Characterisation of Transgene Expression in Human Embryonic Stem Cells by Combining Gene Targeting & Site-Specific Recombination

Alexandra Isabelle Di Domenico

Thesis presented for degree of
Doctor of Philosophy

The University of Edinburgh

2006

Main supervisor: Dr Jim McWhir (Roslin Institute)

University supervisor: Professor John Ansell (Western General Hospital)

Declaration

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

Alexandra Di Domenico

Acknowledgments

I would like to express my sincere gratitude to Dr Jim McWhir who supported me during the course of this project, for his guidance, supervision and his encouragement. I would also like to thank Professor John Ansell, Dr Steve Pells, Dr Ioannis Christodoulou, Dr Alison Thomson, Davina Wojtacha, Dr Helen Priddle, Judy Fletcher, Anthea Spring, Dr Andreas Kolb, Dr Nick Forsyth, Ray Ansell, Dr Margot Marques, Dr Ed Gallagher, Dr Nathalie Beaujean for their invaluable help, discussion and advice. I am also grateful to Geron, and Dr David Russell, for biological material essential for this work.

Finally, I thank my husband Francesco, my mother-in-law Professor Cathy Di Domenico, my whole family and all my friends for their patience, encouragement and love.

This project was funded by Geron.

Para minha mãe querida...

Abstract

Predictable levels of transgene expression will be essential for full exploitation of human ES cells (hES) in basic research and medicine. In practice however, transgene expression in mammalian cells is often found to be silenced, low or variegated presenting a serious drawback in this area of research. Unpredictability of transgene expression arises mostly as a consequence of variable transgene copy number and/or inhibitory effects of surrounding chromatin structures in which the transgene has randomly integrated.

We have addressed this issue in hES cells, by combining gene targeting and site-specific recombination at the *β-casein* and the hypoxanthine phosphoribosyltransferase (*hprt*) genes. A *neomycin* expression cassette flanked by heterospecific *lox* sites is first introduced by homologous recombination (HR). In a subsequent step, recombinase-mediated cassette exchange (RMCE), permits the exchange of the *neomycin* gene with a transgene containing a blasticidin-o*ct-4*-GFP expression cassette flanked by the same set of heterospecific *lox* sites. Using a promoter trap strategy, site-specific integrants can be directly selected with blasticidin.

While we failed to target the *β-casein* gene in hES cells, we obtained homologous-recombinants with targeting frequencies of up to 0.4 % at the *hprt* gene in the male hES cell line (H1). *Hprt*-targeted H1 cells were shown to maintain a normal karyotype; to exhibit common hES cell surface markers; and to differentiate down the osteogenic lineage pathway. Additionally, subsequent targeted correction

of *hprt*-targeted H1 mutant cells was achieved, restoring structure and function of the *hprt* gene.

Following RMCE, site-specific integrant were recovered with 100 % efficiency using a promoter trap strategy at the *hprt* locus. However, site-specific integrants were characterised by a progressive down-regulation of GFP, possibly as a result of transcriptional interference from the *BSD* cassette or inhibitory elements present at the site of integration. Further work will be required to investigate the nature of this silencing.

Targeting transgenes with reliable transgene expression at specific loci in human ES cells will be important for modifying specific ES cell-derived tissues for therapeutic applications in regenerative medicine. Engineering the *hprt* locus should prove very useful for addressing issues such as directed differentiation, immunotolerance as well as protection against tumorigenesis.

Table of Contents

CHAPTER I

General Introduction

1.1. Project Overview

1.2. Transgene Expression in Mammalian Cells

1.3. Gene Targeting and the Cre/*LoxP* System

1.4. Genetic Modification of Human ES Cells

1.5. Project Objectives

1.1. Project Overview

The success of many transgenic applications relies on the stability of integration, expression and inheritance of the transgene or endogenous gene modification. Predictable levels of transgene expression in mammalian cells will be critical for many applications, including the correction of genetic defects in human somatic cells by gene therapy (Kren and Steer, 2002) and the genetic modification of large animals such as pigs for xenotransplantation and sheep for bioreactors (Niemann *et al.*, 2005). In addition, with the recent isolation of human embryonic stem cells (Thomson *et al.*, 1998), reliable transgene expression will provide an important asset to a broader range of transgenic and medical applications.

Very often however, the desired level of transgene expression in mammalian cells is not realised in practice, presenting a serious drawback for transgenic research and medicine. Transgene expression may be silenced or low, and frequently, expression is not seen in all cells-so-called variegated expression (Clark *et al.*, 1994). Unpredictability of transgene expression arises from the use of incomplete regulatory elements, as a consequence of variable transgene copy number or due to inhibitory effects of surrounding chromatin structures in which the transgene has randomly integrated.

One solution to overcome these problems, is to employ gene targeting which uses the cellular DNA repair mechanisms to introduce, by homologous recombination (HR), a transgene (provided that it contains homologous sequence to the target locus) into a specific site in the genome known to be appropriately

regulated. The advantage of using gene targeting at a known locus is that the transgene will be inserted as a single-copy, and will be equipped with the appropriate chromatin structure and regulatory elements therefore eliminating the silencing effects of multi-copy arrays and chromosomal position. Additionally, a site known for high level of transgene expression can be targeted repeatedly by HR providing a useful tool for transgenic approaches (Wallace *et al.*, 2000). Recombinase mediated cassette exchange (RMCE) constitutes an alternative approach to targeting repeatedly a locus with a single-copy transgene (Baer and Bode, 2001). In a first step, a cassette flanked by heterospecific *lox* sites (small prokaryotic palindromic DNA sequences which can only recombine between themselves) is introduced at a specific locus by HR. In the second step, the cassette can be exchanged with another cassette, also flanked by heterospecific *lox* sites, by site-specific recombination catalysed by Cre recombinase.

This PhD will address the possibility of generating a human ES cell line engineered to obtain predictable levels of transgene expression that can be used for a variety of applications. To achieve this aim, the advantages of homologous recombination with site-specific recombination will be combined to introduce a marker gene at the hypoxanthine-guanine phosphorybosyl transferase (*hprt*) or *β-casein* locus by RMCE. Following the desired recombination event, engineered cells will be characterised for their ability to provide reproducible tissue-specific transgene expression in hES cells.

The *hprt* gene, which is ubiquitously expressed, represents a good candidate locus for several reasons. Its disruption can provide a model for Lesh-Nyhan syndrome, which is a severe human genetic disease that is caused by the malfunction

of the *hprt* gene (Stout and Caskey, 1988). Additionally, *hprt* is X-linked and therefore a single targeted mutation in male hES cells, results in a complete loss of function. Finally, the characteristics of the purine metabolism pathway in which the HPRT enzyme is involved, provides a double selection system in which HPRT can be selected for (HAT selection), or against (6-TG selection) its activity. Therefore frequencies for a targeted event or correction event are easy to calculate at a phenotypic level (Albertini, 2001).

β-casein is a milk protein gene with no essential function out-with the mammary epithelium (Kumar *et al.*, 1994). Hence, the disruption of both copies of the gene, by transgene introduction, is very unlikely to have any detrimental effects in other non-mammary cells types. Additionally, high targeting frequencies have been achieved at this locus in mouse ES cells (Kumar *et al.*, 1994; kolb, 2001), suggesting that targeting the *β-casein* gene in human ES cells should be reproducible.

The following introduction will describe the present problems associated with transgene expression in mammalian cells, and some of the approaches that are currently employed to tackle them. Genetic modifications in ES cells by HR, or site-specific recombination, as well as the rationale behind designing gene targeting vectors will be highlighted. Human ES cell characteristics and their wide potential in research and medicine will also be reviewed. Within this context, this introduction will also highlight the implications in research and medicine of engineering specific genes such as the *β-casein* or *hprt* genes to generate a human ES cell line with reliable transgene expression. Finally the main objectives of this project will be outlined within this framework.

1.2. Transgene Expression in Mammalian Cells

1.2.1. Problems Associated with Transgene Expression

1.2.1.1. The Influence of Integration Site
1.2.1.2. The Influence of Copy Number and Methylation

1.2.2. Control of Transgene Expression

1.2.2.1. Inclusion of Regulatory Elements
1.2.2.2. Gene Targeting Approach

1.2.1. Problems Associated with Transgene Expression

Over the last twenty years transgenic technology has revolutionised biological research and a large number of transgenic plant and animal lines have now been created. The production of transgenic mice by injection of transgenes into one pronucleus of a fertilised zygote (Gordon *et al.*, 1980), or by injection of genetically modified embryonic stem (ES) cells into blastocysts or morula (Gossler *et al.*, 1986), has provided a powerful analytical tool for molecular biologists. Animal transgenesis has been applied, for instance, to study the function of housekeeping and developmental genes (reviewed by Melton, 1994; Shastry, 1998), to study gene regulation (reviewed by Porter, 1998), to generate a gain of function mutation (e.g. Ripps *et al.*, 1995), to generate animal models of human diseases (reviewed by Clarke, 1994; Dickinson *et al.*, 1995), to produce recombinant proteins in the milk of farm animals (Houdebine, 2000; Keefer, 2004), or to develop gene therapy strategies (Kren and Steer, 2002). The success of the above transgenic applications relies on the stability of integration, expression and inheritance of the transgene or endogenous gene modification. However, variability of transgene expression is often

observed mostly due to the site of integration in the host chromosome, and variable transgene copy numbers. Transgenes may express poorly or not at all, and in some cases, aberrant patterns of expression may be observed (reviewed by Gallie, 1998). The problems associated with transgene expression present a serious drawback for transgenesis and gene therapy, and the requirement to regulate genes appropriately remains one of the main challenges.

1.2.1.1. The Influence of Integration Site

Transgenes integrated randomly into the host genome of cultured cells (e.g. Feinstein *et al.*, 1982; Butner and Lo, 1986) and animals (Jaenisch *et al.*, 1981; Lacy *et al.*, 1983; Soriano *et al.,* 1986; Al-Shawi *et al.*, 1990) often display poor levels of expression, or exhibit inappropriate patterns of expression in the form of temporally and spatially (ectopic) aberrant expression. The primary reason for such problems is the chromosomal position effect (Wilson *et al.*, 1990; Clark *et al.*, 1994), whereby the particular genetic environment of the integration site is likely to influence the expression of the integrated transgene. Repressive chromosomal position effects are associated with the suppression of DNase hypersensitive sites (DHS) formation (Huber *et al.*, 1994; Festenstein *et al.*, 1996; Pikaart *et al.*, 1998), nucleosome remodelling (Wallrath and Elgin, 1995), and histone deacetylation (Pikaart *et al.*, 1998). Chromosomal position effects form the basis of the "gene trap" technique which is increasingly being used for detecting patterns of gene expression (reviewed by Cecconi and Meyer, 2000; Stanford *et al.*, 2001). The general term gene trap refers to the random integration of a reporter gene construct (gene trap vector) into the genome such that integration events will activate the

reporter gene only if it is brought under the transcriptional regulation of an endogenous gene. While the gene trap approach is commonly used to isolate and mutate endogenous genes (Friedrich and Soriano, 1991; Wurst *et al.*, 1995), it would be relevant to take advantage of these "insertion sites" into genes for other purposes, such as the generation of tagged loci, with reproducible transgene expression, that can be repeatedly manipulated for a broad range of transgenic applications (Cobellis *et al.*, 2005).

Position effects are either stable, where transgenes are expressed in every cell but at different levels (Lewis, 1950); or variegated (Henikoff, 1990), where transgene expression level is maintained but only a proportion of genetically identical cells express the transgene. Position effect variegation (PEV) results from the integration of a transgene into a telomeric or centromeric region of heterochromatin (condensed chromatin), rich in repeated sequences and poor in functional genes (Straub, 2003). Mosaic transgene expression (expression levels vary from cell to cell) arises as a consequence of variable heterochromatin spreading into euchromatic regions (decondensed chromatin structure which contains most of the genes) (reviewed by Henikoff, 1990; Karpen, 1994). The first description of PEV was published over 75 years ago (Muller, 1930). It was first observed as zonal variegation in the color of the facets of the eye of *Drosophila* that occurs when a pigmentation gene is brought into proximity with a region of constitutive heterochromatin by translocation, and was attributed to clonal variegation in the "spreading" of heterochromatinisation into the translocated chromosomal region. Since this early report, variegation has also been reported in plants (Matzke *et al.*,

2001), yeast (Tartof, 1994), fungi (Cogoni *et al.*, 1996), fish (Collas *et al.*, 1999) and mice (Sutherland *et al.*, 2000).

1.2.1.2. The Influence of Copy Number and Methylation

Randomly integrated transgenes are usually present in the form of multicopy arrays mainly in a head-to-tail fashion at one or, rarely, a few chromosomal sites per nucleus (reviewed by Smith, 2001). High copy number is more often associated with transgene silencing or variegated expression, as a result of heterochromatinisation of the transgene array (Dorer and Henikoff, 1997; Garrick *et al.,* 1998; Henikoffs, 1998; Matzke *et al*, 1999; McBurney *et al.*, 2002). In this case, heterochromatin formation arises from the generation of secondary structures through pairing of the copies of the transgene which results in the methylation of cytosines by methyltransferases. Methylation, which is often associated with gene silencing (Cedar, 1988), is involved as part of the genome defense system developed against parasitic sequences, including transposable elements (TEs), viroid and RNA viruses (Matzke *et al.,* 1999, 2000), and also controls the temporal and spatial expression of certain genes (Bird and Wolffe, 1999). Several mouse studies provided some evidence of a negative correlation between the number of transgene copies within an array, and the level of expression of the transgene (e.g. Garrick *et al.*, 1998; McBurney *et al.*, 2002). Garrick and co-workers (1998) used site-specific recombination (see section 1.3.4 for site-specific recombination) to generate transgenic mouse lines in which different copy numbers of a transgene were inserted at the same chromosomal location, thereby eliminating the contribution of position effects and allowing analysis of the effect of copy number alone on transgene

silencing. Using this approach, they showed that reduced transgene arrays, is correlated with increased levels of transcription and expression, and is accompanied by reduced methylation. Silencing of multiple copies has also been documented in fungi (Faugeron, 2000), plants (Flavell, 1994), *Drosophila* (Dorer and Henikoff, 1994), and nematodes (Kelly and Fire, 1998).

1.2.2. Control of Transgene Expression

Several strategies have been devised to overcome position and multi-copy array effects. In some cases, the solution lies with the vector design, for example by ensuring that an appropriate enhancer sequence is included in the transgene, or by including insulator elements to protect a transgene from position effect. Position effect can also be avoided through the use of gene targeting or by the identification of sites permissive for transgene expression. The section below focus on these strategies.

1.2.2.1. Inclusion of Regulatory Elements

Genes are thought to be organised on chromosomes as contiguous but independent units known as expression domains (Dillon and Grosveld, 1994). These expression domains are believed to remain insulated from neighbouring sequences and are thought to include all regulatory elements that are necessary for their correct temporal and spatial gene expression. This concept was confirmed by experiments in which either transgenes harboring the complete set of cis-regulatory elements (Bonifer *et al.,* 1990; 1994; Lien *et al.*, 1997) or large genomic fragments in the form of yeast or bacterial artificial chromosomes were analysed in transgenic mice. In

each case, correct regulation was observed (Schedl *et al.*, 1993; Strauss *et al.*, 1993; Lien *et al.*, 1997; Nielsen *et al.* 1998).

Some regulatory elements are capable of counteracting position and multi-copy array effects, and may be divided into two groups: dominant positive regulatory elements that impose an active chromatin configuration even in heterochromatic surroundings, and boundary elements that can physically block the passage of signals (either positive or negative) from the surrounding chromatin. Locus control regions (LCRs) and strong enhancers constitute the first group of regulatory elements, while matrix attachment region elements (MARs) and chromatin insulators belong to the second group.

LCRs are non-coding genetic elements that confer tissue-specific, position-independent and copy number-dependent expression of linked transgenes in the host genome (reviewed by Li *et al.*, 2002). LCRs were first described in the human *β-globin* genes cluster (Forrester *et al.*, 1987; Grosveld *et al.*, 1987), which is composed of five genes (ε, $^{G}\gamma$, $^{A}\gamma$, δ and β), spread across 32 Kb of DNA, and are arranged in the order of their developmental expression. A 5' and 3' LCR lies at both ends of the cluster. The existence of LCRs first came to attention with data obtained from studies of Hispanic β-thalassemia patients, in which a large deletion (encompassing the 5' LCR) upstream of the cluster led to suppression of the *β-globin* genes expression (Kioussis *et al.*, 1983; Driscoll *et al.*, 1989). LCRs are associated with erythroid-specific DNase I hypersensitive sites (HSs), characterised by distinguishable functions, including transcriptional enhancers, domain-opening and tissue-specific elements. LCRs are either clustered as for the human *β-globin* locus (Grosveld *et al.*, 1987) or scattered throughout the domain as exemplified by the

chicken *lysosyme* locus (Bonifer *et al.*, 1994). Although the human *β-globin* LCR is the most extensively characterised element, other LCRs have also been described in a broad spectrum of mammalian gene systems, including humans, mice, rats, chickens, rabbits, sheep, and goats (reviewed by Li *et al.*, 2002). The discovery of LCR upstream of the human *β-globin* genes, with its ability to confer copy number-dependent gene expression independent of the integration site, originally made this element a promising candidate for incorporation in vectors for transgenic and gene therapy approaches (Grosveld *et al.*, 1987; Sadelain *et al.*, 1995). However, the use of LCR regulatory elements is limited, as they are only known for a few genes, and are tissue-specific. Additionally, their ability to reliably confer position-independent expression of a linked transgene in cells and in mice is not always true (Alami *et al.*, 2000; Ellis and Pannell, 2001).

MARs can be defined as cis-acting elements located at the two boundaries of a transcription unit, which mediate the attachment of the chromatin to the nuclear scaffold. Eukaryotic DNA is organized into higher order chromatin structures, which in turn are periodically attached by MARs to a sub-nuclear matrix to form a series of 30-100 kb loops. Recent data reinforce the idea that MARs are in fact the elements that provide a higher order organization of the eukaryotic genome by matrix attachment (Goetze *et al.* 2003). The binding affinity of MARs to the nuclear matrix correlates with certain AT-rich motifs, which are thought to function as DNA-unwinding elements (Michalowski *et al.*, 1999; Bode *et al.*, 2000a; Goetze *et al.*, 2003). Unlike LCRs, MARs are not bound by known transcription factors, and instead are associated with by topoisomerases I and II, histone H1 and a number of proteins not known in other contexts. The first MAR element to be characterized

was in the chicken *lysosyme* locus (Stief *et al.*, 1989). The chicken lysosyme gene was shown to be expressed in a position-independent manner in transgenic mice (Bonifer *et al.*, 1990). MARs have been shown to increase transgene expression both in stably transfected cells and in transgenic mice, and are believed to collaborate with enhancers to generate an extended domain of accessible chromatin. They have also been shown to decrease *de novo* methylation of transgenes (Bode *et al.*, 2000c; Dang *et al.*, 2000). The advantage of using MARs over LCRs, is that they are not tissue-specific, however they do not confer convincing copy-number dependence (McKnight *et al.*, 1992) and reports have been inconsistent to date.

Insulators are boundary DNA elements that exert their function by shielding genes from either the action of a distal enhancer and/or from the silencing effects of condensed chromatin (reviewed by West *et al.*, 2002; Burgess-Beusse *et al.*, 2002) (figure 1.1). Insulators were first studied in *Drosophila* and have now been described in many other species including yeast, sea urchins, *Xenopus*, chickens, mice and humans (West *et al.*, 2002). In vertebrates, the first best-characterised insulator is the constitutive 1.2 kb 5' HS4, which is located near the 5' end of the chicken *β-globin* locus (Chung *et al.*, 1993). This element is located within a constitutive DNase I hypersensitive site between an open chromatin conformation and 16 kb of condensed chromatin (figure 1.2). Transgenes flanked with copies of the 5'HS4 insulator has proven very useful in the generation of transgenic mice (Wang *et al*, 1997; Potts *et al*, 2000), rabbits (Taboit-Dameron *et al.*, 1999) and cell lines (Pikaart *et al.*, 1998; Inoue *et al.*, 1999; Emery *et al.*, 2000; Rivella *et al.*, 2000; Recillas-Targa *et al.*, 2002) with uniform transgene expression in all tissue type (reviewed by Recillas-Targa *et al.*, 2004).

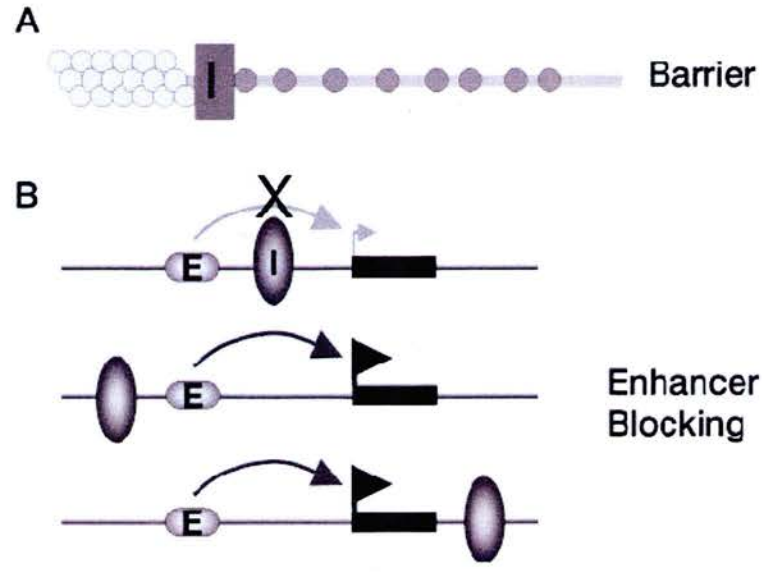


Figure 1.1. Barrier and enhancer blocking function of insulators. **(A)** Some insulators function as barrier against surrounding heterochromatin. **(B)** Some insulators function as enhancer-blocking elements that prevent enhancer action when placed between enhancer and promoter (tagged box) but not otherwise. I: insulator, E: enhancer. (Picture adapted from Burgess-Beusse *et al.*, 2002).

