from the electroporation experiments and optimising the conditions was not as important as for the transient transfections (using Lipofectamine 2000). Consequently, 10μg of the itTA-zeocin containing vector (pZeoSVhtTA; see section 2.3.1) was transfected into 8 million hTERT-RPE1 cells by electroporation (for methods see section 2.8.4). The cells were distributed evenly amongst ten 15cm plates and selected in zeocin for 20 days. Colonies (100+) were successfully generated which indicated that the cells were successfully being transfected using electroporation and these colonies were subsequently analysed for itTA expression.

## 6.2.2 Generating hTERT-RPE1 cells expressing improved transactivator

46 colonies were selected and assayed for their ability to regulate TRE-containing transcripts. This was carried out by transient transfection of pTIGHTLuc (contains TRE-tagged luciferase ORF; Clontech) using Lipofectamine 2000, in either the presence or the absence of tetracycline. The luciferase activity was calculated by carrying out a luciferase assay, as described in the materials and methods (See section

2.7). An induction value was generated for each clone by dividing the luciferase activity in the absence of tetracycline by the activity in the presence of teteracycline. Clone hh28 (hh from **h**umanised itTA in **h**TERT-RPE1 [shown in red, Figure 6.3]) was chosen for further experiments, as it demonstrated a relatively high induction value and a low background of uninduced expression.

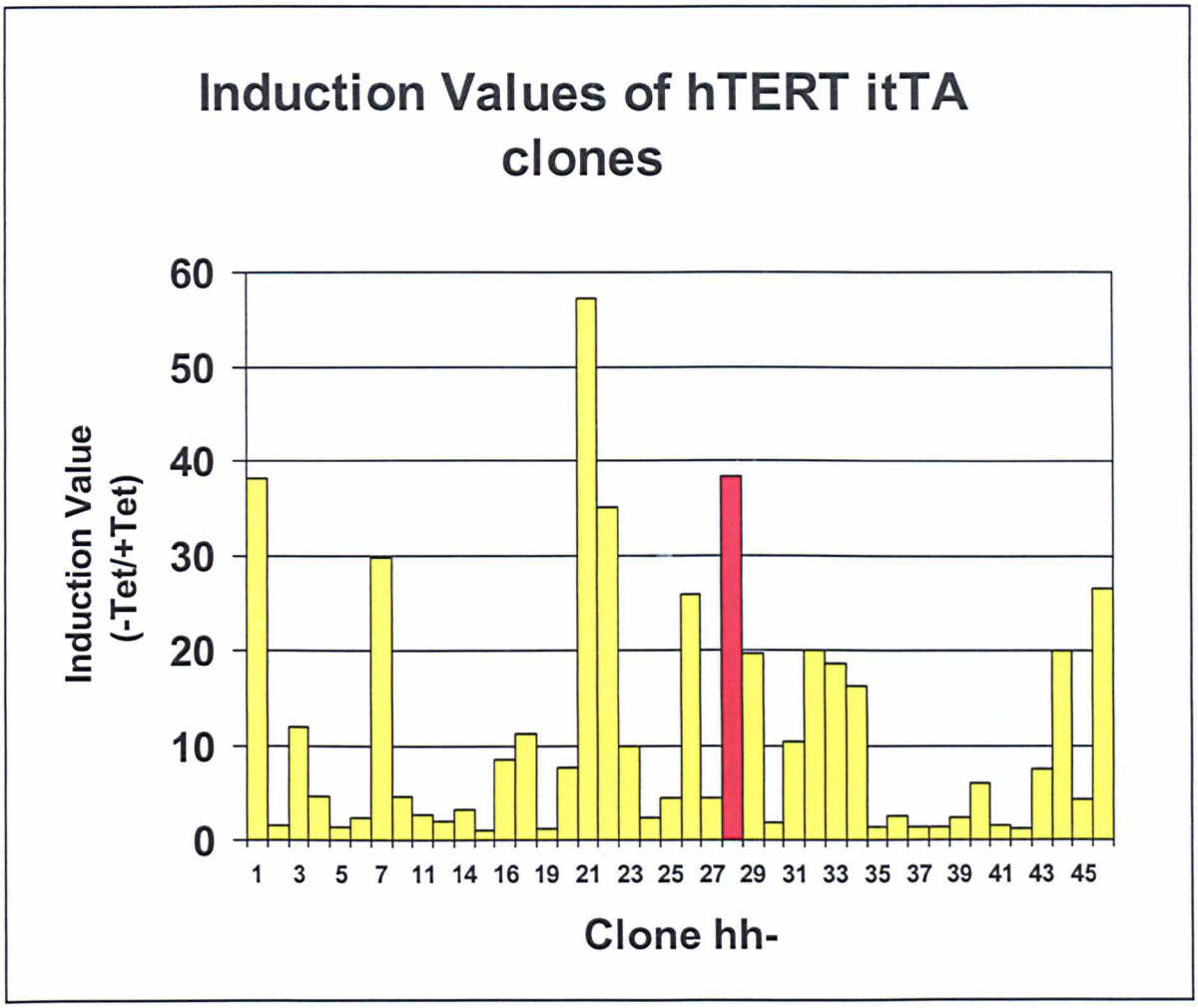



**Figure 6.3 Analysis of induction values in itTA expressing hTERT-RPE1 cells:** 46 itTA expressing hTERT-RPE1 clones were selected in zeocin and once established were placed in either the presence or the absence of tetracycline. After 48 hours, the cells were transiently transfected with pTIGHTLuc (Clontech) and incubated for a further 24 hours. A luciferase assay was carried out to identify those clones that could regulate TRE driven sequences (see section 2.7).

### 6.2.3 Generating tightly regulated GFP expressing hTERT-RPE1 cells

*Co-transfection was used to generate clones with stably integrated 'target' construct*

The FRT containing 'target' construct, pTFRTlox71dGFP (Figure 6.5A), was chosen to generate GFP expressing hh28 cells. This choice of 'target' construct allows the generation of cells that can be used for integration of all the (non-hygromycin containing) 'insertion' constructs discussed in Chapters Four and Five. The hh28 cells were co-transfected with linear 'target' construct and the puromycin-resistance plasmid (pBL-Pur/R) at a 20:1 ratio, respectively (for methods see section 2.8.4). These were the optimal co-transfection conditions determined for HT1080 cells and again co-transfection avoided the need to generate a 'target' construct that included a drug-resistance cassette, as enhancer activity in such cassettes were shown to interfere with expression from the TRE (Figure 3.9; Damke et al., 1995). After 20 days in puromycin selection, 42 colonies were analysed under a UV microscope for GFP fluorescence. The GFP-positive clones were picked (representing 75% of total colonies) and plated in either the presence or the absence of tetracycline for 48 hours and then analysed by FACS. All clones analysed down-regulated their GFP expression upon the addition of tetracycline, however uninduced levels of GFP-expression did not compare to the tightness observed for HT1080 cells (Figure 6.4).

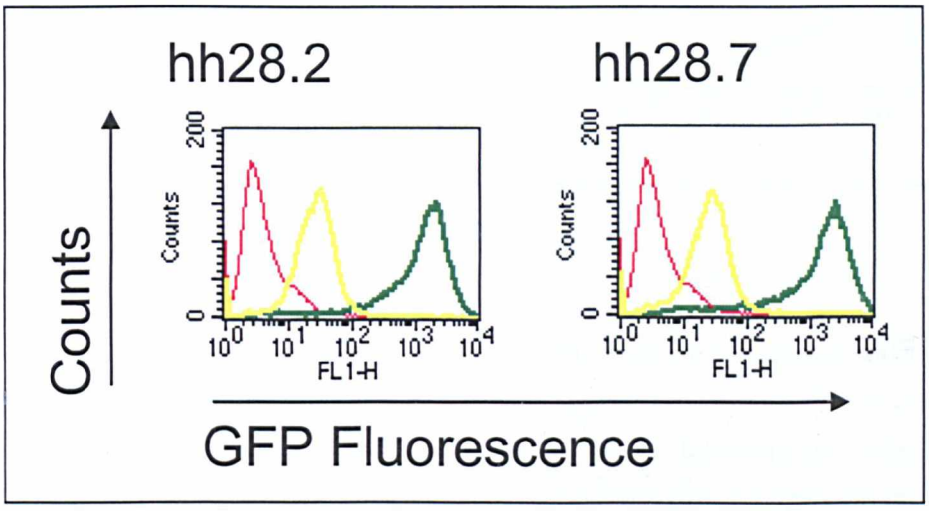



**Figure 6.4 Illustrates the high uninduced GFP-expression observed in hh28 cells after 48 hours in tetracycline:** GFP-expression profiles of clones hh28.2 and hh28.7 illustrating that after 48 hours in tetracycline they were still expressing GFP 10-fold above background non-fluorescing cells. For instructions on how to interpret FACS plots see Appendix IV.

*Tighter regulation was achieved after a longer exposure to tetracycline*

The growth rate of hTERT-RPE1 cells is slower than HT1080 cells; therefore it is possibly they will demonstrate slower GFP kinetics. Consequently, the 33 GFP expressing clones were plated in the presence or absence of tetracycline for 5 days to examine whether the uninduced expression levels in these clones would reduce closer to background. The FACS results revealed that 8 clones (24%) had expression levels that were less than 5-fold greater than non-fluorescing cells (clone hh28). In addition, 9 (27%) of the clones had uninduced expression levels similar to that of background non-fluorescing cells (Figure 6.5B).

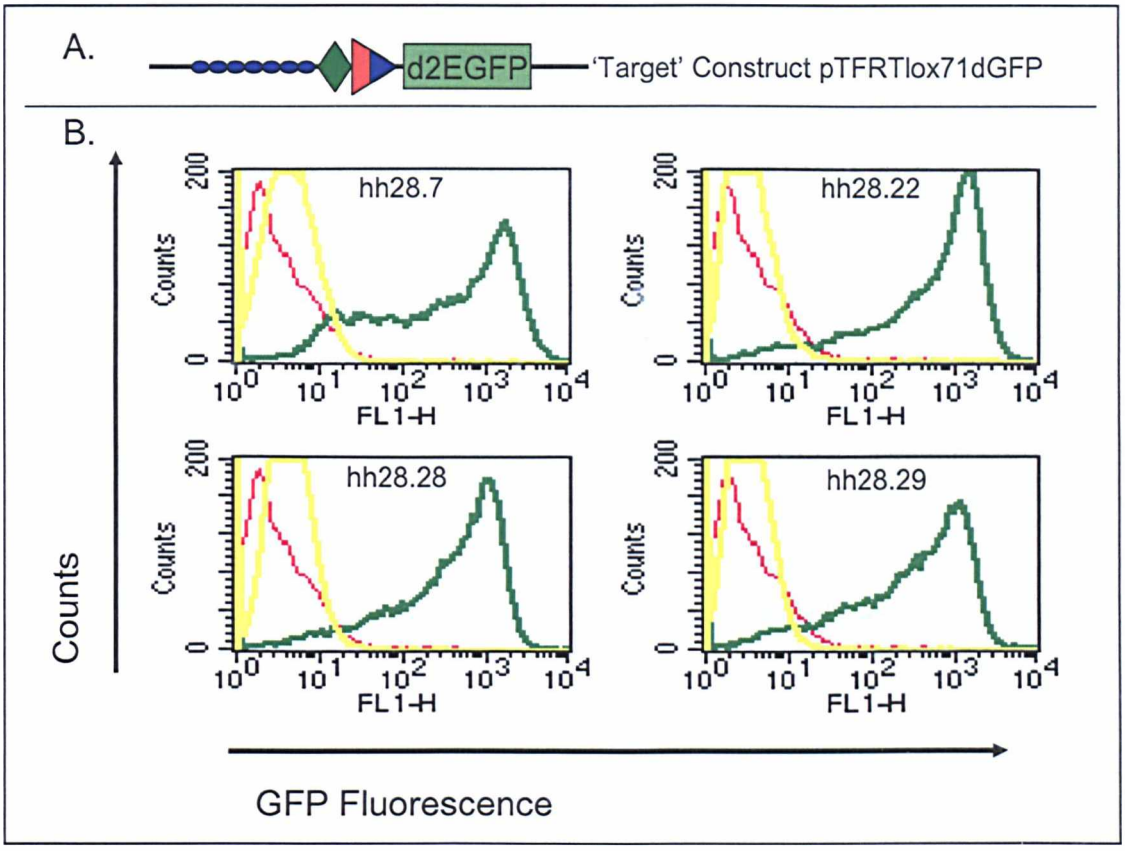



**Figure 6.5 GFP-expression profiles of four pTFRTlox71dGFP-transfected hh28 clones: A.** pTFRTlox71dGFP is illustrated (not to scale). For scale diagram see Appendix II. For key to symbols see Appendix I. **B.** GFP-expression profiles of four clones habouring the 'target' construct, pTFRTlox71dGFP. For instructions on how to interpret the FACS plots see Appendix IV.

The lowest uninduced expression levels, derived from clone hh28, were not quite as low as that seen for HT1080 cells. Clones hh28.22 and h28.29 demonstrated the lowest uninduced profiles (in yellow; Figure 6.5B), with each expressing ~1.5-2-fold greater than background non-fluorescing cells. These values were not quite as low as hoped, however the time was not available to repeat the in-depth search that was carried out for HT1080 cells. The clones depicted in Figure 6.5B were thought to be adequate for preliminary strategy testing. Induction values for the clones demonstrating the lowest uninduced expression ranged from ~250-fold (Clone hh28.28, Figure 6.5B) to ~500-fold (Clone hh28.22, Figure 6.5B). The maximum induction observed for hTERT-RPE1 cells was slightly lower than that seen for HT1080s, however this was due to their slightly elevated uninduced expression levels rather than reduced induction values.

It was also observed that the FACS profile of the GFP expressing clones varied between the two analyses (compare clone hh28.7 in Figure 6.4 and 6.5B). To determine whether this was due to cell confluency or passage number the hh28.7 cells were plated and grown to various densities and then analysed by FACS (Figure 6.6). The results clearly revealed that reduced cell-cell contact and active cell division in hTERT-RPE1 cells was required for uniform GFP expression from the TRE. The ploidy of each of the desirable hh28-GFP clones was analysed and they were all diploid, which confirmed the genetic stability of hTERT-derived cells. Clones hh28.7 and hh28.28 were used in future insertion experiments (Figure 6.5B).

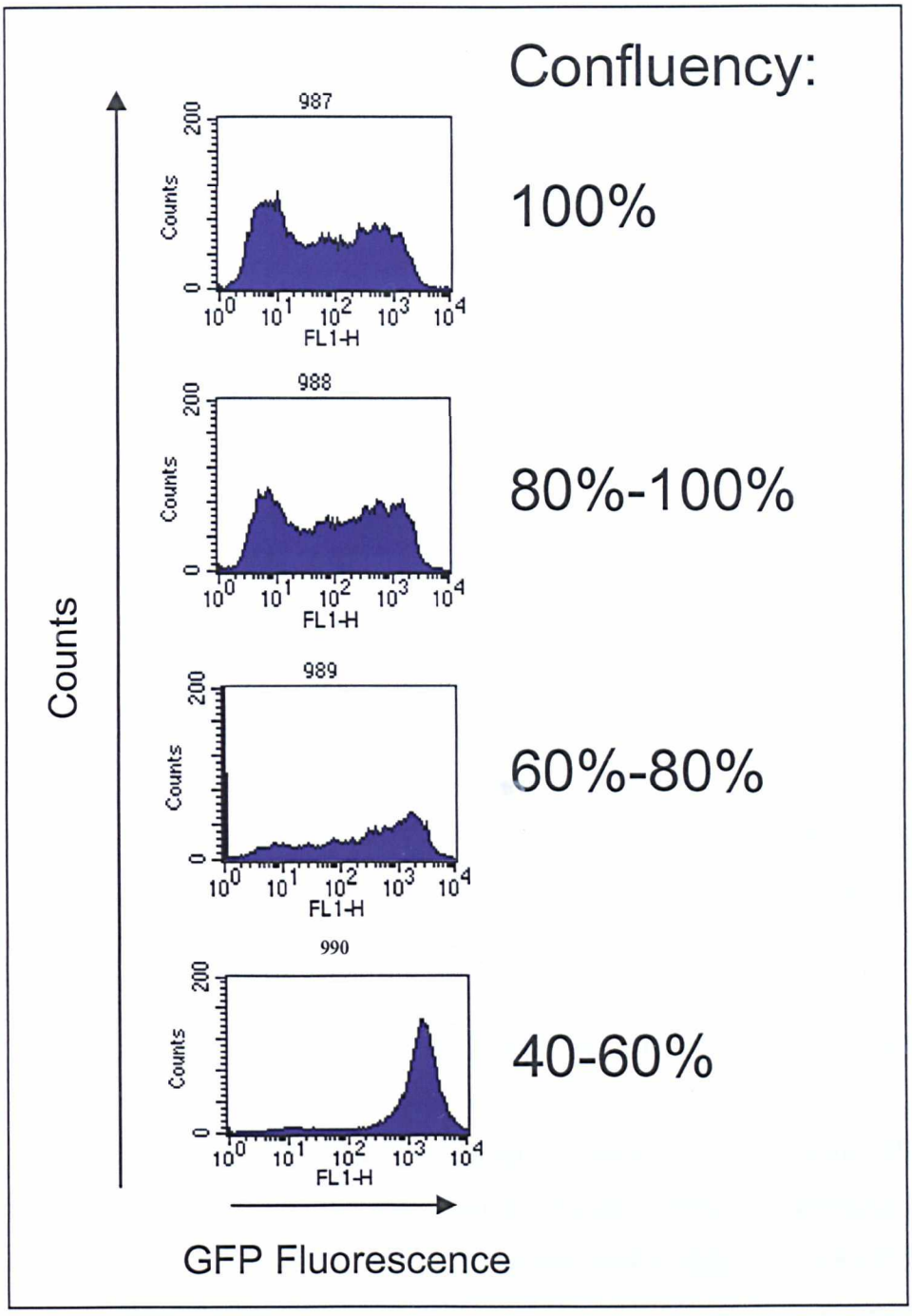



**Figure 6.6 The effects of cell confluency on expression from the TRE in hh28.7 cells:** hh28.7 cells were grown to various (shown) densities and then analysed by FACS.

## 6.2.4 Testing the *Cre*-mediated insertion step in hTERT-RPE cells

*FACS was used to determine the 'target' construct copy number in hh28-GFP clones*

As time was limited, an experiment to quickly distinguish those hh28-GFP clones that most likely contained a single copy of the 'target' construct was carried out. This technique involved transfecting hh28-GFP clones with plox66luc (See Appendix III; See section 4.2.1, Figure 4.4) and pMC-Cre (3μg of each) using Lipofectamine 2000 and analysing for an increase in the proportion of non-fluorescing cells by FACS, when compared to a control transfection (-pMC-Cre), after 72 hours. Clone hh28.7 demonstrated almost a 2-fold increase in GFP-negative cells after the experimental transfection when compared to the control transfection (data not shown). This increase in non-fluorescing cells was higher than what was observed for any of the other hh28-GFP clones. Southern analysis, a more accurate method to examine the copy number in these clones, was also carried out in parallel to confirm the above FACS results; however technical problems delayed the results. To avoid impeding the progress of this study, clone hh28.7 was used in the preliminary test experiment described below, until the problems with the Southern analysis had been resolved.

*A new 'insertion' construct was created to determine the recombination frequency*

To select for immortalised RPE cells expressing telomerase, a hygromycin ORF was used. This therefore meant that the previously generated 'insertion' constructs, which contained a hygromycin cassette (plox66hyg, plox66LucIresHyg; Figure 4.2B and C, respectively; Appendix III), could not be used. In order to test the efficiency and viability of the *Cre*-mediated insertion step, within hTERT-RPE1 cells, a new 'insertion' construct was designed. The new 'insertion' construct, plox66Neo (Figure 6.7), included a *lox66* oligonucleotide proceeded by a promoterless neomycin cassette and was generated using the cloning steps that are described in the materials and methods (See 2.3.3.1). Colonies that had undergone the insertion event after co-transfection with plox66Neo and pMC-Cre (3μg of each) into clone hh28.7 cells were

expected to have tetracycline regulated G418 resistance and hopefully (if a single copy of 'target' locus was present) loss of GFP expression (Figure 6.7).

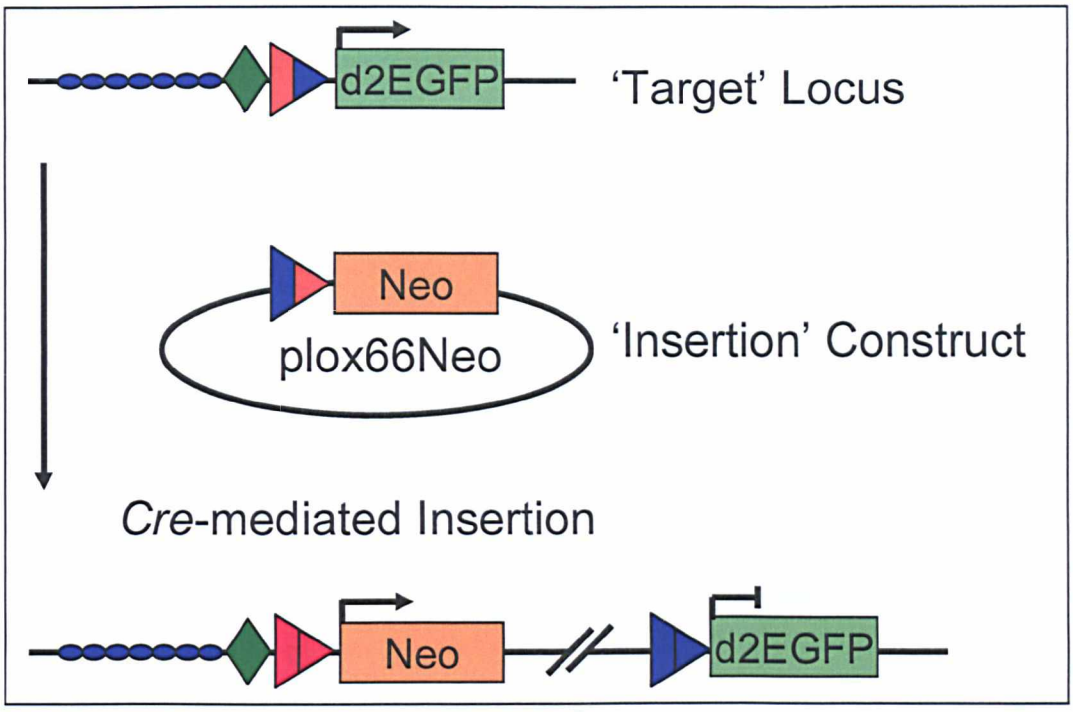



**Figure 6.7 The insertion event using G418 selection:** The expected *Cre*-mediated insertion of plox66Neo (not to scale) downstream of the TRE in hh28-GFP clones is illustrated. Pools of transfected cells were selected in G418 and the colonies generated were analysed for loss of GFP expression and inducible G418 resistance. For key to symbols see Appendix I.

*The recombination frequency for hh28.7 cells was higher than expected*

After co-transfecting 200,000 (40% confluent) clone hh28.7 cells with plox66Neo and pMC-Cre (3µg of each) using Lipofectamine 2000 (see section 2.8.5), 23 G418-resistant colonies were generated (called hh28.7Neo1 to 24 [-17]). Of these, only 9 (39%) were no longer expressing GFP (in yellow; Figure 6.8) and became G418-sensitive upon the addition of tetracycline to the medium. At this stage, it was presumed that the colonies still expressing GFP (in green; Figure 6.8) were random integrants of the plox66Neo construct nearby promoter elements; although this was not consistent with the control experiment (-pMC-Cre), which yielded no (random) G418-resistant colonies. This experiment suggested an absolute recombination frequency in hh28.7 cells of 0.0045%. This was lower than the 0.01% observed when

using plox66Hyg (Section 4.2.2), however more consistent with the experiments involving pINSLuc (0.0018-0.0044%; Section 5.2.1), both using HT1080 cells.

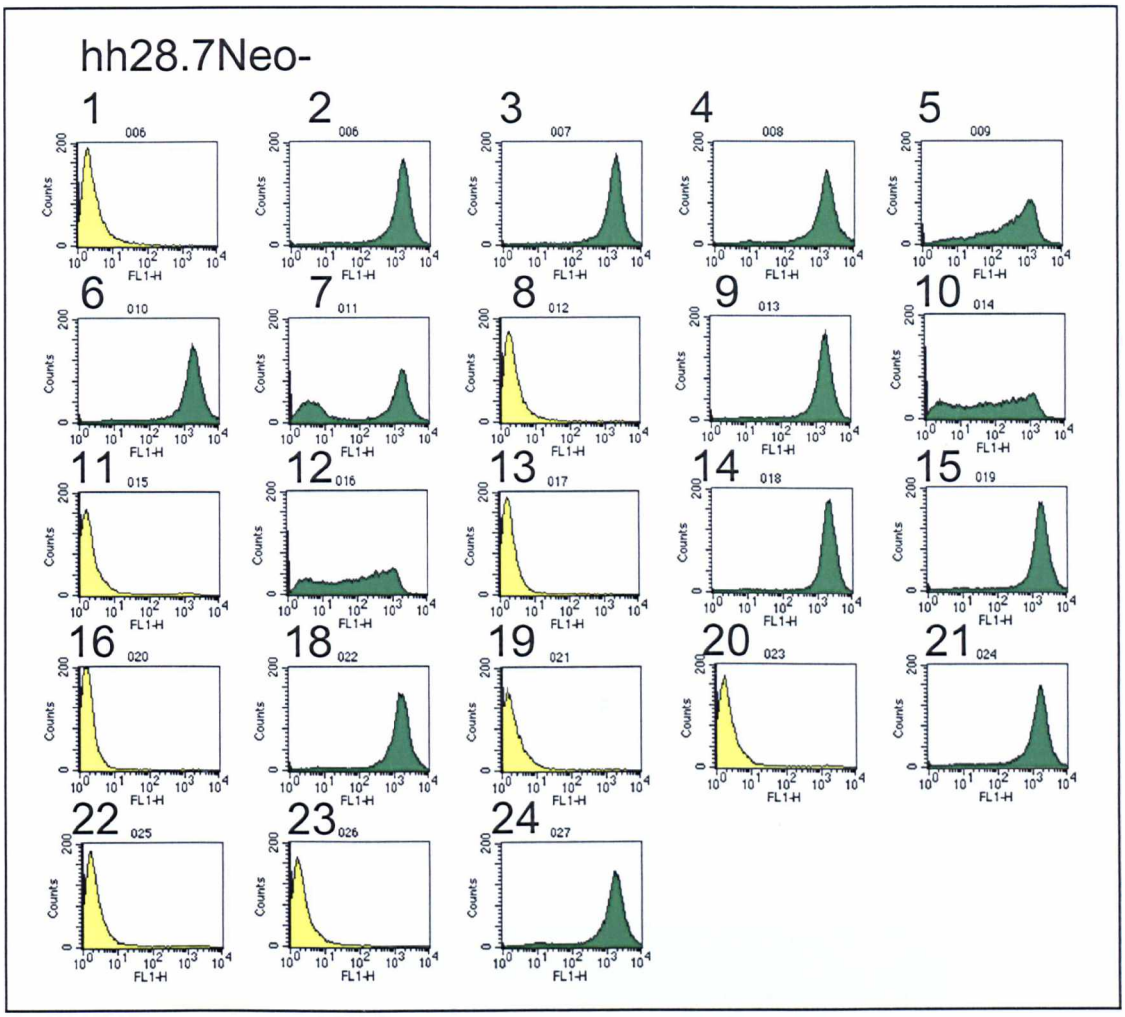



**Figure 6.8 FACS analyses for loss of GFP expression after the insertion event in hh28.7-derived clones:** GFP-expression profiles of 23 G418-resistant colonies generated from the insertion transfection is illustrated. It reveals that 9 clones (39%) were no longer expressing GFP (yellow), consistent with an insertion event. The rest of the colonies, however, were still expressing GFP (green) and were thought to have resulted from random integration of the 'insertion' construct. For instruction on how to interpret FACS plots see Appendix IV.

The number of G418-resistant colonies generated using clone hh28.7 is greater than expected because of their low efficiency of transient transfection when using Lipofectamine 2000. In addition, for optimal transfection conditions a reduced cell confluency is used, which leads to only 200,000 cells being transfected, when using a 6-well plate. Obviously, to generate a greater number of colonies a larger scale transfection (i.e. by using a 9cm plate) could be carried out.

*The expected genome rearrangements are demonstrated by PCR and Southern blot*

Additional confirmation that the desired insertion event has taken place in the GFP-negative clones (Figure 6.8) is demonstrated by PCR amplification of the recombination junction. Cell pellets were treated with pronase (See 2.4) and the primers, TetOF and NeoR (for sequence see Table 2.1), were used to amplify the recombination junction (Figure 6.9A). All five clones tested (hh28.7Neo1, 8, 11, 13 and 16) were positive for this region, as confirmed by the production of a 944bp band (Figure 6.9B). There was no positive control template available for the PCR and the negative control used was an hh28 cell pellet.

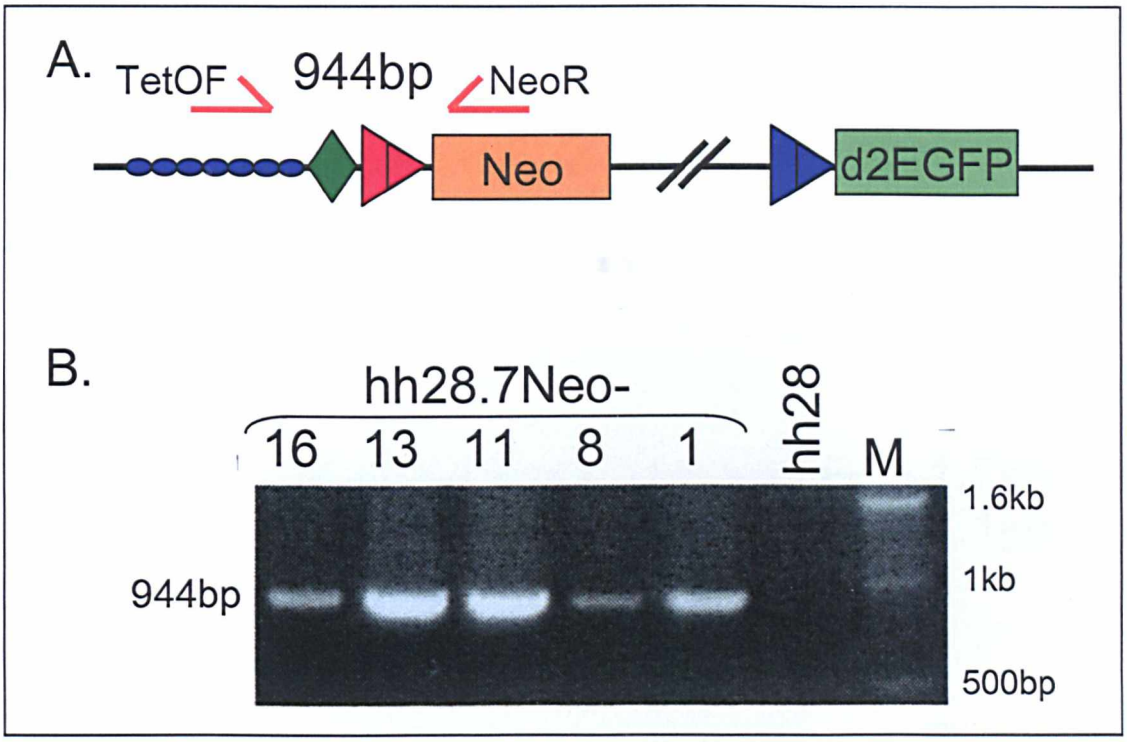



**Figure 6.9 PCR amplification of the recombination junction in hh28.7Neo-clones: A.** The region amplified by PCR (not to scale) to confirm that the recombination event has taken place, is illustrated. A PCR product of 944bp was expected when using primers, TetOF and NeoR (see Table 2.1). **B.** An ethidium bromide-stained agarose gel of PCR products for clones hh28.7Neo1, 8, 11, 13 and 16. A positive 944bp band was produced for each of them. There was no positive control template available and the negative control used was an hh28 cell pellet.