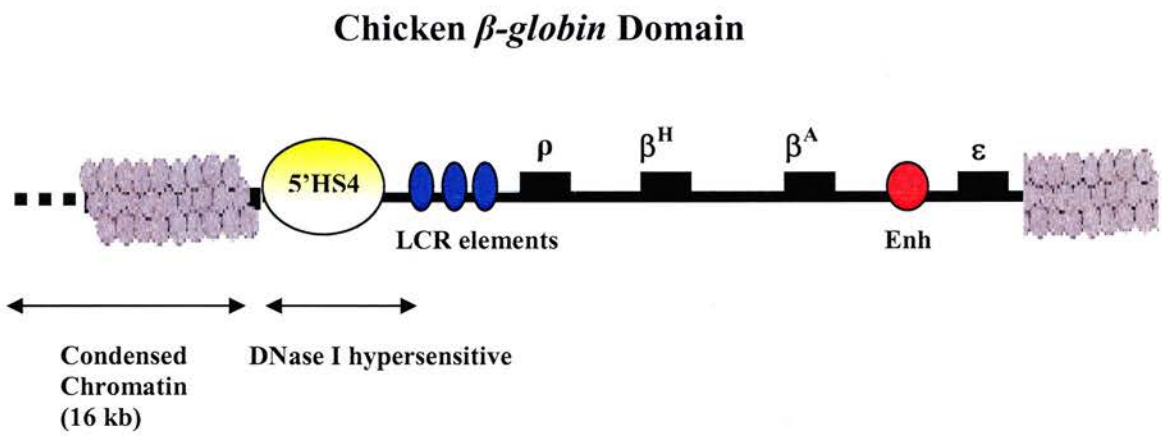


Figure 1.2. Chicken *β-globin* domain. The locus is characterised by four genes (ε, β^A, β^H and ρ), a strong enhancer (enh), LCR elements and the constitutive 5'HS4 insulator flanking 16 kb of condensed chromatin.

There are currently three models of insulator functions (Recillas-Targa and Razin, 2001; West *et al.*, 2002; Burgess-Beusse *et al.*, 2002). (i) The insulator actively recruits histone acetylases (and perhaps other activities associated with active chromatin), maintaining a histone acetylated (open) chromatin configuration within the domain, (ii) the insulator captures or titrates DNA methyltransferases and/or histone deacetylases and methylases, avoiding gene silencing within the protected domain and thus preventing heterochromatin-mediated domain repression, and (iii) the insulator relocates the gene away from heterochromatic regions to an active subnuclear compartment, perhaps in association with the nuclear periphery and nuclear pore components (Gerasimova *et al.*, 2000; Ishii *et al.*, 2002).

Although insulators represent a promising approach in animal transgenesis and gene therapy approaches, it is likely that the use of complementary chromatin elements such as LCRs and MARs, will be necessary to ensure reliable transgene expression (Ramezani et *al.*, 2003; Ma *et al.* 2003).

1.2.2.2. Gene Targeting Approach

An alternative transgenic approach to obtain reproducible transgene expression is the use of gene targeting, in which transgenes can be introduced to a precise location in the genome by homologous recombination (reviewed by Jasin *et al.,* 1996) (see section 1.3.1 for mechanisms). Gene targeting provides the advantage that a single-copy of the transgene is introduced into a specific gene, therefore eliminating the silencing effects of multi-copy arrays (Bronson *et al.*, 1996). Additionally, gene targeting allows the selection of the site of integration and so overcomes the problems associated with position effects owing to random

integration. Thus a site known for high level of transgene expression can be targeted repeatedly providing a useful tool for a broad range of transgenic approaches (Wallace *et al.*, 2000). The ubiquitously expressed hypoxanthine phosphoribosyl transferase (*hprt*) gene played an important role in the initial development of gene targeting strategies in murine ES cells (Thomas and Capecchi, 1987; Doetschman *et al.,* 1987) because this gene provides a positive and negative selection system (see section 1.4.4.2 for more details), and mice containing the disruption are viable (Hooper *et al.*, 1987; Kuehn *et al.*, 1987). The *hprt* gene is located on the X chromosome and is hemizygous in XY ES cells, thus integration of a transgene that disrupts the *hprt* gene can be directly selected for in the appropriate media (Albertini, 2001). Bronson *et al.* (1996) restored an *hprt*-deficient ES cell line, which contains a 5' gene deletion, by inserting a human or mouse *β-actin* promoter driving the murine *bcl-2* cDNA upstream of the *hprt* locus. Both transgenes were highly expressed *in vitro* and in mice, although the human *β-actin* promoter directed consistently a higher level of transgene expression. Although reproducible transgene expression at the *hprt* locus has been observed on several occasions (Vivian *et al.*, 1999; Evans *et al.*, 2000; Guillot *et al.,* 2000; Cvetkovic *et al.,* 2000; Minami *et al.*, 2002; Heaney *et al.*, 2004), variegated transgene expression has also been reported (Shaw-White *et al.*, 1993; Evans *et al.*, 2000).

Gene targeting has also been used to drive reliable transgene expression from endogenous promoters. A promoterless *lacZ* transgene has been frequently targeted into loci to serve as a tag to monitor the expression of the locus (e.g. Mansour *et al.*, 1990; Le Mouellic *et al.*, 1990; Mountford *et al.*, 1994; Tajbakhsh and Buckingham, 1994; Weiler-Guettler *et al.*, 1996). Variations of this

approach include targeting a transgene into an endogenous locus to restrict its expression to a certain cell type (Zhou and Palmiter, 1995), and to replace a mouse gene with its human homologue (Stacey *et al.*, 1994; Detloff *et al.*, 1994).

The site-specific recombination system (Fukushije and Sauer, 1992; Baubonis and Sauer, 1993; Kolb and Siddell, 1996) offers an alternative approach to genetically manipulate ES cells and mice (see section 1.3.4). The introduction in ES cells of a *loxP* site from bacteriophage P1, which is recognised by the site-specific Cre recombinase, allows in a subsequent step (and in the presence of Cre recombinase) site-specific insertion of transgenes which carry flanking *loxP* sites. Used in conjunction with homologous recombination, site-specific recombination allows transgenes to be targeted to a specific site in the genome. Rucker and Piedrahita have exemplified this approach in 1997, with the introduction of a marker into the whey acidic protein gene. In a later report, Kolb and co-workers (1999) used a similar strategy to introduce a single copy of a promoterless luciferase reporter gene under the control of the endogenous *β-casein* promoter in murine ES cells. Recently, Shmerling *et al.* (2005) used recombinase-mediated cassette exchange (see section 1.3.4.1) and showed that several reporter genes could be reliably expressed when placed under the transcriptional control of the endogenous *β-actin* promoter in murine ES cells.

A variation on the gene targeting approach involves an *in vitro* pre-screen of tagged random sites permissive for reproducible transgene expression, followed by the introduction of a single-copy transgene by gene targeting (Wallace *et al.*, 2000) or site-specific recombination (Fukushige and Sauer, 1992). Fukushige and Sauer (1992) introduced a single copy of a promoterless *lox-neo* fusion gene into

two independent random sites in the genome of Chinese hamster ovary (CHO) cells. Subsequent Cre-mediated site-specific recombination with a plasmid containing a *loxP* site, a constitutive promoter and a *β-actin-lacZ* transgene, was detected by the activation of the *neo* marker. Highly reproducible *lacZ* expression was observed among the site-specific integrants obtained at the same site. Variation in levels of transgene expression was however dependent on the locus and the orientation of the *lacZ* gene in the locus.

Similarly, Wallace and co-workers (2000) conducted an *in vitro* screen for sites in mouse ES cells which allowed the expression of a *PGK-hprt* selectable marker lying 5' to an *oct-4-lacZ* transgene. Hundreds of clones were assessed *in vitro* for appropriate down-regulation of the *lacZ* transgene following differentiation (since the *oct-4* promoter is active in ES cells and inactive in differentiated cells). Two clones, containing a single copy of the transgene, were chosen which displayed appropriate and inappropriate *lacZ* expression. In a subsequent step, the *oct-4* promoter was sequentially replaced by three tissue-specific promoters (*thyroglobulin, Hox2.6 and Myf5*) using homologous recombination, to drive expression of *lacZ*. While one site provided reproducible transgene expression with the three promoters *in vitro* and *in vivo*, the other site showed inappropriate transgene expression.

1.3. Gene Targeting and the Cre/*LoxP* System

1.3.1. Models of Homologous Recombination

1.3.2. ES Cell Route to Transgenesis

1.3.3. Gene Targeting

1.3.3.1. Targeting Vectors
1.3.3.2. Screening for Targeted ES Cell Clones
1.3.3.3. Parameters Affecting Targeting Frequency
1.3.3.4. Enrichment Strategies
1.3.3.5. Generation of Subtle Mutations

1.3.4. The Cre/*LoxP* System

1.3.4.1. Cre/*LoxP* Site-Specific Recombination Mechanism
1.3.4.2. Applications

1.3.1. Homologous Recombination in Mammalian Cells

Homologous recombination (HR) refers to any process in which two similar DNA sequences (e.g. a homologous chromosome or a sister chromatid) interact and exchange genetic information. In addition to prokaryotic organisms, HR also occurs in eukaryotic organisms during meiosis, mitosis and DNA repair (reviewed by Bollag, 1989; Hooper, 1992). The process of HR repair provides a mechanism for the error-free removal of damage present in DNA that has replicated (S and G2 phases). Thus HR acts in coordination with the S and G2 checkpoint machinery, to eliminate chromosomal breaks on a DNA molecule using information from the sister DNA molecule before the cell division occurs. HR may occur between two

introduced DNA molecules (extra-chromosomal recombination), between homologous sequences in the chromosomes (intra-chromosomal and inter-chromosomal recombinations), and between an introduced sequence and a chromosomal sequence (gene targeting). In contrast with yeast, where HR is the primary mechanism of double-strand break repair (DSB), mammalian cells favour the non-homologous recombination pathway. This pathway joins the two ends of a DSB without regard to sequence homology. Although this process is more prone to error than HR, it might have evolved to prevent hyperrecombination in the mammalian genome, rich in DNA repeats.

The main current models of HR are based on lower eukaryotes such as yeasts and other fungi (Reviewed by Smith, 2002). All models involve the following events: Initiation, repair and resolution. Initiation involves alignment of homologous sequences, generation of cross-overs and branch migration. Repair takes place after any change or loss of genetic information that occurs prior to or during initiation, and involves mismatch repair or DNA synthesis. Resolution is the excision of the cross-overs and separation of the two recombinant molecules.

The Holliday model (Holliday, 1964) involves homology pairing; single-strand cleavage ("nicking") of both duplexes at a homologous site; strand invasion of free ends between duplexes, such that the crossed strands unite the duplexes in a structure called a holliday junction; branch migration of the holliday junction, moving the crossover point away from its original position; and resolution of the holliday junction by single strand nicking in either of the two senses observed in nature: Cross-overs (splices) or patches (recombinant regions that have not exchanged flanking markers) (figure 1.3). A modification of the holliday model,

Meselson-Radding model, suggests that only one of the two molecules initiates recombination (Meselson and Radding, 1975). The initiator is nicked in one of the strands, thus generating a 3' end for DNA synthesis. This reaction displaces the 5' end, which then invades a homologue. This results in the formation of a holliday junction that is resolved as in Holliday's model, to form splices or patches.

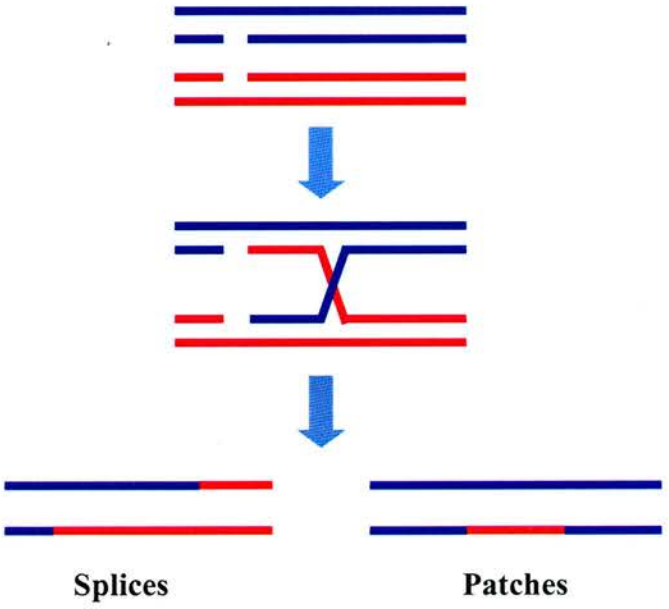


Figure 1.3. The Holliday model. See text for details.

The double-strand break repair model (Szostak *et al.*, 1983) was proposed to account for the observation that in many organisms, DNA DSBs enhance recombination locally. Earlier studies on plasmid-chromosome recombination in yeast (Orr-Weaver *et al.*, 1981, 1983) established the recombinagenic nature of DSBs and demonstrated the occurrence of gene conversion by double-strand gap repair. In this model (figure 1.4), HR is initiated by homology pairing and DSBs of the "receptor" duplex. 5'-3' exonuclease action results in the digestion of one of the strands at both ends of the cut, generating a gap with 3' overhangs. One of these invades the intact duplex ("donor"), creating a heteroduplex and displacing the complementary strand of the "donor" duplex. This generates a "D loop", which is subsequently expanded through DNA synthesis from the 3' end of the invading strand. Eventually, the loop will be large enough to base-pair with the 3' end of the receptor, which can then prime new DNA synthesis to reconstitute the missing strand from the donor template. Migration of the invading 3' end creates a molecule with two holliday junctions, that can be resolved in two ways to give either splices or patches (figure 1.4).

At the moment, the molecular and biochemical details of HR are poorly understood. Therefore it is not yet possible to decide which model is most accurate (initiation by single-strand or by double-strand breaks). Indeed, it has been suggested that multiple recombination pathways may exist (Sedivy and Joyner, 1993). Some mutants of *Drosophila* (Carpenter, 1982) and yeast (Engebrecht *et al.*, 1990) show normal levels of meiotic gene conversion (the correction of one strand of heteroduplex DNA to make it complementary with the other at mistmatch positions)

but reduced levels of crossing over, suggesting that not all gene conversion events are resolved as cross-overs.

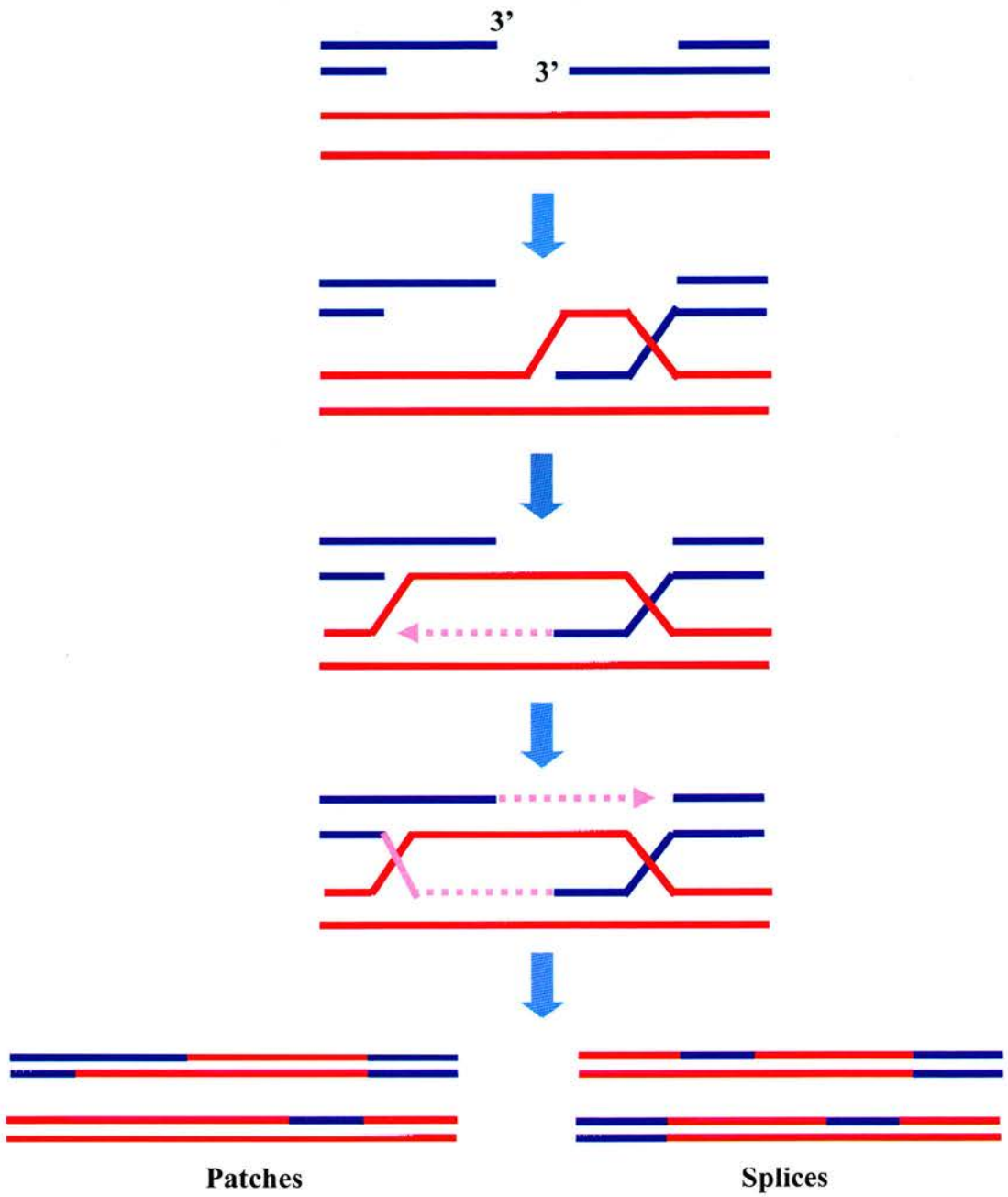


Figure 1.4. Double-Strand Breaks model. See text for details.

1.3.2. ES Cell Route to Transgenesis

The isolation of pluripotent murine ES cells (established *in vitro* from the inner cell mass of explanted blastocyst stage embryos) that can contribute to all tissues of the foetus (Evans and Kaufman, 1981; Martin, 1981) provides a route for precise modification of the mouse genome by gene targeting (for review see Brandon *et al.*, 1995; Babinet *et al.*, 2001). Genetic modification can be achieved in ES cells by taking advantage of HR to target single copy transgenes to specific sites or to modify existing genes *in situ* even down to the level of a single base pair substitution (Hasty *et al.*, 1991b). Genetically modified ES cells can then be returned to the embryo by injection and introduced into surrogate dams where they resume their normal programme of development (Gossler *et al.*, 1986). The donor ES cells can then contribute to a broad range of tissues including the germ cells. Breeding from the chimaera produces heterozygous offspring that can be interbred to produce a pure line of homozygous animals with the engineered modification (Capecchi, 1989). The phenotype of novel mouse lines generated from targeted ES cells gives important information about the function of specific genes for the physiology or development of the mouse. This approach is vital for the generation of mouse models of recessive human diseases and for testing the role of individual gene products in whole animals.

1.3.3. Gene Targeting

1.3.3.1. Targeting Vectors

Two types of targeting vectors have been used to modify target genes in ES cells (figure 1.5). Replacement vectors (Ω type) typically carry a segment of DNA homologous to the target locus, which is interrupted by a positive selection marker gene. This marker is either inserted into the homology region, or alternatively, replaces genomic sequence located between the homology arms. Replacement vectors are linearised outside the region of homology. Such vectors recombine with the target gene by double cross-over events leading to the replacement of the chromosomal sequence with all the components of the vector which are flanked on both sides by the homologous sequences. Heterologous sequences at the ends of the vector homologies are excised from the vector during the process of homologous recombination. In contrast, insertion vectors (O type) are linearised within the region of homology. Such vectors are inserted entirely into the target locus by a single cross-over event at the site of linearisation, resulting in a duplication of the homologous DNA separated by the heterologous sequences in the vector backbone. Positive selectable markers are essential components of both types of targeting vectors, since they allow the isolation of rare transfected cells (typically around 10^{-4} in mouse ES cells) from a predominantly untransfected population. A commonly used selectable marker is *neo*, which confers resistance *in vitro* to the synthetic aminoglycoside G418 (Southern and Berg, 1982), but other markers such as *hyg*, conferring resistance to hygromycin B, and *hprt* minigenes (which function in HPRT-deficient cells) (Selfridge *et al.*, 1992; Detloff *et al.* 1994) are also available.

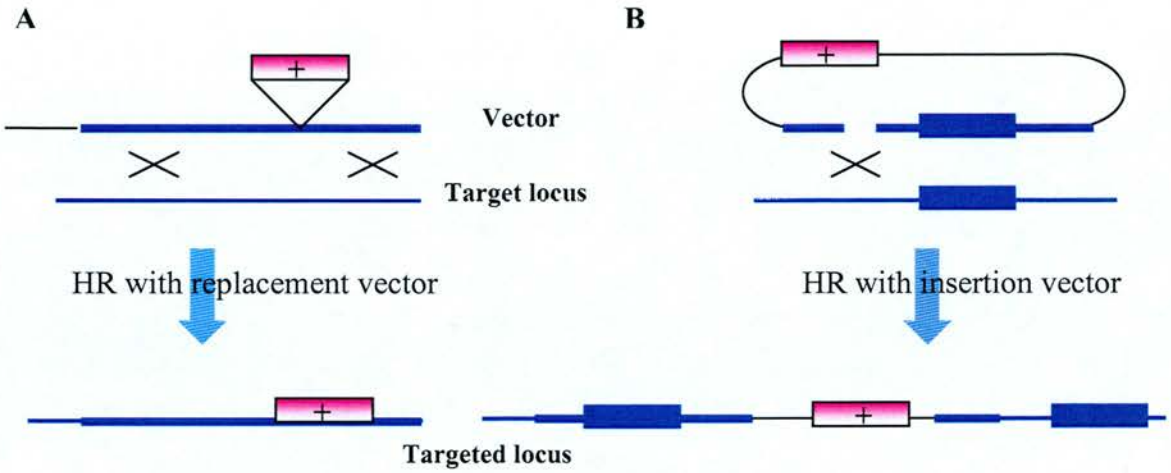


Figure 1.5. Integration patterns for a replacement and an insertion vector. **(A)** Replacement vector**:** The homology (thick blue line) to the target locus (thin blue line) is interrupted with the selectable marker (+) resulting in gene inactivation. The vector is linearised outside the region of homology. A double-crossover event results in an exchange of vector sequence with the chromosomal sequence. **(B)** Insertion vector: The region of homology (thick blue line) to the target locus (thin blue line) is linearised. A single-crossover results in the duplication of the region of homology interrupted by the vector sequence.