Further molecular evidence that the desired insertion event has taken place in all nine G418-resistant clones (hh28.7Neo1, 8, 11, 13, 16, 19, 20, 22 and 23) was demonstrated by Southern blot analysis, by examining for the expected genome

rearrangements. Genomic DNA from each clone was digested with *Dra*I and probed with a portion of the GFP gene (*Nco*I) (Figure 6.10A) (For methods see Section 2.5). Verification that the GFP-negative clones had undergone the insertion event was represented by the conversion of a 1.7kb band to a 2.1 kb fragment (Figure 6.10B).

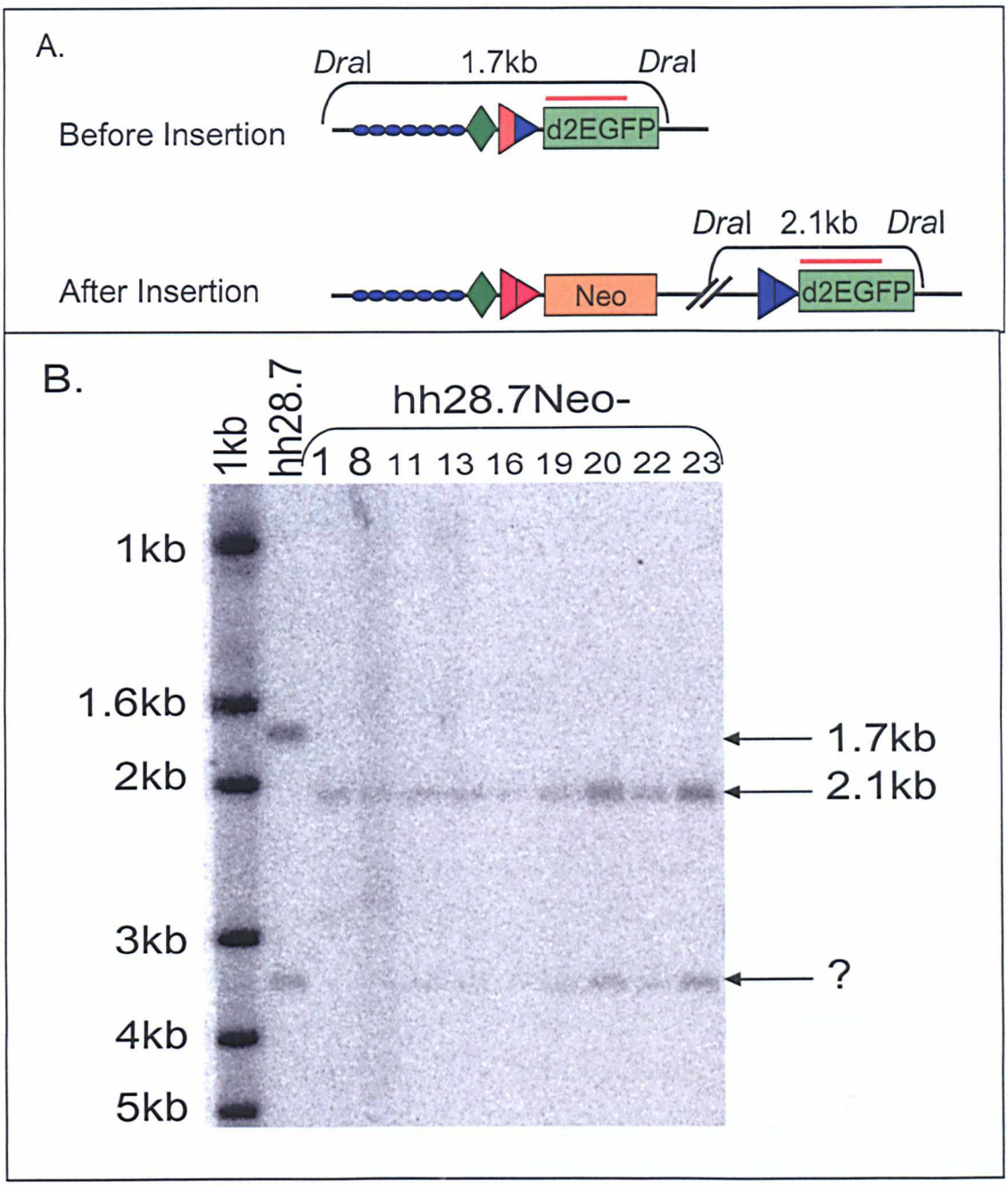



**Figure 6.10 Southern analyses of nine G418-resistant, GFP-negative clones derived from hh28.7:** **A.** A cartoon of the expected restriction pattern when cut with *Dra*I and probed for the GFP gene, before and after the insertion event (not to scale). **B.** The results of the Southern analysis, demonstrating the conversion of the 1.7kb band to a 2.1kb in the clones have undergone the insertion event.

An unexpected additional 3.4kb band was noticed on the Southern blot for all of the clones analysed. The most likely explanation is that clone hh28.7 has more than one copy of the 'target' construct. If this were the case, then the second copy must have been silent, as the G418-resistant clones were no longer expressing GFP (Figure 6.8). There are a number of problems with this hypothesis, however, and they are discussed later (see section 6.3.3).

### 6.2.5 Testing the improved 'screen and insert' strategy on hTERT cells

*Setting up the 'screen and insert' system to test the flp-mediated excision event*

To investigate the *flp*-mediated excision event in hTERT-RPE1 cells the test 'insertion' construct (pINSLuc), from the improved 'screen and insert' strategy (System Two), discussed in chapter five, was used (Figure 6.11). To carry out the *flp*-mediated excision event, the *Cre*-mediated insertion event must firstly be carried out and this involved the isolation of *gpt*-expressing hTERT-RPE1 cells.

*Gpt-expressing hTERT-RPE1 colonies were generated using a gpt-expression vector*

To determine whether it was possible to successfully isolate *gpt*-expressing hTERT-RPE1 colonies in MPA, xanthine and HAT supplemented medium, 8 million hh28.7 cells were electroporated with *Pvu*II linearised pBSgpt (Sullivan *et al.*, 2001). A control transfection was carried out which electroporated the same number of hTERT-RPE1 cells with *Pvu*II linearised empty bluescript vector. After selection for *gpt*-expression for 3 weeks, the number of colonies on the plates from the two transfections was compared and revealed that there were many colonies (100+) on the experimental transfection, compared to 3 colonies on the control plates. These results revealed that it is possible to select for *gpt*-expression in wild-type hTERT-RPE1 cells.

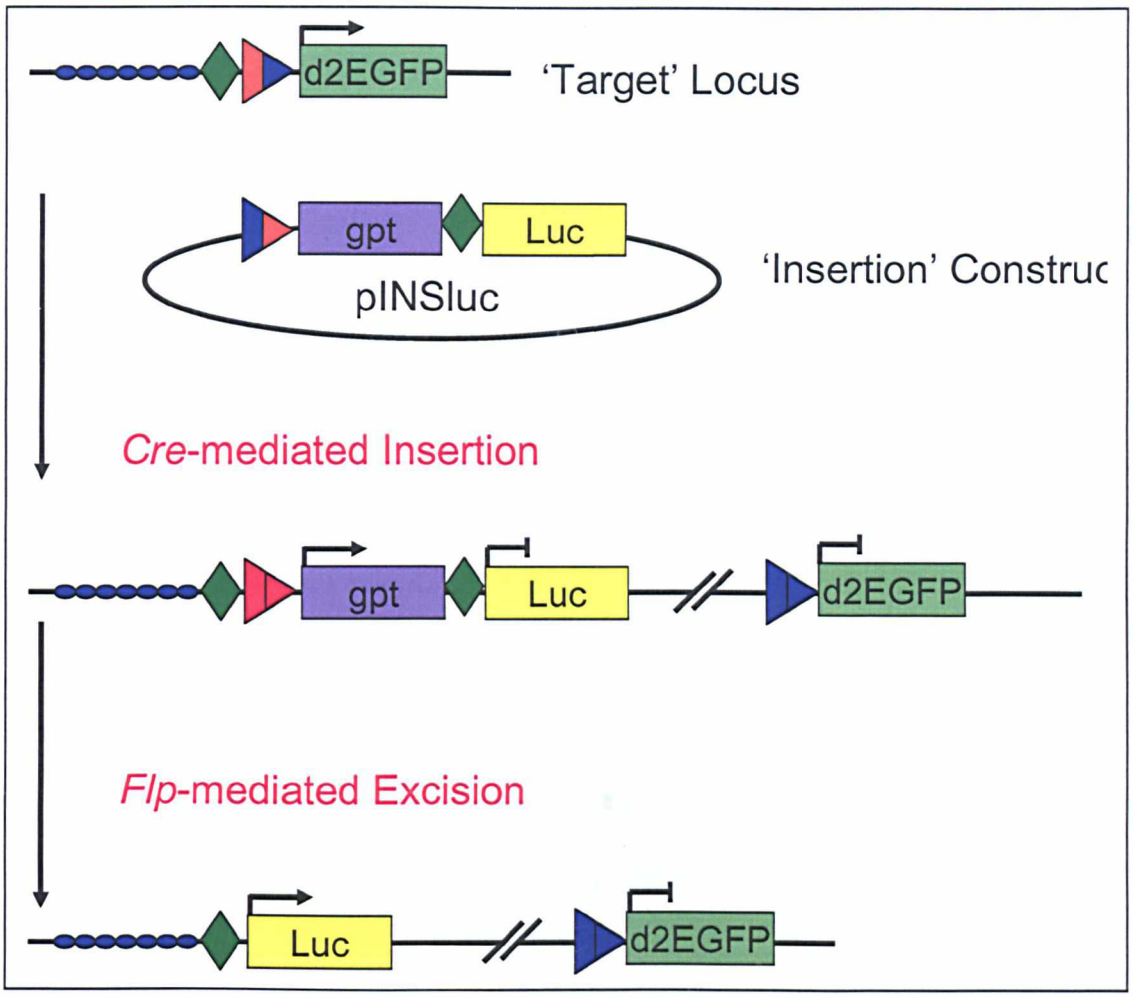



**Figure 6.11 Testing the improved 'screen and insert' strategy on hTERT-RPE1 cells:** The improved 'screen and insert' strategy (System Two), which is discussed for HT1080 cells in Chapter five, is illustrated (not to scale). For key to symbols see Appendix I.

*A single copy 'target' clone was identified by Southern blot analysis*

A single copy 'target' clone was eventually determined by Southern blot analysis. Genomic DNA from 'target' clones (Figure 6.5) were digested with *Dra*I and probed with a portion of the GFP ORF (*Nco*I) (For methods see section 2.5). A single copy number was demonstrated by the presence of a single GFP-positive band and this was demonstrated for clone hh28.29.

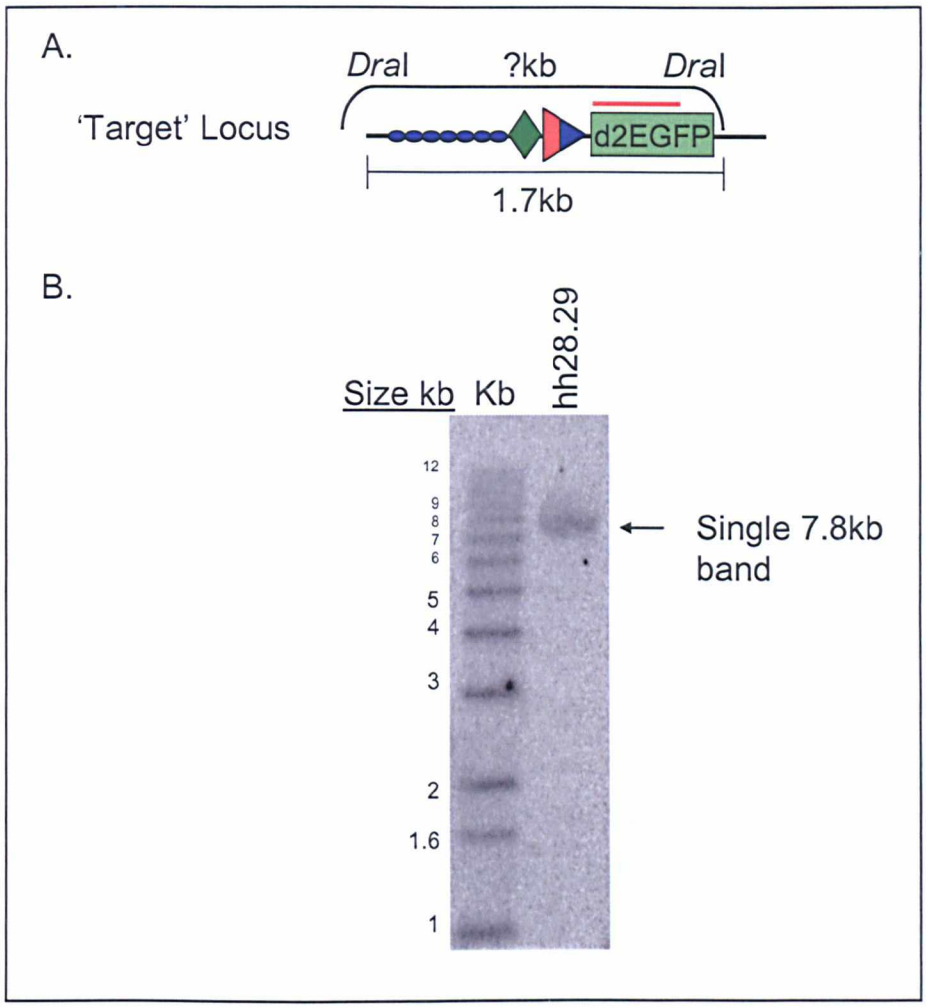



**Figure 6.12 Southern analyses demonstrating a single copy number of the 'target' locus in clone hh28.29: A.** A cartoon illustrating the restriction plan (not to scale) of the 'target' locus when cut with *Dra*I and the minimum size it was expected to be (1.7kb). **B.** Shows that a single band was found at 7.8kb, indicating that clone hh28.29 most likely contained a single copy of the 'target' locus.

*No gpt-expressing hh28.7/29 cells were generated after the transfection with pINSluc*

Both clones hh28.7 (known to have an accessible 'target' locus) and hh28.29 (single copy number) were chosen to set-up and test the improved 'screen and insert' strategy (Figure 6.11). After co-transfecting 200,000 hh28.7 and hh28.29 cells with pINSLuc and pMC-Cre (3µg of each) by Lipofectamine 2000 (for methods see section 2.8.5), no colonies were achieved for either experiment. The transfection was carried out a further two times and the same result was obtained.

## 6.3 Discussion

### 6.3.1 Testing the Cre-mediated insertion step in hTERT-RPE1 cells

The aim of this chapter was to establish and test the applicability of the developed 'screen and insert' strategies, in the non-transformed cell-line, hTERT-RPE1. These cells were derived from retinal pigment epithelial (RPE) cells which were immortalised by stably expressing the telomerase gene. Establishing the tetracycline inducible system in hTERT-RPE1 cells by stable integration of the itTA and 'target' constructs, was not expected to prove difficult as the cells had previously been successfully transfected stably with the telomerase ORF and selected in hygromycin. It was predicted, however, that switching to a new cell type may create two possible limitations, with regards to their suitability for the 'screen and insert' approach. Previous to this study, the ability to achieve tight gene regulation using the tetracycline system in hTERT-RPE1 cells was unclear and it is known that basal uninduced expression levels vary between cell types. The second limitation may arise whilst attempting to carry out the insertion/excision steps. A reduced capacity to transiently transfect hTERT-RPE1 cells could impair the ability to carry out the recombination steps efficiently. The *Cre*-mediated insertion event not only requires transient expression from pMC-Cre, but also the presence of the 'insertion' construction within the same nucleus for recombination to successfully take place. More importantly, the *flp*-mediated excision event requires a transfection efficiency that allows the desired colonies to be isolated under non-selectable conditions.

### 6.3.2 Successful generation of clones with inducible GFP expression

The generation of hTERT-RPE1 cells expressing itTA by electroporation was successfully achieved and clone hh28 was selected for transfection of the 'target' construct. GFP-positive clones were generated (33 in total), of which 27% had uninduced expression levels similar to that of background non-fluorescing cells. The

hh28-GFP clones did not down-regulate their TRE driven sequences as efficiently as HT1080 cells since the most desirable clone still expressed 1.5-fold greater than background (see clones hh28.22 and hh28.29; Figure 6.5). More tests, to confirm these results are required; to check that a greater tetracycline concentration or a longer exposure time is not needed to completely down-regulate these clones. In addition, analysis of a larger number of GFP-expressing clones may yield a more tightly-regulated one.

### 6.3.3 Successful isolation of recombined clones using G418 selection

*The Cre-mediated insertion frequency was slightly higher than predicted*

Whilst optimising the transfection conditions of hTERT-RPE1 cells, using Lipofectamine 2000, it became obvious that they were not as easy to transfect by this method as HT1080 cells. When using optimum transfection conditions, only 35% of hTERT-RPE1 cells expressed transient GFP (Figure 6.2), in contrast to 93% for HT1080 cells (Figure 4.3). Clone hh28.7 was chosen to test the insertion step. A new simple 'insertion' construct was created that was similar to plox66Hyg (Figure 4.2B) used for HT1080 cells, except it contained a promoterless G418 ORF, instead of hygromycin, downstream of the *lox66* site. Considering that the proportion of transiently transfected cells when using hTERT-RPE1 cells compared to HT1080 cells was approximately one third, it was predicted that 0.0033% (0.01%[insertion frequency using plox66hyg in HT1080s; see table 6.1]/3) of hh28.7 cells would undergo the insertion event. In other words, 6-7 colonies should be generated per 200,000 transfected cells; a frequency that should be high enough to successfully generate colonies for analysis.

23 G418-resistant colonies were actually generated from 200,000 transfected cells, of which 9 (39%) no longer expressed GFP. This result revealed that an absolute recombination frequency of 0.0045% was achieved; which was, as expected, lower than the results using a plox66hyg in HT1080 cells (0.01%) (Table 6.1). Although,

the absolute insertion frequency was slightly higher than the predicted 0.0033% mentioned above (Table 6.1).

**Table 6.1 A comparison of the absolute recombination frequency in HT1080 and hTERT-RPE1 cells:** Illustrates the difference in absolute recombination frequency between the two cell types.

| Cell Type | Absolute Recombination Frequency |
|---|---|
| HT1080 (plox66Hyg) | 0.01% |
| hTERT-RPE1 (plox66Neo) | 0.0045% |

*The clone characteristics of an insertion event were successfully demonstrated*

The hh28.7Neo clones were analysed for the expected characteristics of an insertion event. The five clones analysed by PCR successfully amplified the recombination junction and all of the GFP-negative clones analysed demonstrated the expected genome rearrangements.

*A second copy of the 'target' construct appeared to be present in clone hh28.7*

An unexpected additional 3.4kb band was noticed on the Southern blot. Due to the presence of this band in all of the clones it was firstly assumed that the band represented a second silent copy of the 'target' construct. This seemed plausible as it may have helped explain the reason why so many (60%) of the G418-resistant colonies were still expressing GFP (Figure 6.8), despite no random integrants being observed in the control experiment (-pMC-Cre). In other words, these clones could have been generated by insertion of the neomycin cassette into the second silent 'target' locus and expression from the active TRE would have remained unaffected. There is a problem with this hypothesis, however, as integration of plox66Neo into a silent copy of the 'target' locus will not generate G418-resistant clones as there should be no expression from the TRE, as depicted in Figure 6.13. In any situation where G418 resistance is generated, the cells should always express GFP (Figure 6.13).

### 6.3.4 Testing the improved 'screen and insert' approach in hTERT cells

An additional concern when using hTERT-RPE1 cells is that a reduced transfection frequency may hinder the *flp*-mediated excision step of the improved 'screen and insert' approach. This is due to its reliance on ample transient gene expression to generate a high frequency of non-selectable recombination. If the *flp*-mediated excision frequency in HT1080 cells is 14.6% (see section 5.3.3), it could be predicted (as described earlier; see section 6.3.3) that whilst using hTERT-RPE1 cells it will be almost 5% (14.6%/3). Again, if this is the case, then this frequency will be high enough to isolate the desired colonies by the simple PCR screen.

To test the excision step of the improved 'screen and insert' approach in hTERT-RPE1 cells, the 'insertion' construct pINSLuc was used. A problem arose, however, in the initial stages of this strategy by not being able to isolate *gpt*-expressing colonies after the insertion transfection was carried out. It is possible that the frequency of recombination, using pINSluc (Figure 6.11). is too low. The frequency of pINSluc insertion was shown to be reduced in HT1080 cells (by almost 10-fold), when compared to previous experiments involving different 'insertion' constructs (see Table 6.2). As few as 0.0018% of HT1080 cells (for clone Rht14-10) underwent the insertion event when transfected with pINSLuc (Table 6.2). It could therefore be predicted that 0.0006% hTERT-RPE1 cells (due to their 1/3 reduced transfection efficiency) may undergo the insertion event. This would suggest that only 1 cell will undertake the desired event after 200,000 hTERT-RPE1 cells are transfected, a proportion that will likely be lost due to experimental error. A larger scale transfection (e.g. in 9cm plates instead of 6-well) to increase the number of gpt-expressing colonies generated, could be carried out. Incidentally, the reduced capacity to generate recombined clones whilst using pINSluc for both HT1080 and hTERT-RPE1 cells also coincided with additional problems when using the *gpt*-gene. These results and the consequences are discussed in Chapter Seven (See section 7.2.2).

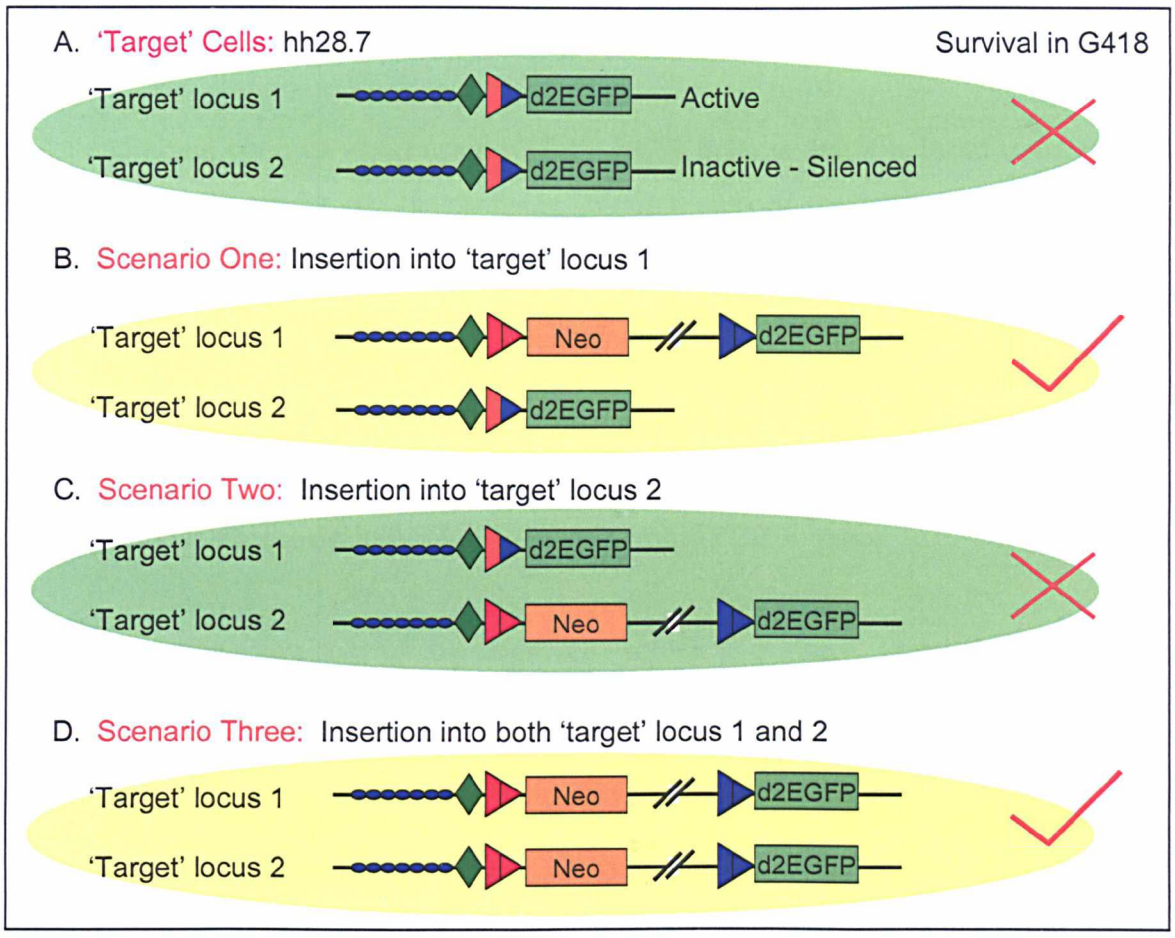



**Figure 6.13 A possible explanation for the additional 'target' locus in hh28.7-derived cells: A.** Two hypothetical 'target' loci, one which is active and the other silent (not to scale) is illustrated. GFP would only be expressed from the active locus. **B.** Scenario 1, is when integration takes place at the active locus. These cells will become G418 resistant and GFP negative. **C.** Scenario 2, is when integration takes place at the silent locus. These cells will become G418-sensitive and GFP-positive. **D.** Scenario 3, is when integration takes place at both the active and the silent 'target' locus. These cells will become G418 resistant and GFP negative.

Alternatively, the second 3.4kb band could just be background signal on the Southern blot, as it was seen in each of the lanes. This would not, however, help explain the reason for the G418-resistant, GFP-positive clones. To help understand this result fully, further analysis on the GFP-positive, G418-resistant clones should be carried out. They will be tested to determine whether their G418 resistance is dependable on tetracycline, whether PCR can amplify the recombination junction and examine what genomic rearrangements have taken place by Southern analysis.

**Table 6.2 A summary of the transfection efficiencies of different 'insertion' constructs in the two cell types:** Illustrates a reduced transfection efficency when pINSluc is used when compared to plox66Hyg/Neo.

| **HT1080 Cells** | **Absolute frequency of *Cre*-mediated Insertion** | **Frequency of *Flp*-mediated Excision** |
|---|---|---|
| Plox66Hyg | 0.01% | - |
| pINSLuc | 0.0018-0.0044% | 14.6% |
| **hTERT-RPE1 Cells** | **Absolute frequency of *Cre*-mediated Insertion** | **Frequency of *Flp*-mediated Excision** |
| Plox66Neo | 0.0045% | - |
| pINSLuc | 0% | ? |

## 6.3.5 Establishing the 'screen and insert' approach in other cell types

Ultimately, establishing the 'screen and insert' system in murine ES cells, to generate transgenic mice expressing tightly regulated GOI, will be desirable. Particularly when uncontrolled overexpression or down-regulation of that GOI results in embryonic or neonatal lethality; as gene function studies at latter stages of development cannot normally be analysed. The system described in this study will therefore be particularly attractive for rescuing mice from embryonic lethality after gene knock-out, by expressing the GOI from the TRE during development. In the healthy (adult) mouse, expression from the TRE can be completely suppressed to generate a null phenotype and the effects at this latter stage of development can be analysed. Alternatively, the system described here could be utilised in transgenic mice when the expression of a GOI generates a background phenotype (e.g. cell death [diphtheria toxin, see section 1.3.2] or *Cre*-recombinase, resulting in uncontrolled excision events), by allowing the complete restriction of a GOIs expression to a particular time point.

# Chapter Seven

# Work in progress - Utilising the 'screen and insert' approach

## 7.1 Introduction

After the successful development of a cell line expressing tight tetracycline regulated luciferase, using the 'screen and insert' approach, it could then be utilised to study the effects of various GOIs that may affect genomic stability and gene targeting frequencies (research interests of the Gene Targeting Group). Candidate genes that were predicted to require tight inducible gene regulation for adequate functional analysis included toxic genes (e.g. *RAD52;* Yanez and Porter, 2002), essential genes (e.g. *RAD51*; Yanez and Porter, 1999a; *RAD50*; Luo *et al.*, 1999; *MreII*; Xiao and Weaver, 1997), enzymes (e.g. *I-SceI, Cre*; Schonig *et al.*, 2002) and shRNA (Matthess *et al.*, 2005). Due to time constraints, the results in this chapter involve conditional gene overexpression studies, rather than conditional gene silencing, by describing the successful generation of HT1080 clones containing tightly regulated, never switched on before, I-*SceI* and *RAD52* genes. In addition, preliminary data involving the hTERT-RPE1 cell line will also be described.

I-*SceI* is a rare-cutting endonuclease from the yeast *S. cerevisiae* which recognises and cuts a unique 18bp recognition site. It is commonly used for gene function studies to generate a double-strand break at a pre-integrated site to encourage gene targeting by up to 5000-fold (Cohen-Tannoudji *et al.*, 1998; Donoho *et al.*, 1998). This is made possible due to the lack of I-*SceI* recognition sites present in the mammalian genome, although it has been suggested that I-*SceI* can cut sites that deviate slightly from the consensus sequence (Colleaux *et al.*, 1988). A blast search of the human genomic sequence database (using www.ensemble.org) did confirm the absence of the 18bp recognition sequence, however, it did reveal a number of sites that contained between 15-17 consecutive base pairs of the I-*SceI* recognition site

(Figure 7.1). It is therefore plausible that sites such as these may act as cryptic sites within the genome and as I-*SceI* is commonly used for gene repair and has been implicated as having potential therapeutic applications (de Piedoue et al., 2005), cryptic sites could potentially lead to some undesirable mutagenic consequences.

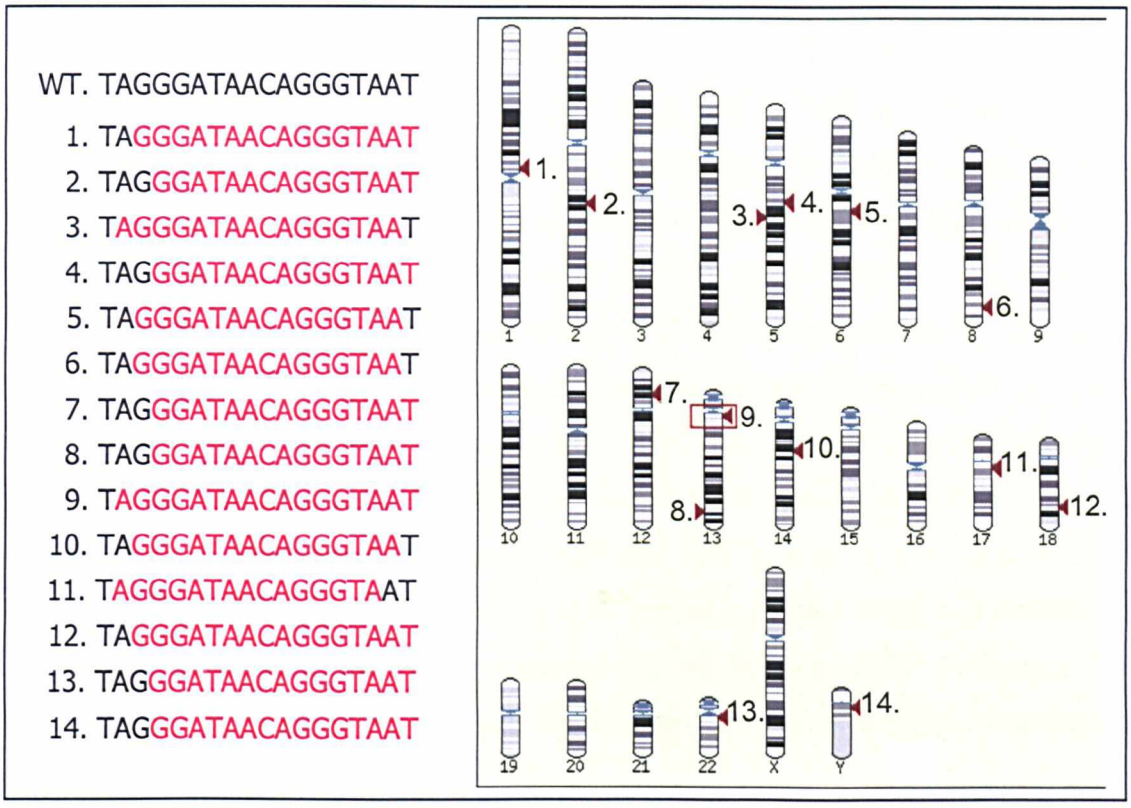



**Figure 7.1 Potential cryptic I-*SceI* sites in the human genome:** The wild-type (WT) 18bp I-*SceI* site is illustrated at the top left of the figure. Below it are 14 possible cryptic sites, highlighting in red are the base-pairs which are present in the wild-type site. The position of each of the potential cryptic I-*SceI* sites on the human chromosomes is also shown (right). Site 9 (chromosome 13) demonstrates the greatest amount of homology, as it contains 17 consecutive base-pairs from the 18bp I-*SceI* site.

Transient I-*SceI* expression could be a feasible method of testing the genome for cyptic I-*SceI* sites, however, the proportion of cells that receive the expression vector would be less then 100% and the timing and amount of I-*SceI* exposure could not be controlled. If it were found that stable I-*SceI* expression did in fact compromise mammalian cell survival, it would be desirable to have a cell line expressing inducible I-*SceI* so that the effects of the enzymes can be controlled. It is also predicted that tight regulation of the I-*SceI* gene is required, as leaky expression could potentially lead to uncontrolled cutting at pre-integrated, or cryptic sites, and their inaccurate repair during the selection or screening process for stable clones. It is also expected

that proving a clone to be tightly regulated with the I-*SceI* transgene will be a technical challenge, due to the insensitive nature of traditional methods (see Section 3.2.4). The 'screen and insert' approach, is therefore an ideal system for this type of analysis as leaky expression will be avoided and the timing of the cells first exposure to the enzyme can be strictly controlled.

Constitutive overexpression of the human *RAD52* gene in human somatic cells has previously been shown to compromise cell viability and its expression was lost over several passages. In addition, it was not possible to achieve large amounts of *RAD52* overexpression and a proportion of the cells within a clone failed to overexpress the RAD52 protein at all (Yanez and Porter, 2002). Overexpression of the yeast homologue of *RAD52*, however, has been shown to be tolerated in human cells and can stimulate gene targeting frequencies by up to 37-fold in HeLa cells (Primio *et al.*, 2005). This is unlike the overexpression phenotype of human *RAD52* which was shown to inhibit gene targeting (Yanez and Porter 2002). It is unclear whether these differences in homologue activities are due to the yeast protein being less harmful to human cells and resulting in higher expression levels being achieved or whether the yeast protein is more effective than the human form at promoting homologous recombination.

The 'screen and insert' system developed in this study will be ideal for re-investigating the effects of human *RAD52* overexpression on homologous recombination and genomic stability in human somatic cells. This approach permits the integration of the transgene into every cell of a clone under restrictive conditions, without posing any detrimental effects. It will also facilitate the forced up-regulation of the transgene to levels not previously achieved for analysis, generating results that will be more comparable to those obtained when using the yeast *RAD52* gene. In addition, the yeast *RAD52* gene could also be overexpressed in parallel in the same tightly regulated clone to provide a valuable and more accurate comparison for analysing the functional differences between the two homologues.

## 7.2 Results

### 7.2.1 Analysing the effects of I-*SceI* expression in mammalian cells

*Transient I-SceI activity was successfully demonstrated in HT1080 cells*

To examine the activity and the effects of I-*SceI* expression in HT1080 cells pTIGHT-ISCEI was constructed (Figure 7.1A), as described in the materials and methods (section 2.3.4.1). This construct contains a TREmod (from pTRE-TIGHT) followed by the I-*SceI* ORF (from pCMV-I-SceI; Rouet *et al.*, 1994). To demonstrate the activity of the expressed I-*SceI* protein in HT1080 cells the construct was transiently transfected (4μg) by lipofection into htTA expressing cells (Rht14-; section 3.2.2) which stably harboured the pDRNeo construct (Liang *et al.*, 1996; Table 2.5 for methods of clone generation). This construct contains a functional hygromycin cassette flanked by two disrupted neomycin genes (Figure 7.2A). The upstream neomycin ORF is interrupted by the integration of an I-*SceI* site and the downstream neomycin genes has some of its 5' sequence removed. When I-*SceI* is expressed in these cells (Rht14-DRNeo) it cuts at the I-*SceI* site to generate a double strand break and it is predicted that the second neomycin ORF will be used as a template for its repair (DNA repair by intrachromosomal recombination [Liang et al., 1996]). In doing so, this produces a functional neomycin gene which can be selected for in G418. If the I-*SceI* ORF is active in HT1080 cells then it was expected that there should be an increase in G418 resistant colonies after 14 days selection when pTRE-ISCEI was used, compared to a control (pTIGHT-Luc) (Figure 7.2A). This was found to be the case and the crystal violet stained colonies are shown in Figure 7.2B (for staining methods see 2.8.6).

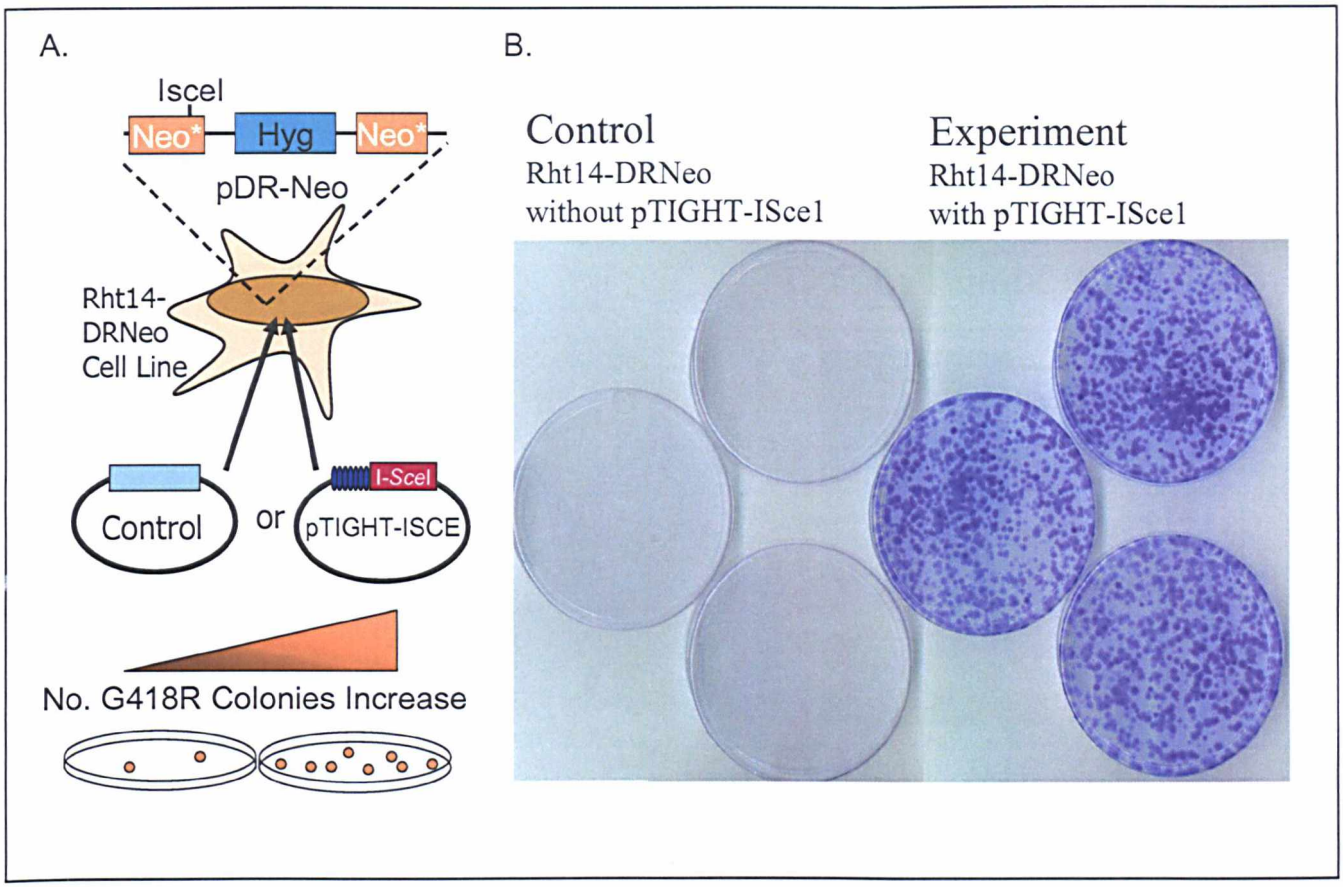



**Figure 7.2 Transient expression of pTIGHT-ISCE1 in Rht14-DRNeo cells: A.** Illustrates the experimental proceedure carried out and the expected results. An increased in the number of intrachromosomal recombination events between the two neomycin genes is expected, resulting in G418 resistant colonies, when pTIGHT-ISceI is used compared to a control construct (pTIGHT-Luc). **B.** Shows the crystal violet stained plates obtained.

*Stable induced expression of I-SceI may compromise HT1080 cell survival*

To investigate the effects of stable expression of I-*SceI* in HT1080 cells and determine whether it was tolerated, 8 million Rht14- cells were co-electroporated with pTIGHT-ISCE and pBL-Puro/R (for transfection conditions see Table 2.5). The transfection was split equally into medium supplemented with or without tetracycline (Figure 7.3A) and after 14 days in puromycin selection the resulting colonies were counted and compared. The results demonstrated a 24% increase in colony numbers when the I-*SceI* enzyme was suppressed during the selection process (Figure 7.3B), therefore suggesting a possible toxicity of the I-*SceI* transgene in HT1080 cells. A similar

experiment was carried out twice in hTERT-RPE1 cells (hh28 cells), however, no stable I-*SceI* expressing colonies were generated (data not shown).

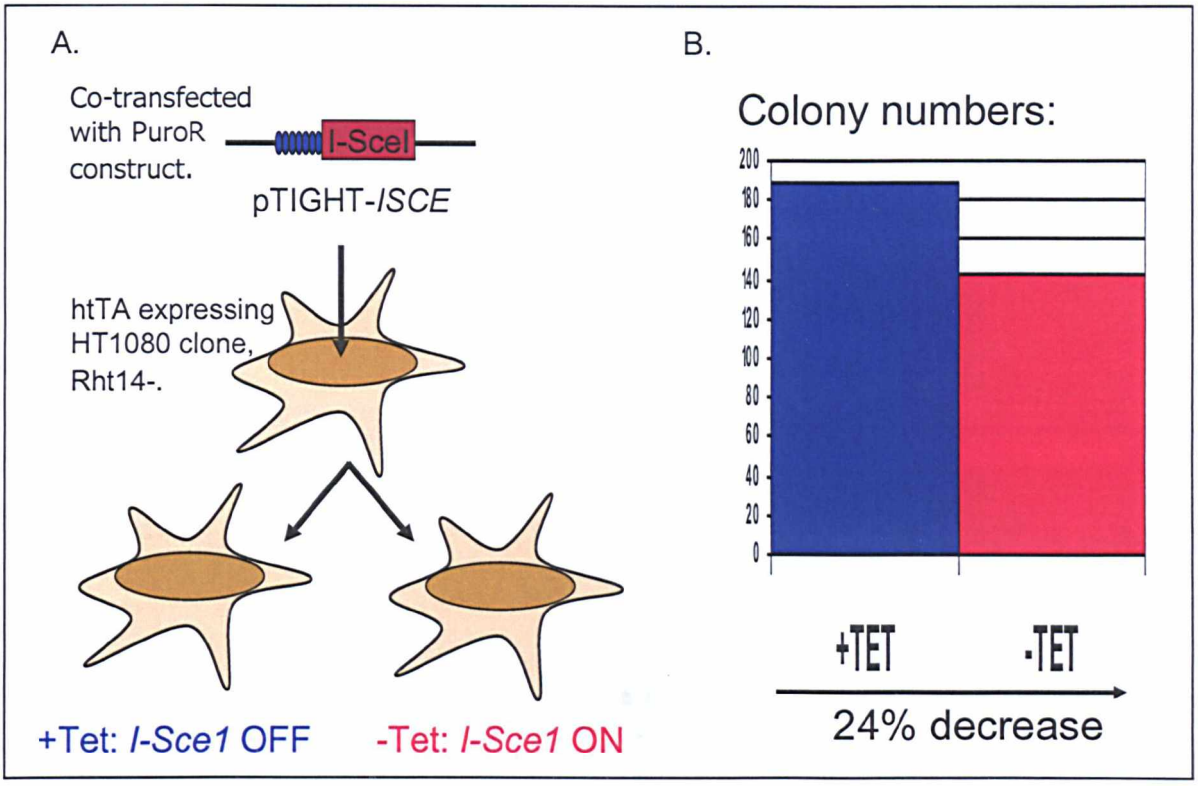



**Figure 7.3 Analysis of stable I-*SceI* expression in HT1080-derived Rht14 cells: A.** A cartoon illustrating the methods used to analyse the effects of stable I-*SceI* expression in HT1080 cells. 8 million Rht14- cells were co-electroporated with pTIGHT-ISCEI using the methods described in Table 2.5. The transfection was split equally into medium supplemented with or without tetracycline. After 14 days in puromycin selection colonies were counted. **B.** The graphs shows the number of colonies obtained for each culture condition (with or without tetracycline).

It was expected that not every puromycin-resistant colony would contain the TRE-I-*SceI* cassette, as they were generated by co-transfection. Consequently, 24 colonies from each experiment (with [+Tet] and without tetracycline [-Tet]) were picked and cell pellets (for methods see 2.4.2) were analysed by PCR for the TRE-I-*SceI* region (Figure 7.4A), using the primers TetOF and RevTIGHT1 (for sequence see Table 2.1). It was initially noticed that following picking, +Tet clones had an increased survival and growth rate when compared to some of the –Tet clones. In fact, the 24 -Tet clones picked, soon reduced to 16 prior to PCR analysis due to a failure of some clones to propagate. The PCR results revealed that when tetracycline was continuously present (+Tet), almost every colony (except clone 12) contained the

transgene (Figure 7.4B). The -Tet clones, however, generated fewer positive colonies (6/16; Clone 6, 7, 8, 12, 14 and 24; Figure 7.4C) and as well as this, a secondary unexplained band was observed (~2kb). Again, the PCR data (Figure 7.4C) together with the reduction in colony numbers (Figure 7.3B) and decreased cell propagation may suggest that stable I-*SceI* expression is not favoured by HT1080 cells.

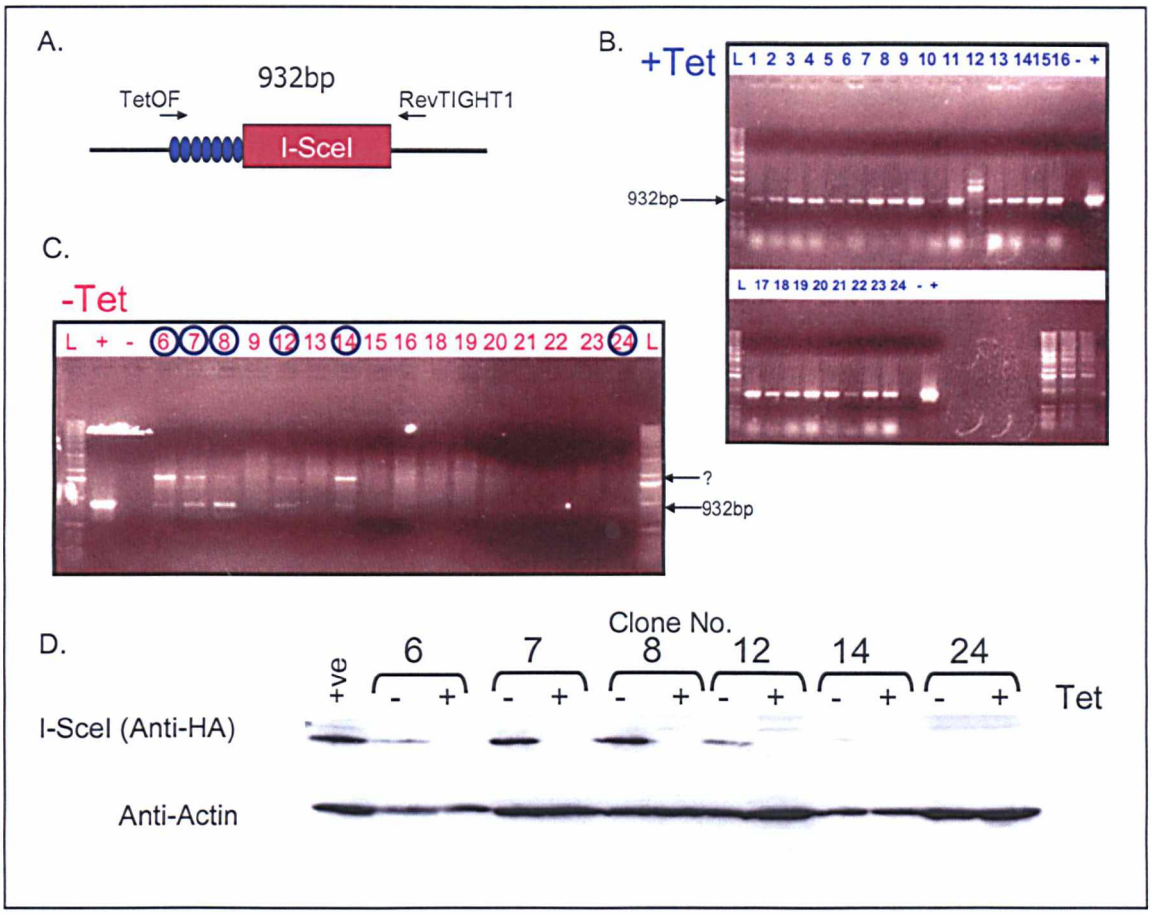



**Figure 7.4 PCR analysis of Rht14 clones stably expressing I-*SceI*: A.** The position of the primers, TetOF and RevTIGHT1, are shown and the expected PCR product size (not to scale). **B.** An ethidium bromide-stained agarose gel of the PCR products from clones continuously exposed to tetracycline. **C.** An ethidium bromide-stained agarose gel of the PCR products generated from clones cultivated without tetracycline. **D.** Western blot analysing for inducible I-*SceI* expression. Six Rht14- clones containing the pTIGHT-ISCEI transgene were grown with or without tetracycline and protein extracts were obtained. SDS-PAGE and immunodetection were performed on these extracts, along with a +ve control (Rht14- cell transiently transfected with pTIGHT-ISCEI). The blot was probed with a primary rat-haemagglutinin antibody (Statagene, 3F10) and detected with a secondary peroxidase-conjugated goat-anti-rat antibody.

Unfortunately, the +Tet clones were later lost (due to contamination) and could not be analysed further. The preferred next step would have been to analyse the healthy +Tet clones by Western blot for inducible I-*SceI* expression; after which the transgene in the desirable clone would be induced and the effects on cell growth could be analysed. Whilst plans to generate fresh +Tet clones were being considered, the six – Tet clones were analysed for inducible I-*SceI* expression. The six clones were treated with or without tetracycline for 72 hours, after which protein samples were produced and a Western blot was carried out (for methods see section 2.6). An antibody against the hemagglutinin (HA) epitope tag, which was fused to the I-*SceI* protein (Rouet *et al.*, 1994), was used (see Table 2.4). The blot revealed that only four of the six clones analysed contained detectable I-*SceI* protein (~33kDa), all of which were inducible (Figure 7.4D).

An additional problem emerged when attempting to determine how tightly regulated the clones studied on the Western blot were. They all appeared to turn off completely; however, this was thought to be unlikely considering the results obtained for clones demonstrating inducible GFP expression in Chapter Three. Consequently, an experiment was set up which was based upon a similar technique carried out by Rouet and colleagues (1994) to help determine whether these clones were tightly regulated. The experiment involved transiently transfecting (for transfection conditions see Table 2.6) the four clones demonstrating inducible I-*SceI* (Clones 6, 7, 8 and 12; Figure 7.4D) with a construct containing an I-*SceI* site (pPurfloxZeo [Figure 7.5]; Yun *et al.*, 2004). After 8 hours, the DNA was extracted from the cells and Southern analysis was carried out by probing for the break region (Figure 7.5). The construct was linearised with *Spe*I either prior to transfection and linearised with *SpeI* along with the genomic DNA after it was extracted from the cells. Both approaches resulted in a blot with high background and even after highly stringent washes, no conclusive results were ever achieved (data not shown).

In hindsight, an additional method could have subsequently been adopted for screening for tight regulation in the four inducible clones, which involved attempting to stably express pDRNeo (discussed above and illustrated in Figure 7.2A) in the presence of tetracycline. The frequency of formation of G418 resistant colonies would have given a representation of the amount of leaky expression. To analyse for

cryptic I-*SceI* sites, however, it was important that the cells had not previously been exposed to the enzyme, due to the possibility that they had already been cut and inaccurately repaired. This was not possible when using cells that had already been cultivated in the absence of tetracycline and the levels of uninduced expression remained questionable. Along with a difficulty proving tight transgene-regulation, these preliminary results have also revealed that uncontrolled I-*SceI* expression may be affecting the generation of healthy colonies and their growth patterns. It was therefore concluded that I-*SceI* was a suitable candidate gene for use with the 'screen and insert' approach.

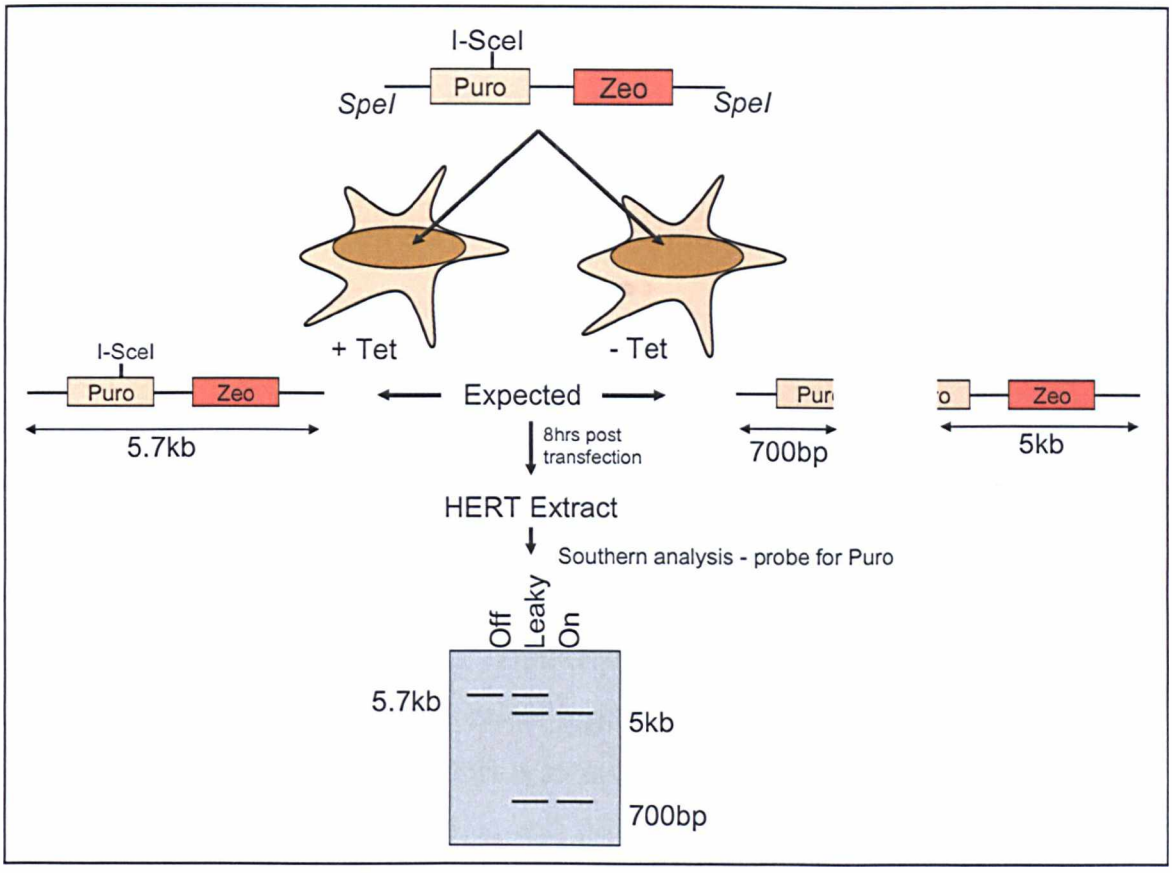


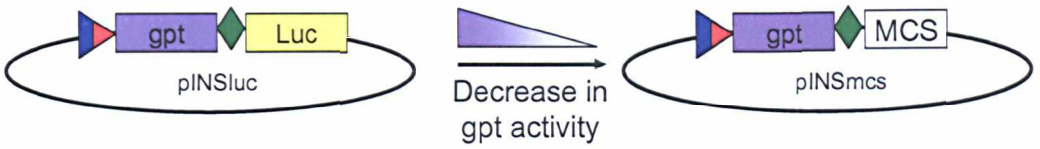

**Figure 7.5 Examining the I-*SceI* expressing clones for tight regulation:** A schematic representation of a method used to analyse the four inducible I-*SceI* expressing clones (Clones 6, 7, 8, 12) for tight uninduced expression. A construct containing an I-*SceI* site (pPurfloxZeo) was transiently transfected into the four clones in either the presence or the absence of tetracycline. After 8 hours, DNA was extracted from the cells (HERT extract) and digested with *SpeI*. A Southern analysis was carried out which probed for the puromycin ORF. It was assumed that expression of I-*SceI* (-Tet) would cut at the I-*SceI* site and generate two distinct bands on the gel. Alternatively, uninduced I-*SceI* (+Tet) would not cut at the site and only one band would be produced. Leaky expression, however, may result in a combination of the two.

### 7.2.2 A difficulty isolating *gpt*-expressing colonies when using pINSmcs

Prior to establishing cell lines expressing inducible I-*SceI* and *RAD52,* the modified insertion construct, pINSmcs, was tested to confirm that the previous cloning steps (insertion of the MCS; see 5.2.3) had not interfered with the vectors ability to undergo the insertion event and generate functional gpt protein. During preliminary tests, however, it soon became apparent that the *gpt* gene was no longer active for the successful isolation of *gpt* expressing colonies. Both target clones Rht14-10 and Rht14-19 were tested and the upstream *lox66* site was confirmed to be correct by sequence analysis. A sequence comparison between what was considered to be a functional *gpt* gene (from pINSluc; Figure 7.6A) and the non-functional version (from pINSmcs; Figure 7.6A) did not demonstrate any differences in the *gpt* ORF. Inconsistencies were observed, however, between the plasmid sequence data and the published *gpt* sequence (from NBCI [M12907]) and these differences are illustrated in Figure 7.6B.

If correct, the sequence differences are almost certainly significant as it not only results in the deletion of eight base pairs, it also truncates the *gpt* ORF by 150bp (or 50 amino acids), due to the formation of a cryptic stop codon (Figure 7.6B). The sequencing reaction was repeated numerous times (six), each time using both a forward and a reverse primer (loxgptF and loxgptR; see Table 2.1) and the same results were generated each time. Furthermore, when pBSgpt (the source of the *gpt* ORF) was examined, which expresses highly functional *gpt*, the published version was obtained (Figure 7.6B). If this is an accurate representation of the *gpt* sequence in the 'insertion' vectors, pINSluc and pINSmcs, and not a consequence of poor sequencing technique, it may suggest an instability of the *gpt* DNA during the cloning procedures. In addition, it may help explain why such a poor *gpt* activity was observed for pINSluc (Section 5.2.1) and the lack of activity for pINSmcs. As a consequence, the *gpt* ORF was replaced with a neomycin cassette to generate pINSmcsNeo (Figure 7.6C), by following the cloning procedures described in the materials and methods (Section 2.3.3.2).

A.




B.

Published *gpt* ORF (NBCI sequence M12907) and pBSgpt:

```
ATGAGCGAAA AATACATCGT CACCTGGGAC ATGTTGCAGA TCCATGCACG
TAAACTCGCA AGCCGACTGA TGCCTTCTGA ACAATGGAAA GGCATTATTG
CCGTAAGCCG TGGCGGTCTG GTACCGGGTG CGTTACTGGC GCGTGAACTG
GGTATTCGTC ATGTCGATAC CGTTTGTATT TCCAGCTACG ATCACGACAA
CCAGCGCGAG CTTAAAGTGC TGAAACGCGC AGAAGGCGAT GGCGAAGGCT
TCATCGTTAT TGATGACCTG GTGGATACCG GTGGTACTGC GGTTGCGATT
CGTGAAATGT ATCCAAAAGC GCACTTTGTC ACCATCTTCG CAAAACCGGC
TGGTCGTCCG CTGGTTGATG ACTATGTTGT TGATATCCCG CAAGATACCT
GGATTGAACA GCCGTGGGAT ATGGGCGTCG TATTCGTCCC GCCAATCTCC
GGTCGCTAA
```

Sequence of *gpt* in pINSluc and pINSmcs:

```
ATGAGCGAAA AATACATCGT CACCTGGGAC ATGTTGCAGA TCCATGCACG
TAAACTCGCA AGCCGACTGA TGCCTTCTGA ACAATGGAAA GGCATTATTG
CCGTAAGCCG TGGCGGTCTG GTACCGGGTG CGTTACTGGC GCGTGAACTG
GGTATTCGTC ATGTCGATAC CGTTTGTATT TCCAGCTACG ATCACGACAA
CCAGCGCGAG CTTAAAGTGC TGAAACGCGC AGAAGGCGAT GGCGAAGGCT
TCATCGTTAT TGATGACCTG GTGGATACCT GCGGTTGCGA TTCGTGAAAT
GTATCCAAAA GCGCACTTTG TCACCATCTT CGCAAAACCG GCTGGTCGTC
CGCTGGTTGA TGACTATGTT GTTGATATCC CGCAAGATAC CTGGATTGAA
CAGCCGTGGG ATATGGGCGT CGTATTCGTC CCGCCAATCT CCGGTCGCTA
A
```

C.