1.3.3.2. Screening for Targeted ES Cell Clones

Two types of screening procedures are used to identify ES cell clones carrying a targeted integration of the construct: Polymerase Chain Reaction (PCR) (Saiki *et al.*, 1988) and Southern blot analysis (Southern, 1975). For PCR analysis, two oligonucleotide primers are used to amplify a specific fragment created by the HR event (Joyner *et al.*, 1989; Soriano *et al.*, 1991). One primer is complementary to sequences unique to the target locus and the other is unique within the targeting construct. PCR amplification of the expected fragment is possible only when these primers are correctly juxtaposed by a HR event. To rule out the possibility of PCR artifacts such as PCR jumping (Paabo and Wilson, 1988), Southern analysis is usually used to confirm a targeted event. With the choice of

restriction digests and probes for hybridization, the wild-type allele can readily be distinguished from the targeted allele since predicted novel restriction fragments are generated by the HR event. Southern analysis is usually performed with a probe that is not contained in the targeting vector (external probe) in order to distinguish the wild-type from the predicted targeted allele. In contrast, an internal probe can result in many false positives, since it will detect the randomly integrated vector and some of these insertions by chance will give the expected restriction fragment size and appear to be targeted insertions. Ideally, both the 5' and 3' aspect of the targeted locus should be analysed by using a 5' and 3' external probes. This will allow to discriminate targeted integration of the entire vector or concatemers of the vector from simple replacement events.

1.3.3.3. Parameters Affecting Targeting Frequency

One major factor affecting targeting frequency is the length of homology to the target locus. Several groups have described a relationship between the length of homology and targeting frequency (Thomas and Capecchi, 1987; Shulman *et al.*, 1990; Hasty *et al.*, 1991a; Deng and Capecchi, 1992; Scheerer and Adair, 1994). As a general rule the greater the length of homology the higher the targeting frequency. However the exact relationship between the length of homology and targeting frequency is yet unclear since differences have been observed between experiments even in the same system. While a linear relationship was observed at the *hprt* locus in murine ES cells using homologies ranging from 1.3 to 6.8 kb (Hasty *et al.,* 1991a), an exponential relationship was observed by other investigators at the same locus using homologies between 2 kb and 18 kb (Deng and Capecchi, 1992).

However, saturation of the mammalian HR apparatus appears to require approximately 14 kb of homologous sequence in the targeting vector (Deng and Capecchi, 1992). An exponential relationship between homology length and targeting frequency was also observed by Scheerer and Adair (1994) at the adenine phosphoribosyltransferase (*aprt*) locus in Chinese hamster ovary (CHO) cells using homologies ranging in size from 896 bp to 3275 bp.

The use of isogenic or non-isogenic constructs is another important determinant of gene targeting efficiency. An evenly dispersed 19 % mismatch reduces the frequency of extrachromosomal recombination up to 15-fold in mouse ES cells (Waldman and Liskay, 1987). Similarly, base pair mismatches have been shown to strongly affect the frequency of HR in bacteria (Rayssiguier *et al.*, 1989), of intrachromosomal recombination in mammalian cells (Bollag *et al.*, 1989) and of gene targeting in embryonic stem cells (Deng and Capecchi, 1992; te Riele *et al.*, 1992; van Deursen and Wieringa, 1992). The number and extent of polymorphic variation between two mouse strains in any given locus is usually unknown and may vary widely from gene to gene. Such variation could account for high targeting frequencies also observed in mouse ES cells without the use of isogenic DNA (Zijlstra *et al.*, 1989; Te Riele *et al.*, 1990). Although it is now clear that the existence of sequence mismatches between the homology in the vector and the target locus usually reduces the targeting frequency in mouse ES cells, it has been suggested that non-isogenic DNA does not affect targeting frequencies in human cells (Sedivy *et al.*, 1999). This is likely because the sequence divergence between two human cell lines is much smaller than the difference between two mouse strains. While single-nucleotide polymorphisms (SNPs) occur at a frequency of one per 500

to 1000 bp in human cells (Wang *et al.*, 1998; Sachidanandam, 2001; Zhao *et al.*, 2003), as many as 17 SNPs may occur in a 150 bp stretch of sequence (e.g. *ckm* gene) between the 129 and BALB/C alleles (Van Deursen and Wieringa, 1992). This would also be consistent with a previous report that suggested that intrachromosomal recombination in mammalian cells increases with the amount of uninterrupted homology available (Waldman and Liskay, 1988).

Locus-to-locus intrinsic variability has been extensively documented to affect the overall targeting frequency. The level of expression of the target gene may account for some of these observations. Early work in *E. coli* showed that HR is highly enhanced by transcription (Ikeda and Matsumoto, 1979). Similarly, it has been shown that transcription enhances HR in eukaryotes (Blackwell *et al.*, 1986; Thomas and Rothstein, 1989; Nickoloff and Reynolds, 1990; Nickoloff, 1992; Prado *et al.*, 1997), and that gene targeting frequency is significantly enhanced (3-fold to >20-fold) in the presence of an agent that stimulates target site transcription in cultured human fibrosarcoma cells (Thyagarajan *et al.,* 1995). The direct correlation between gene expression and HR suggests that actively transcribed hypomethylated genes may be more accessible to the HR machinery (Blackwell *et al.*, 1986). Consistent with this hypothesis, is a recent report that showed a 2-fold increase in gene targeting frequency in murine ES cells lacking the DNA methyltransferase 1 (Dnmt1), an enzyme involved in methylation (Dominguez-Bendala and McWhir, 2004). While there is evidence that transcription can stimulate gene targeting frequency (Thyagarajan *et al.,* 1995), it is also known that untranscribed genes can be disrupted by gene targeting in murine ES cells with frequencies similar to expressed genes (Koller and Smithies, 1989; Kumar *et al.*, 1994; Kolb *et al.*, 1999,

Kolb, 2001). Thus the exact correlation between transcription and targeting frequency (if any) remains to be explored.

1.3.3.4. Enrichment Strategies

The ratio between HR and illegitimate recombination (IR) may be as low as 1:1000 (Bode *et al.*, 2000b), it is therefore very common to use strategies to enrich for targeted events. A widely applied approach includes the use of a positive-negative selection (Thomas and Capecchi, 1987; Mansour *et al.*, 1988) to reduce the number of clones with random integration events in a cell population (figure 1.6). The targeting construct is based on a replacement vector containing a positive selection marker within the region of homology and the negative selection (e.g. HSV*tk*) at one or both ends of the homologous DNA. In a first step, the positive selection is applied to select for cells that have integrated the targeting vector in their genome. In the second step, the negative selection is applied only against cells in which the entire targeting vector has integrated at random, since cells carrying a targeted integration lose the negative selection marker during gene replacement event. The enrichment achieved with a positive-negative selection may vary. While a 2000-fold enrichment for HR events over illegitimate events was initially reported (Mansour *et al.,* 1988), usually enrichments between 2- to 20 fold (Hanson and Sedivy, 1995) are observed. This is likely to be due to exonuclease degradation of the negative selectable marker prior to integration of the plasmid into the genome. Occasional loss of the negative marker in random integrants and metabolic exchange between non-targeted and targeted cells may also account for a low enrichment.

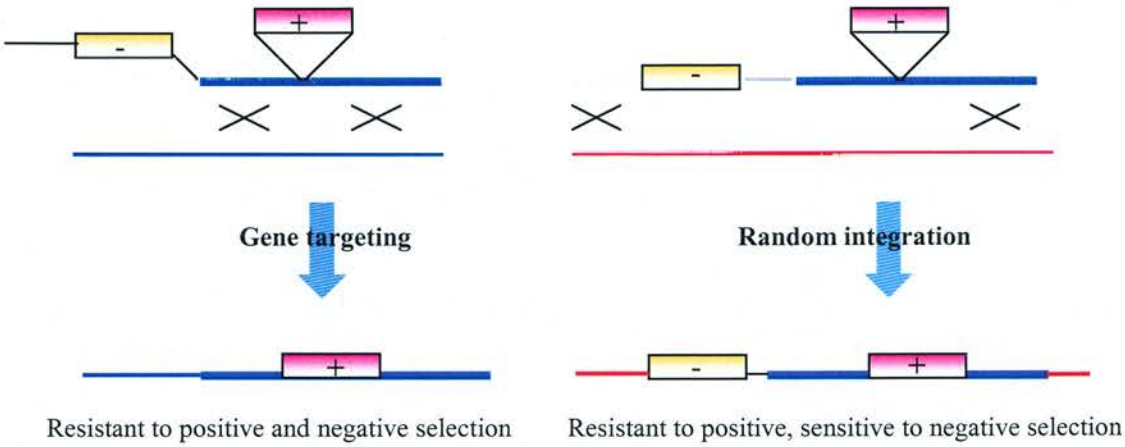


Figure 1.6. Positive-negative selection strategy. A targeted event results in the loss of the negative selection marker (-) since it is located outside the regions of homology, and thus targeted cells are resistant to both positive (+) and negative selections. In contrast, the integration of the targeting vector into a random chromosomal site includes the negative selectable marker. Random integrants are therefore resistant to positive selection but sensitive to negative selection. Thus negative selection permits an enrichment for targeted events.

Enrichment for targeted events may be improved by using a promoter, enhancer, or a polyadenylation trap targeting vector (Doetschman *et al.*, 1988; Dorin *et al.,* 1989; Jasin and Berg, 1988; Sedivy and Sharp, 1989). Promoter and enhancer trap vectors are designed to use the transcriptional activity of the endogenous target gene to drive the positive selection cassette from the targeting construct. Such constructs are thus restricted to genes which are transcriptionally active in the target cells. For promoter trap strategies (figure 1.7), the positive selection cassette is either cloned in-frame with the endogenous translated product, or the cassette may encode its own translation initiation site and is positioned upstream or in place of the nominal translational initiation site. Because in a promoter trap strategy targeting vectors do not contain a promoter, the majority of random integrations will be transcriptionally silent and therefore will not survive positive selection. Promoter

trap selection will yield typically a 100-fold enrichment for targeting events and will work for both insertion and replacement vectors (Schwartzberg *et al.*, 1990). Good candidate genes for a promoter trap strategy are those with high expression levels, although genes with low expression levels have been successfully targeted with promoter trap vectors (Jeannotte *et al.*, 1991). Enhancer trap vectors may also be used to enrich for targeted events. Such vectors use a selection marker with a weak position-dependent promoter which can be activated if integration occurs in the proximity of a transcriptional enhancer element. The enrichment factors achieved with such vectors are however less significant than with promoter trap vectors.

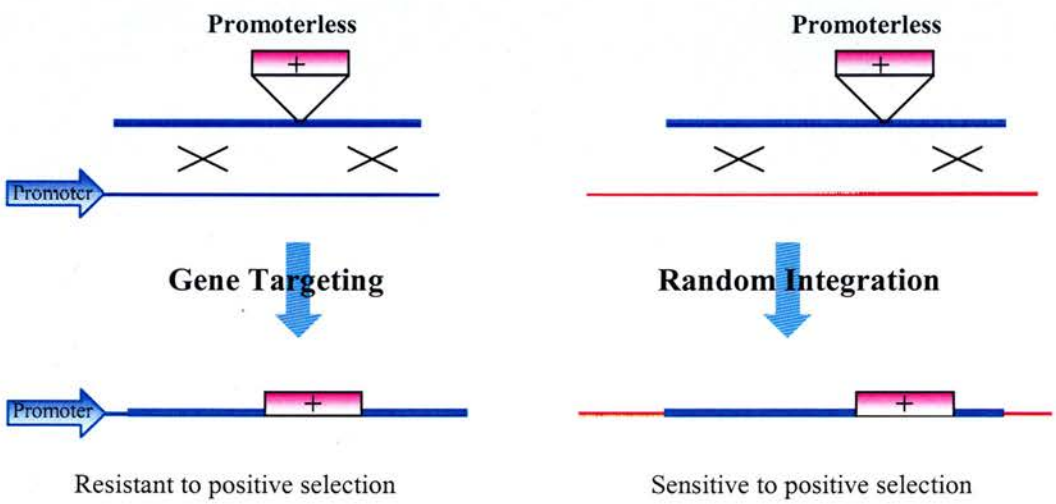


Figure 1.7. Promoter trap strategy. Targeted integration of the promoterless positive selectable marker (+), with a polyadenylation signal, into an actively transcribed gene results in expression of the marker, and thus targeted cells survive positive selection. In contrast, random insertion of the vector does not usually result in expression of the marker, and therefore random integrants do not survive selection.

Polyadenylation trap vectors are designed to use the transcription termination/polyadenylation signals of the target gene to generate a stable hybrid transcript composed of elements from both the target gene and the positive selection marker. Such a vector contains its own promoter but no polyadenylation signal. In contrast to promoter trap strategies, polyadenylation trap strategies may be used for any gene irrespective of whether it is expressed in the cells. The best location for the positive selection marker is in an exon of the target gene so that the splice donor sequence of the disrupted exon can allow normal splicing to occur downstream of the cassette. Typical enrichment factors achieved with such vectors vary between 5-to 50- fold (eg. Donehower *et al.,* 1992) and work for both insertion and replacement vectors.

1.3.3.5. Generation of Subtle Mutations

Several strategies have been designed to introduce subtle mutations (e.g. point mutations) into a target locus without leaving a selectable marker in the target gene, for the purposes of analysing gene function, correcting specific gene defects, or preventing the possible interference of the selectable marker with neighbouring genes (Clark *et al.*, 1997). The "hit and run" or "in and out" approach requires two steps of recombination and is based on an insertion-type vector containing a subtle mutation in the homology and a positive-negative selection marker in the plasmid backbone (Hasty *et al.*, 1991b; Valancius and Smithies, 1991). In the first step ("hit"), HR leads to the complete integration of the vector into the target locus generating a partial duplication interrupted by plasmid sequence and the two selection markers. Negatively-selected intrachromosomal recombination event

("run") between the repeats leads to the correction of the target gene (along with the subtle mutation) and loss of all heterologous sequences. The hit-and run technique has been successfully used to generate mutations in ES cells in the *hprt* locus (Valancius and Smithies, 1991); the *hox-2.6* locus (Hasty *et al.*, 1991b), the β-amyloid precursor protein locus (Gschwind and Huber, 1998); the $G(i2)_2\alpha$ locus (Rudolph *et al.,* 1993) and the cystic fibrosis transmembrane conductance regulator (*Cftr*) locus (Dickinson *et al.*, 2000) (with germ-line transmission of the mutated allele being demonstrated in the latter two examples).

The "tag and exchange" (Askew *et al.*, 1993) or double replacement (Wu *et al.,* 1994) strategy also involves two sequential gene targeting steps but uses two different replacement type vectors. In a first step a selectable marker is inserted in the target locus. In a second step, homologous recombinants are transfected with a targeting vector which replaces the selection cassette in the targeted locus with a non-selectable (e.g. point) mutation. This technique was successfully used to introduce subtle mutations in mouse ES cells in the *Col1α-1* gene (Wu *et al.*, 1994), *α2 Na,K-ATPase* gene (Askew *et al.*, 1993), the Huntington's disease homolog (*Hdh*) gene (Cearley and Detloff, 2001), and to replace the murine β-lactalbumin gene with its human homologue (Stacey *et al.,* 1994). The advantage of this technique is that any transgene can be used in the second step to generate a range of mutation at the same locus. For instance, this technique has been used in ES cells to generate a series of different targeted alterations to the prion protein gene (Moore *et al.*, 1995).

A variance of this strategy is the "plug and socket" approach developed by Detloff and co-workers (1994). In a first step, a functional *neo* and a partial *hprt* minigene ("the socket") are targeted to a specific locus in *hprt*-deficient

ES cells. In a second step, a second targeting vector ("the plug") that supplies the remaining portion of the *hprt* minigene, generates a functional cassette upon HR with the integrated socket. Using this method, the murine *β-globin* gene was replaced by its human equivalent (Detloff *et al.*, 1994).

1.3.4. The Cre/*LoxP* System

1.3.4.1. Cre/*LoxP* Site-Specific Recombination Mechanism

The simplest site-specific recombination systems are comprised of two elements: the recombinase enzyme and a small stretch of DNA specifically recognised by the recombinase. These two elements work together to either delete, insert, invert, or translocate associate DNA. Two such recombinase systems include the Cre-*loxP* system from the bacteriophage P1 and the Flp-*FRT* system from the budding yeast *S. cerevisiae* and have been shown to function in bacteria, fungi, plant and mammalian cells (reviewed by Kilby *et al.*, 1993). The Cre protein (38 kDA) is encoded by the E.*Coli* phage P1 which is present inside E.*Coli* as a single copy, circular DNA plasmid molecule. Cre activity catalyses the exchange and separation of copies of P1 that arise after its replication in order to allow partitioning of the two P1 molecules at each cell division (Sternberg and Hamilton, 1981; Hoess and Abremski, 1984). The target site of Cre is the 34 bp *loxP* (locus of cross-over χ) sequence, characterised by two 13 bp inverted repeats flanking an 8 bp core sequence (figure 1.8.A). Two Cre molecules bind to each *loxP* sites, one on each half of the palindrome to form a tetrameric complex bringing two *lox* sites into proximity (van Duyne, 2001). Cleavage, exchange, and ligation of one DNA strand forms a Holliday intermediate; cleavage, exchange, and ligation of the second strand resolve

the Holliday intermediate into two recombinant molecules (van Duyne, 2001). The final outcome is the exchange of the symmetry arms of the two *loxP* sites and the core regions of the two recombinants heteroduplex. This recombination reaction is carried out with absolute fidelity, such that not a single nucleotide is gained or lost overall. The orientation of a *loxP* site on a segment of DNA directs the type of recombination between *loxP* sites. Cre catalyses both inter-molecular DNA exchanges and intra-molecular excision or inversion (figure1.8.B). Two sites orientated in the same direction on a DNA segment, lead to excision and circularization of the DNA between the sites (Sauer and Henderson, 1989; Baubonis and Sauer 1993) leaving a single *loxP* on the DNA segment and on the resulting circular DNA. The reverse reaction may also occur, with the integration of the circular DNA into a single *loxP* site. Upon targeted integration of the circular DNA, the *loxP* site is however duplicated, leading to the highly favored intra-molecular excision (Araki *et al.*, 1997).

Alternative strategies have subsequently been devised in order to create higher frequency and stability of insertion events. Several mutant *lox* sites (mutation present in the 8 bp core region), that recombine efficiently with an identical *lox* site but very inefficiently with *loxP,* were identified (Lee and Saito, 1998; Langer *et al.*, 2002) (figure 1.8.A). The identification of these mutant *lox* sites led to the development of the recombinase-mediated cassette exchange (RMCE) approach (reviewed by Bode *et al.*, 2000b; Baer and Bode, 2001) in which heterospecific *lox* sites flanking a transgene, are introduced randomly in the genome (Bethke and Sauer, 1997; Bouhassira *et al.*, 1997; Feng *et al.*, 1999; Trinh and Morrison, 2000) or by HR in a specific gene (Soukharev *et al.*, 1999; Kolb, 2001).

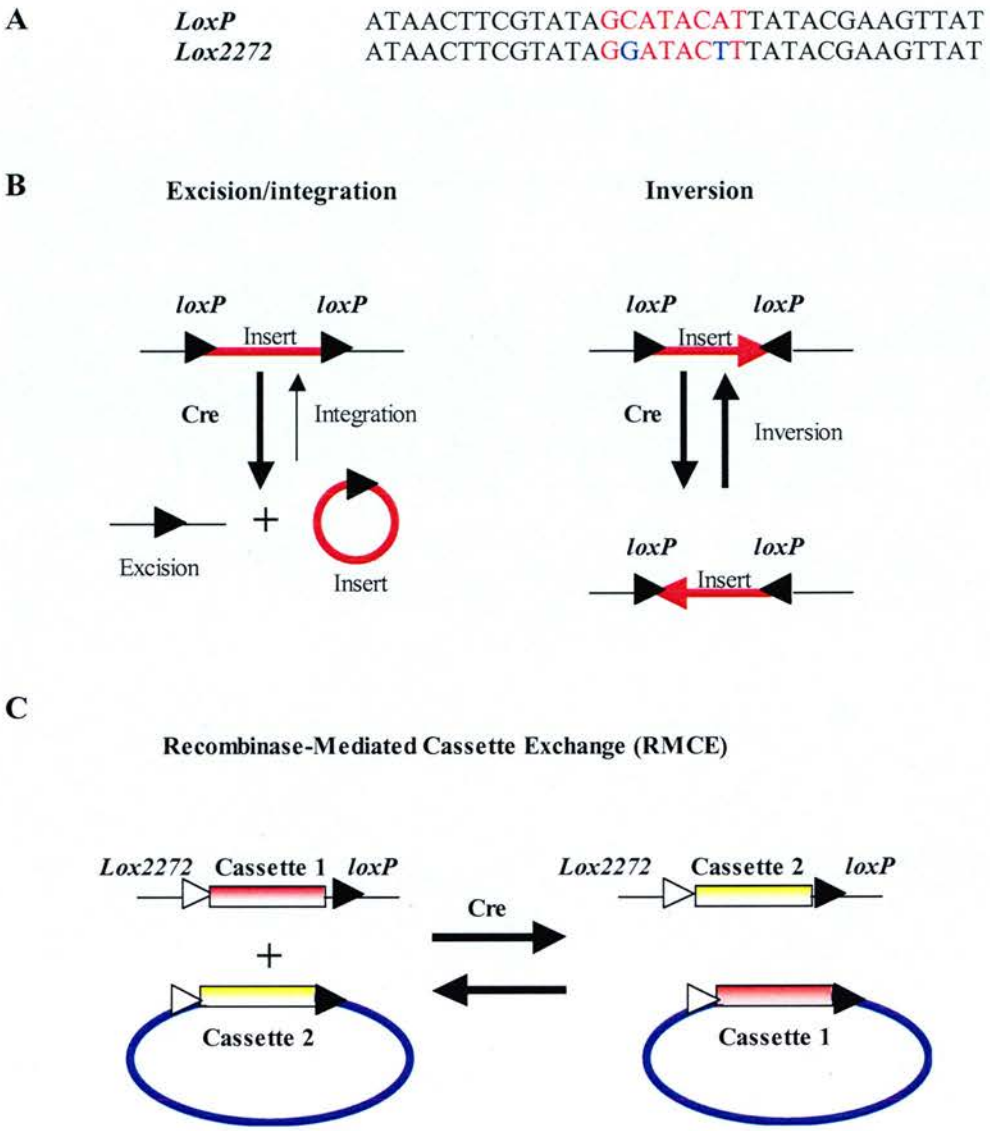


Figure 1.8. The Cre/loxP system. **(A)** DNA sequences of wild type *loxP* site and a mutated *lox* site, *lox2272* site. The 8 bp core is represented in red and the mutations in the *lox2272*, in blue. **(B)** Excision/integration: Cre-mediated recombination between two directly repeated *loxP* sites on a linear DNA molecule leads to the excision and circularisation of the *loxP* flanked DNA insert. One *loxP* site remains on each of the reaction products. In the reverse reaction a *loxP* containing circle is integrated into a linear DNA molecule. Inversion: Cre-mediated recombination between inverted *loxP* sites leads to a 50:50 probability of insert inversion. **(C)** RMCE: Heterospecific *lox* sites (e.g. *loxP* and *lox2272*) flanking an insert can be used to swap a pre-inserted cassette in the genome (cassette 1) for a targeting insert (cassette 2) that is flanked by the same *lox* sites. This double replacement occurs with a 50:50 probability.