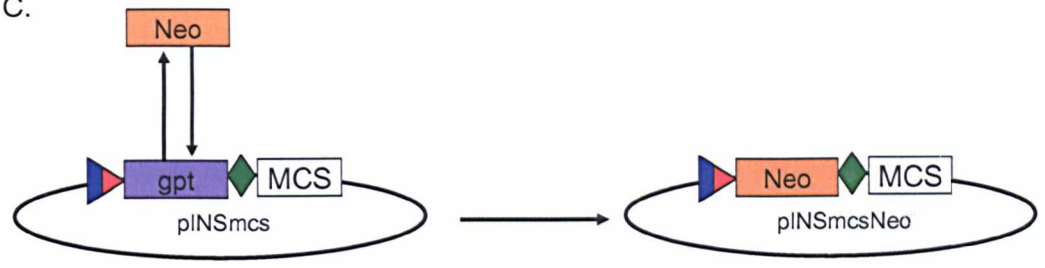



**Figure 7.6 A sequence comparison obtained from various *gpt* expression cassettes: A.** A cartoon showing the two 'insertion' constructs pINSluc and pINSmcs and highlights the direction of reducing *gpt* activity (not to scale). **B.** A sequence comparison of the *gpt* ORF between the published sequence, pBSgpt, pINSluc and pINSmcs (using primers, loxgptF and loxgptR). The pINSluc and pINSmcs sequence appears to be missing 8bp (in purple; top sequence), which has resulted in the formation of a cryptic stop codon (in red; bottom sequence) and therefore the generation of a truncated gpt protein. **C.** Illustrates replacement of the *gpt* ORF with that of neomycin (not to scale).

### 7.2.3 Generating stable uninduced *I-SceI* and *RAD52* HT1080 clones

*Testing pINSmcsNeo for the insertion event in clone Rht14-10 GFP expressing cells*

Again, prior to setting up the I-*SceI* and *RAD52* experiments, the activity of the new insertion construct, pINSmcsNeo, was firstly tested on an inducible GFP expressing clone. After insertion of pINSmcsNeo downstream of the TRE, expression of the neomycin gene can be selected for in G418. If the insertion event has taken place then the drug resistant colonies are expected to no longer express GFP. The 'target' clone chosen to test pINSmcsNeo and subsequently to generate inducible I-*SceI* and *RAD52* expressing clones, was clone Rht14-10 (Figure 7.7A). This clone expresses the FRT containing 'target' construct (pTFRTlox71dGFP) shown in Figure 7.7B. Clone Rht14-10 has uninduced expression levels identical to that of background, non-fluorescing cells and a non-corrected induction value of approximately 1000-fold (Figure 7.7A).

To test the new 'insertion' vector, pINSmcsNeo, it was co-transfected by lipofectamine 2000 into 500,000 Rht14-10 cells (Figure 7.7A), along with pMC-Cre (for conditions see Table 2.6). After 14 days selection in G418, the colonies were counted and analysed under the UV microscope for loss of GFP fluorescence. In total, 188 colonies were counted and of these 96% were no longer expressing GFP. These results suggested an absolute recombination frequency of 0.038%, which is 20-fold greater than what was achieved using the *gpt* expressing construct, pINSluc (Table 7.1, row 1). It therefore appears as though replacing the *gpt* gene with the *neomycin* gene has greatly improved the ability to generate clones which have undergone the insertion event.

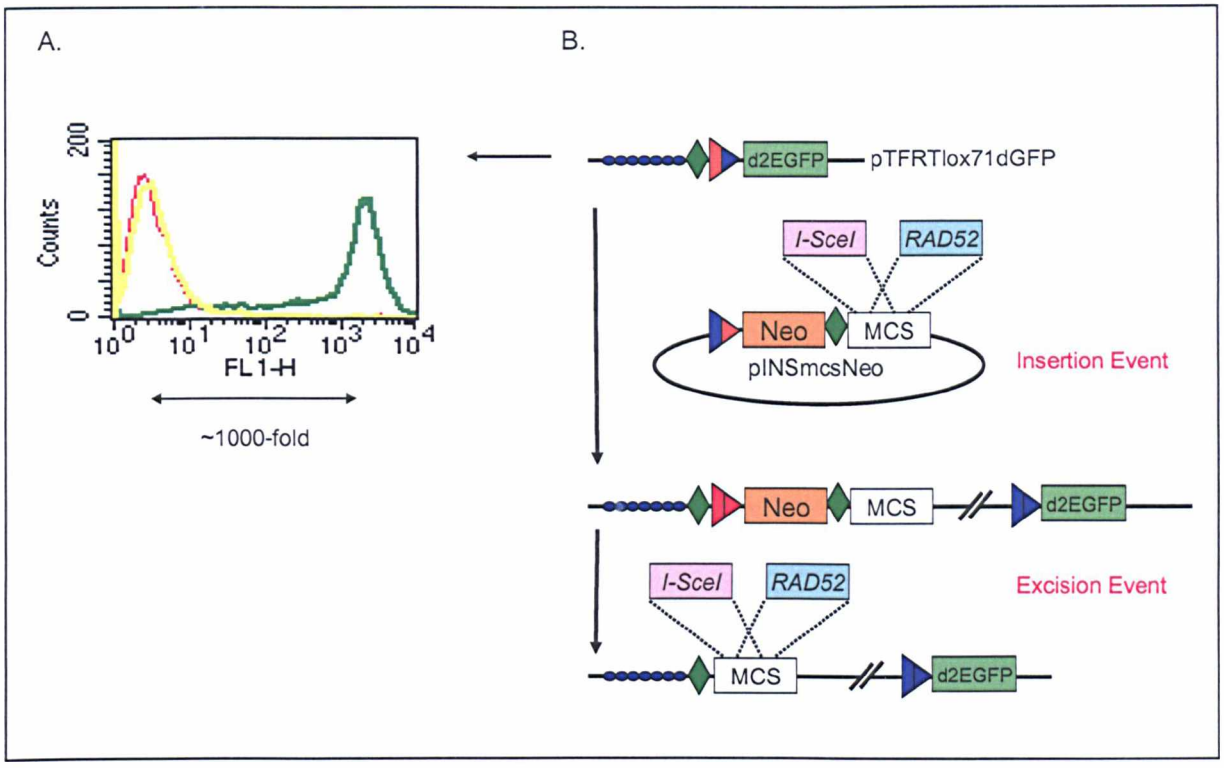



**Figure 7.7 The 'screen and insert' approach using pINSmcsNeo: A.** The FACS profile for clone Rht14-10. For instructions on how to interpret the FACS plot see Appendix IV. **B.** An illustration of the improved 'screen and insert' approach using the new 'insertion' construct, pINSmcsNeo, highlighting the location of the I-*SceI* and *RAD52* genes when they are later integrated (not to scale).

*Insertion of pINSneoISCEI and pINSneoRAD52 into Rht14-10 GFP expressing cells*

Both the I-*SceI* and the *RAD52* ORFs were cloned into pINSmcsNeo (Figure 7.7B), as described in the materials and methods (see 2.3.3.4), to generate pINSneoISCEI and pINSneoRAD52, respectively. To ensure that the correct cloning events had taken place, restriction digests and sequencing of the constructs using the INSF primer (Table 2.1) was carried out (data not shown). Similarly to pINSmcsNeo, after insertion of pINSneoISCEI and pINSneoRAD52 downstream of the TRE, expression of the neomycin gene can be selected for in G418, which should become sensitive to the drug upon the addition of tetracycline to the medium. The neomycin resistant clones ought to also demonstrate loss of GFP expression and the I-*SceI* and *RAD52* genes will remain uninduced as expression from the TRE will terminate after the neomycin cassette.

After transfecting 500,000 Rht14-10 cells with pINSneoISCEI or pINSneoRAD52 and pMC-Cre (for conditions see Table 2.6) and selecting in G418 for 14 days, 169 and 226 colonies were generated, respectively. These colonies were analysed under the UV microscope and the proportion of those clones that were no longer expressing GFP were calculated (Table 7.1, rows 4 and 5). For both experiments involving the 'insertion' constructs, pINSneoISCEI and pINSneoRAD52, 96% of the colonies generated were no longer expressing GFP. These results suggested an absolute insertion frequency of 0.033% for pINSneoISCEI and 0.043% for pINSneoRAD52. These results are consistent with those obtained for the empty vector, pINSmcsNeo (Table 7.1, row 3) and again are much improved when compared to the frequencies obtained for pINSluc (Table 7.1, row 1).

**Table 7.1 Summarises the insertion efficiencies of various 'insertion' constructs using clone Rht14-10:** The blue entries involved the *gpt* containing 'insertion' constructs (pINSluc and pINSmcs) and the red entries involved the neomycin containing 'insertion' constructs (derived from pINSmcsNeo).

| | 'Insertion' Construct | Total Number of Colonies | Number of GFP Negative Colonies | % of GFP Negative Colonies | Absolute Recombination Frequency |
|---|---|---|---|---|---|
| 1. | pINSluc | 11 | 9 | 82% | 0.0018 |
| 2. | pINSmcs | 0 | 0 | 0 | 0 |
| 3. | pINSmcsNeo | 188 | 181 | 96.3% | 0.036 |
| 4. | pINSneoISCEI | 169 | 163 | 96.4% | 0.033 |
| 5. | pINSneoRAD52 | 226 | 217 | 96% | 0.043 |

Three GFP-negative clones, confirmed by FACS analysis (Figure 7.8A), from both the I-*SceI* and the *RAD52* experiments were analysed by PCR for the recombination junction. Cell pellets were treated with pronase (for method see section 2.4) and the primers, TetOF and NeoR (for sequence see Table 2.1), were used. All clones analysed were positive for this region, as confirmed by the production of a 944bp band (Figure 7.8B). Each of the six clones also demonstrated sensitivity to G418 upon the addition of tetracycline to the medium. Due to inadequate time remaining on the project, the final evidence demonstrating the desired insertion event by Southern blot has not yet been produced. Despite this, clones ISCE1 and RAD1 were chosen to carry out the excision step; as it was assumed that enough evidence was provided by their loss-of-GFP expression, their tetracycline inducible G418 resistance and the amplification of their recombination junction.

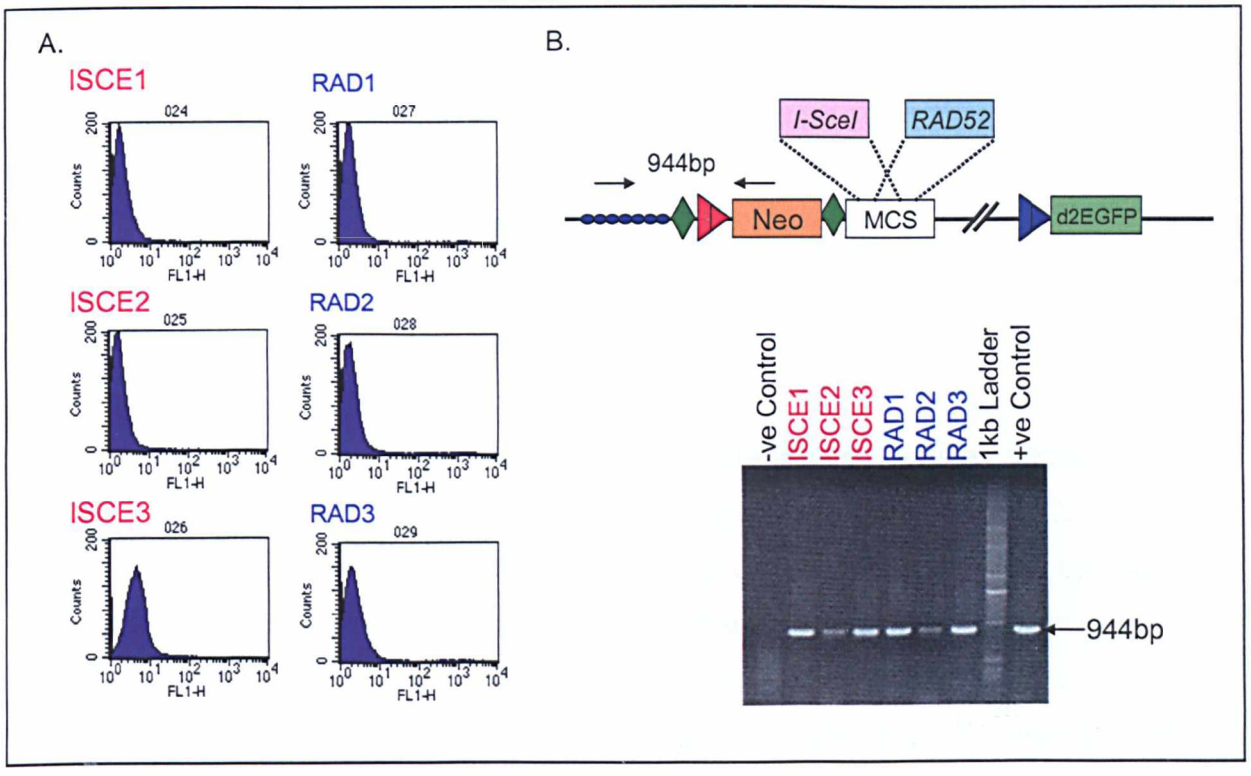



**Figure 7.8 Analysis of G418 resistant Rht14-10 clones: A.** FACS analysis on three clones from both the I-*SceI* (ISCE1-3) and *RAD52* (RAD1-3) insertion experiment; demonstrating loss-of-GFP fluorescence for them all. **B.** The expected genome structure (not to scale) after Cre-mediated insertion is illustrated (Top) and the approximate positions of the PCR primers (TetOF and NeoR) with the expected size of their product after amplification are shown. Shown below is an ethidium bromide-stained agarose gel of the PCR products obtained from clones ISCE1-3 and RAD1-3. The +ve control used were cell pellets from clone hh28.7neo1 (previously discussed; Figure 6.9).

*Executing the excision event on clones Rht14-10-ISCEI and Rht14-10-RAD52*

G418 was removed from the medium five days prior to the excision reaction, as tetracycline was now required to down-regulate the TRE to prevent activating the I-*SceI* and *RAD52* genes (Figure 7.7). 500,000 clone ISCE1 and RAD1 cells were transfected with pCAGGS-flpe using Lipofectamine 2000 (for conditions see Table 2.6). The transfection medium was not supplemented with tetracycline so as not to affect the transfection efficiency, however four hours after the transfection fresh medium containing tetracycline was added to the cells to prevent gene activation from the TRE. 24 hours after the transfection, cells were plated out at low density (50, 100, 250 and 500 cells per 15cm plate) for individual colony formation, again in tetracycline supplemented medium. 10 days post transfection the colonies were big enough to be picked and analysed by PCR. By assuming that the flp-mediated excision frequency was going to be as high as that discussed in Chapter Five (e.g. 14.6%), 24 individual clones were analysed by PCR.

The size changes expected when the excision event takes for clones harbouring the I-*SceI* and *RAD52* genes when using primers, TetOF and ISCE1R or RAD52-2R (for sequence see Table 2.1), are illustrated in Figure 7.9A and C, respectively. For the *flp*-treated ISCE1 clones, clean excised PCR product (427bp) was observed in 10 out of the 24 clones analysed (Clone 1, 6, 7, 12, 14, 15, 17, 18, 20, 22; marked with an * in Figure 7.9B). This therefore suggests that a *flp*-mediated excision frequency of almost 42% was obtained. For the clone containing stably integrated RAD52, RAD1, the excised product was observed in the majority of the clones analysed, however, most were also contaminated with unexcised PCR product (2.596kb). This was most likely caused by a more compact plate of colonies whilst picking, making it more difficult to keep cells from different clones separate. Despite this, four clean excised PCR products (529bp) were obtained (Clone 1, 6, 13, 20; marked with an * in Figure 7.9D) and a *flp*-mediated excision frequency of at least 17% was calculated.

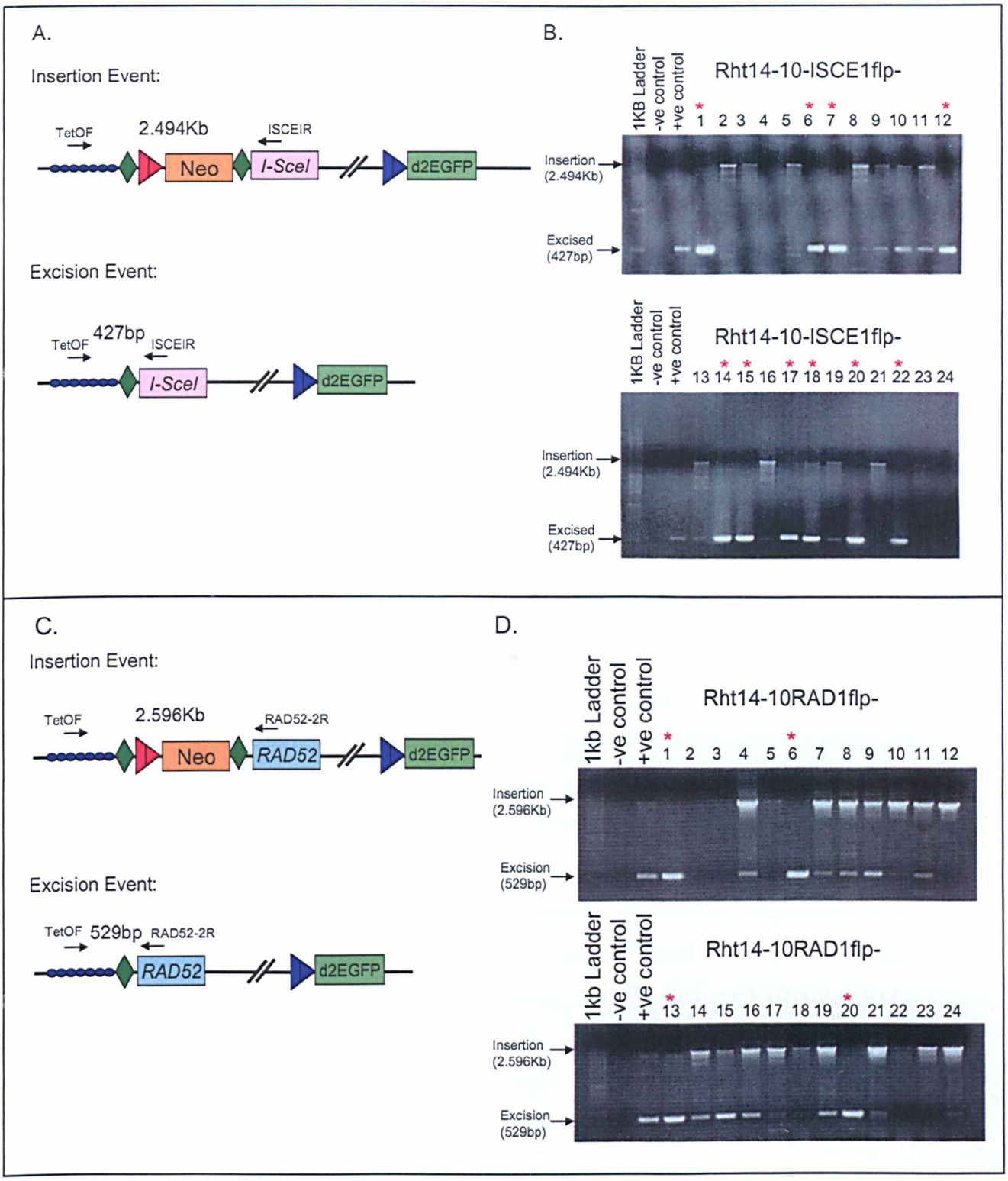



**Figure 7.9 Analysing ISCEI and RAD1 *flp*-treated clones for the excision event by PCR:** The expected genome structure (not to scale) after Cre-mediated insertion of pINSneoISCEI (**A.**) and pINSneoRAD52 (**C.**) and subsequent *flp*-mediated excision is illustrated. The approximate PCR primer positions and the size of the amplified PCR products are shown. **B.** An ethidium bromide-stained agarose gel of PCR products from 24 flp-treated ISCE1 clones. A pool of *flp*-treated ISCEI cells was used as the +ve control. **D.** An ethidium bromide-stained agarose gel of PCR products from 24 *flp*-treated RAD1 clones. A pool of *flp*-treated RAD1 cells was used as the +ve control.

The clones marked by an asterix in Figure 7.9B and D have all been continuously grown up and frozen down in tetracycline containing medium. Clones ISCE1flp1 and RAD1flp1 will be used for the functional analysis of each of these genes as they are thought to be less likely to be contaminated with unexcised cells (as they were picked first). Unfortunately, due to time constraints further testing of these clones was not possible. An account of the planned future direction of this project is discussed at the end of this chapter (see Section 7.3.2 and 7.3.3).

### 7.2.4 Attempting inducible I-*SceI* and *RAD52* expression in hTERT cells

*Successful isolation of G418-resistant colonies using the single copy hh28.29 clone*

Alongside the above experiments the single copy target hTERT-RPE1 clone, hh28.29 (discussed in Chapter Six [See Figure 6.5 and 6.12]), was used to test the 'screen and insert' approach using the new neomycin containing 'insertion' construct (pINSmcsNeo), as the previous 'insertion' vector (pINSluc) failed to isolate any *gpt* positive clones (See section 6.2.5; summarised in Table 7.2, row 2). After co-transfecting 200,000 hh28.29 cells with pINSmcsNeo and pMC-Cre (3μg of each) by Lipofectamine 2000 and selecting in G418 for 16 days, 38 colonies were generated. Of the 38 clones, 87% (33) appeared to demonstrate loss-of-GFP expression when examined under the UV-microscope. This suggested an absolute insertion frequency of 0.0165% (Table 7.2, row 3); which was greater than the frequency obtained when using plox66Neo (0.0045%; Table 7.2, row 2), however that particular experiment involved a different 'target' clone, hh28.7, which was thought to contain two copies of the 'target' locus (one active and one inactive) leading to a reduced insertion frequency.

The encouraging results involving pINSmcsNeo led to the trials of pINSneoISCEI and pINSneoRAD52 on clone hh28.29. Clone hh28.29 demonstrates the lowest uninduced expression levels, similar to that of background, out of all the GFP expressing hTERT-RPEI clones tested and the gene expression can be induced by up to ~300-fold (uncorrected). 200,000 hh28.29 cells were co-transfected with pMC-Cre

and pINSneoISCEI or pINSneoRAD52 (3μg of each) by Lipofectamine 2000. After selection for 16 days in G418 the colonies that were generated were counted and the proportion that no longer expressed GFP were calculated. For the experiment involving the I-*SceI* gene, 28 colonies were created, of which 79% (22) were GFP-negative and subsequently an absolute recombination frequency of 0.011% was calculated (Table 7.2, row 4). For pINSneoRAD52, 33 colonies were produced, of which 76% (25) were no longer expressing GFP and an absolute recombination frequency of 0.0125% was calculated (Table 7.2, row 5).

**Table 7.2 A summary of the insertion frequencies achieved using the hTERT-RPE1 clones:** The entries in blue are results which have previously been discussed in Chapter Six. The entries in red are the insertion efficiency results for the neomycin containing 'insertion' constructs.

| | 'Insertion' Construct | Total Number of Colonies | Number of GFP Negative Colonies | % of GFP Negative Colonies | Absolute Recombination Frequency |
|---|---|---|---|---|---|
| 1. | ploxNeo (hh28.7) | 23 | 9 | 39% | 0.0045% |
| 2. | pINSluc (hh28.7/28.29) | 0 | 0 | 0 | 0 |
| 3. | pINSmcsNeo (hh28.29) | 38 | 33 | 87% | 0.0165% |
| 4. | pINSneoISCEI (hh28.29) | 28 | 22 | 79% | 0.011% |
| 5. | pINSneoRAD52 (hh28.29) | 33 | 25 | 76% | 0.0125% |

Four G418 resistant clones from each experiment were picked and analysed by PCR for the recombination junction. Cell pellets were treated with pronase (see section 2.4.2) and the primers, TetOF and NeoR (for sequence see Table 2.1), were used by adopting the conditions described in section 2.4.3. All clones analysed were positive for the recombination junction, as confirmed by the production of a 944bp band (Figure 7.10B). Each of the eight clones also demonstrated sensitivity to G418 upon the addition of tetracycline to the medium. Again, due to time constraints the final evidence demonstrating the desired insertion event by Southern blot has not yet been produced. Despite this, clones hhISCE1 and hhRAD2 were chosen to attempt the excision step; as it was assumed that enough evidence was provided by their loss-of-GFP expression (Figure 7.10A), their tetracycline inducible G418-resistance and the amplification of their recombination junction.

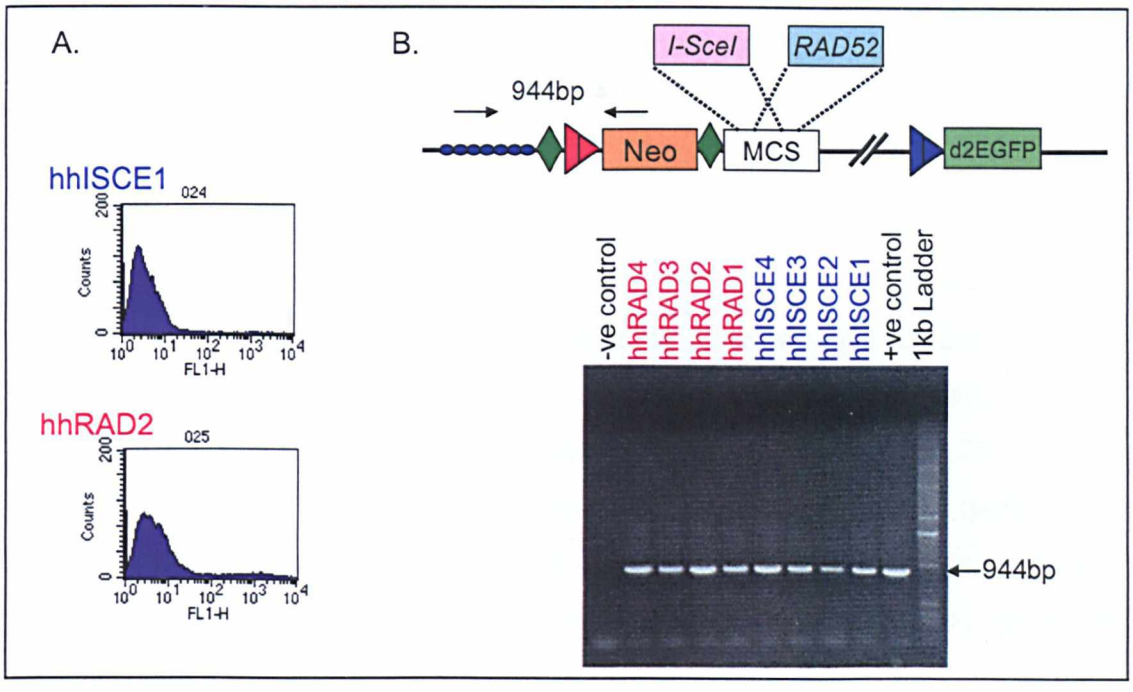



**Figure 7.10 Analysis of G418 resistant hh28.29 clones:** **A.** FACS of clones hhISCE1 and hhRAD2 demonstrating a loss-of-GFP expression. **B.** The expected genome structure (not to scale) after Cre-mediated insertion is illustrated (Top) and the approximate positions of the PCR primers (TetOF and NeoR) with the expected size of their PCR product after amplification is shown. An ethidium bromide-stained agarose gel of the PCR products obtained from clones hhISCE1-4 and hhRAD1-4 (bottom). The +ve control used were cell pellets from the HT1080 clone ISCEI (discussed above).

G418 was removed from the medium a week prior to the excision reaction, as tetracycline was now added to down-regulate the TRE before insertion of I-*SceI* and *RAD52* (Figure 7.7). 200,000 clone hhISCE1 and hhRAD2 cells were transfected with 6μg of pCAGGS-flpe using Lipofectamine 2000. The transfection medium was not supplemented with tetracycline so as not to affect the transfection efficiency, however it was added four hours after the start of the transfection to keep the TRE uninduced. 24 hours after the transfection, cells were plated out at low density (50, 100, 250 and 500 cells per 15cm plate) for individual colony formation in tetracycline supplemented medium. 18 days post transfection the colonies were big enough to be picked and analysed by PCR.

48 colonies from each experiment were picked and the cell pellets were treated with pronase and a PCR reaction was carried out using the conditions described in section 2.4. The size changes expected when the excision event takes place for clones harbouring I-*SceI* and *RAD52* when using primers, TetOF and ISCE1R or RAD52-2R (for sequence see Table 2.1), are described above (Figure 7.9A and C). For the *flp*-treated ISCE1 clones, no excised PCR products (427bp) were observed for any of the 48 clones. For the RAD52 experiment, 2 colonies (hhRAD2flp11 and hhRAD2flp24) appeared to demonstrate both the excised (529bp) and the un-excised (2.596Kb) recombination products (Figure 7.11B). Replating these two cell populations at low density and rescreening the colonies by PCR may isolate a pure excised clone. Isolating clones that have undergone the excision event in the hTERT-RPE1 cell line is most likely limited by their low transfection efficiency when using Lipofectamine 2000 (discussed in section 6.2.1). Further optimisation of the transfection technique is now required and a discussion of the proposed methods are discussed below (see Section 7.3.4).

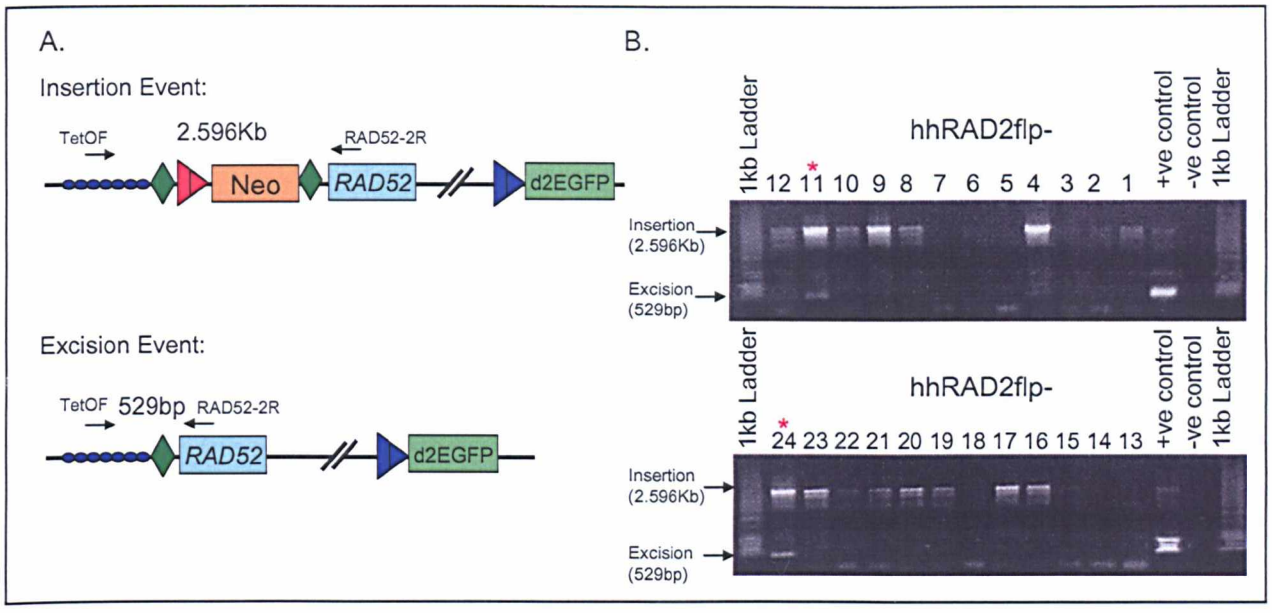



**Figure 7.11 Analysing hhRAD2 *flp*-treated clones for the excision event by PCR: A.** The expected genome structure (not to scale) after Cre-mediated insertion of pINSneoRAD52 and subsequent flp-mediated excision is illustrated. The approximate PCR primer positions and the size of the amplified PCR products are shown. **B.** An ethidium bromide-stained agarose gel of PCR products from 24 of the 48 flp-treated hhRAD2 clones analysed. A pool of flp-treated hhRAD2 cells were used as the +ve control.

## 7.3 Discussion

### 7.3.1 Successful generation of inducible I-*SceI* and *RAD52* HT1080 cells

Replacing the *gpt* gene with the *neomycin* ORF greatly improved the ability to isolate clones that had undergone the insertion event by over 20-fold (Table 7.1). In addition, the percentage of drug resistant clones that had lost their GFP-fluorescence increase to 96%, compared to 82% when using pINSluc (Table 7.1). It is clear from these results that the *gpt* ORF used in pINSluc had a reduced capability to isolate recombined clones, rather than a decrease in the insertion frequency itself. In fact, a sequence comparison between the *gpt* ORFs observed an 8 base-pair deletion between the highly functional (pBSgpt) and the less active forms of the gene (pINSluc and pINSmcs). The deletion of the 8 base pairs appears to have taken place during the

cloning procedures, therefore suggesting a possible instability of the gene during such manipulations.

Both the I-*SceI* and *RAD52* containing 'insertion' constructs, pINSneoISCEI and pINSneoRAD52, were successfully inserted downstream of the TRE in clone Rht14-10. The excision reaction was also carried out with much success, with an excision frequency of 42% for the I-*SceI* experiment and 17% for the *RAD52* experiment; the latter frequency, however, was limited by contaminating excised cells when picking clones. Despite this, both frequencies were still higher than the 14% observed when using Rht14-10 and the 'insertion' construct, pINSluc, in Chapter Five. Further work is still required to produce a Southern blot demonstrating that both recombination steps have taken place. Following this, Western analysis is needed to confirm that tight, inducible I-*SceI* and *RAD52* expression can be obtained using these clones when plated with or without tetracycline. Regardless of lacking these final results, it has been demonstrated in this chapter that each step of the improved 'screen and insert' approach is working extremely efficiently to generate HT1080 clones expressing an inducible GOI (I-*SceI* and *RAD52*) as soon as four weeks after the start of the Cre-mediated insertion step.

### 7.3.2 Future plans for examining the inducible I-*SceI*-HT1080 clones

After preliminary experiments involving the stable I-*SceI* expression in HT1080 cells, it was concluded that this gene would be a good candidate for use with the 'screen and insert' approach, for analysing for cryptic I-*SceI* sites within the mammalian genome. Stable I-*SceI* expression has previously been reported to have no effects on the cells survival in murine 3T3 cells, however, no direct comparison between I-*SceI* expressing and non-expressing clones was ever carried out (Rouet *et al.*, 1994). If cryptic I-*SceI* sites do exist in the mammalian genome then it is predicted that upon stable expression of the enzyme these sites will be cut and inaccurately repaired to prevent the continuous generation of DNA breaks. If these sites are inaccurately repaired so as not to affect the function of the gene in which they are located (non-mutagenic), then the host cells will appear normal. Alternatively, if repair of the

cryptic sites alter the function of the affected proteins, then a mutagenic phenotype may be observed, possibly resulting in impaired growth or cell death. It is expected that a mixture of these events will take place, if cryptic I-*SceI* sites do exist, within a cell population expressing stable I-*SceI*.

During both conditions (non-mutagenic and mutagenic), it should be possible to observe any effects caused by the generation of double-strand breaks by analysing for changes in cell-cycle and cell growth upon the cells first exposure to the I-SceI protein. The first experiment that will be carried out will involve inducing the I-SceI protein and analyse the effects on cell cycle at various time intervals after induction (possibly 2 hourly intervals), using propidium iodide nuclear staining (for methods see section 2.10). The plots will be compared to cells samples that remain uninduced throughout the experiment and to those which have been irradiated or exposed to a chemical which induces double strand breaks (e.g. etoposide). If the expression of I-*SceI* results in the generation of double strand breaks then it would be expected that cells would temporarily stop in the growth phase of the cell cycle (i.e. G1 and G2) whilst the damage is repaired. In addition, whilst DNA damage is being repaired then a reduced cell growth rate would be expected, therefore comparing the growth curves of both induced and uninduced I-*SceI* containing clones will also be carried out.

Finally, the generation of double strand breaks can be visualised by confocal microscopy using the anti-histone H2AX antibody. Histone H2AX has been shown to act as a sensor of DNA damage and localises to the site of double strand breaks (Yoshida and Morita, 2004). By comparing the number of anti-H2AX foci generated in the induced and uniduced I-*SceI* expressing HT1080 cells, using cells exposed to irradiation or genotoxic reagents as a positive control, will also indicate whether or not cryptic I-*SceI* sites are present. The timing of cell fixing for confocal analysis after the induction of I-*SceI,* will most likely be calculated from the results observed from the cell cycle FACS analysis. A limitation of using HT1080 cells, however, is that a high background of anti-H2AX foci formation is commonly achieved (Carpenter, personal communication) making any slight increase in foci formation difficult to determine visually, however, by using Photoshop software (Adobe) the background can be calculated (from uninduced cells as an average of fluorescence per

$\mu m^2$) and subtracted from the induced and drug treated cells, indicating whether or not an increase in double strand breaks is obtained.

### 7.3.3 Future plans for examining the inducible *RAD52*-HT1080 clones

Initially, the effects of *RAD52* overexpression, in such high amounts (expected to be up to 1000-fold), on HT1080 cells will be examined. Such high levels of *RAD52* overexpression is not likely to be tolerated in HT1080 cells and it is expected that the cell viability will be greatly reduced (Yanez and Porter, 2002). Again, the effects on cell cycle and cell growth will be analysed. Also, any effects on genomic stability will also be examined using the anti-H2AX antibody (as discussed for I-*SceI*), as an increase in double strand breaks may suggest an inability to repair such DNA damage. During previous investigations of human *RAD52* in HT1080 cells it was unclear as to whether the negative effect on gene targeting frequencies was due to the inability to achieve high levels of overexpression (Yanez and Porter, 2002). The homologous recombination assays will therefore be repeated on such clones in both the induced and uninduced states to compare any differences in gene targeting efficiencies. After the induction of *RAD52* for a homologous recombination assay, the transgene can immediately be down-regulated to alleviate any stresses caused on the cell and this itself may potentially have an effect on the recombination frequencies obtained.

In addition, if the yeast *RAD52* gene, which can elevate gene targeting efficiencies in HeLa cells by up to 37-fold, were to be inserted into clone Rht14-10 using the 'screen and insert' approach, a direct and accurate comparison of the two homologues can be carried out. This method would rule out any clonal variations in expression levels; as the expression of each gene should be induced by the same amount. In addition, a sequence comparison of the two homologues may identify a particular domain in the yeast *RAD52* which is responsible for the elevated gene targeting frequencies. If the yeast *RAD52* has the same effect on gene targeting in HT1080 cells, then the 'screen and insert' approach could be used to investigate whether linking the yeast *RAD52* gene (by an IRES) to another candidate gene will elevate gene targeting frequencies further (e.g. human or yeast *RAD51*). In addition, this system could be used to limit

gene expression to gene targeting experiments, so as not to permanently affect the host cells DNA repair machinery.

### 7.3.4 Plans to improve the transfection efficiency of hTERT-RPE1 cells

The Cre-mediated insertion step was successfully carried out consistently at a frequency of over 0.01%, using hh28.29 cells, using the 'insertion' constructs pINSmcsNeo, pINSneoISCEI and pINSneoRAD52. The limiting factor whilst using this cell line was their reduced transient transfection efficiency; which restricted the number of *flp*-mediated excision events. This made isolating excised clones difficult to isolate under non-selective conditions. In Chapter Six an optimum transient transfection efficiency, when using Lipofectamine 2000, of 35% was achieved (see Figure 6.2). Similar efficiencies were obtained when a colleague tested the FuGene transfection reagent on this cell line (Roche; personal communication). In order to optimise the *flp*-mediated excision step of the 'screen and insert' approach when using hTERT-RPEI cells a method for increased transient gene expression is required.

Future work will involve analysing the optimum transient transfection efficiency of electroporation in hTERT-RPE1 cells. If this method does not succeed in generating promising results (a transient transfection efficiency greater than 50%), then a nucleofector (Amaxa) approach will be adopted. A nucleofector is a new technology which claims to be capable of transfecting hard-to-transfect cell lines, such as primary cells, using an arrangement of electrical restrictions and cell type specific solutions. The conditions have already been optimised for many cell types, however, hTERT-RPEI cells are not yet listed. Similar cell-types, such as human mammary epithelial cells (HMRC), have achieved up to 73% transient transfection efficiency whilst using the nucleofector system.

# Chapter Eight

# Final Discussion

## 8.1 The basic 'screen and insert' approach and thesis objectives

### 8.1.1 A summary of the fundamental 'screen and insert' approach

This dissertation describes attempts to develop a system that can isolate clones expressing tightly regulated GOI, more easily. Two 'screen and insert' approaches were developed, which were adapted from preliminary findings using two simple 'insertion' vectors (plox66Luc and plox66Hyg (Appendix III); Chapter Four). In summary, the first system involves a single *Cre*-mediated insertion step of an IRES-Hyg containing 'insertion' construct (Chapter Four) and the second improved system incorporates an additional recombination step to facilitate the isolation of desirable clones under restrictive conditions (+ Tet, Chapter Five). Both 'screen and insert' systems utilise cells expressing tetracycline regulated *loxP*-tagged GFP, which operates as a marker of TRE-regulation and its expression can be sensitively measured using flow cytometry. After tightly regulated clones are isolated, the GFP ORF is then replaced with any GOI by site-specific recombination. The expression of the inserted GOI will represent that observed for the parental GFP expressing clone.

### 8.1.2 Thesis goals

The goals of this project and how they were executed are summarised in this chapter and any potential applications or limitations of the designed 'screen and insert' strategies will be examined. The progress made for three main objectives will be discussed and these include:

1. The generation of cell lines with tightly regulated *loxP*-tagged GFP
2. Designing an efficient 'screen and insert' approach
3. Utilising the proposed system for controlled expression of a GOI

## 8.2 Generating clones expressing tightly regulated GFP

### 8.2.1 TRE-regulation was accurately measured using flow cytometry

The generation of clones expressing tightly regulated GFP expression, is discussed in Chapter Three for HT1080 cells and Chapter Six for hTERT-RPE1 cells. The ability to monitor gene regulation from the TRE using GFP as a marker in cell culture was extremely successful. It not only facilitated the isolation of clones habouring the 'target' construct at a genomic location assisting tight regulation, it also allowed the identification of unsuitable clones, which could have otherwise been chosen for analysis. Such unsuitable clones include those that demonstrate 'leaky' expression when uninduced and clones with mosaic induced expression profiles (i.e. variations in TRE-expression within a clone). In both circumstances, traditional methods for analysing inducible gene regulation (e.g. Western blot) would have failed to detect these undesirable characteristics.

### 8.2.2 Tight GFP-regulation was achieved using optimised Tet technology

Recent advances in technology within the tetracycline system field, has helped towards achieving tightly regulated clones. Initial attempts, using the original TRE, did not attain tight regulation in HT1080 cells. It is believed to be due to additional factors recognising and binding at the TRE, resulting in 'leaky' uninduced expression (Gould and Chernajovsky, 2004). In fact, TRE-driven luciferase expression has been observed in non-tTA expressing HT1080 cells (personal communication, Carpenter). The modified TRE (TREmod; Clontech), on the other hand, did not demonstrate 'leaky' expression when uninduced in 6% of the clones analysed; and therefore proved that it was superior to the original TRE at generating clones with tightly regulated transgenes. These results have since been confirmed by Tomac and colleagues (2004), when regulated expression was compared using three versions of the TRE (Backman et al., 2004).

An additional limitation of the tetracycline system is the tendency for cells to lose their regulative capacity due to tTA silencing; as the tTA can be toxic when expressed at high levels (squelching [Gil and Ptashne, 1988; Gallia and Khalili, 1998]) and it is also vulnerable for epigenetic silencing (Pikaart *et al.,* 1998). The tTA toxicity is almost certainly a result of the VP16 activation domain (fused to TetR) sequestering the host cells transcriptional machinery when expressed within cells at high amounts (Barron *et al.*, 1997). Transcriptional silencing of the tTA is likely to be caused by its non-mammalian genetic code (from E. coli [TetR] and Herpes simplex virus [VP16]); resulting in it being recognised by the host cell as foreign DNA for gene silencing by methylation (Yoder *et al.*, 1997). Wells and colleagues were the first to demonstrate improved TRE-expression by using a codon-optimised TetR sequence (Wells *et al.*, 1999).

A mosaic pattern of induced GFP-expression in the isolated clones provided the first indication that tTA silencing was taking place in this study. Again, this feature would have been hard to detect by Western analysis. Various 'humanised' transactivators were already reported, although most accounts describe optimised rtTAs (Tet-ON system), rather than tTAs (Tet-OFF system) (Urlinger *et al.*, 2000; Anastassiadis *et al.,* 2002). Seeberg and colleagues (2004), however, have produced an 'improved' tTA (itTA) which was kindly donated to us by Rolf Sprengel (Department of Molecular Neurobiology, Germany; Krestrel *et al.,* 2004). The itTA is made up of a codon-optimised TetR, fused to three synthetic activation domains (F domains). Subsequent GFP-expressing clones, habouring the itTA, demonstrated stable, non-mosaic and long-term expression.

After resolving the encountered technical difficulties, HT1080 clones expressing tightly-regulated, highly inducible, *loxP*-tagged GFP expression were successfully identified by flow cytometry, when combining the actions of the TREmod with the itTA. Two clones in particular, Rht14-10 and Rht14-19, were chosen to test the 'screen and insert' approaches. Clone Rht14-10 has an induction value of nearly 1000-fold and uninduced expression as low as that of background. Clone Rht14-19 has an induction of 500-fold and also has uninduced expression as low as that of background. Tight(ish) GFP regulation in hTERT-RPE1 cells was also demonstrated, although it could be argued that the selected clones (hh28.7 and hh28.29)

demonstrated slight 'leakiness' from the uninduced TRE. Induction values for such clones range from 250- to 500-fold (hh28.29 and hh28.7, respectively). Examination of a larger number of clones and the suppression kinetics of hTERT-RPE1 cells will most definitely identify more clones mimicking the regulation achieved for HT1080 cells.

### 8.2.3 Further work

In this study it was demonstrated that controlling the amount of TRE-expression by varying the amount of tetracycline was limited, due to a rapid exchange between the on and the off states (Section 3.2.3). Generally, induction values for tightly regulated clones varied from 200- to 1000-fold in HT1080 cells and 200- to 500-fold in hTERT-RPE1 cells. It would therefore be desirable to generate a bank of inducible clones, each tightly regulated but with varying induction values, where the amount of induced expression can be chosen to suit any GOI (Figure 8.1). Cell-sorting could be carried out to isolate those tightly regulated clones that only induce GFP by, for e.g., 10- and 100-fold; as it is believed that UV microscopy fails to detect low GFP-expressing clones. Such a choice would be attractive for limiting the levels of a GOIs expression when it is known that high levels of overexpression lead to detrimental effects and may limit phenotypic analysis (e.g. the expression of VP16 fused proteins, such as the tTA, to reduce 'squelching').

It would also be interested to study the regions of the genome which facilitate tight regulation, in clones such as Rht14-10 and Rht14-19, as it would allow the identification of a few rare genomic regions that permit long-term expression along with reduced leakiness. This could be achieved by a process known as inverse PCR (Ochman *et al.,* 1990). Inverse PCR is a method for rapidly amplifying sequences that flank a region of known sequence (in this case the TRE-GFP cassette). This method utilises the polymerase chain reaction (PCR), but it has primers orientated in the reverse direction of the usual orientation. The template for the reverse primers is a restriction fragment that has been ligated upon itself to form a circle. Sequencing of

the resulting PCR products and a database search for the 'unfamiliar' sequence will reveal where in the genome the construct has integrated.

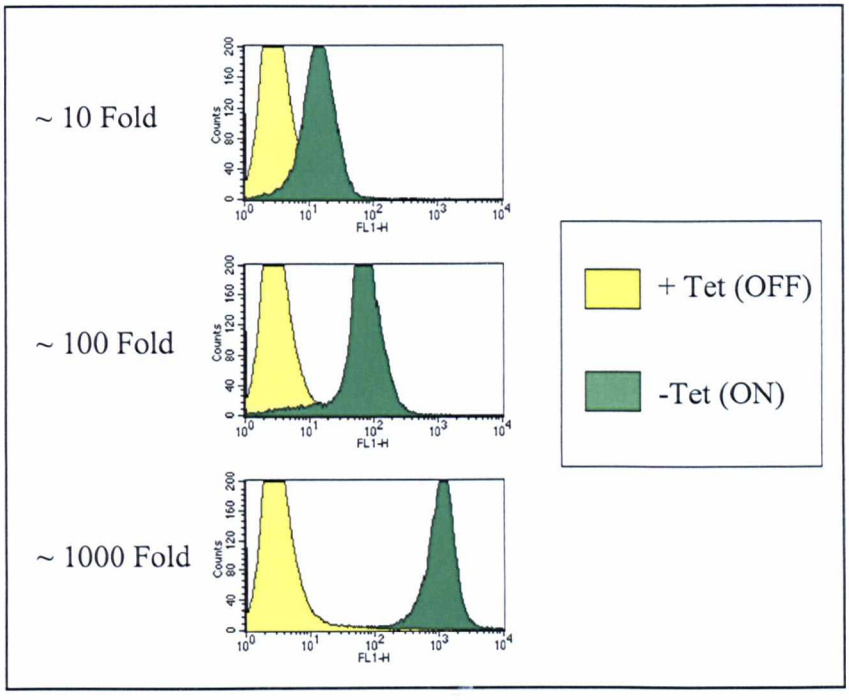



**Figure 8.1 Generating a bank of tightly regulated clones with various induction values:** A cartoon of GFP expression profiles of tightly regulated clones demonstrating 10-, 100-, or 1000-fold induction values, is illustrated. For instructions on how to interpret FACS plots see Appendix IV.

Examination of the kinetics of GFP-regulation in HT1080 cells (Clone Rht14-10 and Rht14-19) revealed that it took 96 hours to completely suppress gene expression and 48 hours to resume gene expression. These results suggest a potential limitation of the tetracycline system for applications that require rapid switch on/off (e.g. for expression only during a particular experiment or point in the cell-cycle). Potentially, the tetracycline system could be modified further to facilitate fast gene regulation by combining it with additional regulatory mechanisms (e.g. a ligand binding domain). The mutated hormone binding domain of the oestrogen receptor, ERT2, has been successfully used for the regulation of Cre in mice, for inducible site-specific recombination. In the absence of tamoxifen the Cre-ERT2 fusion is cytoplasmic and upon tamoxifen administration Cre can translocate to the nuclease where it is active. The system is limited, however, by leakiness when uninduced (Zhang *et al.*, 1995; Casanova *et al.*, 2002).

If the tetracycline system were combined with the ERT2 system in the 'screen and insert' approach, by fusing the GOI in the 'insertion' construct to ERT2 (Feil et al., 1997); dual control over the GOIs activity would be achieved. The TRE can firstly be induced (-Tet), however the GOI-ERT2 fusion will remain in the cytoplasm until tamoxifen is added and it is likely that faster kinetics of the GOI activity would be achieved (as the transcriptional/translational steps would be avoided), along with the tight regulation provided by the tetracycline system (Figure 8.2). This approach may be limited, however, by the unknown effects that fusing the ERT2 to the GOI will have on GOI activity and regulation. In fact, a reduced activity was reported for Cre recombinase when it was fused to ERT2, compared to its wild-type function (Zhang *et al.*, 1996).

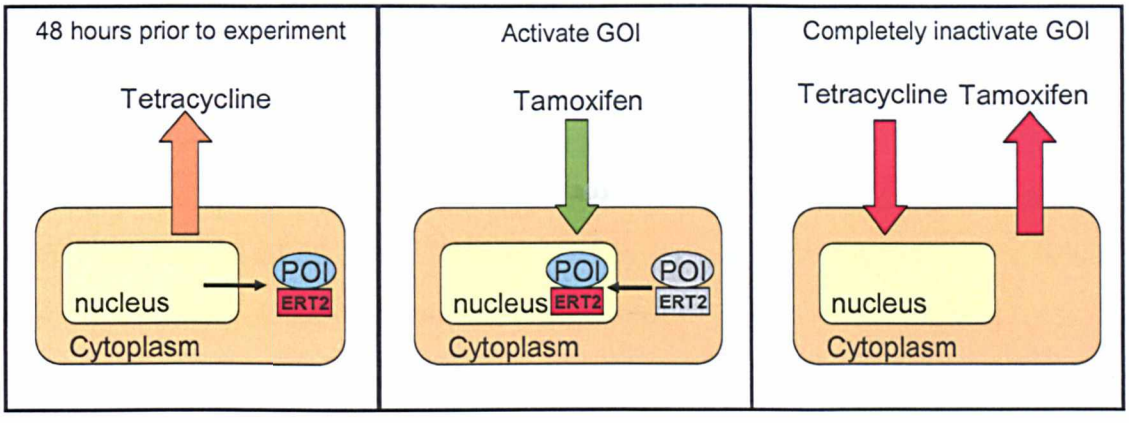



**Figure 8.2 The GOI under the dual control from the tetracycline and the ERT2 system:** 48 hours prior to the experiment, the GOI-ERT2 fusion is induced, removing tetracycline, to produce the protein fusion, P(protein)OI-ERT2, which sequesters in the cytoplasm (left). When the activity of the POI is required tamoxifen is added to the cells, which translocates the POI into the nucleaus, where it is active (middle). To suppress the POI activity tetracycline is added to the cells and tamoxifen is removed (right).

Ultimately, establishing the 'screen and insert' system in murine ES cells, to generate transgenic mice expressing tightly regulated GOI, will be desirable. Particularly when uncontrolled overexpression or down-regulation of that GOI results in embryonic or neonatal lethality; as gene function studies at latter stages of development cannot normally be analysed (See section 6.3.5).

## 8.3 The 'screen and insert' strategies

### 8.3.1 'Screen and insert' System One

The first 'screen and insert' strategy tested in HT1080 (Rht14-19) cells was System One (discussed in Chapter Four; Figure 8.3). Preliminary experiments revealed that Cre-mediated insertion of a GOI into the GFP-expressing cells required a method of selection. The inclusion of promoter elements on the 'insertion' construct was avoided by cloning an IRES-hygromycin cassette downstream of the *loxP*-tagged GOI (luciferase was used for testing purposes), to generate plox66LucIresHyg (Appendix III). 98% of the hygromycin-resistant clones generated were GFP-negative and the absolute frequency of insertion of plox66LucIresHyg was 0.01% in Rht14-19 cells (Table 8.1). Further molecular evidence that the desired insertion event had taken place in at least 10 of the GFP-negative clones (e.g. inducible luciferase, PCR and Southern analysis) was successfully demonstrated. This system could not be tested on the GFP-expressing hTERT-RPE1 clones as they were already hygromycin-resistant (used for selection for stable telomerase expression).

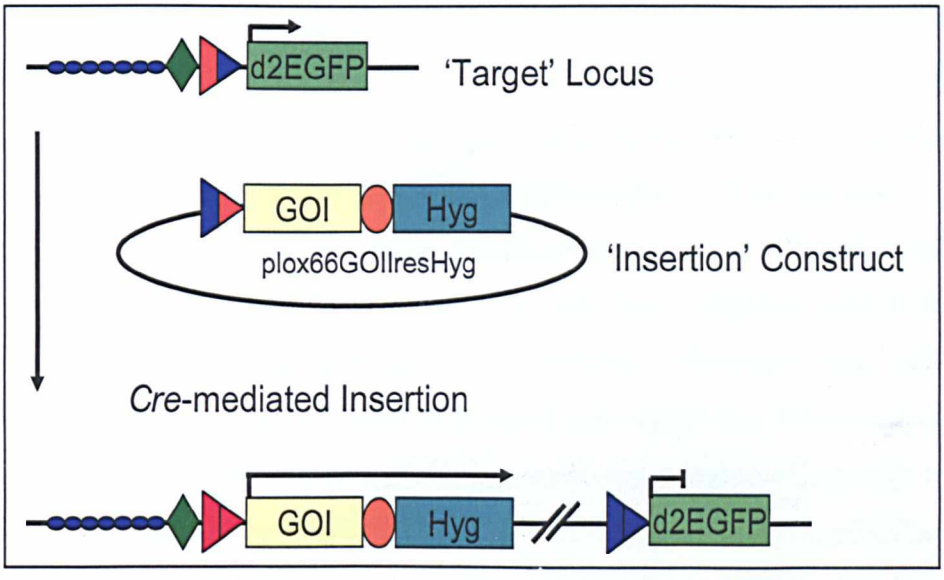



**Figure 8.3 'Screen and insert' Strategy One:** The strategy is illustrated for a GOI in an 'insertion' construct (plox66GOIIresHyg) that undergoes Cre-mediated insertion into the 'target' locus (DNA - not to scale). For key to symbols see Appendix I.

This IRES-base 'screen and insert' strategy is limited, however, by the need to express the GOI whilst selection for the desired insertion event (in hygromycin) is taking place. It would be more desirable if the system facilitated selection for the insertion event whilst the GOI is completely suppressed. An additional limitation may be the unknown effects on the levels of gene expression achieved when produced as part of a bicistronic message (using an IRES), as conflicting reports claim that an IRES favours expression of both the first (Mizuguchi et al., 2000) and the second cistron (Houdebine and Attal, 1999). As the 'screen and insert' approach relies on faithful, reproducible expression from the TRE, after the insertion of the GOI, use of an IRES-based approach may have its limitations. Consequently, a second improved 'screen and insert' strategy (System Two) was developed which facilitates IRES-free selection of the insertion event without expressing the GOI.

### 8.3.2 'Screen and insert' System Two

In Chapter Five, an improved 'screen and insert' strategy is discussed (System Two). This system contains an additional step, following Cre-mediated insertion, before the GOI can be expressed. The extra step uses *flp*-recombinase to delete a section of DNA to link the GOI to the TRE (Figure 8.4). It was predicted that the flp-mediated excision frequency would be high enough to allow the isolation of desirable clones by a simple PCR screen, in the presence of tetracycline. Cre-mediated insertion of pINSluc (Appendix III) into clone Rht14-19 generated a reduced recombination efficiency (Table 8.1), when compared to the 'insertion' construct used in System One (plox66LucIresHyg; Appendix III). In addition, GFP-expressing hTERT-RPE1 clones (hh28.7 and hh28.29) failed to generate *gpt*-expressing colonies at all (Chapter Six; section 6.2.5). Successive problems involving a reduced capacity to generate *gpt*-positive clones in both HT1080 and hTERT-RPE1 cells using pINSluc, and later pINSmcs (Appendix III), led to the replacement of the *gpt* ORF with the neomycin-resistance cassette (Figure 8.4; Chapter Seven; section 7.2.2).

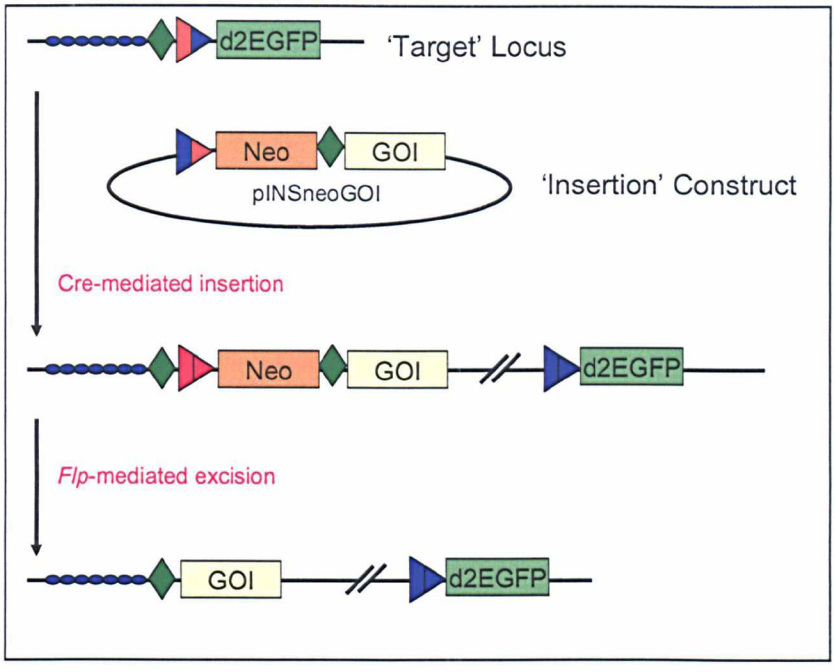



**Figure 8.4 Outline of the improved 'screen and insert' strategy:** The improved 'screen and insert' strategy is illustrated (not to scale). 'Target' clones which undergo the insertion event are selected for neomycin expression in G418. In selected clones, the GOI is placed downstream of the silent TRE using flp-mediated excision. For key to symbols see Appendix I.

Flp-mediated excision in Rht14-10 and Rht14-19-derived clones was successfully executed (Table 8.1). The frequency, however, was dramatically elevated in the Rht14-10-derived clones (at least 14.6%). The low frequency of excision seen for clone Rht14-19 (1.4%) was most likely due to the tetracycline supplementing the transfection medium, during lipofection. The protocol does suggest removing all antibiotics from the medium to obtain optimum transfection efficiencies (Invitrogen). Again, further molecular evidence that the desired insertion events have taken place were demonstrated by inducible luciferase expression and Southern blot.