In a subsequent Cre-mediated recombination step, transgenes flanked by the same set of heterospecific *lox* sites can be exchanged by a double-reciprocal recombination event with the previously introduced transgene (figure 1.8.C). Once a locus is tagged with heterospecific *lox* sites, RMCE can be used to generate a broad range of mutation at the same locus.

1.3.4.2. Applications

Site-specific recombination mediated by Cre/*loxP* approach allows researchers to control not only site-specific integration and copy number of transgene, but also to replace or delete precisely any sequence within a targeting locus (Gu *et al.*, 1994; Kolb *et al.*, 1999; Kolb, 2001). For instance, the cre/*loxP* system can be used to remove selection markers, after conventional gene targeting in order to avoid interference with the expression of the targeted gene or a neighbouring transgene (Wang *et al.*, 1999; Kolb, 2001). Moreover, Cre-mediated site-specific recombination can be used to manipulate gene expression in a cell type specific or inducible manner at the stage of development, which needs to be investigated (Schwenk *et al.,* 1995; Sauer, 1998; Metzger and Feil, 1999). Another application of this system has been to induce a variety of chromosome rearrangements (deletions, inversions, duplications, and translocations) in ES cells that can be selected and then used to produce ES-derived mice (Smith *et al.*, 1995; van Deursen *et al.*, 1995; Ramirez-Solis *et al.*, 1995). This technology known as "chromosome engineering", has successfully generated numerous mouse models that accurately recapitulate human chromosomal rearrangements (e.g. Lindsay *et al.*, 1999; Tsai *et al.*, 1999; Buchholz *et al.*, 2000; Collins *et al.*, 2000).

1.4. Genetic Modification of Human ES Cells

1.4.1. Human ES Cells

1.4.1.1. Origin
1.4.1.2. Properties
1.4.1.3. Applications
1.4.1.4. Ethical Considerations

1.4.2. Genetic Modification of HES Cells: Applications

1.4.2.1. Directed Differentiation and Purification
1.4.2.2. Protection Against Tumorigenesis
1.4.2.3. Long-Term Survival and Functionality of the Graft
1.4.2.4. Immunotolerance
1.4.2.5. *In Vitro* Modeling of Human Diseases
1.4.2.6. Gene Correction

1.4.3. Methods for the Genetic Manipulation of HES Cells

1.4.3.1. Additive Transgenics in Human ES Cells
1.4.3.2. Gene Targeting in Human ES Cells

1.4.4. Engineering *β-casein* and *hprt* Genes in HES Cells

1.4.4.1. *β-Casein* Gene as a Target Locus: Characteristics & Advantages

1.4.4.2. *Hprt* Gene as a Target Locus: Characteristics & Advantages

1.4.1. Human ES Cells

1.4.1.1. Origin

The derivation of hES cells arose from the graduated progress which started more than 40 years ago with the isolation of embryonic carcinoma (EC) cells from mouse testicular teratocarcinomas (Kleinsmith and Pierce, 1964). This early report led to the first derivation of murine ES cells (Evans and Kaufman, 1981; Martin, 1981), followed by human EC cells (Andrews *et al.*, 1984; Thompson *et al.*, 1984), murine embryonic germ (EG) cells (Matsui *et al.*, 1992) and primate ES cells (Thomson *et al.*, 1995). Although these cell lines differ in their phenotype, they all share common properties of multipotency and immortality.

Human embryonic stem (hES) cells were initially isolated by J. Thomson and co-workers in 1998 (Thomson *et al.*, 1998). Human ES cells are derived from the inner cell mass (ICM) (cells that are destined to develop into the embryo proper) of the blastocyst of excess embryos produced for clinical *in vitro* fertilisation. After mechanical or immunosurgical (Solter and Knowles, 1975) separation of the ICM from the trophoblastic cells, the isolated ICMs are subsequently plated onto a supportive matrix, such as mouse embryonic fibroblasts, or human embryonic fibroblasts. Colonies of appropriate undifferentiated morphology are subsequently selected and expanded to form established human ES cell lines. Since this first report, a large number of hES cell lines have been isolated (e.g. Reubinoff *et al.*, 2000; Lanzendorf *et al.*, 2001; Richards *et al.,* 2002; Hovatta *et al.*, 2003; Mitalipova *et al.*, 2003; Pickering *et al.*, 2003; Cowan *et al.*, 2004).

1.4.1.2. Properties

The most important property of hES cells is their pluripotency, a common characteristic shared with mES cells and human EG cells (Shamblott *et al.*, 1998). When grafted into severe combined immunodeficiency disease (SCID) beige mice, hES cells give rise to teratomas containing derivatives of all three primary embryonic germ layers (Thomson *et al.*, 1998; Reubinoff *et al.*, 2000). These include gut epithelium (endoderm); cartilage, bone, smooth muscle and striated muscle (mesoderm); and neural epithelium, embryonic ganglia and stratified squamous epithelium (ectoderm) (reviewed by Przyborski, 2005). In addition, hES cells in suspension can also differentiate spontaneously into cells of the three germ layers when allowed to aggregate and form spheroid clumps termed embryoid bodies (EBs) (Itskovitz-Eldor *et al.*, 2000). After several days of differentiation, the EBs contain cells from various tissues including muscle, bone, kidney, blood cells, skin, liver (Schuldiner *et al.*, 2000), β cells of the pancreas (Schuldiner *et al.*, 2000; Assady *et al.*, 2001), and a large number of neurons (Reubinoff *et al.*, 2000; Itskovitz-Eldor *et al.*, 2000; Schuldiner *et al.*, 2001; Carpenter *et al.*, 2001). In some cases, the cells have been shown to be functional as in the case of cardiomyocytes (Itskovitz-Eldor *et al.*, 2000; Kehat *et al.*, 2001), neurons (Carpenter *et al.*, 2001) or endothelial cells (Levenberg *et al.*, 2002). Spontaneous differentiation into various cell types may also occur without aggregation into EBs (Reubinoff *et al.*, 2000; Reubinoff *et al.*, 2001).

Like their murine counterparts, hES cells have also the capacity of self-renewal and can be propagated in culture indefinitely while maintaining a normal karyotype (Thomson *et al.*, 1998; Amit *et al.*, 2000). Recent reports however

showed that chromosomal abnormalities such as translocations, deletions and duplications, may occur in prolonged culture (Draper *et al.*, 2004; Inzunza *et al.*, 2004; Rosler *et al.*, 2004; Mitalipova *et al.*, 2005; Maitra *et al.*, 2005; McWhir *et al.*, unpublished data). Undifferentiated hES cells are characterised by the expression of several ES specific molecular markers such as the transcription factor Oct-3/4 (Nichols *et al,* 1998), the cell surface markers SSEA-3, SSEA-4, Tra-1-60 and Tra-1-81, telomerase and alkaline phosphatase (Thomson *et al.*, 1998; Xu *et al.*, 2001; Amit *et al.*, 2000; Draper *et al.*, 2002). Undifferentiated proliferation of hES cells relies mostly on the presence of a mouse (Thomson *et al.*, 1998) or human (Richards *et al.*, 2002; 2003; Amit *et al.*, 2003; Cheng *et al.*, 2003; Hovatta *et al.*, 2003; Lee *et al.*, 2005; Yoo *et al.*, 2005) fibroblast feeder layer, which produces a matrix required to allow adherence of the cells, and secretes factors maintaining pluripotency and self-renewal potential (Thomson *et al.,* 1998). Alternatively, the feeder cell layer can be replaced with matrigel, an extracellular matrix (ECM) preparation, and conditioned medium from the feeder cell layer (Xu *et al.*, 2001). In addition to the important role of the feeder cells in maintaining pluripotency, is the use of basic fibroblast growth factor (bFGF) (Thomson *et al.,* 1998) for the propagation of hES cells. Despite the wide knowledge gained on the importance of leukemia inhibitory factor (LIF) (Moreau *et al.*, 1988; Smith *et al.*, 1988) in maintenance of mES cells, evidence from hES cell lines indicates that LIF does not play a critical role in sustaining undifferentiated growth (Thomson *et al.*, 1998, Reubinoff *et al.*, 2000).

1.4.1.3. Applications

Given the properties of hES cells, the potential applications are many (figure 1.9). Human ES cells will be useful for the study of gene function and for the development of *in vitro* human disease modelling (see section 1.4.2.5). These cells could also provide new information about the cellular and molecular basis of commitment and differentiation events during embryonic human development and disorders, such as birth defects and embryonal tumours, as these issues cannot be addressed in the human embryo for ethical reasons. Additionally, hES-derived cell lineages could provide a novel source for cell replacement therapy addressing the chronic shortage of tissue for transplantation in the treatment of a wide range of diseases such as type I diabetes, cardiovascular disease, Parkinson's disease, and blood cell diseases. Human ES cells could also be a potential answer to an end to the use of immunosuppressive therapy in transplantation, if cloning techniques can be used to derive stem cells from a patient's own tissue or if genetic modification strategies are established to render hES derivatives invisible to the immune system (see section 1.4.2.4). Human ES cells will also have widespread applications in the areas of drug discovery and drug development (reviewed by Pouton and Haynes, 2005). Cell types such as cardiomyocytes and hepatocytes generated from hES cells could provide an ideal source of cells for predictive toxicology. These cells could reveal the toxicity of certain drugs that might not be detected using conventional assays that rely on animal models. Other applications could be made possible such as an investigational tool for the discovery of novel growth and differentiation factors that might find application in tissue regeneration and repair; and new delivery systems for gene therapy.

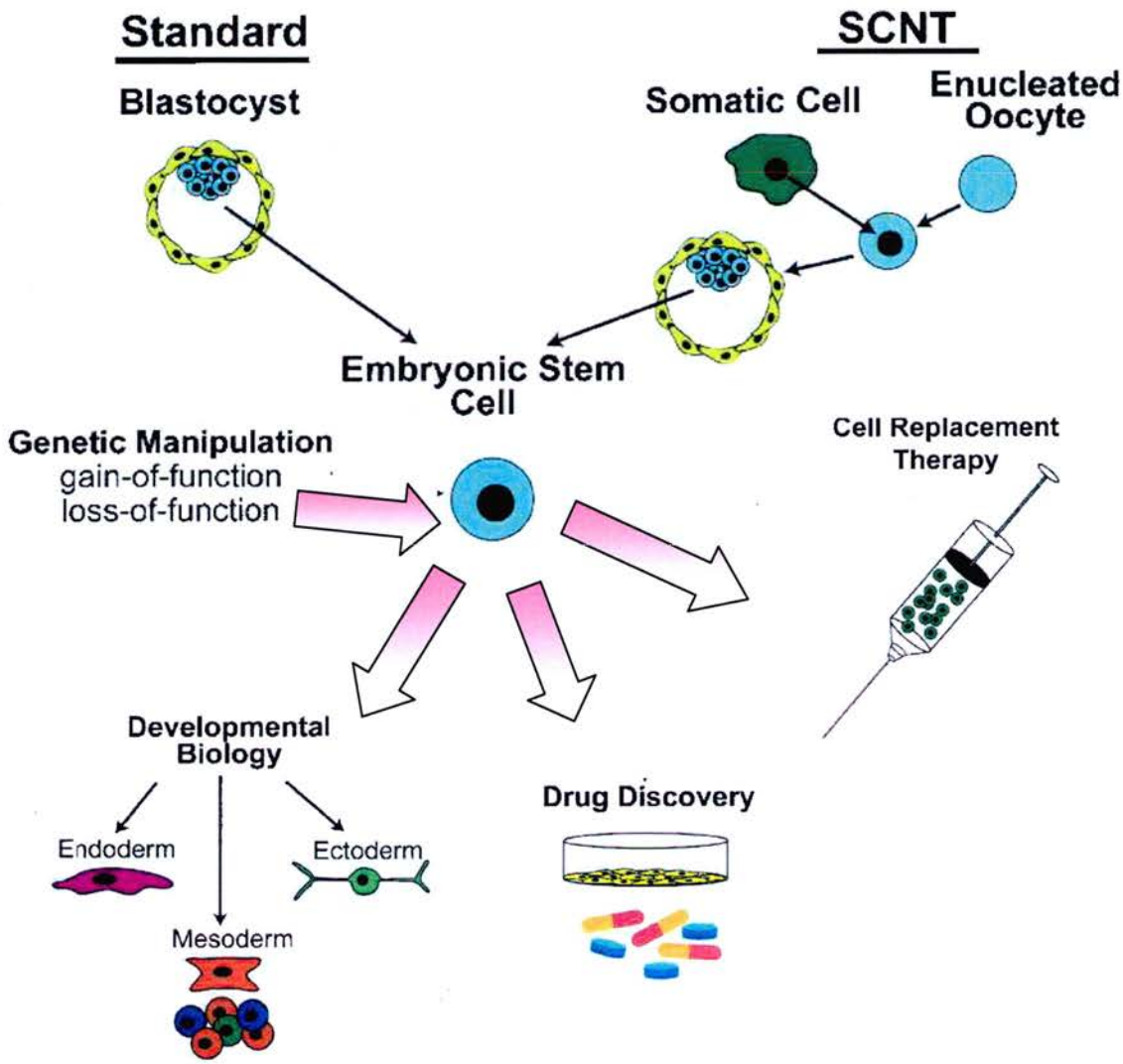


Figure 1.9. Potential use of hES cells in basic research and medicine. See text for details. SCNT: somatic cell nuclear transfer (cloning).

1.4.1.4. Ethical Considerations

Despite the potential applications of hES cells mentioned above, these cells are surrounded by a number of controversies, the extent of which is partly dependent on their source (Frankel, 2000). Human ES cell lines are derived from "excess" embryos from *in vitro* fertilization (IVF) clinics. This is not acceptable to many people including ethicists, religious and political leaders who consider full human life as starting with conception (i.e. the formation of a zygote). The establishment of human somatic cell nuclear transfer (SCNT) technology (transfer of a nucleus of a somatic cell into an enucleated egg) and the derivation of hES cell lines may lead to the reality of "therapeutic cloning" (Colman and Kind, 2000). This raises even more ethical concerns because embryos will be specifically created and destroyed for the purpose of deriving hES cell lines for different applications (e.g. "autotransplant").

The use of human adult stem cells in cell-based therapy would be a solution to the present ethical concerns. Several investigators have reported the isolation of primitive human adult multipotent cells from different tissues. Subpopulations of human mesenchymal stem cells (hMSCs) exhibiting features of primitive adult pluripotent or multipotent stem cells have been described (Young *et al.*, 1998; Pittenger *et al.*, 1999; Colter *et al.*, 2000; Reyes *et al.*, 2001; Jiang *et al.*, 2002; Gronthos *et al.*, 2003). Adult stem cells have also been found in umbilical cord blood (Erices *et al.*, 2000), peripheral blood (Huss *et al*, 2000; Kuci *et al.*, 2003), adipose tissue (Zuk *et al.*, 2001), skeletal muscle and dermis (Young *et al.*, 2001), as well as other tissues (Asakura and Rudnicki, 2002). Although these primitive adult cell subpopulations have been shown to give rise to multiple lineages,

they are not pluripotent and have usually less self-renewal ability than hES cells. Recently however, Catherine Verfaillie and colleagues (Schwartz *et al.*, 2002; Jiang *et al.*, 2002a; Jiang *et al.,* 2002b) have isolated multipotent adult progenitor cells (MAPCS) from mesenchymal cell cultures obtained from rodent and human bone marrow. These MAPCS were capable of in excess of a hundred population doublings and could be induced to differentiate into mesenchyme, endothelia, neuroectoderm (neurons, astrocytes and oligodendrocytes) and endoderm (hepatocytes) lineages. Similarly, D'Ippolito *et al.* (2004) described a unique subpopulation of human marrow stromal cells, MIAMI (marrow-isolated adult multilineage inducible) that could give rise *in vitro* to cell types of the three germ layers and that could be expanded for more than 50 population doublings. The advantage of these adult bone marrow cells over hES cells is that the former can be derived from autologous adult bone marrow, circumventing the problem of immune rejection in the context of transplantation, and ethical issues raised by the use of hES cells. Although these data raised hopes that these cells would be as useful for therapeutic purposes as hES cells, many questions remain concerning the nature and status of adult stem cells both *in vivo* and *in vitro* and their proliferation and differentiation/transdifferentiation capacity (Wurmser and Gage, 2002; DeWitt and knight, 2002; Holden and Vogel, 2002; Wells, 2002; D'Amour and Gage, 2002). One issue in particular has been raised based on *in vitro* co-culture experiments with mES cells and adult neural stem cells (Ying *et al.*, 2002; Pells *et al.,* 2002) or with mES cells and adult bone marrow stem cells (Terada *et al.*, 2002) suggesting that fusion between adult bone marrow stem cells and differentiated cells rather than transdifferentiation, occurs *in vivo*. Regardless of the present issues, the potential

applications of both human adult and embryonic stem cells cannot be ignored, and it is likely that both fields of research will be mutually beneficial.

1.4.2. Genetic Modification of HES Cells: Applications

1.4.2.1. Directed Differentiation and Purification

While some present differentiation protocols have been shown to enrich for a particular cell type such as mature neurons (Reubinoff *et al.*, 2001; Zhang *et al.*, 2001; Schuldiner *et al.*, 2001; Carpenter *et al.*, 2001), motor neurons (Singh Roy *et al.*, 2005; Shin *et al.*, 2005; Li *et al.*, 2005), astrocytes, oligodendrocytes (Zhang *et al.*, 2001), osteoblasts (Sottile *et al.* 2003; Karp *et al.*, 2005), skin cells, liver cells, and adrenal cells (Schuldiner *et al.*, 2000), they do not induce a single differentiation pathway. Thus the main challenge will be to purify homogenous populations of terminally differentiated functional cells or their progenitors. One way to address this issue would be to genetically modify hES to express master genes that promote commitment to a desired lineage pathway. For instance, directed differentiation towards a chondrocyte phenotype was accomplished in murine embryonic mesenchymal cells by expressing the human bone morphogenetic protein-2 (BMP-2) cDNA (Carlberg *et al.*, 2001). Similarly, expression of HoxB4 transgene in mouse ES cells was shown to induce differentiation towards haematopoietic stem cells (Kyba *et al.*, 2002). Likewise, *in vitro* differentiation towards a hepatocyte phenotype was achieved in mES cells by overexpressing the hepatocyte nuclear factor 3 beta (Kanda *et al.*, 2003).

Human ES cells could also be genetically modified to express a lineage-specific selectable marker which would allow the purification of a defined

cell population using drug ablation (Klug *et al.*, 1996) or FACS/MACS separation (Li *et al.*, 1998; Muller *et al.*, 2000). This strategy was first demonstrated by Klug *et al.*, in 1996 who engineered mES cells to express a *neo* marker under the control of a cardiac-specific promoter. Following differentiation, an enrichment of more than 99% of cardiomyocytes was achieved by subjecting the mixed differentiated cell populations to G418. Similarly, Muller *et al.* (2000) showed that 97 % of mouse ES-derived cardiomyocytes, engineered with a cardiac-specific promoter driving a GFP marker, could be purified using percoll gradient centrifugation and subsequent fluorescence-activated cell sorting. Harvesting a single population of cells could be further improved by using an antibody against a tissue-specific membrane protein present in the differentiating cells of interest. Such a procedure was demonstrated for the purification of hES-derived endothelial cells using a fluorescent-activated cell sorter (FACS) (Levenberg *et al.*, 2002) and hES-derived hematopoietic cells using Magnetic Column Separation (MACS) (Kaufman *et al.*, 2001).

1.4.2.2. Protection Against Tumorigenesis

The ability of undifferentiated hES cells to form tumours in histocompatible animals is well documented (e.g. Thomson *et al.*, 1998; Reubinoff *et al.*, 2000). Therefore, efforts need to be made to ensure that the cell graft is depleted of any undifferentiated hES cells before transplantation. Elimination of undifferentiated hES cells could be achieved by expression of a negative selectable marker from an undifferentiated hES cell-specific promoter. A "suicide" gene therapy described for treating cancer (Freytag *et al.*, 2002) could be adapted for hES cells with less complexity as the cells would be modified *in vitro* to express the

suicide gene without the requirement of targeted delivery *in vivo*. Towards this end, hES cells have already been modified to constitutively express the herpes simplex virus thymidine kinase (HSV-*tk*) suicide gene which resulted in the ablation of HSV-tk^+ hES cells *in vitro* and their derivatives *in vivo* in SCID mice with the drug ganciclovir (Schuldiner *et al.* 2003). For this application a suicide gene targeted to a locus known to be permissive for transgene expression, would be the method of choice to ensure the complete elimination of the potentially tumorigenic undifferentiated hES cells from a differentiated cell population.

An alternative approach would be to target a tetracycline "on/off " inducible *tk* transgene (Lewandoski, 2001) to a locus known to be permissive for transgene expression. This strategy would allow the engineered-derived tissue of interest, grafted into a patient, to be eliminated with tetracycline in the event that tumorigenic cells developed within.

1.4.2.3. Long-Term Survival and Functionality of the Graft

Success in regenerative medicine will depend upon the functional integration of the graft into host tissue. This will require the long-term survival of the grafted cells, the presence of a critical tissue mass, the appropriate alignment of the donor cells and their structural and functional integration within the host tissue. Several questions and issues remain to be addressed in this area. Cell death occurring after engraftment will be a major problem as it is believed to have a negative impact on the size of the graft (Zhang *et al.*, 2001; Muller-Ehmsen *et al.*, 2002). Hence, cell survival following transplantation may depend on adequate vascularisation of the graft, which may require additional revascularisation

procedures or induction of angiogenesis. In that respect, the ability to genetically modify hES cells may be used to produce derivatives that are more resistant to ischemia or apoptosis, that can display larger proliferative capacity, that can secrete angiogenic growth factors such as VEGF (Rio *et al.*, 1999), or may be coupled with ES cell-derived endothelial progenitors.

1.4.2.4. Immunotolerance

Rejection of stem cell-derived grafts by an allogeneic host could be circumvented in a number of ways (reviewed by Bradley *et al.*, 2002). One solution may be to render hES cells and their derivatives "invisible" to the immune system. To this end, hES cells could be genetically modified to create a "universal donor" hES cell line, which could escape or inhibit the host immune response. HLA-A, -B and -C antigens, responsible for immunogenicity, could be suppressed by gene targeting to prevent expression of major histocompatibility class (MHC) I molecules (Bradley *et al,* 2002). In the mouse model, skin from mice that are deficient for both MHC class I and II antigens grafted to recipients that differ by just a few MHC antigens survives for a longer period of time, but eventually is rejected (Grusby *et al.*, 1993). This is likely because MHC antigens initiate rejection, although residual MHC antigens might contribute also (Lee *et al.*, 1997). A drawback to this strategy, however, is the risk of susceptibility to tumorigenesis, as some of the most aggressive malignancies are those that evade the immune system by a lack of MHC expression (Restifo *et al.*, 1993; Cromme *et al.,* 1994; Hicklin *et al.,* 1998). Human ES-derived tissues lacking MHC class-I-dependent defence system would also be more susceptible to viral infection as there would be no MHC class I molecules to

present viral peptides to virus-specific cytotoxic T cells (Zeidler *et al.*, 1997). Nevertheless, the host defence has evolved to counter these survival strategies with the presence of NK cells that effectively lyse cells that have no MHC class-I-mediated protection (Storkus *et al.*, 1987). Thus protection of hES-derived "universal-donor" cells from NK cells will be required, possibly by harnessing immune-evasion strategies. For instance, hES cells could be further genetically modified with transgenes that express certain viral genes which have evolved to escape the host immune system (Arase *et al.*, 2002).

While the idea of creating a MHC-deficient cell line might sound attractive, such hES cells-derived grafts would likely still encounter rejection just as it has been described for tissue grafts from MHC-/- mice (Grusby *et al.*, 1993). Thus alternative strategies, such as the generation of individualised ES cell lines through SCNT and tolerisation approaches, may be more realistic. In the former approach, the cells used for transplantation would be identical to those of the patient, thereby overcoming the problem of graft rejection (Colman and Kind, 2000). An ethical and practical limitation of SCNT is the requirement of oocytes. The derivation of oocytes from ES cells may however provide a solution to this problem (Hubner *et al.*, 2003). The latter approach involves tolerisation of the patient's immune system to the graft by the establishment of a chimeric immune system upon transplantation of hES cell-derived hematopoietic stem cells (Bradley *et al.*, 2002).

1.4.2.5. *In Vitro* Modeling of Human Diseases

Modifying specific genes for *in vitro* studies in hES cells will be important for learning more about the pathogenesis of diseases for which mouse models are not appropriate. In the context of this project for instance, the *hprt* gene, whose dysfunction in humans leads to Lesh-Nyhan Syndrome (LNS), would be a good candidate. LNS is a severe genetic disease syndrome which is characterised by excessive levels of uric acid in the blood, which eventually leads to kidney damage (Stout and Caskey, 1988). LNS may also lead to physical and mental retardation, nephrolithiasis, urinary tract stones and may impact brain development. It is still unclear how the HPRT-deficiency leads to the neurological symptoms, although it has been reported that dopaminergic neurons are affected in LNS patients (Visser *et al.*, 2000; Saito and Takashima, 2000). Urbach *et al.* (2004), have recently demonstrated that a targeted mutation in the *hprt* gene in hES cells, mimic some clinical features associated with LNS such as the abolishment of HPRT activity, and a higher rate of uric acid accumulation (a characteristic never observed in the mouse model). Thus their data point to the relevance of hES cells in complementing murine models for the study of human genetic diseases. *In vitro* studies of HPRT-deficient hES-derived dopaminergic neurons as well as transplantation of these cells to an animal model should therefore provide useful clues regarding the etiology of the neurological deficit, and may provide new means of drug development for the neurological symptoms.

The study of genetic diseases, in which mutations are associated with a gain of function such as the *SOD1* gene, in the case of motoneuron disease (also known as amyotrophic lateral sclerosis (ALS)) (Rosen, 1993; Shaw *et al.*, 1998),

could be further improved by introducing a disease-linked mutation into a chromosomal locus known for reproducible transgene expression in hES cells, followed by *in vitro* differentiation and subsequent analysis of the affected cell type.

1.4.2.6. Gene Correction

Breakthroughs in adult somatic nuclear transfer (Campbell *et al.*, 1996) and human ES cell derivation (Thomson *et al.*, 1998) have raised the possibility of using both techniques to generate unlimited sources of autologous cells for tissue repair. Known as "therapeutic cloning", it might be possible to combine this technique with gene therapy to derive hES cells from a patient with a specific genetic defect, correct the gene by gene targeting, and return the derivatives of the genetically-matched corrected cells to the body of the patient, such that repopulation by the corrected cells repairs damaged and diseased tissues (reviewed by Colman and Kind, 2000). This strategy has previously been exemplified in immune-deficient *Rag2*$^{-/-}$ (recombination-activating gene 2) mice (Rideout *et al.*, 2002). These authors used immune-deficient *Rag2*$^{-/-}$ mice as nuclear donors for transfer into enucleated oocytes, and isolated ES cells from the NT-derived blastocysts. One of the mutated alleles in the *Rag2*$^{-/-}$ ES cells was genetically repaired by HR, and hematopoietic precursors were derived by *in vitro* differentiation and engrafted into *Rag2*$^{-/-}$ mice. Mature myeloid and lymphoid cells and immunoglobulins became detectable 3-4 weeks after transplantation, demonstrating the potential of this therapeutic approach. In the face of the ethical and practical problems associated with therapeutic cloning and subsequent gene correction in humans, it seems likely however, that a transplant derived from normal, but allogeneic hES cells, will be less problematic.

1.4.3. Methods for the Genetic Manipulation of HES Cells

1.4.3.1. Additive Transgenics in Human ES Cells

Developing improved technology for the genetic manipulation of hES cells will be crucial for their effective application in research. The generation of stable transfectants has now been achieved by electroporation (Eiges *et al.*, 2001; Zwaka and Thomson, 2003), Chemical methods (Eiges *et al.*, 2001; Urbach *et al.*, 2004), or viral transduction with lentiviral (Pfeifer *et al.*, 2002; Gropp *et al.*, 2003; Ma *et al.*, 2003) and adenoviral (Smith-Arica *et al.*, 2003) vectors. Transfection efficiencies are however not as high as with mES cells which likely reflect a lack of optimal culture conditions. Different cell culture regimes have been employed by laboratories developing transfection procedures. These include mechanical disaggregation of hES cells (Gropp *et al.,* 2003), trypsin-based disaggregation (Eiges *et al.,* 2001), collagenase-based disaggregation (Thomson *et al.*, 1998), feeder-dependent (Eiges *et al.,* 2001), and feeder-free cultivation (Xu *et al.*, 2001). Problems generating single-cell suspensions, poor plating efficiencies, and low cloning efficiencies have hampered progress in developing efficient transfection procedures. However, more laboratories are using trypsin-based disaggregation of hES cells, allowing single-cell suspensions which are more effective for electroporation.

1.4.3.2. Gene Targeting in Human ES Cells

Gene targeting has now been reported in hES cells at the hypoxanthine phosphoribosyl transferase (*hprt*) (Zwaka and Thomson, 2003; Urbach *et al.*, 2004), and the *oct-4* genes (Zwaka and Thomson, 2003) with frequencies similar to those observed in mouse ES cells. The work described in this thesis began one year before these reports. For targeting *hprt*, Zwaka and Thomson used a targeting construct containing 12 kb of homology, achieving transfection efficiencies of 2.3×10^{-5} and a targeting frequency (measured as a proportion of transfectants) of 2 % using an electroporation method in a subclone of the male H1 hES cell line. Urbach *et al.* (2004) used a targeting construct containing 8.5 kb of homology, resulting in transfection efficiencies of 10^{-6} and a targeting frequency of 1 % using a chemical method in the male H13 hES cell line. Gene targeting has also been recently exemplified by adeno-associated virus (AAV) transduction at the *hprt* locus in the H1 hES cell line with frequencies much greater than with plasmid-based vectors (Thomson *et al.*, unpublished data). Karyotypic stability, pluripotency, and stemness of the hES-targeted cell lines are however issues that are yet to be addressed for these cells to fulfill their role in research and medicine.

Gene targeting can also be achieved indirectly by targeting the transcript through the use of short interfering (si) RNA. The new field of RNA interference has provided a powerful tool for assessing the consequences of the down-regulation of specific genes. Several recent reports have now shown that this technique can be used to knockdown gene expression in hES cells (Hay *et al.*, 2004; Vallier *et al.*, 2004; Liu *et al.*, 2005). A downside of this approach, however, is the transient nature of siRNA permitting only temporary gene knockdown.

1.4.4. Engineering *β-casein* and *hprt* Genes in HES Cells

Many of the applications described in section 1.4.2, such as the modeling of human diseases (in the case of LNS), reduction of immunogenicity of hES cell-derived transplants, protection against tumorigenesis and directed differentiation, could be achieved at the *hprt* and *β-casein* genes in hES cells.

1.4.4.1. *β-Casein* Gene as a Target Locus: Characteristics & Advantages

β-casein is a major protein component of mammalian milk and is assembled into micelles in conjunction with other caseins (for review see Ginger and Grigor, 1999). Caseins are a heterogenous group of phosphoproteins whose primary function is to provide the suckling infant with essential amino acids, calcium and phosphate, which are crucial elements for bone development in neonates. Caseins are encoded by single-copy genes comprising a multigene family. Although caseins have the same biological function in all mammals, their numbers vary widely among species; bovine milk, for instance, contains four caseins: α_{S1}, α_{S2}, β and κ, whereas mouse milk contains five caseins: α, κ, γ, δ and β. Human milk, on the other hand, contains only three caseins: β, κ and α_{S1}. β-casein is the major casein and represents about 30% of the protein mass in human milk. In contrast, the amount of α-CN in human milk is very little (Rasmussen *et al.*, 1995).

The *β-casein* gene represents an attractive candidate locus for targeting transgenes as it was previously shown that the disruption of the *β-casein* gene by gene targeting in mouse ES cells generated viable and fertile homozygous *β-casein* mutant mice, and that homozygous females can lactate and successfully rear young, although the growth of pups feeding on their milk, was reduced when

compared to those feeding on the milk of wild-type mice (Kumar *et al.,* 1994). These results demonstrated that the *β-casein* gene has no essential function out-with the mammary gland. In addition, the *β-casein* gene is only expressed in the lactating mammary gland. Hence its dysfunction in non-mammary cell types by transgene introduction is unlikely to have detrimental effects. The integration of transgenes into pre-selected loci such as *β-casein* also presents the advantage that both alleles can be genetically modified without risks of insertional mutagenesis of essential genes such as house keeping genes that are ubiquitously expressed. The attraction of having two copies of the transgene lies in the low frequency with which both copies could be lost or rendered dysfunctional. This is a particularly important consideration where transgenes are introduced to provide protection against tumorigenesis (eg. Tet-inducible *tk* transgene). Although unexpressed genes might be more difficult to target in hES cells than expressed genes, it has been achieved in mES cells (Johnson *et al.,* 1989; Koller and Smithies, 1989) including the *β-casein* gene (Kumar *et al.*, 1994; Kolb *et al.*, 1999; Kolb, 2001), suggesting that it might not be an issue in hES cells.

1.4.4.2. *Hprt* Gene as a Target Locus: Characteristics & Advantages

The *hprt* gene is a housekeeping gene, which is expressed at a low level in most tissues with highest expression in brain and testes (Watts *et al.,* 1987). It is one of the three enzymes involved in the inosinate cycle (figure 1.10) (reviewed by Stout and Caskey, 1988). The enzyme acts by salvaging hypoxanthine and guanine by transfer of a phosphoribosyl group from the 5' phosphoribosyl-1-pyrophosphate (PRPP) to generate inosine monophosphate (IMP) or guanosine monophosphate (GMP), respectively. These mononucleotides exert a strong feedback control on de novo synthesis pathway, which will be borrowed if purines are not provided in the medium or if *hprt* is not functional. HPRT-deficiency results in uric acid overproduction due to the inability to recycle either hypoxanthine or guanine, which interrupts the inosinate cycle producing a lack of feedback control of synthesis, accompanied by rapid catabolism of these bases to uric acid. PRPP not utilised in the salvage reaction of the inosinate cycle provide an additional stimulus to the *de novo* synthesis and uric acid overproduction.

The *hprt* gene has proved to be a very useful system in the development of targeting strategies as well as in the study of HR in m ES cells (Thomas and Capecchi, 1987; Doetschman *et al.*, 1987). Two features make the *hprt* gene a great tool for gene targeting strategies. *Hprt* is X-linked, therefore the disruption of a single allele is sufficient to lead to complete loss of function in male cell lines. Additionally, *hprt* is a selectable gene as its function may be selected either for (HAT selection) or against (6-TG) in cultured cells by chemical selection rendering targeting frequencies easy to calculate at a phenotypic level (Albertini,

2001). The drug aminopterin is known to block the *de novo* pathway as well as the thymidilate synthase reaction. Therefore in HAT (hypoxanthine-aminopterin-thymidine) selection only wild type cells will be able to survive selection as they can use the salvage pathway (Szybalski & Szybalska, 1962). In contrast, HPRT-deficient cells have both routes affected and will die under HAT selection. HPRT is also capable of phosphoribosylating the toxic guanine analogue 6-thioguanine (6-TG), which will be fatal for wild type cells if this is added to the medium (Stutts & brockman, 1963). Since HPRT-deficient cells are unable to salvage free purine bases, they are not affected when 6-TG is added to the culture medium.

The *hprt* gene also provides the advantage that transgenes can be introduced into a desirable chromatin environment with the appropriate regulatory elements. Reproducible transgene expression has been previously demonstrated at the *hprt* gene in a variety of cell lines including mES cells (Evans *et al.*, 2000; Guillot *et al.,* 2000; Cvetkovic *et al.,* 2000; Minami *et al.*, 2002; Heaney *et al.*, 2004) and should therefore be a good candidate gene in hES cells for targeting transgene with predictable levels of expression.

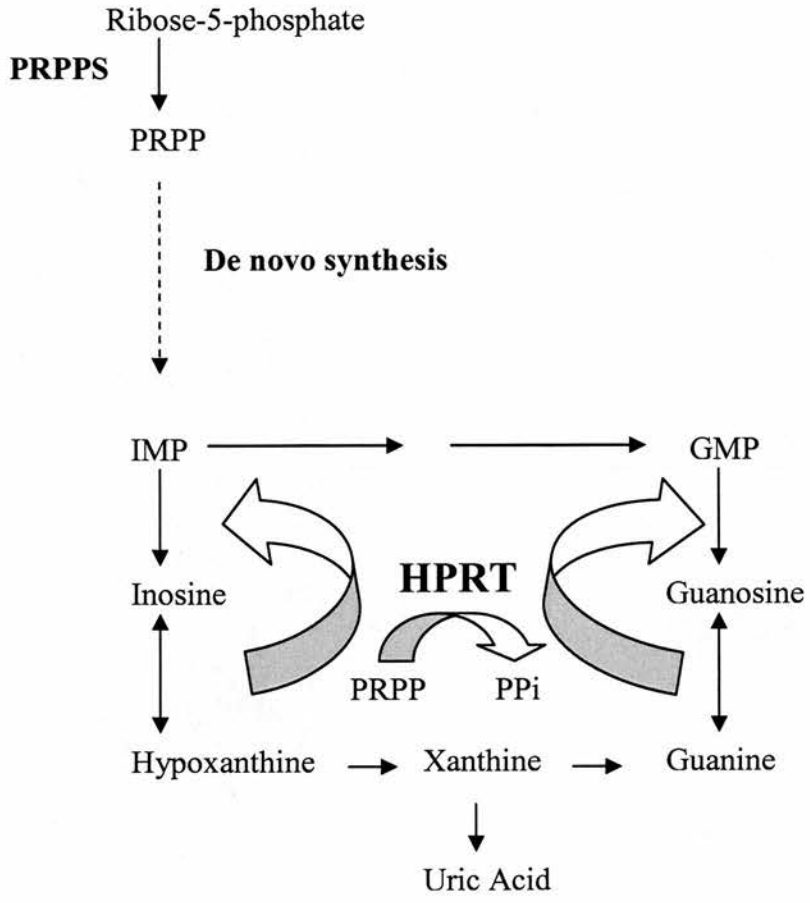


Figure 1.10. HPRT in the purine salvage pathway. See text for details. PRPPS: phosphoribosyl-1-pyrophosphate synthetase. PRPP: phosphoribosyl-1-pyrophosphate. PPi: pyrophosphate. IMP: inosine monophosphate. GMP: guanosine monophosphate.

1.5. Project Objectives

The aim of this project is to obtain a human ES cell line with predictable levels of transgene expression at the *β-cas*ein and *hprt* loci. Such cell lines will be a critical tool in research and medicine. To achieve this aim, the objectives below are to be followed:

1- To introduce heterospecific *lox* sites by homologous recombination into the *β-casein* and *hprt* genes in hES cells (Chapters 3 and 4).

2- To investigate whether a hES cell line targeted at the *β-casein* or *hprt* genes, retains a normal karyotype, can differentiate *in vitro* towards the three germ layers, and express markers characteristic of hES cells (Chapter 5).

3- To correct, by homologous recombination, the targeted mutation previously generated in hES cells at the *hprt* gene, in a functional and structural levels (Chapter 6).

4- To introduce a series of lineage-specific markers by site-specific recombination into the *β-casein* and *hprt* genes in hES cells (Chapter 7).

5- To characterise transgene expression at both *β-casein* and *hprt* loci following differentiation (Chapter 7).

CHAPTER II

Materials & Methods

2.1. DNA Manipulation

2.2. DNA Analysis

2.3. Plasmid and Probes Construction

2.4. Cell Culture

2.1. DNA Manipulation

2.1.1. DNA Extraction

2.1.1.1. Extraction of Genomic DNA from Human ES Cells
2.1.1.2. Extraction of Plasmid DNA from Bacteria
Minipreps
Maxipreps

2.1.2. DNA Purification

2.1.2.1. Ethanol Precipitation
2.1.2.2. Gel Extraction
2.1.2.3. DNA Clean Up from Enzymatic Reactions

2.1.3. DNA Quantitation

2.1.3.1. UV Spectrophotometer
2.1.3.2. Gel Quantification

2.1.4. Cloning

2.1.4.1. TopoTA Cloning
2.1.4.2. Blunting
2.1.4.3. Dephosphorilation
2.1.4.4. Ligation
2.1.4.5. Transformation

2.1.1. DNA Extraction

2.1.1.1. Extraction of Genomic DNA from Human ES Cells

Genomic (g) DNA was extracted using a mammalian DNA isolation procedure adapted from Laird *et al.* (1991). Lysis buffer was prepared by mixing 100 mM Tris-HCl pH8.5 (50 ml from 1 M Tris pH 8.5 stock), 5 mM EDTA pH 8.4 (5 ml from 0.5 M EDTA stock) and 200 mM NaCl (33.3 ml from 3 M NaCl stock), for a total volume of 490 ml in distilled water. The solution was autoclaved and then supplemented with 10 ml of 10% SDS (0.2 % final concentration (f.c)). Cells were treated with lysis buffer (3 ml per 25 cm^2 flasks) supplemented with Proteinase K (Promega) (100 µg/ml f.c) prior to use. Lysates were subsequently transferred to a suitable tube (depending on the sample volume) and incubated overnight at 55 °C. The next day, lysates were incubated at 37 °C for 1 hour with ribonuclease A diluted 1:100 (32.5 mg/ml stock, Sigma) and a further 2 hours with 100 µg/ml Proteinase K at 55 °C. One volume of isopropanol was added to each lysate and samples were mixed until a white DNA precipitate was visible. For cell cultures with surface area ≥ 10 cm^2, gDNA was recovered in an Eppendorf tube by fishing the precipitate with an inoculating loop and needle (Nunc). Alternatively DNA precipitates, obtained from culture vessels below a surface area of 10 cm^2, were recovered by centrifuging for 5 minutes at 14,000 x g in a table-top microcentrifuge (miniSpin plus, Eppendorf). In both cases, gDNA was subsequently washed with 70% ethanol by mixing the samples and centrifuged for 5 minutes at 14,000 x g. DNA pellets were air-dried for 5 minutes, dissolved in a suitable volume of TE buffer (10mM Tris-HCl, pH 7.5, 1mM EDTA) and incubated at 65 °C for several hours until the DNA was dissolved properly.

2.1.1.2. Extraction of plasmid DNA from bacteria

Minipreps

Small-scale extraction of plasmid DNA from bacteria were performed using the QIAprep Spin Miniprep kit (Qiagen). Individual bacterial colonies were picked and incubated overnight at 37 °C in 15 ml Falcon tubes containing 3ml of autoclaved LB (Luria-Bertani broth) medium (10 g tryptone, 5 g yeast extract, 5 g NaCl, 1 litre of distilled water, pH 7.5) supplemented with 100 μg/ml of ampicillin or 50 μg/ml of kanamycin depending on the plasmid antibiotic resistance required. 1.5 ml of overnight cultures were transferred to Eppendorf tubes and centrifuged for 5 minutes at 14,000 x g. Bacterial pellets were resuspended in 250 μl Buffer P1 (50 mM Tris-HCl, pH8.0; 10 mM EDTA, 100 μg/ml RNase A) and 250 μl Buffer P2 (200 mM NaOH, 1% SDS) was added to lyse the cells. Lysis reactions were stopped with the addition of 350 μl of Buffer N3 (3.0 M potassium acetate, pH5.5), gently mixed and centrifuged for 10 minutes at 17,900 x g. Resulting supernatants were applied to a QIAprep Spin Column and centrifuged for 1 minute at 14,000 x g. Columns were washed by adding 750 μl Buffer PE (1.0 M NaCl; 50 mM MOPS, pH7.0; 15% isopropanol (v/v)) centrifuged for 1 minute at 14,000 x g. The flow-through was discarded and columns were centrifuged for an additional 1 minute to remove residual wash buffer. Columns were then placed in Eppendorf tubes and DNAs were eluted with the addition of 50 μl Buffer EB (10 mM Tris-HCl, pH8.5) or distilled water by centrifuging 1 minute at 14,000 x g.