### 8.3.3 Comparing the 'screen and insert' System Two to other strategies

In Chapter One (Section 1.3.3), three similar systems that have been developed by other groups for the easy generation of clones expressing tightly regulated transgenes, are described (these include the Flp-In T-REx [Invitrogen], the 'select and swap' strategy [Sullivan *et al.*, 2001] and the 'targeted transgenesis' strategy [Puttini *et al.*,

2005]). Limitations of the reported strategies include the failure to efficiently screen clones for tight tetracycline regulation prior to the integration of the GOI (Flp-In T-REx) and the need for repeated screening of tightly regulated clone for every new GOI ('select and swap'). The most relevant system for comparison with the improved 'screen and insert' approach, System Two, is the 'targeted transgenesis' strategy (Puttini *et al.*, 2005; Figure 1.22, Section 1.3.3). It facilitates pre-screening of tetracycline regulation at the particular site of random integration using SEAP as a reporter gene, after which I-*Sce*I-mediated gene targeting inserts the GOI downstream of the tightly regulated, suppressed TRE. An advantage of this system over the improved 'screen and insert' approach is that it only involves one insertion step (instead of an insertion and an excision step). The 'targeted transgenesis' strategy is limited, however, by a low targeting efficiency and unpredictable TRE expression after targeting, due to the insertion of an enhancer downstream of the TRE (see Figure 1.22 in Section 1.3.3).

### 8.3.4 A comparison of site-specific recombination efficiencies

*Reported Cre-mediated insertion efficiencies in cell culture*

The Cre/*loxP* system has primarily been utilised in transgenic mouse models and on most occasions the Cre-mediated excision reaction is carried out (Gu *et al.*, 1994; Hennet *et al.*, 1995). Fewer reports regarding Cre-mediated insertion are available when compared to the excision reaction; furthermore, they typically involve murine ES cells (Metzger and Feil, 1999). The insertion reaction is used less often because it is more difficult to control, since the excision event is kinetically favoured. One of the first demonstrations of Cre-mediated insertion of plasmid DNA into a *loxP*-tagged genomic locus was in murine ES cells; where an absolute insertion frequency of 0.00037% was described (Rucker and Piedrahita, 1997). As a result of the low Cre-mediated insertion frequency, mutant *loxP* sites have been engineered (such as those used in this study [*lox66* and *lox71*]) to increase the insertion frequency, by preventing subsequent excision reactions (Araki *et al.*, 1997; Albert *et al.*, 1995; Ow, 2002).

Prior to the work carried out for this thesis, the frequency of Cre-mediated insertion in HT1080 and hTERT-RPE1 cells, using the mutant *loxP* system, was not clear. In mouse ES cells, however, an absolute insertion frequency of 0.002% whilst using the same mutant *loxP* sites (*lox66* and *lox71*) was reported. The use of the mutant *loxP* system therefore appeared to improve the insertion frequency by almost 10-fold, when compare to using wild-type *loxPs*. The absolute Cre-mediated insertion frequency observed in this study, when using the *lox66/71* system, varied from 0.01 to 0.043% in HT1080 cells and 0.0045 to 0.0165% in hTERT-RPE1 cells (omitting the results for the *gpt*-containing 'insertion' vectors [red]; Table 8.1). The reduced insertion frequency observed for RPE1 cells is likely to be due to their reduced transfection efficiency; despite this, both cell types achieved greater insertion frequencies than previous reports (using the *lox66/71* system [Araki *et al.*, 1997]). More recently, Peidrahita and colleagues have identified a combination of mutant *loxP* sites which exhibit Cre-mediated insertion at levels 300-fold greater than the previously reported *lox71/lox66* model (Thomson *et al.*, 2003). Variations in insertion frequencies can also depend upon the accessibility of the *loxP*-tagged genomic locus, to both the 'insertion' vector and the Cre-recombinase protein (Baubonis and Sauer, 1993).

**Table 8.1 The efficiency of Cre-mediated insertion:** Cre-mediated insertion efficiencies obtained in this study for both HT1080 and hTERT-RPE1 (*italic*) cells. Clones derived using selection for *gpt* are in red. * RPE1 clone hh28.7 has a multicopy 'target' locus.

| 'Insertion' Construct | Clone | Selection | % of GFP-negative, drug resistant clones | Absolute insertion frequency |
|---|---|---|---|---|
| plox66Hyg | Cd46 | Hyg | 93% | 0.01% |
| plox66Neo | *hh28.7** | *G418* | *39%* | *0.0045%* |
| plox66LucIresHyg | Rht14-19 | Hyg | 98% | 0.01% |
| pINSLuc | Rht14-10<br>Rht14-19<br>*hh28.7**<br>*hh28.29* | gpt | 82%<br>69%<br>-<br>- | 0.0018%<br>0.0044%<br>*0%*<br>*0%* |
| pINSmcs | Rht14-10 | gpt | - | 0% |
| pINSmcsNeo | Rht14-10<br>*hh28.29* | G418 | 96.3%<br>*87%* | 0.036%<br>*0.0165%* |
| pINSneoISCE1 | Rht14-10<br>*hh28.29* | G418 | 96.4%<br>*79%* | 0.033%<br>*0.011%* |
| pINSneoRAD52 | Rht14-10<br>*hh28.29* | G418 | 96%<br>*76%* | 0.043%<br>*0.0125%* |

*Reported recombinase-mediated excision frequencies in cell culture*

As mentioned above, the site-specific excision reaction is highly favoured over the insertion reaction; therefore, in this study it was hoped that the frequency of excision would be high enough in HT1080 and hTERT-RPE1 cells to facilitate the selection-free isolation of the desirable clones in tetracycline (Figure 8.4). Typically, Cre-mediated excision of *loxP*-flanked DNA sequences has been carried out in ES cells and mice (Garcia and Mills, 2002). The efficiency of Cre-mediated excision in ES cells has ranged from ~20% (Gu *et al.*, 1993) to 63% (Taniguchi *et al.*, 1998). Excision events, however, have been executed in other cells types, such as the leukemic K562 cells, which demonstrated a Cre-mediated excision efficiency approaching 80%, when Cre mRNA was directly transfected into the cells (Van den Plas *et al.*, 2003).

*Flp*-mediated excision, on the other hand, is limited by its reduced activity at 37°C, when compared to Cre-recombinase; with *flp*-mediated excision of FRT-flanked DNA typically taking place in less than 1% of transfected ES cells (at $37^0$C; Buchholz *et al.*, 1996). Alterations to the *flp*-ORF to improve its thermostability, generated *flpe* (Buchholz *et al.*, 1998); which was subsequently cloned into an expression vector to create pCAGGS-flpe (Cambion). pCAGGS-flpe is reported to be capable of achieving *flp*-mediated excision of FRT-flanked DNA in 6% of transiently transfected ES cells (Schaft *et al.*, 2001). In this study, *flp*-mediated excision was observed in up to 42% of HT1080 cells transfected with pCAGGS-flpe (Table 8.2), a value which is markedly higher than previous reports (Schaft *et al.*, 2001). Excision efficiencies were greatly reduced in hTERT-RPE1 cells; however, this was most likely due to their reduced transfection capacity using Lipofectamine 2000, when compared to HT1080 cells. Further work to improve the 'screen and insert' strategy in hTERT-RPE1 cells would involve improving the transfection efficiency to a high enough level to allow the easy isolation of excised clones.

**Table 8.2 The frequency of *flp*-mediated excision:** Flp-mediated excision efficiencies obtained in this study for both HT1080 and hTERT-RPE1 (*italic*) clones.

| **'Insertion' Construct** | **Clone** | **Absolute *flp*-medated excision frequency** |
|---|---|---|
| pINSluc | Rht14-10IN5<br>Rht14-19IN5 | at least 14.6%<br>1.4% |
| pINSneoISCEI | Rht14-10ISCE1<br>*hh28.29ISCE1* | 42%<br>*?* |
| pINSneoRAD52 | Rht14-10RAD1<br>*hh28.29ISCE2* | 17%<br>*4.2%* |

## 8.3.5 Potential limitations of the improved 'screen and insert' approach

A potential limitation of the improved 'screen and insert' approach (System Two) when compared to other developed systems (see section 1.3.3), is that it involves two steps, which some may translate as lengthy. In fact, it only takes four weeks for tightly regulated, GFP-expressing HT1080 clones (e.g. Rht14-10 and Rht14-19) to be transformed into tightly-regulated, GOI-expressing HT1080 clones, if the steps are carried out sequentially. Compared to some applications in biological research (e.g. mouse transgenics), a month is relatively quick. If desired, this limitation could possibly be overcome by adopting a different experimental approach, such as that described by Meyer and colleagues (2002). They developed an alternative method of Cre-mediated cassette exchange (see section 1.1.4; Figure 1.9 for details) to swap a drug-resistance cassette with a GOI under non-selectable conditions in ES cells, by using a single vector that expresses both Cre- and *flpe*-recombinase (Lauth *et al.*, 2002). The design of the system they described is similar to the improved 'screen and insert' approach, however, by using the double expression vector it can be executed in a single step (Figure 8.5).

The minimal FRT sequence (used in both Meyer's study and this thesis) is refractive to integration (Jayaram, 1985); therefore the Cre-mediated insertion event primarily takes place to create an intermediate arrangement (Figure 8.5). Subsequently, one of two reactions is possible; either the insertion construct is excised to create the original structure in transient Cre, or *flpe* catalyses excision of the FRT-flanked DNA to

remove the drug resistance cassette (termed floxing; Figure 8.5). Up to 2.8% of Cre/*flp*-treated target-containing clones underwent the Cre-mediated insertion event. Of these, up to 97% had also undergone the *flp*-mediated excision event. In this study, it appeared as though *flp*-mediated excision had occurred predominantly over Cre-mediated excision. If conversion of the improved 'screen and insert' approach into a single step strategy was at any stage considered desirable, then this double recombinase-expression approach could be utilised. The conditions would have to be optimised for HT1080 cells and it is highly likely that this approach would be limited in hTERT-RPE1 cells due to their reduced transfection efficiency.

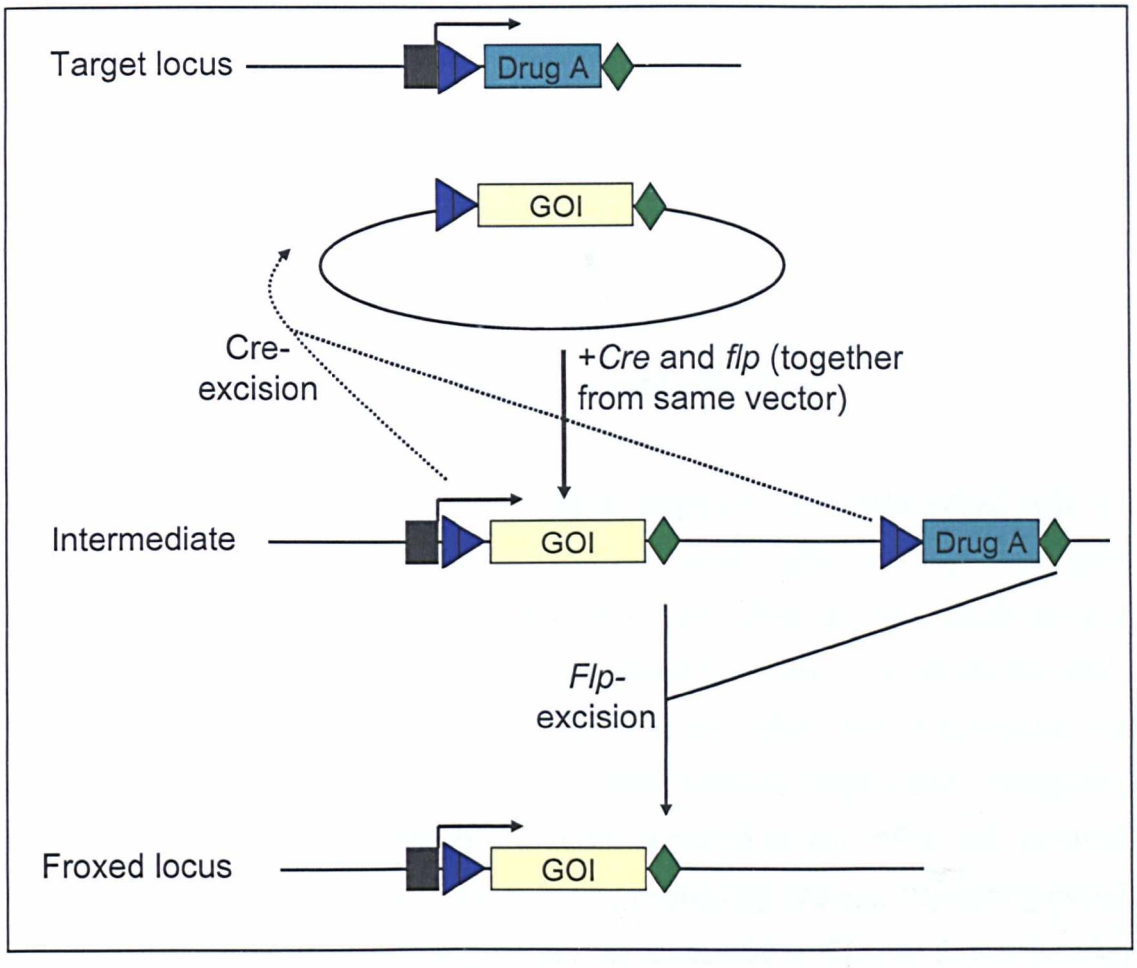



**Figure 8.5 A system for recombinase-mediated cassette exchange in non-selectable conditions:** A system described by Meyer and colleagues (2002) which facilitates the exchange of a pre-integrated drug-resistance gene with a GOI at a high a enough frequency that it can be executed under non-selectable conditions, is illustrated (Lauth *et al.*, 2002).

## 8.4 Utilising the 'screen and insert' approach

### 8.4.1 The 'screen and insert' approach was successfully deployed

As described in Chapter Seven, both I-*SceI* and *RAD52* were successfully placed downstream of the tightly regulated TRE in the HT1080 clone, Rht14-10, under restrictive conditions (+ Tet). In the hTERT-RPE1-derived clone, hh28.29, the 'insertion' constructs (pINSneoISCE1 and pINSneoRAD52; Appendix III) have been integrated into the 'target' locus, but the excision event remains inefficient due to their low transfection capacity. Again, optimising the amount of transient expression achieved in hTERT-RPE1 cells will be the primary concern when working with these cells following this study. This will be executed by testing the efficiency of electroporation and nucleofection (Amaxa) to obtain elevated transient expression (at least 50% would be desirable), as discussed in Chapter Seven (Section 7.3.4).

### 8.4.2 Future plans for the I-*SceI*-expressing clones

The generation of HT1080 clones habouring suppressed inducible I-*SceI,* will not only facilitate the study of DNA double-strand break repair and encourage gene targeting at pre-integrated recognition sites, it will also allow the mammalian genome to be screened for cryptic I-*SceI* sites. Preliminary experiments involving the stable expression of I-*SceI* in HT1080 cells, suggested that I-*SceI* may compromise cell survival; although at present, there is no conclusive data to support this. Previously, stable I-*SceI* expression in murine 3T3 cells appeared to not effect cell survival; however, no direct comparison between I-*SceI* expressing and non-expressing clones was ever carried out (Rouet *et al.*, 1994). As discussed in Chapter Seven (section 7.3.2), if cryptic I-*SceI* sites were present in the genome of mammalian cells then it is expected that cell growth rate and cell-cycle alterations would transpire upon I-*SceI* induction. In addition, it should be possible to visualise the increase in formation of double-strand breaks during I-*SceI* exposure by treating cells with the anti-H2AX antibody, which localises at sites of DNA damage (Yoshida and Morita, 2004), and analysing exposed cells using confocal microscopy.

### 8.4.3 Future plans for the *RAD52*-expressing clones

Previously, elevated levels of human *RAD52* overexpression in somatic cells has been difficult to achieve due to possible silencing of the transgene whilst in culture (Yanez and Porter, 2002). The generation of HT1080 cells habouring suppressed inducible *RAD52* will therefore facilitate the forced upregulation of the transgene (by up to 1000-fold) in every cell of a clonal population, where the effects on cell viability and homologous recombination can be analysed. Additionally, the 'screen and insert' system will allow the direct comparison of the human and the yeast *RAD52* genes function if the latter ORF were to also be integrated downstream of the TRE in clone Rht14-10. This method will avoid any clonal variations in transgene expression, which may lead to different phenotypic conclusions (for further discussion see section 7.3.3).

### 8.4.4 A few of many applications for the 'screen and insert' system

*Inducible Cre-expression*

Additional applications where tight, highly inducible gene regulation may be required include the stable expression of Cre-recombinase for catalysing conditional site-specific recombination. Many reports regarding inducible Cre-expression already exist, especially applications using transgenic mice (for e.g. see Zhang *et al.*, 2005; Yu *et al.*, 2005; Guo *et al.*, 2005). There are concerns, however, about 'leaky' uninduced expression resulting in a background of recombination events (Mahonen *et al.*, 2004). This has also become of particular concern since the discovery of cryptic 'pseudo-*loxP*' sites in the mammalian genome (Schmidt *et al.*, 2000). In tissue culture, the Cre-expression construct is often transiently transfected into the cells to stimulate the recombination reaction, however, this can be limited by low transfection efficiencies and the proportion of cells receiving the vector and expressing the recombinase protein is always less than 100%. Guaranteed tightly-regulated Cre-expressing cells (especially ES cells), therefore, would greatly facilitate transgenic research.

*Inducible RNA interference*

Constitutive knockdown of gene expression by RNAi (see Section 1.1.3 for more details) of a target that is essential for the cells survival is not plausible. Under circumstances such as these it is more desirable to use a conditional shRNA expression system to facilitate controlled gene down-regulation. The last two years has seen a rapid development of numerous techniques that permit inducible shRNA expression, however these methods typically deploy RNA-polymerase III U6 (Lee *et al.*, 2002) and H1 promoters (Brummelkamp *et al.*, 2002). The tetracycline system (as well as the other systems described in section 1.2) utilise the polymerase-II CMVmin promoter, which normally produces longer DNA transcripts that are not recognised by the cellular RNAi machinery (Paddison *et al.*, 2002).

Alterations to the response elements of various inducible systems to incorporate a RNA polymerase-III promoter have been successfully described (see Table 8.3 [Czauderna *et al.*, 2003; Gupta *et al.*, 2004; Higuchi *et al.*, 2004]). Nonetheless, it has also been reported that shRNA expression can now been achieved using a polymerase-II CMVmin promoter, such as that found for the TREmod, if the shRNA is cloned into an exact position coinciding with the transcriptional start site and it is followed by a truncated polyA (Xia *et al.*, 2003; Zhou *et al.*, 2005). This suggests that it should be possible to express shRNA from the TREmod if cloned into the correct position and the polyA is altered. In addition, reports have indicated that significant regulation of an shRNA expression is required to prevent a background down-regulation of the target gene whilst uninduced, especially when working with potent shRNA (Lin *et al.*, 2004).

If shRNA expression from the TREmod could be achieved, then the 'select and insert' approach could potentially provide the tight regulation required for adequate shRNA regulation. In fact, attempts were been made (by myself; not discussed elsewhere) to express shRNA, targeting the GFP gene, from the TREmod by reproducing the cloning procedures carried out by Xu and colleagues (Xia *et al.*, 2003). After stable integration of the TREmod-shGFP-containing construct into GFP expressing HT1080 cells, no gene down-regulation was achieved in any of the clones analysed. If the trial had been successful, utilising the 'screen and insert' approach for shRNA-regulation

would have come across a second technical difficulty. After the excision step of the 'screen and insert' strategy the remaining FRT site, located upstream of the GOI (Figure 8.4), would be incorporated into the shRNA sequence as it lies downstream of the transcriptional start. This would most likely render the shRNA non-functional. To overcome this limitation, the FRT site must be re-located so as not to become part of the shRNA transcript. A report by Jacks and colleagues described the successful incorporation of an altered, but still functional, *loxP* site into the TATA box region of the promoter without affecting the recombination efficiency or shRNA gene silencing capability (Ventura *et al.*, 2004). If shRNA expression from the TREmod is ever achieved, perhaps a similar method could be adopted for the 'screen and insert' approach to generate efficient, inducible gene silencing.

Finally, an alternative technique for producing conditional gene silencing using the 'screen and insert' approach would be to express micro RNAs (miRNA). These are structurally very similar to siRNA and are processed by the host cell for gene silencing using a similar mechanism (Mourelatos *et al.*, 2002). Mature miRNA can be generated from polymerase-II (Table 8.3) transcribed mRNAs by containing extraneous sequences in addition to the predicted pre-miRNA precursor sequence (Zeng *et al.*, 2002). This feature is highly desirable, as it suggests that miRNA can be regulated from the already established polymerase-II-containing inducible promoters, such as the TREmod; although reports demonstrating such regulation have yet to be published.

**Table 8.3 Published vector-based expression systems for RNAi:** Adapted from Amarzguioui et al., 2005.

| Type of RNAi | Promoter | Inducible System | Reference |
|---|---|---|---|
| Chimeric miRNA | CMV | Not inducible | Zeng *et al.*, 2002 |
| shRNA | U6 | Not inducible<br>Tetracycline<br>Ecdysone<br>Cre-loxP | Sui *et al.*, 2002<br>Czauderna *et al.*, 2003<br>Gupta *et al.*, 2004<br>Kasim *et al.*, 2004 |
| | H1 | Not inducible<br>Tetracycline<br>Lac | Brummelkamp *et al.*, 2002<br>Van der Wetering *et al.*, 2003<br>Higuchi *et al.*, 2004 |
| siRNA | U6 | Not inducible | Lee *et al.*, 2002 |

# References

Ackland-Berglund, C.E. and Leib, D.A.: Efficacy of tetracycline-controlled gene expression is influenced by cell type. Biotechniques 18 (1995) 196-200.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D., Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor, G.L., Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., et al.: The genome sequence of Drosophila melanogaster. Science 287 (2000) 2185-95.

Albert, H., Dale, E.C., Lee, E. and Ow, D.W.: Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. Plant J 7 (1995) 649-59.

Amarzguioui, M., Rossi, J.J. and Kim, D.: Approaches for chemically synthesized siRNA and vector-mediated RNAi. FEBS Lett (2005).

Anastassiadis, K., Kim, J., Daigle, N., Sprengel, R., Scholer, H.R. and Stewart, A.F.: A predictable ligand regulated expression strategy for stably integrated transgenes in mammalian cells in culture. Gene 298 (2002) 159-72.

Anderson, W.F.: Prospects for human gene therapy. Science 226 (1984) 401-9.

Andrews, B.J., Proteau, G.A., Beatty, L.G. and Sadowski, P.D.: The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences. Cell 40 (1985) 795-803.

Araki, K., Araki, M. and Yamamura, K.: Targeted integration of DNA using mutant lox sites in embryonic stem cells. Nucleic Acids Res 25 (1997) 868-72.

Araki, K., Araki, M. and Yamamura, K.: Site-directed integration of the cre gene mediated by Cre recombinase using a combination of mutant lox sites. Nucleic Acids Res 30 (2002) e103.

Araki, K., Imaizumi, T., Sekimoto, T., Yoshinobu, K., Yoshimuta, J., Akizuki, M., Miura, K., Araki, M. and Yamamura, K.: Exchangeable gene trap using the Cre/mutated lox system. Cell Mol Biol (Noisy-le-grand) 45 (1999) 737-50.

Arbones, M.L., Austin, H.A., Capon, D.J. and Greenburg, G.: Gene targeting in normal somatic cells: inactivation of the interferon-gamma receptor in myoblasts. Nat Genet 6 (1994) 90-7.

Askew, D.S., Bartholomew, C. and Ihle, J.N.: Insertional mutagenesis and the transformation of hematopoietic stem cells. Hematol Pathol 7 (1993) 1-22.

Aylon, Y. and Kupiec, M.: New insights into the mechanism of homologous recombination in yeast. Mutat Res 566 (2004) 231-48.

Aylon, Y. and Kupiec, M.: DSB repair: the yeast paradigm. DNA Repair (Amst) 3 (2004) 797-815.

Azizan, A., Cong, Y.S., Shain, K., Tsai, S.C., Yao, Y.L., Olashaw, N. and Seto, E.:

Transactivation by expression of the hepatitis B virus X protein with an inducible system. Mol Biol Rep 25 (1998) 231-6.

Backman, C.M., Zhang, Y., Hoffer, B.J. and Tomac, A.C.: Tetracycline-inducible expression systems for the generation of transgenic animals: a comparison of various inducible systems carried in a single vector. J Neurosci Methods 139 (2004) 257-62.

Baim, S.B., Labow, M.A., Levine, A.J. and Shenk, T.: A chimeric mammalian transactivator based on the lac repressor that is regulated by temperature and isopropyl beta-D-thiogalactopyranoside. Proc Natl Acad Sci U S A 88 (1991) 5072-6.

Banbura, M., Ackland-Berglund, C., Lee, S.H., Hamernik, D. and Jones, C.: Analysis of transcriptional activation of a cyclic AMP response element by 2,6,10,14-tetramethylpentadecane (pristane) in JB6 mouse epidermal cells. Mol Carcinog 11 (1994) 204-14.

Baron, U., Gossen, M. and Bujard, H.: Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. Nucleic Acids Res 25 (1997) 2723-9.

Baubonis, W. and Sauer, B.: Genomic targeting with purified Cre recombinase. Nucleic Acids Res 21 (1993) 2025-9.

Belshaw, P.J., Ho, S.N., Crabtree, G.R. and Schreiber, S.L.: Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. Proc Natl Acad Sci U S A 93 (1996) 4604-7.

Bezzubova, O., Silbergleit, A., Yamaguchi-Iwai, Y., Takeda, S. and Buerstedde, J.M.: Reduced X-ray resistance and homologous recombination frequencies in a RAD54-/- mutant of the chicken DT40 cell line. Cell 89 (1997) 185-93.

Blackburn, E.H.: Structure and function of telomeres. Nature 350 (1991) 569-73.

Blot-Chabaud, M., Coutry, N., Laplace, M., Bonvalet, J. and Farman, N.: Role of protein phosphatase in the regulation of Na+-K+-ATPase by vasopressin in the cortical collecting

duct. J Membr Biol 153 (1996) 233-9.

Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S. and Wright, W.E.: Extension of life-span by introduction of telomerase into normal human cells. Science 279 (1998) 349-52.

Branda, C.S. and Dymecki, S.M.: Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. Dev Cell 6 (2004) 7-28.

Brown, M., Figge, J., Hansen, U., Wright, C., Jeang, K.T., Khoury, G., Livingston, D.M. and Roberts, T.M.: lac repressor can regulate expression from a hybrid SV40 early promoter containing a lac operator in animal cells. Cell 49 (1987) 603-12.

Brummelkamp, T.R., Bernards, R. and Agami, R.: A system for stable expression of short interfering RNAs in mammalian cells. Science 296 (2002) 550-3.

Buchholz, F., Ringrose, L., Angrand, P.O., Rossi, F. and Stewart, A.F.: Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. Nucleic Acids Res 24 (1996) 4256-62.

Buchholz, F., Angrand, P.O. and Stewart, A.F.: A simple assay to determine the functionality of Cre or FLP recombination targets in genomic manipulation constructs. Nucleic Acids Res 24 (1996) 3118-9.

Buchholz, F., Angrand, P.O. and Stewart, A.F.: Improved properties of FLP recombinase evolved by cycling mutagenesis. Nat Biotechnol 16 (1998) 657-62.

Buerstedde, J.M. and Takeda, S.: Increased ratio of targeted to random integration after transfection of chicken B cell lines. Cell 67 (1991) 179-88.

Carpenter, A.J. and Porter, A.C.: Construction, characterization, and complementation of a conditional-lethal DNA topoisomerase IIalpha mutant human cell line. Mol Biol Cell 15 (2004) 5700-11.

Casanova, E., Fehsenfeld, S., Lemberger, T., Shimshek, D.R., Sprengel, R. and

Mantamadiotis, T.: ER-based double iCre fusion protein allows partial recombination in forebrain. Genesis 34 (2002) 208-14.

Cech, T.R. and Brehm, S.L.: Replication of the extrachromosomal ribosomal RNA genes of Tetrahymena thermophilia. Nucleic Acids Res 9 (1981) 3531-43.

Cech, T.R., Zaug, A.J. and Grabowski, P.J.: In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell 27 (1981) 487-96.

Chandrasegaran, S. and Smith, J.: Chimeric restriction enzymes: what is next? Biol Chem 380 (1999) 841-8.

Cheng, Z., Lin, Q. and Shen, Z.: [Study on association of vascular endothelial growth factor with the pathogenesis of pregnancy induced hypertension]. Zhonghua Fu Chan Ke Za Zhi 36 (2001) 72-5.

Cheng, K.W., Cheng, C.K. and Leung, P.C.: Differential role of PR-A and -B isoforms in transcription regulation of human GnRH receptor gene. Mol Endocrinol 15 (2001) 2078-92.

Chiba, H., Chambon, P. and Metzger, D.: F9 embryonal carcinoma cells engineered for tamoxifen-dependent Cre-mediated site-directed mutagenesis and doxycycline-inducible gene expression. Exp Cell Res 260 (2000) 334-9.

Chopra, I. and Anderson, M.A.: Inhibition of beta-lactamase synthesis in Staphylococcus aureus by minocycline. J Antimicrob Chemother 16 (1985) 17-21.

Chung, S., Andersson, T., Sonntag, K.C., Bjorklund, L., Isacson, O. and Kim, K.S.: Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines. Stem Cells 20 (2002) 139-45.

Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D. and Jaisser, F.: I-SceI-induced gene replacement at a natural locus in embryonic stem cells. Mol Cell Biol 18 (1998) 1444-8.

Colleaux, L., D'Auriol, L., Galibert, F. and Dujon, B.: Recognition and cleavage site of the intron-encoded omega transposase. Proc Natl Acad Sci U S A 85 (1988) 6022-6.

Colleaux, L., D'Auriol, L., Galibert, F. and Dujon, B.: Recognition and cleavage site of the intron-encoded omega transposase. Proc Natl Acad Sci U S A 85 (1988) 6022-6.

Curcio, L.D., Bouffard, D.Y. and Scanlon, K.J.: Oligonucleotides as modulators of cancer gene expression. Pharmacol Ther 74 (1997) 317-32.

Czauderna, F., Santel, A., Hinz, M., Fechtner, M., Durieux, B., Fisch, G., Leenders, F., Arnold, W., Giese, K., Klippel, A. and Kaufmann, J.: Inducible shRNA expression for application in a prostate cancer mouse model. Nucleic Acids Res 31 (2003) e127.

Damke, H., Gossen, M., Freundlieb, S., Bujard, H. and Schmid, S.L.: Tightly regulated and inducible expression of dominant interfering dynamin mutant in stably transformed HeLa cells. Methods Enzymol 257 (1995) 209-20.

de Piedoue, G., Maurisse, R., Kuzniak, I., Lopez, B., Perrin, A., Negre, O., Leboulch, P. and Feugeas, J.P.: Improving gene replacement by intracellular formation of linear homologous DNA. J Gene Med 7 (2005) 649-56.