Maxipreps

The QIAGEN Plasmid maxi kit was used for the large scale extraction of plasmid DNA from bacteria. Plasmid DNA prepared by this method was used for stable and transient transfections and for the construction of the Southern DNA probes. Typically, one individual bacterial colony was picked from a selective plate into a 15 ml Falcon tube containing 5 ml of autoclaved LB medium supplemented with the appropriate selective antibiotic and incubated for 8 hours at 37 °C at 225 rpm in a orbital shaker. Starter culture was diluted (1/500) into 200 ml of LB complemented with the selective antibiotic in a 1 litre conical flask and grown overnight at 37°C in a orbital shaker at 225 rpm. After 12 to 16 hours growth, the bacterial culture was transferred into a 250 ml dry-spin™ bottle (Sorvall) and the cells were harvested by centrifugation in a Sorvall RC-5B rotor at 6000 x *g* for 15 minutes at 4°C. After removing all traces of supernatant, the bacterial pellet was resuspended in 10 ml of a Resuspension Buffer P1 (50 mM Tris.Cl pH 8.0, 10 mM EDTA, 100 µg/ml RNase) and 10 ml of a Lysis Buffer P2 (200 mM NaOH, 1% SDS) was added with thorough mixing. The lysis reaction was carried out at room temperature for 5 minutes. 10 ml of chilled Neutralization Buffer P3 (3.0 M potassium acetate, pH 5.5) was added and the sample was incubated on ice for 20 minutes. The resulting precipitate was then centrifuged at 20,000 x g for 30 minutes at 4°C. The supernatant containing the plasmid DNA was carefully filtered through a filter paper (Whatman) into a 50 ml test tube. The resulting clear supernatant was then applied to a QIAGEN-tip 500 previously equilibrated with 10 ml of Buffer QBT (750 mM NaCl, 50 mM MOPS pH 7.0, 15 % isopropanol, 0.15% Triton X-100) and was allowed to enter the resin by gravity flow. The QIAGEN-tip was washed twice

with 30 ml of Buffer QC (1.0 M NaCl, 50 mM MOPS pH 7.0, 15 % isopropanol) to remove all contaminants of plasmid DNA preparation and the DNA was eluted with 15 ml of Elution Buffer QF (1.25 M NaCl, 50 mM Tris.Cl pH 8.5, 15 % isopropanol) previously heated at 55°C. The eluted plasmid DNA collected in Corex tubes was precipitated by adding 10.5 ml (0.7 volumes) of room-temperature isopropanol and centrifuged at 15,000 x g for 30 minutes. The supernatant was decanted carefully and the DNA pellet was washed with 5 ml of room-temperature 70% ethanol and centrifuged at 15,000 x g for 10 minutes. The supernatant was decanted carefully and the DNA pellet was then air-dried for 10 minutes, dissolved in 500 µl of TE buffer and transferred into a 1.5 ml centrifuge tube.

2.1.2. DNA Purification

2.1.2.1. Ethanol Precipitation

Purification of plasmids used for transfection experiments was carried out as follows: Plasmid DNA in solution was precipitated by the addition of 1/10 volume of 3 M sodium acetate and 2.5 volumes of 100% ethanol. The sample was then centrifuged for 10 minutes at 14,000 x g in a table-top microcentrifuge and the resulting DNA pellet was washed in 500 µl 70% ethanol by centrifuging for 5 minutes at 14,000 x g. Excess ethanol was removed in a tissue culture hood to avoid contamination and the pellet was air-dried and resuspended in distilled water to get a final DNA concentration of 1µg/µl.

2.1.2.2. Gel Extraction

DNA was extracted and purified from standard or low-melting point agarose (Invitrogen) gels using the QIAquick gel extraction kit protocol (QIAGEN). A cut was made in the gel immediately in front of the fragment to be purified with a clean, sharp scalpel. The gel slice was transferred to an Eppendorf tube and weighed. Three volumes of Buffer QG were added to one volume of gel (100 mg ~ 100 µl). The sample was incubated at 50°C for ten minutes in a water bath to melt the gel slice. One gel volume isopropanol was added to the sample and mixed. The sample was applied to a QIAquick spin column that was placed in a 2 ml collection tube and centrifuged at 14,000 x g for one minute to bind the DNA. After discarding the flow-through and placing the QIAquick column back in the same collection tube, 500 µl of Buffer QG was added to the column to remove all traces of agarose. The column was centrifuged for 1 minute at 14,000 x g. 750 µl of Buffer PE complemented with 100% ethanol was added to the column to wash the DNA and the column was then centrifuged for 1 minute at 14,000 x g. The flow-through was discarded and the column was centrifuged for an additional 1 minute in order to remove residual ethanol from Buffer PE. The column was then placed into a sterile 1.5-ml Eppendorf microfuge tube and the DNA was eluted with 50 µl of elution buffer EB by centrifuging the column for 1 minute at 14,000 x g.

2.1.2.3. DNA Clean Up from Enzymatic Reactions

DNA clean up from enzymatic reactions were performed using the same QIAquick gel extraction kit protocol but 3 volumes of Buffer QG and 1 volume of isopropanol was directly added to the enzymatic reaction. Subsequent steps were carried out as described above.

2.1.3. DNA Quantitation

2.1.3.1. UV Spectrophotometer

Typically, 1:200 DNA dilution was used to determine the concentration. Quantitation was conducted in an UNICAM 5625 UV/VIS spectrophotometer at a wavelength of λ = 260 nm. Since 1 unit of OD corresponds to approximately 50 μg/ml of double-stranded DNA, concentrations could be determined using the following formula:

$$\mu g\ DNA/\ \mu l = OD \times \text{dilution factor} \times 50$$

2.1.3.2. Gel Quantification

When there was not sufficient material for spectrophotometer analysis, the DNA was quantified by loading an aliquot on an agarose gel supplemented with ethidium bromide along with 5 μl Bioline DNA hyper ladder molecular marker. DNA concentration could be determined by comparing the intensity of the bands with the corresponding size of the molecular marker.

2.1.4. Cloning

2.1.4.1. TopoTA cloning

All PCR products produced for the construction of plasmids were first cloned into the Invitrogen pCR2.1-Topo TA vector (appendix 1, section 2.1) prior to subsequent cloning steps. *Taq*-amplified PCR products are characterised by a single 3'-adenosine (A) overhangs on both ends. The linearised pCR2.1-Topo TA vector contains 3'-thymidine (T) overhangs and two topoisomerase I covalently bound to the vector (referred to as "activated" vector) that enable the direct ligation of PCR inserts. Typically, 4 µl of purified PCR product was used in a Topo cloning reaction. 1 µl of salt solution and 1 µl of pCR2.1-Topo TA vector were added to obtain a final volume of 6 µl. Reactions were mixed and incubated for 5 minutes at room temperature (RT), and 2 µl of each reaction was used for subsequent transformations of DH5α competent cells (see section 2.1.4.5).

2.1.4.2. Blunting

When required, blunt ends were generated with the DNA polymerase I large fragment (klenow). Klenow fragment (New England Biolabs) is a proteolytic product of E.*coli* DNA polymerase I which retains 5'-3' polymerisation and 3'-5' exonuclease activity but lacks 5'-3' exonuclease activity (Joyce and Grindley, 1983). Typically, 1 to 5 units of klenow and 1.65 µl of 1 mM dNTP mix (33 µM f.c) was added to the purified DNA and the reaction was incubated at room temperature for 15 minutes. The reaction was subsequently inactivated by incubation at 75 ° C for 20 minutes and DNA was purified using the QIAquick gel extraction kit protocol .

2.1.4.3. Dephosphorylation

Prior to ligation between the vector and insert DNA, the linearised vector DNA was dephosphorylated with shrimp alkaline phosphatase (SAP) (Roche), to prevent self-religation of DNA ends. 1 unit of SAP was used to catalyse the dephosphorylation of up to 1 pmol 5'-terminal phosphorylated 5'-protruding or 5'-recessive ends DNA vectors, and up to 0.2 pmol 5'-terminal phosphorylated blunt-ended DNA vectors for 60 minutes at 37 ° C. The reaction was subsequently inactivated by incubation at 65° C for 15 minutes and the vector DNA was purified using the QIAquick gel extraction kit protocol.

2.1.4.4. Ligation

Ligations of sticky-end and blunt-end DNA fragments (200 ng maximum) were performed using the Rapid DNA ligation kit (Roche). As most plasmid construction involved blunting of both vector DNA and insert DNA, the molar ratio of vector to insert DNA was usually between 1:3 to 1:10. For instance for a 1:3 ligation, 100 ng of purified dephosphorilated vector DNA (10 kb) and 45 ng of purified insert DNA (1.5 kb) were dissolved in 1× DNA dilution buffer in a total volume of 10 µl. After thorough mixing, 10 µl of 2× T4 DNA ligation buffer was added to the vial and mixed thoroughly before adding 1 µl of T4 DNA ligase (5 units). Mixed ligation reactions were subsequently incubated for 5 minutes at room temperature. Occasionally, when the total volume of DNA solution in 1× DNA dilution buffer was greater than 10 µl, the volume of all reagents in the reaction was increased accordingly and the ligation reactions were incubated for 30 minutes at room temperature. A maximum of one-tenth of the volume ligation reaction was

used for the transformation of Subcloning Efficiency (SE) competent cells of the DH5α bacterial strain (Life Technologies) to avoid an inhibition by a surplus of DNA.

2.1.4.5. Transformation

The *Escherichia Coli* Subcloning Efficiency (SE) DH5α strain was used for the routine propagation of all plasmids and DNA from ligation reactions. SE DH5α competent cells were thawed on ice and typically, 50µl of cells were incubated with 50-500 ng of plasmid DNA solution in a 1.5 ml pre-chilled microcentrifuge tube for 30 minutes. Cells were heat-shocked for 20 seconds at 37°C and incubated on ice for 2 minutes. 900 µl of autoclaved room-temperature LB was added to the mixture and the cells were grown for 1 hour at 37ºC in a orbital shaker at 225 rpm. Cultures were diluted as necessary and plated onto 90 mm agar plates supplemented with 100 µg/ml of ampicillin or 50 µg/ml of kanamycin (depending on the plasmid selective antibiotic) next to a Bunsen Burner to avoid contamination. Plates were inverted and incubated at 37°C overnight.

2.2. DNA Analysis

2.2.1. Agarose Gel Electrophoresis

Agarose gels were prepared by dissolving agarose (BDH) at 0.8-1.5 % (w/v) in 1X TAE buffer (e.g. 980 ml dH_20 + 20 ml 50 X TAE buffer (2.42 g Tris base + 57.1 ml glacial acetic acid + 100 ml 0.5M EDTA pH 8.0, in 800 ml dH_20)) by heating the suspension in a microwave oven. Ethidium bromide (0.5 μg/ml) (Sigma) was added to the solution and poured into a gel mould. Once set, the gel was submerged in 1X TAE buffer in a gel tank and the DNA samples, containing 20% DNA loading buffer, were loaded into the wells. Electrophoresis was carried out either for 2 hours (80-120 Volts) or overnight (15-30 Volts). DNA was visualised by illumination on a long wave UV light box and photographed. The size of the DNA fragment was examined by comparison of their mobility to that of restriction fragments of known size, typically DNA Hyperladder 100 lanes (Bioline).

2.2.2. Restriction Enzymes Analysis

Restriction enzyme digestions of plasmid and gDNA were performed according to the manufacturer's instructions (Roche or New England Biolabs). DNA was usually digested with a 10-fold excess of enzyme (typically 10 units of enzyme which digests 10 μg of DNA in 1 hour was used to cut 1μg of DNA) with a final

volume of glycerol not exceeding 5% (v/v). The following example describes the typical proportions of reagents and incubation time in a standard plasmid and gDNA restriction analysis.

Plasmid DNA	1µl (1µg/µl)
10 × Buffer	2µl
Enzyme	1µl (10 units: 1 unit digests 1 µg of DNA in 1 hour)
Distilled water	16µl
Total volume	30 µl
Incubation	1-4 hours at 37 °C

Genomic DNA	10 µl (1 µg/µl)
10 x Buffer	4 µl
Enzyme concentrated	1-2 µl (40-80 units)
Distilled water	25µl
Total volume	40 µl
Incubation	overnight at 37 °C

For analysis of restriction digests, reactions were stopped by the addition of DNA loading buffer to 20% (v/v) or by heat inactivation according to the manufacturer's instructions.

2.2.3. Polymerase Chain Reaction (PCR)

Polymerase chain reaction was performed using the Expand High Fidelity PCR system (Roche) to amplify DNA fragments up to 6 kb. This system is composed of a unique enzyme mixture containing thermostable *Taq* DNA polymerase and *Tgo* DNA polymerase with proof reading activity, ensuring high fidelity of gDNA (Barnes, 1994). Although PCR conditions had to be adjusted for every individual experiment, a typical setting for DNA amplification of fragments of up to 6 kb can be summarized as follows:

10 mM dNTP mix	1.0 µl (0.2 mM final concentration)
10 µM downstream primer	1.5 µl (300 nM final concentration)
10 µM upstream primer	1.5 µl (300 nM final concentration)
Template DNA	1.0 µl (gDNA 500 ng; plasmid, 10 ng)
Expand HF 10 × buffer with 15 mM MgCl2	5.0 µl
Autoclaved distilled water	39.25 µl
Expand High Fidelity PCR enzyme mix	0.75 µl (2.6 units)

The reaction mix and the DNA were mixed together in 0.2 ml thermo-strip tubes (Abgene) and the samples were placed in a Hybaid gradient thermocycler, which was typically set using the following cycling parameters:

1× denaturating step	94 ° C for 5 minutes
10× cycle 1	Denaturation, 94 ° C for 20 seconds; annealing , 45-70° C for 30 seconds; elongation, 72 ° C or 68 ° C (for PCR products longer than 3 kb) for 1 to 5 minutes
20× cycle 2	Denaturation , 94 ° C for 15 seconds; annealing , 45-70° C for 30 seconds; elongation, 72 ° C or 68 ° C for 1 to 5 minutes (+ 5 seconds increment/cycle to ensure proper elongation as DNA concentration increases overtime)
1× final elongation	72 ° C for 7minutes
Cooling	4 ° C for unlimited time

Annealing temperatures are dependent on the melting temperatures of the primers used and are usually set 5 ° C below that of the melting point. There are several ways to estimate the melting point temperatures of primers and one of them is to add up 4 °C for every G+C and 2 °C for every A+T base pairs. The elongation time was calculated depending on the length of the fragment to be amplified. Typically, a 2 minutes elongation time was used to amplify a 3 kb DNA fragment.

For the generation of longer PCR products, the Expand Long Template PCR system (Roche) was used to amplify DNA fragments above 3 kb. This system is composed of an enzyme mix, containing thermostable *Taq* polymerase and *Pwo* DNA polymerase with proofreading activity, which is designed to amplify DNA with high fidelity. The PCR conditions usually used were as follows:

10 mM dNTP mix	1.0 µl (0.2 mM final concentration)
10 µM downstream primer	1.5 µl (300 nM final concentration)
10 µM upstream primer	1.5 µl (300 nM final concentration)
Template DNA	1.0 µl (gDNA, 500 ng; plasmid, 10 ng)
10 × buffer system 1 (with 17.5 mM MgCl2)	5.0 µl
Autoclaved distilled water	39.25 µl
Expand long template PCR enzyme mix	0.75 µl (2.6 units)

Cycling parameters were as described above with the exception that elongation temperatures were always maintained at 68 ° C.

Table 2.1. List of the PCR primers used in the course of this project. Primers were designed (with the exception of T7 and T3 primers) using the Primer 3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and ordered through MWG.

Gene	Size (bp)	Primer name	Sequence (5' to 3')	Application
β-casein	2371	5' Cas-F 5' Cas-R	ctccttatctatcttccaccctg caatgactggaggacagtgatac	Generation of the human 5' *β-casein* homology
β-casein	5449	3' Cas-F 3' Cas-R	ctacattcacaggacttagtagc gcatcatatttccagtctcagtc	Generation of the human 3' *β-casein* homology
β-casein	348	β-cas probe-F β-cas probe-R	ccgtctttgtatcacgttaccc aaagaggccgattaggtttga	Generation of the human *β-casein* Southern probe
Hprt	7091	5'ext Hprt 3'ext Hprt	ggagatgctcaataggtactaac gcatgagccatcatcaaatcaag	Generation of the human *hprt* homology
Neo *Hprt*	6712	Neo 5'ext Hprt	gatccggtgaacaagatggattg ggagatgctcaataggtactaac	Detection of a targeting event at the *hprt* locus
Hprt GFP	896	HprtloxM GFP	ctgtaggactgaacgtcttgctcg catcaaggtgaacttcaagatcc	Transient Cre recombination assay between pfBSD-OctGFP and pHPRT-EfN
EF1-α *BSD*	197	EF1α BSD	cgtgaggaattagcttggtac gagggtggattcttcttgagac	Detection of a site-specific recombination event at the *hprt* locus
T7 *T3*	227	T7 T3	taatacgactcactataggg attaaccctcactaaaggga	*In vitro* Cre recombination assay Between pCas-EfN and pB-RMCE2272

2.2.4. Southern Blotting Analysis

Southern blotting (Southern, 1975) allows DNA fragments corresponding to a particular probe to be identified directly from a restriction digest of gDNA. Since DNA fragments cannot be handled directly in an agarose gel, it is necessary to transfer them to a membrane (nitrocellulose or nylon) which provides a suitable support. After immobilisation, the DNA can be subjected to hybridisation analysis, enabling bands with identical sequence to a labelled probe to be identified. The following protocol was adapted from Sambrook *et al.* (1989).

gDNA digests (5-10 μg) were loaded in a 0.8% agarose gel and run either for 3-4 hours at 100-120 Volts or overnight at 30-40 Volts. A UV photograph of the gel next to a ruler (to determine the distance the fragments had run following autoradiography) was taken before proceeding to denaturation, an essential step for subsequent hybridisation of single strands DNA probes. Denaturing solution was prepared by dissolving 87.66 g of NaCl (1.5 M final concentration) and 20 g of NaOH (0.5 M final concentration) in 1 litre of distilled water (solution A); and 87.66 g of NaCl (1.5 M final concentration), and 500 ml of 1 M Tris HCl pH 7.5 (0.5 M final concentration) and 2ml of 0.5 M EDTA pH 8.0 (1 mM final concentration) in one litre of distilled water (solution B). DNA fragments in agarose gels were denatured by gently shaking the gel 2x15 minutes in solution A and 2x15 minutes in solution B. The gel was then transferred to the blotting apparatus which was assembled as follows: A glass plate was placed on top of a tank filled with 20 x SSC buffer (876.5 g of NaCl and 441.0 g of tri-sodium citrate added to 4 litres of dH_2O; pH 7.0). Two sheets of 3 mm Whatman paper were cut (as the width of the gel) and placed over the glass plate such that the edges were submerged in 20 x SSC. The gel

was carefully transferred wells side down onto the 3 mm paper and surrounded with strips of parafilm such that the buffer is drawn from the tanks by capillary transfer only through five 3 mm Whatman papers (two in direct contact with the buffer and three gel-sized papers on top of the gel). One sheet of positively charged hybond nylon membrane (Amersham International plc) was cut to the size of the gel, rinsed in 3 x SSC using forceps and placed over the gel. Three sheets of 3 mm paper cut to the size of the gel, comprising one sheet rinsed with 3 x SSC and two dry sheets, were placed on top of the stack. Air bubbles were removed carefully between each step. In the following order, paper towels evenly stacked, a glass plate and a 500 ml filled bottle used as a weight were placed on top of the stack. The following day, the nylon membrane was rinsed in 3 x SSC with the DNA transferred to it side up. The membrane was then air dried on 3 mm paper, crossed linked in a UV stratalinker 1800 (Stratagene), wrapped in Saran wrap and stored at room temperature.

The membrane carrying the DNA was pre-hybridised for 1 to 2 hours at 65°C in a hybridisation tube with 25 ml of a pre-hybridisation solution. Prehybridisation solution was prepared with 50 ml of 1 M sodium phosphate (0.5 M final concentration), 35 ml of 20% SDS (7% final concentration), 200 µl of 0.5 M EDTA (1 mM final concentration) and 15 ml of deionised water.

DNA probes for hybridisation (~25 ng) were dissolved in 45 µl of 10 mM Tris HCl pH8.0 containing 1 mM EDTA (TE buffer), boiled for five minutes for denaturation and snap cooled on ice for five minutes. The denatured DNA was added to the labelling reaction tube (amersham pharmacia biotech) which consists of a buffered solution of dATP, dGTP, dTTP, exonuclease free klenow enzyme and random primers in a dried, stabilised form. Denatured probes were radioactively

labelled with 5 µl of α ^{32} P-dCTP (Amersham International plc) for 10 to 30 minutes at 37 ° C.

To separate labelled probes from unincorporated α ^{32} P-dCTP nucleotides, the labelled reaction mixture was applied to a Pharmacia NICK column, in which the gel was previously equilibrated with 3 ml of TE buffer pH7.5, and 400 µl of TE buffer pH7.5 was added to the column. The first sample was collected into a test tube and discarded into the radioactive waste. The purified sample was eluted with an additional 400 µl of TE buffer into a fresh 1.5 ml Eppendorf tube to which 40 µl of 5M NaOH (40g/mol) was added for denaturation. The resulting hybridisation solution was added to the tube containing the filter in prehybridisation solution and was incubated at 65° C overnight in an oven with rotor.

The following day, the hybridisation solution was discarded and the filter was washed at 65 ° C 2×15 minutes in solution 1 (2 x SSC and 0.1 % SDS) and 2×15 minutes in solution 2 (0.2 x SSC and 0.1 % SDS). The filter was carefully wrapped in saran wrap and placed into a phosphor image cassette overnight. The next day, the exposed screen was placed face down onto a PhosphorImager to reveal the autoradiography. Alternatively, the membrane was exposed to X-ray film (Fuji) at –80° C during one week which was revealed in an autoradiograh.

2.2.5. Cre Recombinase *In Vitro* Assay

Cre recombinase is a type I topoisomerase from bacteriophage P1 that catalyses the site-specific recombination of DNA between *loxP* sites (Abremski and Hoess, 1984). The enzyme requires no energy cofactors and Cre-mediated recombination quickly reaches equilibrium between substrate and reaction products. Recombination products depend on the location and relative orientation of the *lox* sites. Cre recombinase can catalyse (i) the excision of DNA between two identical *lox* sites in circular form, (ii) the fusion of two DNA species containing single *lox* sites, (iii) the inversion of DNA between opposite *lox* sites with respect to external sequences, and (iv) the reciprocal exchange of DNA flanked by heterospecific *lox* sites between two DNA species provided that these are flanked by the same set of heterospecific *lox* sites.

Prior to transfection, all plasmids, containing heterospecific *lox* sites, used for recombinase-mediated cassette exchange experiments were tested *in vitro* with Cre recombinase in the following manner: Two plasmids with the same set of heterospecific *lox* sites were linearised with an appropriate restriction enzyme, purified with the QIAquick gel extraction kit protocol and gel quantified. Typically, 100 ng of each DNA species were incubated along with 5 µl (1 unit/µl) of Cre recombinase (New England Biolabs), 5 µl of 10 × Cre Recombinase buffer (New England Biolabs) in a total reaction volume of 50 µl. Reactions were incubated for 30 minutes at 37 ° C followed by 10 minutes at 70 ° C to inactivate the enzyme. To determine recombination between the two DNA plasmids (cassette exchange), 5 to 10 µl of the reaction was used as template in a PCR reaction using a primer set specific to one of the recombination product only (table 2.1). The presence of a PCR

product with the predicted size, indicating recombination, was subsequently determined by agarose gel analysis.