Deng, C. and Capecchi, M.R.: Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol 12 (1992) 3365-71.

Di Primio, C., Galli, A., Cervelli, T., Zoppe, M. and Rainaldi, G.: Potentiation of gene targeting in human cells by expression of Saccharomyces cerevisiae Rad52. Nucleic Acids Res 33 (2005) 4639-48.

Doetschman, T., Gregg, R.G., Clarke, A., Melton, D.W., Thompson, S. and Smithies, O.: Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature 330 (1987) 576-578.

Dominguez-Bendala, J., Priddle, H., Clarke, A. and McWhir, J.: Elevated expression of exogenous Rad51 leads to identical increases in gene-targeting frequency in murine

embryonic stem (ES) cells with both functional and dysfunctional p53 genes. Exp Cell Res 286 (2003) 298-307.

Donoho, G., Jasin, M. and Berg, P.: Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol 18 (1998) 4070-8.

Dujon, B.: Group I introns as mobile genetic elements: facts and mechanistic speculations--a review. Gene 82 (1989) 91-114.

Dymecki, S.M.: Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. Proc Natl Acad Sci U S A 93 (1996) 6191-6.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T.: Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411 (2001) 494-8.

Ellis, N.A., Lennon, D.J., Proytcheva, M., Alhadeff, B., Henderson, E.E. and German, J.: Somatic intragenic recombination within the mutated locus BLM can correct the high sister-chromatid exchange phenotype of Bloom syndrome cells. Am J Hum Genet 57 (1995) 1019-27.

Episkopou, V., Arkell, R., Timmons, P.M., Walsh, J.J., Andrew, R.L. and Swan, D.: Induction of the mammalian node requires Arkadia function in the extraembryonic lineages. Nature 410 (2001) 825-30.

Essers, J., Hendriks, R.W., Swagemakers, S.M., Troelstra, C., de Wit, J., Bootsma, D., Hoeijmakers, J.H. and Kanaar, R.: Disruption of mouse RAD54 reduces ionizing radiation resistance and homologous recombination. Cell 89 (1997) 195-204.

Feederle, R., Delecluse, H.J., Rouault, J.P., Schepers, A. and Hammerschmidt, W.: Efficient somatic gene targeting in the lymphoid human cell line DG75. Gene 343 (2004) 91-7.

Feil, R., Wagner, J., Metzger, D. and Chambon, P.: Regulation of Cre recombinase

activity by mutated estrogen receptor ligand-binding domains. Biochem Biophys Res Commun 237 (1997) 752-7.

Feng, Y.Q., Seibler, J., Alami, R., Eisen, A., Westerman, K.A., Leboulch, P., Fiering, S. and Bouhassira, E.E.: Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. J Mol Biol 292 (1999) 779-85.

Figge, J., Wright, C., Collins, C.J., Roberts, T.M. and Livingston, D.M.: Stringent regulation of stably integrated chloramphenicol acetyl transferase genes by E. coli lac repressor in monkey cells. Cell 52 (1988) 713-22.

Filmus, J., Remani, J. and Klein, M.H.: Synergistic induction of promoters containing metal- and glucocorticoid-responsive elements. Nucleic Acids Res 20 (1992) 2755-60.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C.: Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391 (1998) 806-11.

Fukushige, S. and Sauer, B.: Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. Proc Natl Acad Sci U S A 89 (1992) 7905-9.

Gallia, G.L. and Khalili, K.: Evaluation of an autoregulatory tetracycline regulated system. Oncogene 16 (1998) 1879-84.

Garcia, E.L. and Mills, A.A.: Getting around lethality with inducible Cre-mediated excision. Semin Cell Dev Biol 13 (2002) 151-8.

Geller, B.L.: Antibacterial antisense. Curr Opin Mol Ther 7 (2005) 109-13.

Geller, B.L., Deere, J., Tilley, L. and Iversen, P.L.: Antisense phosphorodiamidate morpholino oligomer inhibits viability of Escherichia coli in pure culture and in mouse peritonitis. J Antimicrob Chemother 55 (2005) 983-8.

Gil, J., Bernard, D., Martinez, D. and Beach, D.: Polycomb CBX7 has a unifying role in

cellular lifespan. Nat Cell Biol 6 (2004) 67-72.

Gil, J., Kerai, P., Lleonart, M., Bernard, D., Cigudosa, J.C., Peters, G., Carnero, A. and Beach, D.: Immortalization of primary human prostate epithelial cells by c-Myc. Cancer Res 65 (2005) 2179-85.

Gill, G. and Ptashne, M.: Negative effect of the transcriptional activator GAL4. Nature 334 (1988) 721-4.

Gingrich, J.R. and Roder, J.: Inducible gene expression in the nervous system of transgenic mice. Annu Rev Neurosci 21 (1998) 377-405.

Goffeau, A., Park, J., Paulsen, I.T., Jonniaux, J.L., Dinh, T., Mordant, P. and Saier, M.H., Jr.: Multidrug-resistant transport proteins in yeast: complete inventory and phylogenetic characterization of yeast open reading frames with the major facilitator superfamily. Yeast 13 (1997) 43-54.

Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W. and Bujard, H.: Transcriptional activation by tetracyclines in mammalian cells. Science 268 (1995) 1766-9.

Gossen, M. and Bujard, H.: Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A 89 (1992) 5547-51.

Gossen, M., Bonin, A.L. and Bujard, H.: Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. Trends Biochem Sci 18 (1993) 471-5.

Gossen, M. and Bujard, H.: Anhydrotetracycline, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells. Nucleic Acids Res 21 (1993) 4411-2.

Gould, D.J., Berenstein, M., Dreja, H., Ledda, F., Podhajcer, O.L. and Chernajovsky, Y.: A novel doxycycline inducible autoregulatory plasmid which displays "on"/"off" regulation suited to gene therapy applications. Gene Ther 7 (2000) 2061-70.

Gould, D.J. and Chernajovsky, Y.: Endogenous GATA factors bind the core sequence of the tetO and influence gene regulation with the tetracycline system. Mol Ther 10 (2004) 127-38.

Gregory, T.R.: Genome size and developmental complexity. Genetica 115 (2002) 131-46.

Griffiths, A.J.F.: An ¤introduction to genetic analysis, 5. ed. W.H. Freeman, New York, 1993.

Groth, A.C., Olivares, E.C., Thyagarajan, B. and Calos, M.P.: A phage integrase directs efficient site-specific integration in human cells. Proc Natl Acad Sci U S A 97 (2000) 5995-6000.

Groth, A.C. and Calos, M.P.: Phage integrases: biology and applications. J Mol Biol 335 (2004) 667-78.

Gu, H., Zou, Y.R. and Rajewsky, K.: Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell 73 (1993) 1155-64.

Gu, H., Marth, J.D., Orban, P.C., Mossmann, H. and Rajewsky, K.: Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. Science 265 (1994) 103-6.

Guerrero, I., Villasante, A., Corces, V. and Pellicer, A.: Activation of a c-K-ras oncogene by somatic mutation in mouse lymphomas induced by gamma radiation. Science 225 (1984) 1159-62.

Guo, Z.M., Xu, K., Yue, Y., Huang, B., Deng, X.Y., Zhong, N.Q., Hong, X., Chen, X.G. and Xiao, D.: Temporal control of Cre recombinase-mediated in vitro DNA recombination by Tet-on gene expression system. Acta Biochim Biophys Sin (Shanghai) 37 (2005) 133-8.

Gupta, S., Schoer, R.A., Egan, J.E., Hannon, G.J. and Mittal, V.: Inducible, reversible, and stable RNA interference in mammalian cells. Proc Natl Acad Sci U S A 101 (2004)

1927-32.

Hammond, S.M., Bernstein, E., Beach, D. and Hannon, G.J.: An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404 (2000) 293-6.

Harley, C.B., Futcher, A.B. and Greider, C.W.: Telomeres shorten during ageing of human fibroblasts. Nature 345 (1990) 458-60.

Hasty, P., Rivera-Perez, J. and Bradley, A.: The length of homology required for gene targeting in embryonic stem cells. Mol Cell Biol 11 (1991) 5586-91.

Hasty, P., Bradley, A., Morris, J.H., Edmondson, D.G., Venuti, J.M., Olson, E.N. and Klein, W.H.: Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. Nature 364 (1993) 501-6.

Hasty, P., Ramirez-Solis, R., Krumlauf, R. and Bradley, A.: Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. Nature 350 (1991) 243-6.

Hasty, P., Rivera-Perez, J. and Bradley, A.: The length of homology required for gene targeting in embryonic stem cells. Mol Cell Biol 11 (1991) 5586-91.

Hatanaka, A., Yamazoe, M., Sale, J.E., Takata, M., Yamamoto, K., Kitao, H., Sonoda, E., Kikuchi, K., Yonetani, Y. and Takeda, S.: Similar effects of Brca2 truncation and Rad51 paralog deficiency on immunoglobulin V gene diversification in DT40 cells support an early role for Rad51 paralogs in homologous recombination. Mol Cell Biol 25 (2005) 1124-34.

Henikoff, S.: Conspiracy of silence among repeated transgenes. Bioessays 20 (1998) 532-5.

Henikoff, S. and Comai, L.: Trans-sensing effects: the ups and downs of being together. Cell 93 (1998) 329-32.

Hennet, T., Hagen, F.K., Tabak, L.A. and Marth, J.D.: T-cell-specific deletion of a

polypeptide N-acetylgalactosaminyl-transferase gene by site-directed recombination. Proc Natl Acad Sci U S A 92 (1995) 12070-4.

Herrlinger, U., Pechan, P.A., Jacobs, A.H., Woiciechowski, C., Rainov, N.G., Fraefel, C., Paulus, W. and Reeves, S.A.: HSV-1 infected cell proteins influence tetracycline-regulated transgene expression. J Gene Med 2 (2000) 379-89.

Higuchi, M., Tsutsumi, R., Higashi, H. and Hatakeyama, M.: Conditional gene silencing utilizing the lac repressor reveals a role of SHP-2 in cagA-positive Helicobacter pylori pathogenicity. Cancer Sci 95 (2004) 442-7.

Hillen, W. and Berens, C.: Mechanisms underlying expression of Tn10 encoded tetracycline resistance. Annu Rev Microbiol 48 (1994) 345-69.

Hinnen, A., Hicks, J.B. and Fink, G.R.: Transformation of yeast. Proc Natl Acad Sci U S A 75 (1978) 1929-33.

Hoess, R.H., Ziese, M. and Sternberg, N.: P1 site-specific recombination: nucleotide sequence of the recombining sites. Proc Natl Acad Sci U S A 79 (1982) 3398-402.

Hoess, R.H., Wierzbicki, A. and Abremski, K.: The role of the loxP spacer region in P1 site-specific recombination. Nucleic Acids Res 14 (1986) 2287-300.

Houdebine, L.M. and Attal, J.: Internal ribosome entry sites (IRESs): reality and use. Transgenic Res 8 (1999) 157-77.

Howe, J.R., Skryabin, B.V., Belcher, S.M., Zerillo, C.A. and Schmauss, C.: The responsiveness of a tetracycline-sensitive expression system differs in different cell lines. J Biol Chem 270 (1995) 14168-74.

Hu, M.C. and Davidson, N.: Targeting the Escherichia coli lac repressor to the mammalian cell nucleus. Gene 99 (1991) 141-50.

Hug, B.A., Wesselschmidt, R.L., Fiering, S., Bender, M.A., Epner, E., Groudine, M. and Ley, T.J.: Analysis of mice containing a targeted deletion of beta-globin locus control

region 5' hypersensitive site 3. Mol Cell Biol 16 (1996) 2906-12.

Hunter, T., Hunt, T., Jackson, R.J. and Robertson, H.D.: The characteristics of inhibition of protein synthesis by double-stranded ribonucleic acid in reticulocyte lysates. J Biol Chem 250 (1975) 409-17.

Itzhaki, J.E. and Porter, A.C.: Targeted disruption of a human interferon-inducible gene detected by secretion of human growth hormone. Nucleic Acids Res 19 (1991) 3835-42.

Izumi, M. and Gilbert, D.M.: Homogeneous tetracycline-regulatable gene expression in mammalian fibroblasts. J Cell Biochem 76 (1999) 280-9.

Jantz, D., Amann, B.T., Gatto, G.J., Jr. and Berg, J.M.: The design of functional DNA-binding proteins based on zinc finger domains. Chem Rev 104 (2004) 789-99.

Jasin, M.: Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet 12 (1996) 224-8.

Jayaram, M.: Two-micrometer circle site-specific recombination: the minimal substrate and the possible role of flanking sequences. Proc Natl Acad Sci U S A 82 (1985) 5875-9.

Ji, P., Agrawal, S., Diederichs, S., Baumer, N., Becker, A., Cauvet, T., Kowski, S., Beger, C., Welte, K., Berdel, W.E., Serve, H. and Muller-Tidow, C.: Cyclin A1, the alternative A-type cyclin, contributes to G1/S cell cycle progression in somatic cells. Oncogene 24 (2005) 2739-44.

Jiang, X.R., Jimenez, G., Chang, E., Frolkis, M., Kusler, B., Sage, M., Beeche, M., Bodnar, A.G., Wahl, G.M., Tlsty, T.D. and Chiu, C.P.: Telomerase expression in human somatic cells does not induce changes associated with a transformed phenotype. Nat Genet 21 (1999) 111-4.

Jimenez, J., Alphey, L., Nurse, P. and Glover, D.M.: Complementation of fission yeast cdc2ts and cdc25ts mutants identifies two cell cycle genes from Drosophila: a cdc2 homologue and string. Embo J 9 (1990) 3565-71.

Johnson, R.D. and Jasin, M.: Double-strand-break-induced homologous recombination in mammalian cells. Biochem Soc Trans 29 (2001) 196-201.

Jones, J., Nivitchanyong, T., Giblin, C., Ciccarone, V., Judd, D., Gorfien, S., Krag, S.S. and Betenbaugh, M.J.: Optimization of tetracycline-responsive recombinant protein production and effect on cell growth and ER stress in mammalian cells. Biotechnol Bioeng 91 (2005) 722-32.

Kasim, V., Miyagishi, M. and Taira, K.: Control of siRNA expression using the Cre-loxP recombination system. Nucleic Acids Res 32 (2004) e66.

Keyvani, K., Baur, I. and Paulus, W.: Tetracycline-controlled expression but not toxicity of an attenuated diphtheria toxin mutant. Life Sci 64 (1999) 1719-24.

Kilby, N.J., Snaith, M.R. and Murray, J.A.: Site-specific recombinases: tools for genome engineering. Trends Genet 9 (1993) 413-21.

Kolb, A.F.: Genome engineering using site-specific recombinases. Cloning Stem Cells 4 (2002) 65-80.

Kothary, R., Clapoff, S., Darling, S., Perry, M.D., Moran, L.A. and Rossant, J.: Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. Development 105 (1989) 707-14.

Krestel, H.E., Shimshek, D.R., Jensen, V., Nevian, T., Kim, J., Geng, Y., Bast, T., Depaulis, A., Schonig, K., Schwenk, F., Bujard, H., Hvalby, O., Sprengel, R. and Seeburg, P.H.: A genetic switch for epilepsy in adult mice. J Neurosci 24 (2004) 10568-78.

Kuhn, R., Schwenk, F., Aguet, M. and Rajewsky, K.: Inducible gene targeting in mice. Science 269 (1995) 1427-9.

Kuo, W.L., Gehm, B.D., Rosner, M.R., Li, W. and Keller, G.: Inducible expression and cellular localization of insulin-degrading enzyme in a stably transfected cell line. J Biol Chem 269 (1994) 22599-606.

Labow, M.A., Baim, S.B., Shenk, T. and Levine, A.J.: Conversion of the lac repressor into an allosterically regulated transcriptional activator for mammalian cells. Mol Cell Biol 10 (1990) 3343-56.

Lakso, M., Sauer, B., Mosinger, B., Jr., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L. and Westphal, H.: Targeted oncogene activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci U S A 89 (1992) 6232-6.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al.: Initial sequencing and analysis of the human genome. Nature 409 (2001) 860-921.

Langston, L.D. and Symington, L.S.: Opposing roles for DNA structure-specific proteins Rad1, Msh2, Msh3, and Sgs1 in yeast gene targeting. Embo J 24 (2005) 2214-23.

Lauth, M., Spreafico, F., Dethleffsen, K. and Meyer, M.: Stable and efficient cassette exchange under non-selectable conditions by combined use of two site-specific recombinases. Nucleic Acids Res 30 (2002) e115.

Lee, N.S., Dohjima, T., Bauer, G., Li, H., Li, M.J., Ehsani, A., Salvaterra, P. and Rossi, J.: Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. Nat Biotechnol 20 (2002) 500-5.

Liang, F., Romanienko, P.J., Weaver, D.T., Jeggo, P.A. and Jasin, M.: Chromosomal double-strand break repair in Ku80-deficient cells. Proc Natl Acad Sci U S A 93 (1996) 8929-33.

Lim, D.S. and Hasty, P.: A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. Mol Cell Biol 16 (1996) 7133-43.

Lin, X., Yang, J., Chen, J., Gunasekera, A., Fesik, S.W. and Shen, Y.: Development of a tightly regulated U6 promoter for shRNA expression. FEBS Lett 577 (2004) 376-80.

Liu, H.S., Scrable, H., Villaret, D.B., Lieberman, M.A. and Stambrook, P.J.: Control of Ha-ras-mediated mammalian cell transformation by Escherichia coli regulatory elements. Cancer Res 52 (1992) 983-9.

Ludwig, T., Fisher, P., Murty, V. and Efstratiadis, A.: Development of mammary adenocarcinomas by tissue-specific knockout of Brca2 in mice. Oncogene 20 (2001) 3937-48.

Luetke, K.H., Zhao, B.P. and Sadowski, P.D.: Asymmetry in Flp-mediated cleavage. Nucleic Acids Res 25 (1997) 4240-9.

Luo, G., Yao, M.S., Bender, C.F., Mills, M., Bladl, A.R., Bradley, A. and Petrini, J.H.: Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. Proc Natl Acad Sci U S A 96 (1999) 7376-81.

Lupton, S.D., Brunton, L.L., Kalberg, V.A. and Overell, R.W.: Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. Mol Cell Biol 11 (1991) 3374-8.

Mahonen, A.J., Airenne, K.J., Lind, M.M., Lesch, H.P. and Yla-Herttuala, S.: Optimized

self-excising Cre-expression cassette for mammalian cells. Biochem Biophys Res Commun 320 (2004) 366-71.

Mansour, S.L., Thomas, K.R. and Capecchi, M.R.: Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature 336 (1988) 348-52.

Mansour, S.L.: Gene targeting in murine embryonic stem cells: introduction of specific alterations into the mammalian genome. Genet Anal Tech Appl 7 (1990) 219-27.

Matsukura, S., Jones, P.A. and Takai, D.: Establishment of conditional vectors for hairpin siRNA knockdowns. Nucleic Acids Res 31 (2003) e77.

Matthess, Y., Kappel, S., Spankuch, B., Zimmer, B., Kaufmann, M. and Strebhardt, K.: Conditional inhibition of cancer cell proliferation by tetracycline-responsive, H1 promoter-driven silencing of PLK1. Oncogene 24 (2005) 2973-80.

Metzger, D. and Feil, R.: Engineering the mouse genome by site-specific recombination. Curr Opin Biotechnol 10 (1999) 470-6.

Meyer-Ficca, M.L., Meyer, R.G., Kaiser, H., Brack, A.R., Kandolf, R. and Kupper, J.H.: Comparative analysis of inducible expression systems in transient transfection studies. Anal Biochem 334 (2004) 9-19.

Migliaccio, A.R., Bengra, C., Ling, J., Pi, W., Li, C., Zeng, S., Keskintepe, M., Whitney, B., Sanchez, M., Migliaccio, G. and Tuan, D.: Stable and unstable transgene integration sites in the human genome: extinction of the Green Fluorescent Protein transgene in K562 cells. Gene 256 (2000) 197-214.

Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E. and Hayakawa, T.: IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. Mol Ther 1 (2000) 376-82.

Mohammadi, S., O'Malley, M., Etemad, A., Wang, Z., Xiao, X. and Lotze, M.T.: Second-generation tetracycline-regulatable promoter: repositioned tet operator elements optimize

transactivator synergy while shorter minimal promoter offers tight basal leakiness. J Gene Med 6 (2004) 817-28.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M. and Dreyfuss, G.: miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev 16 (2002) 720-8.

Muller, U.: Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. Mech Dev 82 (1999) 3-21.

Mulligan, R.C. and Berg, P.: Selection for animal cells that express the Escherichia coli gene coding for xanthine-guanine phosphoribosyltransferase. Proc Natl Acad Sci U S A 78 (1981) 2072-6.

Niederlander, C., Walsh, J.J., Episkopou, V. and Jones, C.M.: Arkadia enhances nodal-related signalling to induce mesendoderm. Nature 410 (2001) 830-4.

No, D., Yao, T.P. and Evans, R.M.: Ecdysone-inducible gene expression in mammalian cells and transgenic mice. Proc Natl Acad Sci U S A 93 (1996) 3346-51.

Ochman, H., Ajioka, J.W., Garza, D. and Hartl, D.L.: Inverse polymerase chain reaction. Biotechnology (N Y) 8 (1990) 759-60.

Ow, D.W.: Recombinase-directed plant transformation for the post-genomic era. Plant Mol Biol 48 (2002) 183-200.

Paddison, P.J. and Hannon, G.J.: RNA interference: the new somatic cell genetics? Cancer Cell 2 (2002) 17-23.

Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. and Conklin, D.S.: Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes Dev 16 (2002) 948-58.

Paddison, P.J., Caudy, A.A. and Hannon, G.J.: Stable suppression of gene expression by RNAi in mammalian cells. Proc Natl Acad Sci U S A 99 (2002) 1443-8.

Paulus, W., Baur, I., Oberer, D.M., Breakefield, X.O. and Reeves, S.A.: Regulated expression of the diphtheria toxin A gene in human glioma cells using prokaryotic transcriptional control elements. J Neurosurg 87 (1997) 89-95.

Pease, S. and Williams, R.L.: Formation of germ-line chimeras from embryonic stem cells maintained with recombinant leukemia inhibitory factor. Exp Cell Res 190 (1990) 209-11.

Pham, C.T., MacIvor, D.M., Hug, B.A., Heusel, J.W. and Ley, T.J.: Long-range disruption of gene expression by a selectable marker cassette. Proc Natl Acad Sci U S A 93 (1996) 13090-5.

Pierce, E.A., Liu, Q., Igoucheva, O., Omarrudin, R., Ma, H., Diamond, S.L. and Yoon, K.: Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. Gene Ther 10 (2003) 24-33.

Pierce, A.J., Johnson, R.D., Thompson, L.H. and Jasin, M.: XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev 13 (1999) 2633-8.

Pikaart, M.J., Recillas-Targa, F. and Felsenfeld, G.: Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev 12 (1998) 2852-62.

Porter, A.: Controlling your losses: conditional gene silencing in mammals. Trends Genet 14 (1998) 73-9.

Porteus, M.H. and Carroll, D.: Gene targeting using zinc finger nucleases. Nat Biotechnol 23 (2005) 967-73.

Puttini, S., Ouvrard-Pascaud, A., Palais, G., Beggah, A.T., Gascard, P., Cohen-Tannoudji, M., Babinet, C., Blot-Chabaud, M. and Jaisser, F.: Development of a targeted transgenesis strategy in highly differentiated cells: a powerful tool for functional genomic analysis. J Biotechnol 116 (2005) 145-51.

Qu, Z., Thottassery, J.V., Van Ginkel, S., Manuvakhova, M., Westbrook, L., Roland-

Lazenby, C., Hays, S. and Kern, F.G.: 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 (2004) 61-73.

Rang, A., Heise, T. and Will, H.: Lack of a role of the interferon-stimulated response element-like region in interferon alpha -induced suppression of Hepatitis B virus in vitro. J Biol Chem 276 (2001) 3531-5.

Rang, A. and Will, H.: The tetracycline-responsive promoter contains functional interferon-inducible response elements. Nucleic Acids Res 28 (2000) 1120-5.

Rasheed, S., Nelson-Rees, W.A., Toth, E.M., Arnstein, P. and Gardner, M.B.: Characterization of a newly derived human sarcoma cell line (HT-1080). Cancer 33 (1974) 1027-33.

Rennel, E. and Gerwins, P.: How to make tetracycline-regulated transgene expression go on and off. Anal Biochem 309 (2002) 79-84.

Rickert, R.C., Roes, J. and Rajewsky, K.: B lymphocyte-specific, Cre-mediated mutagenesis in mice. Nucleic Acids Res 25 (1997) 1317-8.

Rigby, S.P., Griffiths, D.J., Weiss, R.A. and Venables, P.J.: Human retrovirus-5 proviral DNA is rarely detected in salivary gland biopsy tissues from patients with Sjogren's syndrome. Arthritis Rheum 40 (1997) 2016-21.

Rijkers, T., Van Den Ouweland, J., Morolli, B., Rolink, A.G., Baarends, W.M., Van Sloun, P.P., Lohman, P.H. and Pastink, A.: Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. Mol Cell Biol 18 (1998) 6423-9.

Rodriguez, C.I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A.F. and Dymecki, S.M.: High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet 25 (2000) 139-40.

Rohlmann, A., Gotthardt, M., Willnow, T.E., Hammer, R.E. and Herz, J.: Sustained

somatic gene inactivation by viral transfer of Cre recombinase. Nat Biotechnol 14 (1996) 1562-5.

Rouet, P., Smih, F. and Jasin, M.: Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14 (1994) 8096-106.

Rouet, P., Smih, F. and Jasin, M.: Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A 91 (1994) 6064-8.

Rucker, E.B. and Piedrahita, J.A.: Cre-mediated recombination at the murine whey acidic protein (mWAP) locus. Mol Reprod Dev 48 (1997) 324-31.

S, A.M. and Hawkins, R.E.: Efficient transgene regulation from a single tetracycline-controlled positive feedback regulatory system. Gene Ther 5 (1998) 76-84.

Sambrook, J., Maniatis, T. and Fritsch, E.F.: Molecular cloning : a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y, 1989.

Santerre, R.F., Allen, N.E., Hobbs, J.N., Jr., Rao, R.N. and Schmidt, R.J.: Expression of prokaryotic genes for hygromycin B and G418 resistance as dominant-selection markers in mouse L cells. Gene 30 (1984) 147-56.

Sauer, B. and Henderson, N.: The cyclization of linear DNA in Escherichia coli by site-specific recombination. Gene 70 (1988) 331-41.

Sauer, B.: Inducible gene targeting in mice using the Cre/lox system. Methods 14 (1998) 381-92.

Sawicki, J.A., Monks, B. and Morris, R.J.: Cell-specific ecdysone-inducible expression of FLP recombinase in mammalian cells. Biotechniques 25 (1998) 868-70, 872-5.

Schaft, J., Ashery-Padan, R., van der Hoeven, F., Gruss, P. and Stewart, A.F.: Efficient FLP recombination in mouse ES cells and oocytes. Genesis 31 (2001) 6-10.

Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S. and Capecchi, M.R.: Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. Proc Natl Acad Sci U S A 97 (2000) 13702-7.

Schonig, K., Schwenk, F., Rajewsky, K. and Bujard, H.: Stringent doxycycline dependent control of CRE recombinase in vivo. Nucleic Acids Res 30 (2002) e134.

Sclimenti, C.R., Thyagarajan, B. and Calos, M.P.: Directed evolution of a recombinase for improved genomic integration at a native human sequence. Nucleic Acids Res 29 (2001) 5044-51.

Scrable, H. and Stambrook, P.J.: Activation of the lac repressor in the transgenic mouse. Genetics 147 (1997) 297-304.

Sedivy, J.M. and Sharp, P.A.: Positive genetic selection for gene disruption in mammalian cells by homologous recombination. Proc Natl Acad Sci U S A 86 (1989) 227-31.

Semionov, A., Cournoyer, D. and Chow, T.Y.: 1,5-isoquinolinediol increases the frequency of gene targeting by homologous recombination in mouse fibroblasts. Biochem Cell Biol 81 (2003) 17-24.

Senecoff, J.F., Bruckner, R.C. and Cox, M.M.: The FLP recombinase of the yeast 2-micron plasmid: characterization of its recombination site. Proc Natl Acad Sci U S A 82 (1985) 7270-4.

Shaked, H., Melamed-Bessudo, C. and Levy, A.A.: High-frequency gene targeting in Arabidopsis plants expressing the yeast RAD54 gene. Proc Natl Acad Sci U S A 102 (2005) 12265-9.

Shmerling, D., Danzer, C.P., Mao, X., Boisclair, J., Haffner, M., Lemaistre, M., Schuler, V., Kaeslin, E., Korn, R., Burki, K., Ledermann, B., Kinzel, B. and Muller, M.: Strong and ubiquitous expression of transgenes targeted into the beta-actin locus by Cre/lox cassette replacement. Genesis 42 (2005) 229-35.

Shockett, P., Difilippantonio, M., Hellman, N. and Schatz, D.G.: A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. Proc Natl Acad Sci U S A 92 (1995) 6522-6.

Smith, M.C. and Thorpe, H.M.: Diversity in the serine recombinases. Mol Microbiol 44 (2002) 299-307.

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. and Kucherlapati, R.S.: Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature 317 (1985) 230-4.

Sonntag, K.C., Simantov, R., Kim, K.S. and Isacson, O.: Temporally induced Nurr1 can induce a non-neuronal dopaminergic cell type in embryonic stem cell differentiation. Eur J Neurosci 19 (2004) 1141-52.

Sonoda, E., Sasaki, M.S., Buerstedde, J.M., Bezzubova, O., Shinohara, A., Ogawa, H., Takata, M., Yamaguchi-Iwai, Y. and Takeda, S.: Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. Embo J 17 (1998) 598-608.

Soukharev, S., Miller, J.L. and Sauer, B.: Segmental genomic replacement in embryonic stem cells by double lox targeting. Nucleic Acids Res 27 (1999) e21.

Southern, P.J. and Berg, P.: Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J Mol Appl Genet 1 (1982) 327-41.

Stafa, A., Svetec, I.K. and Zgaga, Z.: Inactivation of the SGS1 and EXO1 genes synergistically stimulates plasmid integration in yeast. Food Technology and Biotechnology 43 (2005) 103-108.

Stemmer, W.P., Crameri, A., Ha, H.D., Brennan, T.M. and Heyneker, H.L.: Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene 164 (1995) 49-53.

Sternberg, N. and Hamilton, D.: Bacteriophage P1 site-specific recombination. I.

Recombination between loxP sites. J Mol Biol 150 (1981) 467-86.

Strathdee, C.A., McLeod, M.R. and Hall, J.R.: Efficient control of tetracycline-responsive gene expression from an autoregulated bi-directional expression vector. Gene 229 (1999) 21-9.

Struhl, K.: The new yeast genetics. Nature 305 (1983) 391-7.

Struhl, K., Cameron, J.R. and Davis, R.W.: Functional genetic expression of eukaryotic DNA in Escherichia coli. Proc Natl Acad Sci U S A 73 (1976) 1471-5.

Sui, G., Soohoo, C., Affar el, B., Gay, F., Shi, Y. and Forrester, W.C.: A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. Proc Natl Acad Sci U S A 99 (2002) 5515-20.

Sullivan, M.J., Carpenter, A.J. and Porter, A.C.: A 'select and swap' strategy for the isolation of clones with tightly regulated transgenes. Eur J Biochem 268 (2001) 1605-12.

Taniguchi, M., Sanbo, M., Watanabe, S., Naruse, I., Mishina, M. and Yagi, T.: Efficient production of Cre-mediated site-directed recombinants through the utilization of the puromycin resistance gene, pac: a transient gene-integration marker for ES cells. Nucleic Acids Res 26 (1998) 679-80.

Thomas, K.R. and Capecchi, M.R.: Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51 (1987) 503-12.

Thomson, J.G., Rucker, E.B., 3rd and Piedrahita, J.A.: Mutational analysis of loxP sites for efficient Cre-mediated insertion into genomic DNA. Genesis 36 (2003) 162-7.

Thorpe, H.M. and Smith, M.C.: In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proc Natl Acad Sci U S A 95 (1998) 5505-10.

Trott, D.A., Cuthbert, A.P., Todd, C.M., Themis, M. and Newbold, R.F.: Novel use of a selectable fusion gene as an "in-out" marker for studying genetic loss in mammalian

cells. Mol Carcinog 12 (1995) 213-24.

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

Urlinger, S., Helbl, V., Guthmann, J., Pook, E., Grimm, S. and Hillen, W.: The p65 domain from NF-kappaB is an efficient human activator in the tetracycline-regulatable gene expression system. Gene 247 (2000) 103-10.

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D. and Holmes, M.C.: Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435 (2005) 646-51.

van de Wetering, M., Oving, I., Muncan, V., Pon Fong, M.T., Brantjes, H., van Leenen, D., Holstege, F.C., Brummelkamp, T.R., Agami, R. and Clevers, H.: Specific inhibition of gene expression using a stably integrated, inducible small-interfering-RNA vector. EMBO Rep 4 (2003) 609-15.

Van den Plas, D., Ponsaerts, P., Van Tendeloo, V., Van Bockstaele, D.R., Berneman, Z.N. and Merregaert, J.: Efficient removal of LoxP-flanked genes by electroporation of Cre-recombinase mRNA. Biochem Biophys Res Commun 305 (2003) 10-5.

Vasquez, K.M., Marburger, K., Intody, Z. and Wilson, J.H.: Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci U S A 98 (2001) 8403-10.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts,

R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., et al.: The sequence of the human genome. Science 291 (2001) 1304-51.

Ventura, A., Meissner, A., Dillon, C.P., McManus, M., Sharp, P.A., Van Parijs, L., Jaenisch, R. and Jacks, T.: Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci U S A 101 (2004) 10380-5.

Waldman, B.C. and Waldman, A.S.: Illegitimate and homologous recombination in mammalian cells: differential sensitivity to an inhibitor of poly(ADP-ribosylation). Nucleic Acids Res 18 (1990) 5981-8.

Wang, W., Seki, M., Narita, Y., Sonoda, E., Takeda, S., Yamada, K., Masuko, T., Katada, T. and Enomoto, T.: Possible association of BLM in decreasing DNA double strand breaks during DNA replication. Embo J 19 (2000) 3428-35.

Wang, X., Zou, L., Zheng, H., Wei, Q., Elledge, S.J. and Li, L.: Genomic instability and endoreduplication triggered by RAD17 deletion. Genes Dev 17 (2003) 965-70.

Wang, Y., Krushel, L.A. and Edelman, G.M.: Targeted DNA recombination in vivo using an adenovirus carrying the cre recombinase gene. Proc Natl Acad Sci U S A 93 (1996) 3932-6.

Wassmann, K. and Benezra, R.: Mitotic checkpoints: from yeast to cancer. Curr Opin Genet Dev 11 (2001) 83-90.

Wells, K.D., Foster, J.A., Moore, K., Pursel, V.G. and Wall, R.J.: Codon optimization, genetic insulation, and an rtTA reporter improve performance of the tetracycline switch. Transgenic Res 8 (1999) 371-81.

Welsh, S. and Kay, S.A.: Reporter gene expression for monitoring gene transfer. Curr Opin Biotechnol 8 (1997) 617-22.

Wigler, M., Sweet, R., Sim, G.K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. and Axel, R.: Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell 16 (1979) 777-85.

Williams, T. and Tjian, R.: Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. Genes Dev 5 (1991) 670-82.

Williams, T. and Tjian, R.: Characterization of a dimerization motif in AP-2 and its function in heterologous DNA-binding proteins. Science 251 (1991) 1067-71.

Wong, E.A. and Capecchi, M.R.: Homologous recombination between coinjected DNA sequences peaks in early to mid-S phase. Mol Cell Biol 7 (1987) 2294-5.

Wood, K.V.: Bioluminescence and chemiluminescence current status. John Wailey and Sons LTD (1991).

Wutz, A. and Jaenisch, R.: A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell 5 (2000) 695-705.

Xia, X.G., Zhou, H., Ding, H., Affar el, B., Shi, Y. and Xu, Z.: An enhanced U6 promoter for synthesis of short hairpin RNA. Nucleic Acids Res 31 (2003) e100.

Xia, H., Mao, Q., Paulson, H.L. and Davidson, B.L.: siRNA-mediated gene silencing in vitro and in vivo. Nat Biotechnol 20 (2002) 1006-10.

Xiao, Y. and Weaver, D.T.: Conditional gene targeted deletion by Cre recombinase demonstrates the requirement for the double-strand break repair Mre11 protein in murine embryonic stem cells. Nucleic Acids Res 25 (1997) 2985-91.

Xu, X., Wagner, K.U., Larson, D., Weaver, Z., Li, C., Ried, T., Hennighausen, L., Wynshaw-Boris, A. and Deng, C.X.: Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat Genet 22 (1999) 37-43.

Yamaguchi-Iwai, Y., Sonoda, E., Buerstedde, J.M., Bezzubova, O., Morrison, C., Takata, M., Shinohara, A. and Takeda, S.: Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. Mol Cell Biol 18 (1998) 6430-5.

Yamamoto, A., Hen, R. and Dauer, W.T.: The ons and offs of inducible transgenic technology: a review. Neurobiol Dis 8 (2001) 923-32.

Yamana, Y., Maeda, T., Ohba, H., Usui, T., Ogawa, H.I. and Kusano, K.: Regulation of homologous integration in yeast by the DNA repair proteins Ku70 and RecQ. Mol Genet Genomics 273 (2005) 167-76.

Yanez, R.J. and Porter, A.C.: Therapeutic gene targeting. Gene Ther 5 (1998) 149-59.

Yanez, R.J. and Porter, A.C.: Gene targeting is enhanced in human cells overexpressing hRAD51. Gene Ther 6 (1999a) 1282-90.

Yanez, R.J. and Porter, A.C.: Influence of DNA delivery method on gene targeting frequencies in human cells. Somat Cell Mol Genet 25 (1999b) 27-31.

Yanez, R.J. and Porter, A.C.: A chromosomal position effect on gene targeting in human cells. Nucleic Acids Res 30 (2002) 4892-901.

Yanez, R.J. and Porter, A.C.: Differential effects of Rad52p overexpression on gene targeting and extrachromosomal homologous recombination in a human cell line. Nucleic Acids Res 30 (2002) 740-8.

Yao, T.P., Forman, B.M., Jiang, Z., Cherbas, L., Chen, J.D., McKeown, M., Cherbas, P. and Evans, R.M.: Functional ecdysone receptor is the product of EcR and Ultraspiracle genes. Nature 366 (1993) 476-9.

Yao, F., Svensjo, T., Winkler, T., Lu, M., Eriksson, C. and Eriksson, E.: Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. Hum Gene Ther 9 (1998) 1939-50.

Ye, K., Compton, D.A., Lai, M.M., Walensky, L.D. and Snyder, S.H.: Protein 4.1N binding to nuclear mitotic apparatus protein in PC12 cells mediates the antiproliferative actions of nerve growth factor. J Neurosci 19 (1999) 10747-56.

Yoder, J.A., Walsh, C.P. and Bestor, T.H.: Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13 (1997) 335-40.

Yoshida, K. and Morita, T.: Control of radiosensitivity of F9 mouse teratocarcinoma cells by regulation of histone H2AX gene expression using a tetracycline turn-off system. Cancer Res 64 (2004) 4131-6.

Yu, H.M., Liu, B., Chiu, S.Y., Costantini, F. and Hsu, W.: Development of a unique system for spatiotemporal and lineage-specific gene expression in mice. Proc Natl Acad Sci U S A 102 (2005) 8615-20.

Yu, T.S., Dandekar, M., Monteggia, L.M., Parada, L.F. and Kernie, S.G.: Temporally regulated expression of Cre recombinase in neural stem cells. Genesis 41 (2005) 147-53.

Yun, S., Lie, A.C.C. and Porter, A.C.: Discriminatory suppression of homologous recombination by p53. Nucleic Acids Res 32 (2004) 6479-89.

Zamecnik, P.C. and Stephenson, M.L.: Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proc Natl Acad Sci U S A 75 (1978) 280-4.

Zeng, Y., Wagner, E.J. and Cullen, B.R.: Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol Cell 9 (2002) 1327-33.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E. and Filipowicz, W.: Human Dicer

preferentially cleaves dsRNAs at their termini without a requirement for ATP. Embo J 21 (2002) 5875-85.

Zhang, Z. and Lutz, B.: Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein. Nucleic Acids Res 30 (2002) e90.

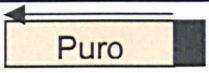
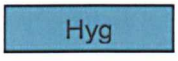
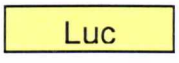
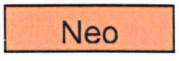
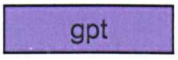
Zhang, X., Shu, M.A., Ross, H.E. and Kennedy, T.G.: Regulation of plasminogen activator in rat endometrial stromal cells: the role of prostaglandin E2. Biol Reprod 54 (1996) 1046-51.

Zhang, Y., Riesterer, C., Ayrall, A.M., Sablitzky, F., Littlewood, T.D. and Reth, M.: Inducible site-directed recombination in mouse embryonic stem cells. Nucleic Acids Res 24 (1996) 543-8.

Zhang, H., Fujitani, Y., Wright, C.V. and Gannon, M.: Efficient recombination in pancreatic islets by a tamoxifen-inducible Cre-recombinase. Genesis 42 (2005) 210-7.

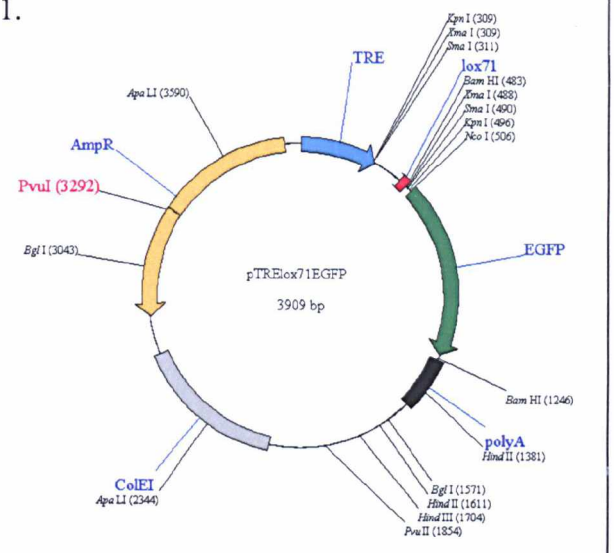
Zhou, H., Xia, X.G. and Xu, Z.: An RNA polymerase II construct synthesizes short-hairpin RNA with a quantitative indicator and mediates highly efficient RNAi. Nucleic Acids Res 33 (2005) e62.

# Appendix I: Key to commonly used symbols

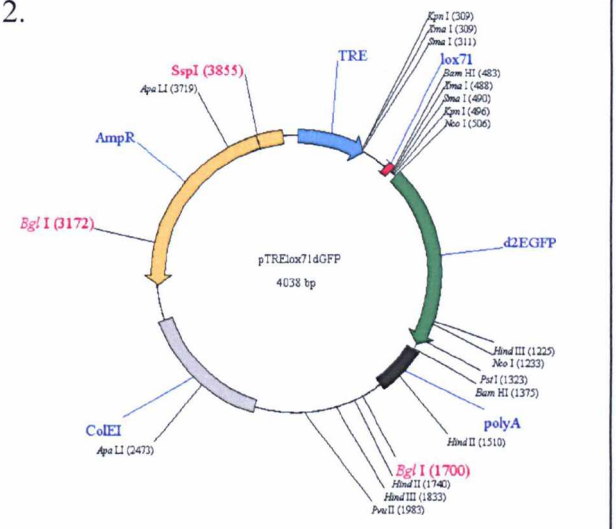
| Symbol: | Key: | Symbol: | Key: |
|---|---|---|---|
|  | Tetracycline Response Element (TRE) |  | Puromycin resistance gene |
|  | Modified Tetracycline Response Elements (TREmod) |  | Hygromycin resistance gene |
|  | *lox71* site |  | Gene coding for a gene of interest |
|  | *lox66* site |  | Multiple Cloning Site |
|  | Inactive mutant loxP site |  | Firefly *luciferase* gene |
|  | Wild-type *loxP* site |  | Neomycin resistance gene |
|  | GFP gene |  | Gene coding for *gpt* |
|  | Destabilised GFP gene |  | CMV minimal promoter |
|  | Internal Ribosomal Entry Site (IRES) |  | Tetracycline repressor |
|  | FRT site |  | Tetracycline transactivator |
|  | Gene not Expressed |  | Reverse transactivator |
|  | Gene expressed |  | Tetracycline |

# Appendix II: 'Target' Vectors

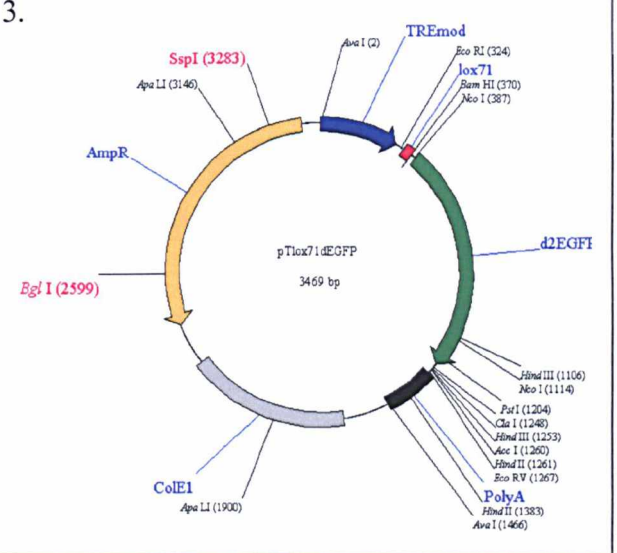
1.



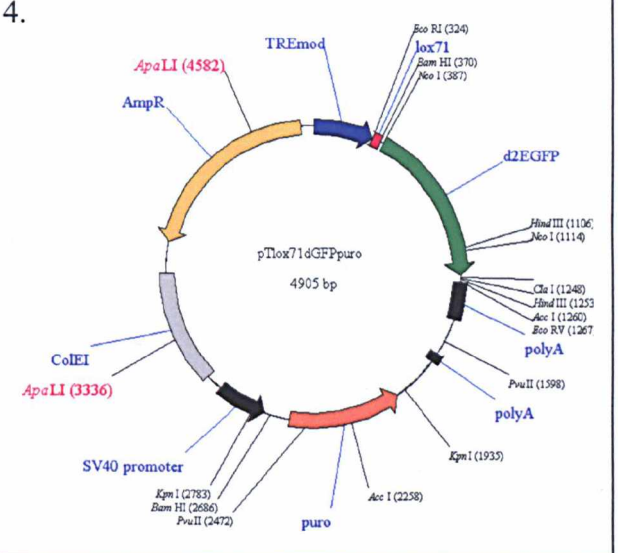

2.



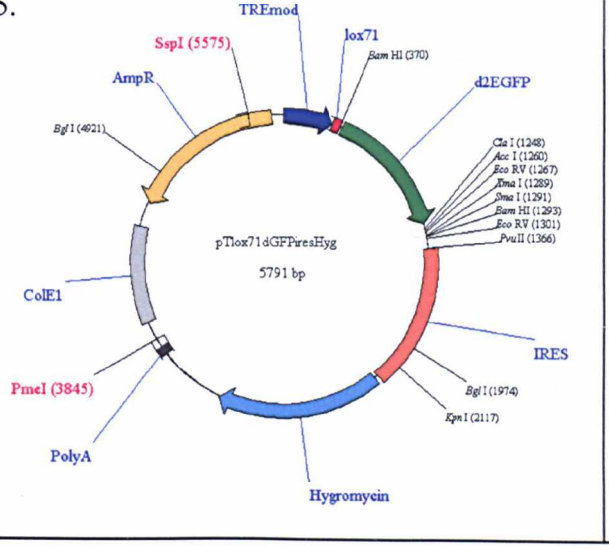

3.



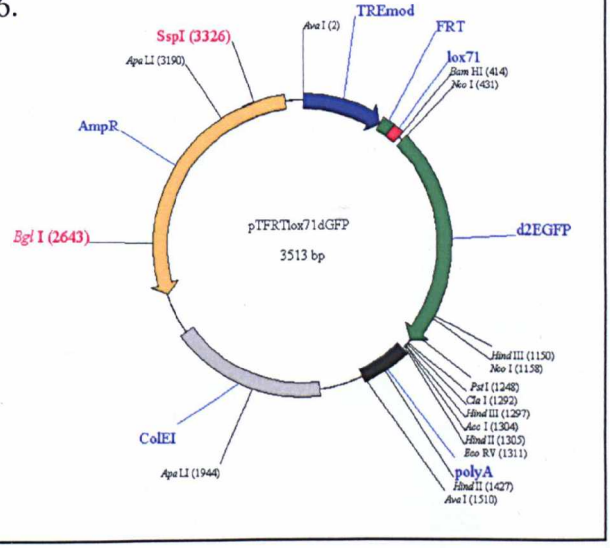

4.



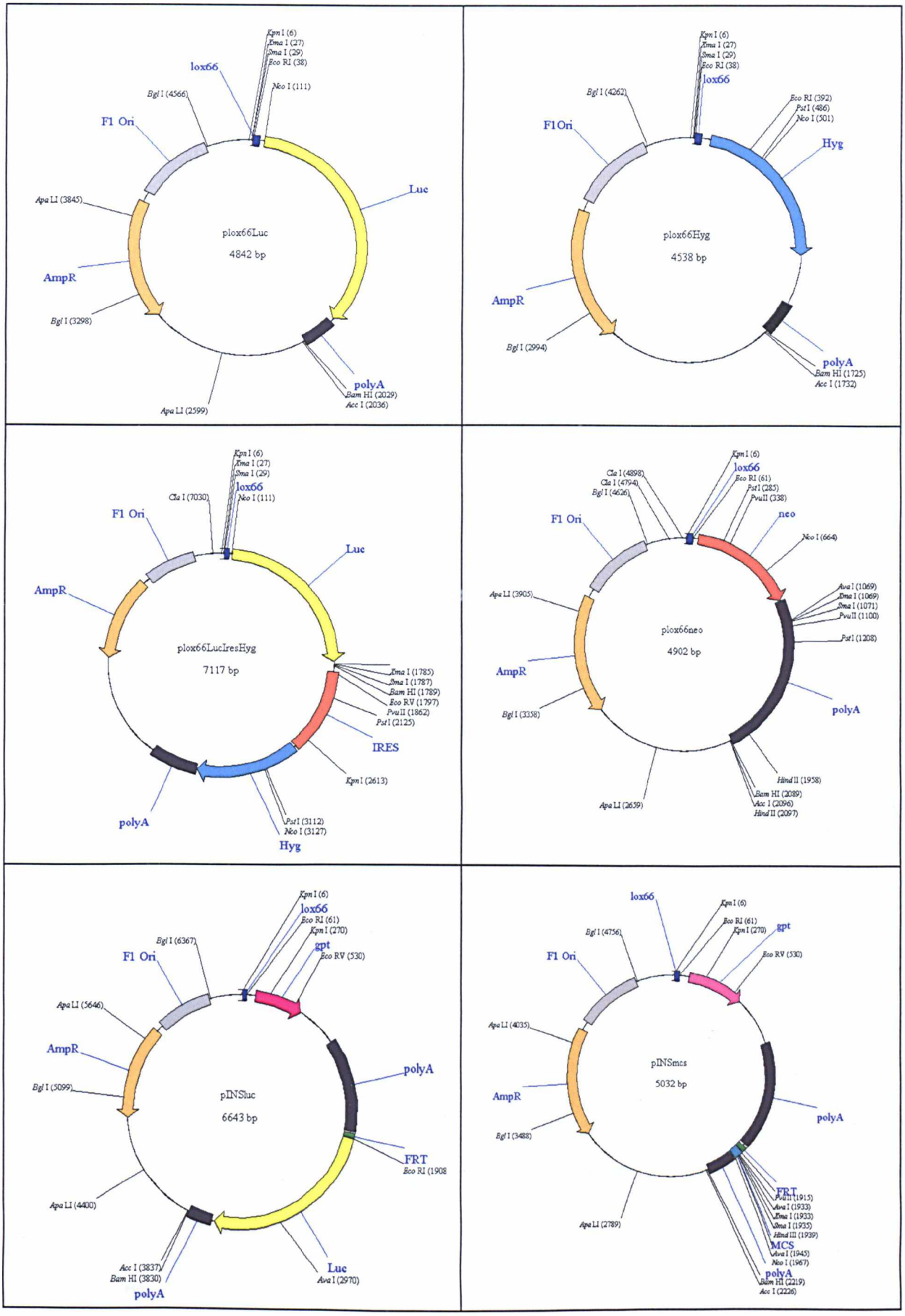

5.

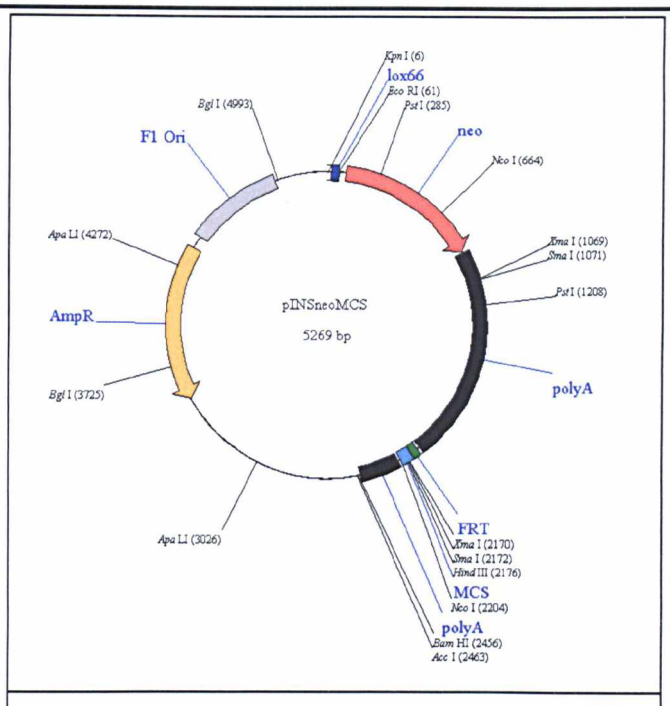


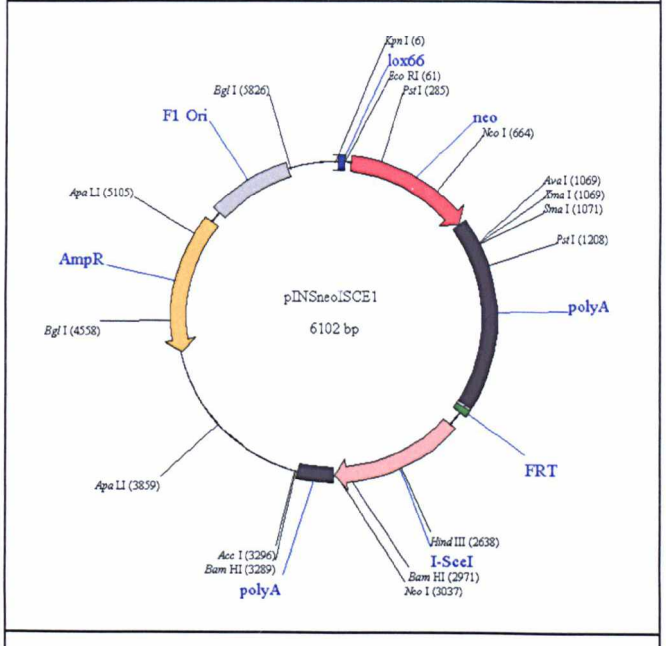

6.

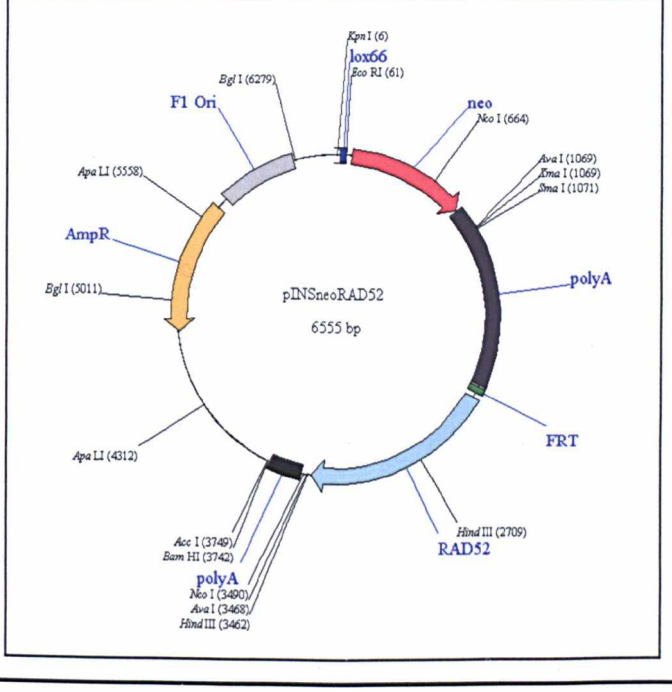


# Appendix III: 'Insertion' Vectors

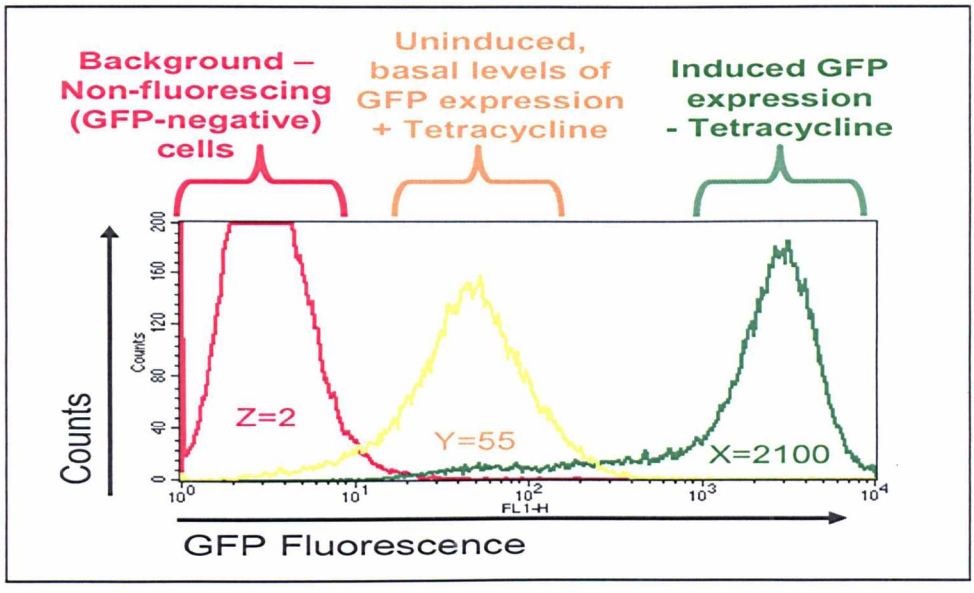



pINSneoMCS
5269 bp
Kpn I (6)
loxP
Eco RI (61)
Pst I (285)
neo
Nco I (664)
Xma I (1069)
Sma I (1071)
Pst I (1208)
polyA
FRT
Xma I (2170)
Sma I (2172)
Hind III (2176)
MCS
Nco I (2204)
polyA
Bam HI (2456)
Acc I (2463)
Apa LI (3026)
Bgl I (3725)
AmpR
Apa LI (4272)
F1 Ori
Bgl I (4993)


pINSneoISCE1
6102 bp
Kpn I (6)
loxP
Eco RI (61)
Pst I (285)
neo
Nco I (664)
Ava I (1069)
Xma I (1069)
Sma I (1071)
Pst I (1208)
polyA
FRT
Hind III (2638)
I-SceI
Bam HI (2971)
Nco I (3037)
Acc I (3296)
Bam HI (3289)
polyA
Apa LI (3859)
Bgl I (4558)
AmpR
Apa LI (5105)
F1 Ori
Bgl I (5826)


pINSneoRAD52
6555 bp
Kpn I (6)
loxP
Eco RI (61)
neo
Nco I (664)
Ava I (1069)
Xma I (1069)
Sma I (1071)
polyA
FRT
Hind III (2709)
RAD52
Acc I (3749)
Bam HI (3742)
polyA
Nco I (3490)
Ava I (3468)
Hind III (3462)
Apa LI (4312)
Bgl I (5011)
AmpR
Apa LI (5558)
F1 Ori
Bgl I (6279)

# Appendix IV: Interpreting a FACS plot




## Methods for calculating a clones GFP induction value

There are two methods to calculate the induction value:

1. Non-corrected
2. Corrected

$$\text{Non-Corrected method} = \frac{\text{Mean Induced GFP Expression (X)}}{\text{Mean Uninduced GFP Expression (Y)}} = \frac{2100}{55} = 38.2\text{-fold}$$

$$\text{Corrected Method} = \frac{\{X - \text{background (Z)}\}}{\{Y - \text{background (Z)}\}} = \frac{\{2100 - 2\}}{\{55 - 2\}} = \frac{1098}{53} = 39.6\text{-fold}$$

The difference between the two methods is minimal for a FACS plot such as the one above, however, for a clone which turns completely off when uninduced (i.e. Y=2)….

$$\text{Non-Corrected method} = \frac{\text{Mean Induced GFP Expression (X)}}{\text{Mean Uninduced GFP Expression (Y)}} = \frac{2100}{2} = 1050\text{-fold}$$

$$\text{Corrected Method} = \frac{\{X - \text{background (Z)}\}}{\{Y - \text{background (Z)}\}} = \frac{\{2100 - 2\}}{\{2 - 2\}} = \frac{2098}{0} = \infty \text{ (infinity) -fold}$$

……. the corrected method reaches infinity and it becomes difficult to compare the induction values between clones.

** During this study only non-corrected induction values were calculated **