2.3. Plasmid and Probes Construction

2.3.1. Plasmid Material
2.3.2. Plasmid Construction

2.3.2.1. pCas-EfN
2.3.2.2. pHPRT-EfN
2.3.2.3. pTopo-HPRT
2.3.2.4. pfBSD-OctGFP
2.3.2.5. pEfBSD-OctGFP

2.3.3. Construction of probes

2.3.3.1. 3' External *β-Casein* Probe
2.3.3.2. 3' Internal *Hprt* Probe

2.3.1. Plasmid Material

The following commercial plasmids were used in this project: pBlueScript II SK +/- (Stratagene) and pCR2.1-TOPO (Invitrogen).

The plasmids listed below were obtained from non-commercial sources: pMMPoNeo (given by M. Marques), pEFGIN (given by Geron), pB-RMCE2272 (given by A. Kolb), phOCT-4-EGFP-1 (given by D. Zhao), pLoxBSD1B and pPGKGFP-SV40Neo (given by A. Thomson), pHS4PGKGFP-SV40Neo (given by T. Lacroix), pNI-CD (given by A. West), pBSPGKNeo (given by E. Gallagher) pE3ΔNeo (provided by A. Bradley), pTurboCre (gift from T. Ley), pHP-Hind/Xho (given by D. Russell).

2.3.2. Plasmid Construction

2.3.2.1. pCas-EfN

The targeting vector pCas-EfN (figure 2.1) was built as follows: The 5'arm of homology was generated by PCR amplification (from gDNA extracted from the H9 female human ES cell line) of a 2371 bp fragment (3' end of the promoter region) of the *β-Casein* gene, using the 5' Cas-F and 5' Cas-R primers (table 2.1). The resulting PCR product was subcloned into the pCR2.1-Topo TA vector (appendix 1, section 2.1) to give pTopoCas5'. Similarly, the 3' arm of homology was generated by PCR amplification of a 5449 bp fragment of the open reading frame region of the gene (encompassing exon 2 to start of exon 8), using the 3' Cas-F and 3' Cas-R primers (table 2.1). The resulting PCR product was subsequently subcloned into the pCR2.1-Topo TA vector to produce pTopoCas3'. A 2388 bp EcoRI fragment of pTopoCas5' containing the 5' genomic region of homology was subcloned in the desired orientation into a EcoRI site in the multicloning site of the pBlueScript II SK +/- (Stratagene) (appendix 1, section 2.2) to give pCas5'. A 5538 bp KpnI/XhoI fragment of pTopoCas3' containing the 3' arm of homology of the *β-Casein* gene was subsequently subcloned into a KpnI/XhoI-digested pCas5' in the same transcriptional orientation as the 5' arm of homology, resulting in pCas5'/3'. In a third step, the human *EF1-α* promoter was extracted from the plasmid pEFGIN (appendix 1, section 2.3) as a 1511 bp NheI/SalI fragment, blunted with the DNA polymerase I Klenow fragment (New England Biolabs), and subcloned into the XhoI-linearised and blunted pCas5'/3', to produce pCasEF1α. As a final step, a *neoHGHpA* cassette flanked by heterospecific *lox* sites (*loxP* and a *lox2272*) was extracted as a 1775 bp NotI/Asp718 (KpnI isoschizomer) from the plasmid

pMMpoNeo (appendix 1, section 2.5), blunted with Klenow, and subcloned into SalI-linearised and blunted pCasEF1-α, to generate the final version of the *β-Casein* targeting construct, pCas-EfN (14.2 kb).

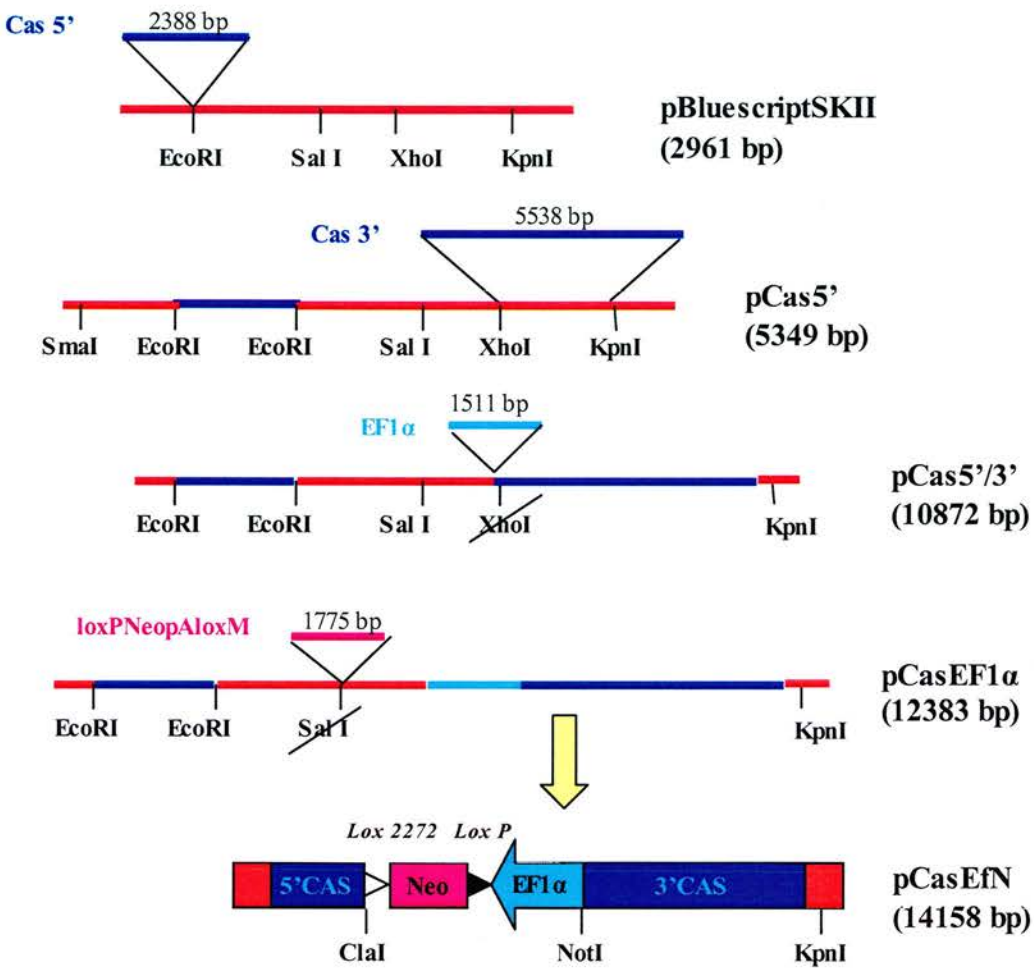


Figure 2.1. Construction of the human *β-Casein* construct, pCas-EfN. The 5' and 3' regions of homology of the *β-Casein* gene (7909 bp in total), the *EF1α* promoter, the *neo* cassette flanked by heterospecific *lox* sites, and the pBlueScript II SK +/- backbone are indicated as dark blue, light blue, pink and red lines, respectively. *LoxP* and *lox2272* are represented as plain and white arrows, respectively. The XhoI and SalI restrictions sites were eliminated following blunting, as described in the text. Linearisation restriction sites are indicated in the final version of the *β-Casein* targeting construct.

2.3.2.2. pHPRT-EfN

The human *hprt* targeting vector pHPRT-EfN (figure 2.2) is derived from the pE3ΔNeo plasmid (appendix 1, section 2.6). The pE3ΔNeo plasmid (gift from A. Bradley) contains a 6.9 kb HindIII *hprt* fragment (Edwards *et al.*, 1990) comprising exons 2 and 3, intron 2 and fragments of introns 1 and 3. A MC1*neo*pA cassette (Thomas & Capecchi, 1987) inserted at the XhoI site in exon 3, in the opposite transcriptional orientation, divides the *hprt* homology into a 4.8 kb 5' and 2.1 kb 3' elements. To generate pHPRT-EfN, the MC1*neo* cassette from pE3ΔNeo was extracted by restriction digestion with NcoI and XhoI. A 3.2 kb ClaI/NotI Klenow blunted EF1α-*LoxP-neopALox2272* from pCas-EfN (figure 2.1) was subsequently subcloned into the NcoI/XhoI-linearised and blunted pE3 (pE3ΔNeo lacking the MC1*neo*pA cassette) to produce pHPRT-EfN.

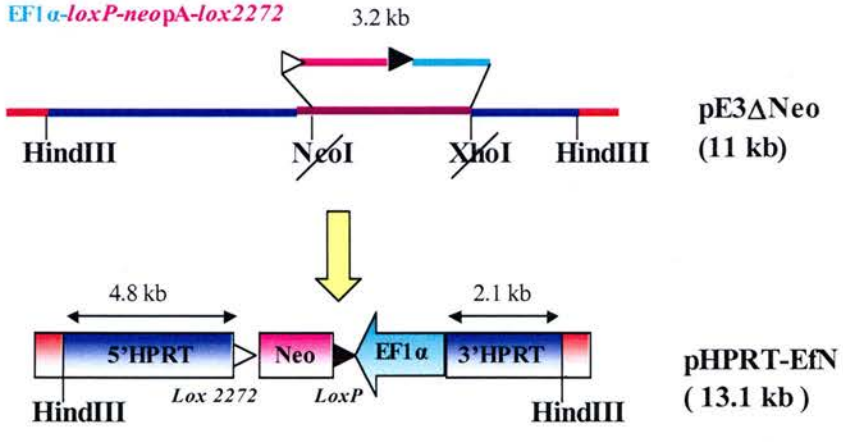


Figure 2.2. Construction of the human *hprt* targeting construct, pHPRT-EfN. The 5' and 3' regions of homology of the *hprt* gene (6909 bp in total), the floxed *neo* expression cassette, and the pT218R backbone are indicated as dark blue, pink, and red boxes, respectively. The *EF1-α* promoter is shown as plain light blue arrow. The *loxP* and *lox2272* sites are represented as small plain and white arrows, respectively. The MC1*neo*pA (in brown) was eliminated from pE3ΔNeo by XhoI and NcoI restriction digestions, and the resulting pE3 plasmid was blunted with Klenow, as described in the text. HindIII digestion of pHPRT-EfN results in the relevant targeting sequence used for the series of targeting experiments.

2.3.2.3. pTopo-HPRT

To generate the pTopo-HPRT targeting vector, a 7.1 kb fragment of the *hprt* gene was amplified by PCR from gDNA extracted from the H1 male human ES cell line, using the 5'ext *hprt* and 3'ext *hprt* primers (table 2.1). The resulting PCR product was subcloned into the pCR2.1-Topo TA vector to give pTopo-HPRT. The 7.1 kb of *hprt* homology comprises the 6.9 kb HindIII *hprt* fragment included in the pHPRT-EfN targeting construct (figure 2.2).

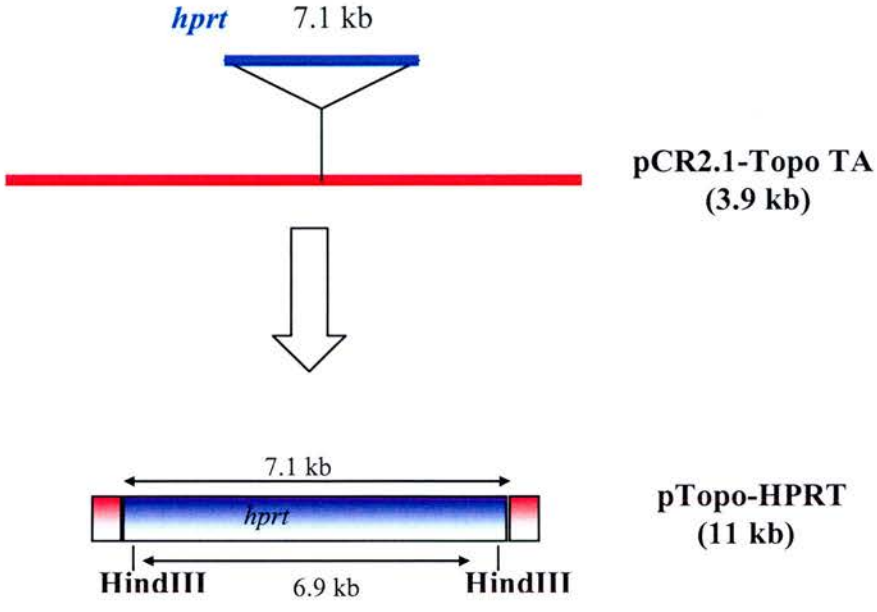


Figure 2.3. Construction of the human *hprt* targeting construct, pTopo-HPRT. The *hprt* homology (7.1 kb) and the backbone of pCR2.1-TopoTA are indicated as dark blue and red boxes, respectively. HindIII digestion of pTopo-HPRT results in the relevant targeting sequence used for the targeting experiment.

2.3.2.4. pfBSD-OctGFP

The ploxBSD1B plasmid (appendix 1, section 2.7), containing a *BSD-SV40pA* under the transcriptional control of the *SV40* promoter and flanked by a *loxP* and *lox2272* sites, was used to generate the pfBSD-OctGFP plasmid (figure 2.4). In a first step, the *SV40* promoter was removed by HindIII digestion, and the resulting ploxBSD1BSV40$^-$ was religated. In a second step, a 5 kb BglII/SspI *oct-4*-GFP cassette was extracted from the phOCT-4-EGFP-1 plasmid, blunted with the DNA polymerase I Klenow fragment, and subcloned downstream of the *BSD-SV40pA* (in the same orientation) into the BamHI-linearised and blunted ploxBSD1BSV40$^-$, to produce pfBSD-OctGFP.

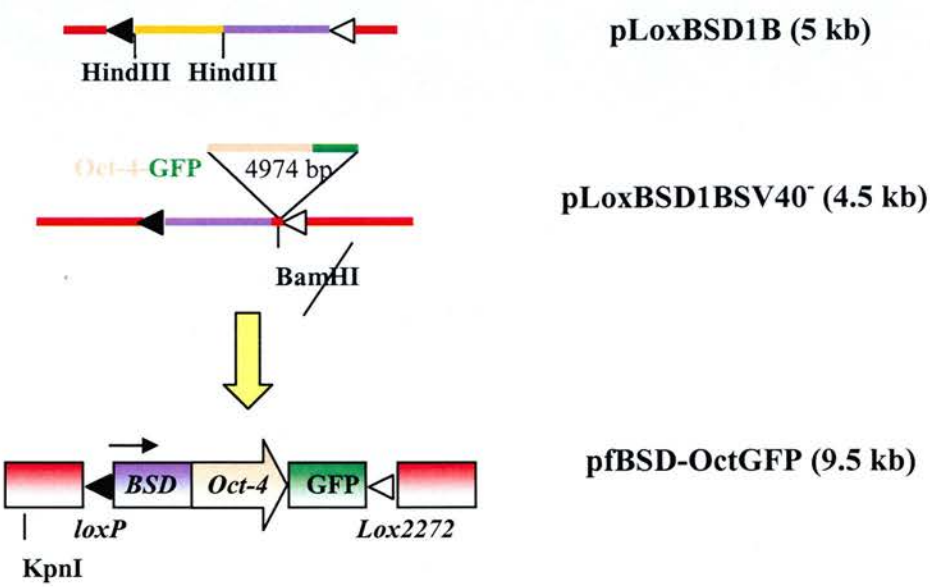


Figure 2.4. Construction of the RMCE construct, pfBSD-OctGFP. The *SV40* promoter, the *BSD* gene, the *oct-4* promoter, the GFP, and the plasmid pBlueScript II KS +/- backbone are indicated as yellow, purple, pink, green, and red lines, respectively. *LoxP* and *lox2272* are represented as plain and white arrows, respectively. The KpnI linearisation restriction site is indicated in the final version of the pfBSD-OctGFP construct.

2.3.2.5. pEfBSD-OctGFP

The plasmid pEfBSD-OctGFP was generated by subcloning the human *EF1-α* promoter upstream of the *loxP* site in pfBSD-OctGFP, in the same transcriptional orientation as the *BSD* gene (figure 2.4). The *EF1-α* promoter was obtained by extraction from the plasmid pEFGIN (appendix 1, section 2.3) as a 1511 bp NheI/SalI fragment, blunted with the DNA polymerase I Klenow fragment (New England Biolabs), and subcloned into the KpnI-linearised and blunted pfBSD-OctGFP (figure 2.4), to give pEfBSD-OctGFP (figure 2.5).

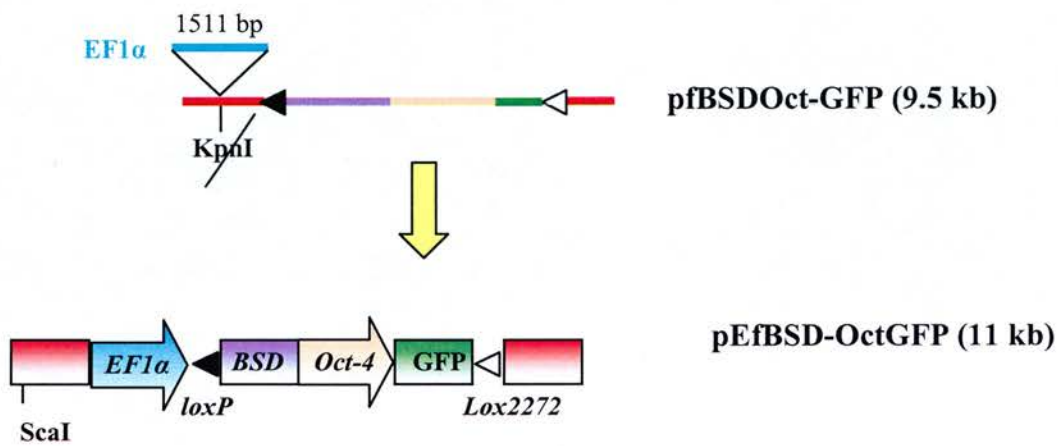


Figure 2.5. Construction of the pEfBSD-OctGFP plasmid. The *BSD* gene, the *oct-4* promoter, the *EF1-α* promoter, the GFP, and the plasmid pBlueScript II KS +/- backbone are indicated as purple, pink, blue, green, and red lines, respectively. *LoxP* and *lox2272* are represented as plain and white arrows, respectively. The ScaI linearisation restriction site is indicated in the final version of the pEfBSD-OctGFP construct.

2.3.3. Contruction of Probes

2.3.3.1. 3' External *β-Casein* Probe

To generate the *β-casein* probe (figure 2.6) for Southern Blot analysis, a 348 bp (encompassing region of exon 8 and 3' untranslated region of the *β-casein* gene) was amplified by PCR from H9 gDNA using the β-cas probe-F and β-cas probe-R primers (table 2.1), and subcloned into the pCR2.1-TopoTA vector to generate pTopo-Cas. The probe was excised as a 365 bp EcoRI fragment, with the QIAquick gel extraction kit protocol, and gel quantified using 5 µl of Bioline hyper ladder molecular marker.

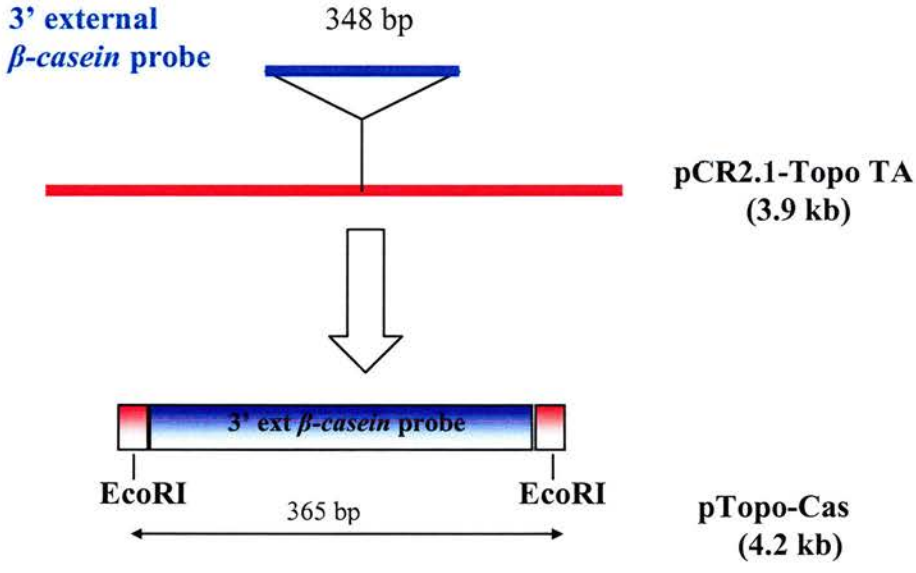


Figure 2.6. Construction of the human 3' external *β-casein* probe. The *β-casein* sequence (348 bp) and the backbone of pCR2.1-TopoTA are indicated as dark blue and red boxes, respectively. EcoRI digestion of pTopo-Cas results in the production of the 365 bp 3' external *β-casein* probe used for the Southern Blot analysis.

2.3.3.2. 3' Internal *Hprt* Probe

The 3' internal *hprt* probe (encompassing regions of exon 3 and intron 3) used for Southern blot analysis was excised from a 1 % agarose gel as a 1193 bp BamHI fragment from the plasmid pHP-Hind/Xho (gift from D. Russell), purified with the QIAquick gel extraction kit protocol and gel quantified using 5 µl of Bioline hyper ladder molecular marker.

2.4. Cell Culture

2.4.1. Cell Lines
2.4.2. Maintenance of Mouse Embryonic Fibroblasts

2.4.2.1. Preparation of MEF medium
2.4.2.2. Isolation of MEFs
2.4.2.3. Passage of MEFs

2.4.3. Maintenance of Human ES Cells

2.4.3.1. Feeder-Free Human ES Cell Culture
Conditioned Medium Preparation
Matrigel Coating for Human ES Cultures

2.4.3.2. Passage Regime

2.4.4. Maintenance of Other Cell Lines
2.4.5. Cryopreservation and Thawing of Cell Lines
2.4.6. Subcloning, Cryopreservation and Thawing of HES Clones
2.4.7. *In Vitro* Differentiation of HES Cells

2.4.7.1. Generation of Embryoid Bodies
2.4.7.2. Osteogenic Differentiation

2.4.8. Transfections

2.4.8.1. Eppendorf Multiporation
2.4.8.2. BioRad Electroporation
2.4.8.3. Lipofection

2.4.9. Analysis

2.4.9.1. Colony Staining
2.4.9.2. Mitotic Spreads
2.4.9.3. Immunocytochemistry
2.4.9.4. Flow Cytometry
2.4.9.5. Alizarin-Red S Staining
2.4.9.6. Calcium Assay
2.4.9.7. Statistical Analysis

2.4.1. Cell Lines

H1 and H9 cell lines were obtained from Geron (San Francisco, USA). Human embryonic kidney (HEK) epithelium cells 293 were obtained from W. Cui (Roslin Institute, Scotland) and the HPRT-deficient mouse ES cell line, HM1 was provided by J. McWhir (Roslin Institute, Scotland). Mouse embryonic fibroblasts (MEFs) were generated in the lab from 13 days postcoitum mouse embryos.

2.4.2. Maintenance of Mouse Embryonic Fibroblasts

2.4.2.1. Preparation of MEF Medium

MEF medium was prepared by filtering in a 500 ml filter unit (0.22 μm, corning, cellulose acetate, low protein binding), 450 ml Dulbecco's modified Eagle's medium (DMEM, Gibco-BRL), 50 ml foetal calf serum (FCS, Gibco-BRL), 5ml L-glutamine (Gibco-BRL) (2 mM f.c) and 5 ml non-essential amino acids (Gibco-BRL) (0.1 mM f.c). Medium was then stored at 4° C for up to one month.

2.4.2.2. Isolation of MEFs

Pregnant female mice were killed on the thirteenth or fourteenth day of pregnancy. The abdomens were swabbed with 70 % ethanol and the abdominal cavities were dissected to expose the uterine horns. Uterine horns were then placed into a 10 cm bacteriological Petri dish containing 10 ml of Phosphate-Buffered Saline (PBS) (Oxoid) supplemented with 0.25 ml penicillin (5000 u/ml) / streptomycin (5000 μg/ml) (Gibco-BRL) and the embryos were released into the saline by cutting open using fine scissors. The embryos were then

removed together with the associated placentas and foetal membranes into a fresh dish containing PBS. Placentas and membranes were removed using two pairs of forceps and soft tissues (e.g. liver, heart and the other viscera) were separated from individual embryos. The embryos were washed in two further changes of PBS and placed individually into bijou tubes containing 2ml of TEG (3.15 g NaCl, 0.06 g NA_2HPO_4, 0.108 g KH_2PO_4, 0.166 g KCl, 0.45 g D-Glucose, 1.35 g Tris, 0.45 ml 1% Phenol Red; add up to 400 ml distilled water; 50 ml of 10 x trypsin 2.5 %, 0.2 g EGTA, 0.05 g PVA; adjust pH up to 7.6 and add distilled water up to 500 ml). Tubes were incubated for 10 minutes at 37 ° C and the content was vortexed, incubated for a further 10 minutes at 37 ° C and vortexed one more time. 3 ml of MEF medium was added to inactivate the trypsin, the tubes were vortexed for a few seconds and individual supernatants containing the cells were transferred into 75 cm^2 flasks containing 12 ml PEF medium. Flasks were incubated at 37 ° C and the medium was changed the following day to remove the cellular debris.

2.4.2.3. Passage of MEFs

Subconfluent 75 cm^2 flasks were washed once with PBS and incubated with 2 ml of TEG at 37 ° C (5 % C02) for 2-3 minutes. 9 ml of MEF medium was added to the cells and those were split 1:2 into a 150 cm^2 flask and MEF medium was added up to obtain a final volume of 30 ml. MEFs could be expanded for up to 4 passages before being discarded.

2.4.3. Maintenance of Human ES Cells

2.4.3.1. Feeder-Free Human ES Cell Culture

Conditioned Medium Preparation

MEFs (150 cm^2 flask) were used to prepare conditioned medium (CM) by removing MEF medium and adding 50 ml of serum-free human ES medium (80 % Knock-out DMEM (KO-DMEM, Gibco), 20 % serum replacement (SR, Gibco), 0.1 mM non-essential amino acids, 2mM L-glutamine, 0.1 mM β-mercaptoethanol and 4 ng/ml basic fibroblast growth factor (bFGF, Invitrogen); filter in a 500 ml filter unit (0.22 μm, corning, cellulose acetate, low protein binding)). CM was collected from the feeder flasks the following day and replaced with a further 50 ml of human ES medium. The same flask of feeders was used for 4-5 days, to generate 200-250 ml of CM. collected CM was either stored at –20° C or supplemented with 4 ng/ml bFGF, filtered, and used for daily human ES cultures maintenance.

Matrigel Coating for Human ES Cultures

Matrigel (Becton Dickinson) was thawed at 4 ° C for at least 2 hours to avoid the formation of a gel. Matrigel was then diluted 1:100 in cold KO-DMEM using a chilled pipette and typically 25 cm^2 flasks were coated with 3 ml of the solution. Flasks were either incubated at RT for 1-2 hours before use or stored at 4 ° C for no longer than 2 weeks. The matrigel solution was removed from the flasks and washed with KO-DMEM before use.

2.4.3.2. Passage Regime

When H1 and H9 cells had reached 80-90 % confluence, the CM was removed and cells were washed once with KO-DMEM. Typically, for a 25 cm^2 flask, cultures were incubated with 1ml of TEG at 37 ºC for 3-5 minutes and cells were detached from the surface by shaking the flasks vigorously. 5 ml of KO-DMEM was added to the cells and a single-cell suspension was obtained by pipetting the cell suspension vigorously. Cells were then transferred to 15 ml falcon tubes and centrifuged at 200 x g for 5 minutes. Supernatants were discarded and the pellets were resuspended in fresh CM supplemented with 4 ng/ml bFGF and seeded onto matrigel-coated 25 cm^2 flasks as required (splitting usually between 1:3 and 1:4). Medium was changed the following day.

2.4.4. Maintenance of Other Cell Lines

HM1 cells were grown at 37 ºC (5 % CO_2) on 0.1 % gelatin-coated flasks and fed daily with mouse ES medium. ES medium was prepared with 450 ml Glasgow modified Eagle's medium (GMEM, Gibco), 28 ml FCS, 28 ml new born calf serum (NBCS, Gibco), 5ml L-glutamine (2 mM f.c), 5 ml non-essential amino acids (0.1 mM f.c), 5 ml sodium pyruvate (Gibco, 1 mM f.c), 1.1 ml 2-mercaptoethanol (0.1 mM f.c) and 280 µl of Leukaemia Inhibitory Factor (ESGRO-LIF, Chemicon, at 1 x 10^6 U/ml).

For passaging, subconfluent HM1 cells were washed with PBS, trypsinised with TEG for 1-3 minutes at 37 ºC. After incubation, flasks were shaken vigorously, an appropriate volume of ES medium was added to the cells, and a single-cell suspension was obtained by gently pipetting. Cells were centrifuged at 200 x g for 5

minutes and pellets were resuspended in ES medium and plated out onto 0.1 % gelatin-coated flasks as required. HEK 293 cells were maintained as described above but in MEF medium.

2.4.5. Cryopreservation and Thawing of Cell Lines

Cells were trypsinised and collected by centrifugation for 5 minutes at 200 x g. Pellets were resuspended in the appropriate volume of medium and placed on ice. One volume of ice-cold 2x freezing mix* was added progressively and the suspension was gently mixed and aliquoted into pre-chilled cryovial tubes (1ml per tube). Vials were stored at – 80° C for 24 hours and then transferred to – 150° C. Cells were thawed by placing the vials in a 37° C water bath. 5 ml of appropriate cell line medium was added to the contents of a vial in a 15 ml falcon tube and centrifuged at 200 x g for 5 minutes. Supernatants were discarded and pellets were resuspended in an appropriate volume of medium for plating out.

* For human ES cells, 2 x freezing mix was made by mixing 50 % CM, 30 % SR and 20 % dimethyl sulfoxide (DMSO, Sigma). For MEFs, 2X freezing mix was made by mixing 50 % MEF medium, 30 % FCS and 20 % DMSO. Freezing mixes were filtered and stored at – 20° C.

2.4.6. Subcloning, Cryopreservation and Thawing of HES Clones

Plates containing H9 G418-resistant colonies, from *β-casein* targeting experiments, were washed with KO-DMEM and 10 µl of TEG was added onto each colony. Plates were incubated at 37 ºC for few minutes and the incubation was stopped when the cells in the colonies started lifting up. Colonies were broken into small clusters by adding 10 µl of CM and gently pipetting up and down. Clumps were transferred into matrigel-coated 96-well plates containing CM supplemented with bFGF. When reaching confluence after few days, colonies were trypsinised with 25 µl of TEG, incubated at 37 ºC for few minutes, and single cell suspension was obtained by adding 50 µl of CM and pipetting up and down. Cell clones were transferred into matrigel-coated 48-well plates containing 0.5 ml CM/bFGF. This procedure was repeated until the clones were expanded up to two wells of a 12-well plate for freezing and gDNA extraction. Using a protocol optimised by H. Priddle (Roslin Institute) for *in situ* freezing and thawing of hES cells, subconfluent individual colonies in 12-well plates were trypsinised with 200 µl TEG and a single suspension was obtained by subsequently adding 400 µl of cold quench medium (70 % KO-DMEM and 30 % SR). 600 µl of cold fresh 2x freezing mix was added to each well and the plates were swirled gently until the DMSO was well diluted. Plates were stored at –80 ºC overnight and then transferred at 150 ºC. When required, H9 G418-resistant clones were thawed by adding 1.5 ml of warm CM/bFGF to the well. Cell clones were subsequently transferred to two wells of a matrigel-coated 12-well plate containing 1.5 ml of fresh CM/bFGF. For H1 colonies obtained from *hprt* targeting and RMCE experiments, cells were isolated and

expanded the same way but clones were frozen and thawed as described in section 2.4.5.

2.4.7. *In Vitro* Differentiation of HES Cells

2.4.7.1. Generation of Embryoid Bodies

75 cm^2 flasks of subconfluent H9 and H1 cells received fresh medium 1 hour before use. Cells were then washed with KO-DMEM, trypsinised and centrifuged as described in section 2.4.3.2. Each individual pellet was resuspended in 5 ml CM/bFGF and transferred to a 10 cm bacteriological petri dish containing 20 ml of CM supplemented with bFGF. Resulting embryoid bodies (EBs) were collected 48 hours later by transferring the contents of the dish into a universal tube and allowing the EBs to sink to the bottom. The supernatant was discarded and EBs were resuspended in 20 ml EB medium (90% KO-DMEM, 10% FCS, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 0.1 mM β-mercaptoethanol) and transferred into a new bacteriological petri dish. EBs were maintained in suspension for a further 2-3 days and then transferred onto 0.1 % gelatin-coated 24-well plates or chamber well slides (~ 10 EBs per well) and cultured for an additional 21 days before immunocytochemistry analysis.

2.4.7.2. Osteogenic Differentiation

H1 and H1-derived cells were plated out at a cell density of 10^5 cells and 10^4 cells per well in matrigel-coated 6-well (for Alizarin-Red S staining, see section 2.4.9.5) and 96-well plates (for calcium assay, see section 2.4.9.6), respectively. HEK 293 cells were seeded at a density of 10^4 cells per well in 0.1 %

gelatin-coated 6-well plates and 10^3 cells per well in a 0.1 % gelatin-coated 96-well plate. To induce osteogenic differentiation (work performed by Davina Wojtacha), the CM was replaced 48 hours following plating with EB medium containing osteogenic supplements (50 μM ascorbic acid phosphate, Wako; 10 mM β-glycerophosphate, Sigma; and 100 nM dexamethasone, Sigma). Medium was changed every 2 days and cultures were treated with or without osteogenic supplements for the indicated time periods (9, 13, 16 and 20 days) before analysis.

2.4.8. Transfection

2.4.8.1. Eppendorf Multiporation

Exponentially growing cells received fresh medium 1 hour before electroporation. Human ES cultures were trypsinised, resuspended in 10 ml CM, and counted using a Neubauer haemacytometer (Weber, UK). 10^6 to $2x10^6$ cells were resuspended in 750 μl of hypososmolar buffer (Eppendorf) and incubated for 20 minutes at RT. Linearised or supercoiled (for RMCE experiments) plasmids DNA (50-138 μg) were diluted in 50 μl of distilled water and mixed with the cells in a 0.4 cm electrode Gene Pulser Cuvette (BioRad). Cells were electroporated with the Eppendorf Multiporator at 300 V and 100 μs and incubated for 5 minutes at RT. Cells were then resuspended in CM/bFGF, plated at a density of $5x10^5$ cells per matrigel-coated 10 cm or 15 cm dish, and incubated at 37 ° C in a humid 5% CO_2 atmosphere. Medium was replaced the following day and selection was applied 48 hours following transfection.

2.4.8.2. BioRad Electroporation

Cells were prepared as described in section 2.4.8.1 but in this protocol, cells were resuspended in 750 µl of PBS. Linearised or supercoiled (for RMCE experiments) plasmids DNA (50-138 µg) were diluted in 50 µl of distilled water and mixed with the cells in a 0.4 cm electrode Gene Pulser Cuvette (BioRad). Cells were electroporated with a BioRad Gene Pulser II at 200 V / 950 µF or 320 V / 200 µF. Cells were incubated for 10 minutes at RT and plated at a density of $5x10^5$ cells per matrigel-coated 15 cm dish.

2.4.8.3. Lipofection

Lipofections were performed using the Lipofectamine™ 2000 kit (Invitrogen). 2 to 4 days prior to transfections, $1x10^5$ to $4x10^5$ cells were seeded per well in a 6-well plate. One hour before transfection, 50-95 % confluent cultures had the medium changed and a typical lipofection mixture was prepared as follows:

Solution A: 2.5 µl purified uncut DNA (1µg/µl) or 2.5 µl dH_20 (Mock) + 122.5 µl Opti-Mem I (Invitrogen)

Solution B: 6 µl Lipofectamine + 119 µl Opti-Mem I

Solution A and B were incubated for 5 minutes at room temperature, mixed together and incubated for 20-30 minutes at room temperature to allow the DNA-lipofectamine complexes to form. The CM was then removed from the wells and replaced with 2.25 ml of fresh CM in which the lipofection mixture was added. Plates were gently mixed and the cells were incubated at 37°C (5% CO_2). After 6-12 hours incubation, cells were washed with KO-DMEM and the medium was

replaced with fresh CM. Bright field and fluorescent images of GFP-expressing cells were captured 48 hours following transfection with a Nikon microphot SA microscope, camera and Digital Pixel software.

2.4.9. Analysis

2.4.9.1. Colony Staining

Colonies were washed with PBS and fixed for 10 minutes in methanol. Methanol was removed and colonies were stained 10-15 minutes in a 10 % Giemsa R-66 solution. After staining, plates were gently washed with water and air-dried.

2.4.9.2. Mitotic Spreads

Confluent human ES cells in 25 cm^2 flasks were split 1:5, and 50 µl of 10 µg/ml karyoMax colcemid solution (Invitrogen) was added per 5 ml culture 22 to 24 hours later (to ensure that most cells are in exponential growth). Colcemid arrests dividing cells at metaphase of mitosis by preventing the formation of the spindle apparatus responsible for cell division. This results in an accumulation of metaphases. The medium was collected 2 hours later in a 15 ml falcon tube and the cells were washed with 2 ml PBS which was collected in the same 15 ml falcon tube. Cells were trypsinised and resuspended with the CM/PBS solution previously collected and centrifuged for 5 minutes at 200 x g. The supernatant was discarded and cells were resuspended in 8-10 ml fresh hypotonic solution (0.56 % w/v KCl in dH_20) using a vortex. Cells were incubated for 8-10 minutes at room temperature

and centrifuged for 5 minutes at 200 x g. The supernatant was decanted and the cells were fixed by adding slowly, down the side of the tube, 2-3 ml of freshly made ice-cold fixative (18 ml methanol and 6 ml acetic acid) and vortexing simultaneously. Cells were then resuspended in a final volume of 8-10 ml of fixative. Tubes were centrifuged at 200 x g for 5 minutes, supernatant was discarded and cells were washed twice more in fixative. Pellets were then resuspended in 1 ml of fixative for mitotic spreads analysis.

Glass slides were cleaned in cold 70 % ethanol, polished and air-dried before use. A Pasteur pipette was used to release a single drop of suspension onto each slide. Fixative was allowed to spread and was air-dried, and bright field images of mitotic spreads were obtained using a Zeiss microscope and the AxioVision software (Carl Zeiss, Jena, Germany).

2.4.9.3. Immunocytochemistry

Human EBs and undifferentiated hES cells were washed twice with PBS, fixed with 4% paraformaldehyde (PFA, Sigma) (dissolve 4 g PFA in 100 ml PBS by moderate stirring on a hotplate/stirrer at 60 °C, filter and store at –20 °C) and incubated at RT for at least 20 minutes. Samples were washed three times with PBS, permeabilised with 0.2 % Triton X-100 (Sigma) in PBS for 10-15 minutes at RT, and washed twice with PBS. EBs were then incubated with 2% BSA/PBS (13.3 ml 7.5 % BSA and 36.6 ml PBS) for 1 hour at 37 ° C to block non-specific sites. Samples were incubated with a 2 % BSA/PBS diluted primary antibody against Troponin T (1: 50 from a 0.2 μg/μl stock, NeoMarkers), or α-fetoprotein (1:400 from

a 9 μg/μl stock, Sigma), or β-III-tubulin (1:200 from a 0.6 μg/μl stock, Sigma) for 1 hour at 37 ° C or overnight at 4 ° C. Cells were washed once quickly in PBS and then twice by gently agitating for 20 minutes in PBS supplemented with 0.05 % tween 20 (Sigma). Cells were then covered with 2 % BSA/PBS diluted secondary fluorescein anti-mouse antibody (1:1000 from a 1.5 μg/μl stock, Vector Laboratories) and incubated for 1 hour at 37 ° C in a dark place. Cells were washed once quickly in PBS and twice for 20 minutes (with gentle agitation) in PBS/0.05 % Tween 20. Samples were mounted with Vectashield mounting medium (Vector Laboratories) for fluorescence microscopy analysis, or analysed directly without the mounting medium (EBs grown in 24 well-plates). The bright field and fluorescent images were obtained using a Zeiss microscope and the AxioVision software (Carl Zeiss, Jena, Germany).

For *oct-4* expression analysis, subconfluent H1 and H1-derived cells in chamber slides were washed twice with PBS, incubated with 4% PFA for 20 minutes at RT and washed three times with PBS. Cells were then incubated with ethanol for 2 minutes for permeabilisation, washed twice with PBS for 5 minutes and blocked for 1 hour at RT with blocking buffer (10% NGS/PBS/0.1% Tween 20). Samples were incubated with a 0.1% NGS/PBS/Tween 20 diluted primary antibody against OCT-4 (1: 200 from a 0.2 μg/μl stock, Santa Cruz) for 1 hour at RT. Cells were washed twice with PBS for 5 minutes. Cells were then covered with the PBS diluted secondary anti-mouse antibody conjugated to FITC (1:100, from a 5 μg/μl stock, Sigma) and incubated for 1 hour at RT in a dark place. Cells were washed twice in PBS for 5 minutes. Samples were mounted with Vectashield mounting medium with 4'-6' Diamidino-2-phenylindole (DAPI) (Vector Laboratories).

Fluorescent and merged images were acquired with an Axiovert S100 microscope (Carl Zeiss), and the AQMe imaging acquisition software using a high-resolution black and white ORCa1 camera.

For flow cytometry analysis of human ES cell surface markers, H1 and H1.HPRT-1 cells were trypsinised and resuspended to approximately $5x10^5$ cells/sample in 50 µl blocking buffer (40% heat inactivated rabbit serum in staining buffer). Blocking was carried out for 15 minutes on ice. Primary antibody and appropriate isotype-matched control diluted in staining buffer (2% FCS in PBS and 2mM EDTA) was added to each sample and incubation was carried out on ice for 30 minutes. Mouse IgG3 (1 µg/µl stock, Sigma), mouse IgM (1 µg/µl stock, Sigma), SSEA-4 (DSHB), SSEA-1 (DSHB), Tra-1-60 (1 µg/µl stock, Chemicon) and Tra-1-81 (1 µg/µl stock, Chemicon) were diluted 1:100, 1:3, 1:100, 1:5, 1:12 and 1:20, respectively. Cells were washed with staining buffer, spin at 200 x g for 5 minutes, resuspended with the secondary antibody diluted with staining buffer and incubated for 30 minutes on ice in the dark. Staining with secondary antibody, was carried out with fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG3 (1:100 from a 5 µg/µl stock, Caltag laboratories) for IgG3 and SSEA-4, and with phycoerythrin (PE)-conjugated goat anti-mouse IgM (1:100 from a 0.2µg/µl stock, Caltag laboratories) for IGM, SSEA-1, Tra-1-60 and Tra-181. CD9 staining was performed with PE-conjugated mouse-anti human CD9 (1: 5 dilution, BD PharMingen). Cells were washed with PBS, spin for 5 minutes at 200 x g and resuspended in 1ml PBS/0.1% PFA for flow cytometry analysis.

2.4.9.4. Flow Cytometry

Flow cytometry was performed on a 488 nm laser FACScan (Becton Dickinson). A 530/30 nm band pass filter was used to measure EGFP fluorescence intensity or FITC-conjugated antibodies. A 585/42 nm band pass filter was used to measure phycoerythrin-conjugated antibodies. For GFP analysis, detector settings were adjusted with untransfected cells. For cell surface markers analysis, detector settings were adjusted with both untreated and isotype controls. Analysis was performed on CellQuestPro software (BD Biosciences). At least 10,000 "live" cells were acquired for each sample. Forward size and side scatter plots were used to exclude dead cells and debris from the histogram analysis plots.

2.4.9.5. Alizarin-Red S Staining

Cells in 6-well plates analysed for osteogenic differentiation were washed twice with PBS and fixed in 95 % methanol (staining performed by Davina Wojtacha). Alizarin-Red Staining, which forms a complex with calcium in a chelation process, was performed for mineralised nodule detection as follows: Fixed cells were incubated with a 1 % Alizarin-Red S (sigma) solution for 10 minutes and washed twice in water. High-power and low-power bright field images were obtained using a Zeiss microscope and the AxioVision software (Carl Zeiss, Jena, Germany).

2.4.9.6. Calcium Assay

Matrix-associated calcium deposition of H1, H1-derived, and HEK cells was examined in a 96-well plate using a calcium colorimetric assay kit (Randox) in quadruplet wells (assay and analysis performed by Davina Wojtacha). The principle of this assay is as follows: Calcium in alkaline medium forms a purple-red complex with ortho-cresolphthalein. Intensity of the developed colour is proportional to the calcium concentration in the sample.

At each time point (9, 13, 16 and 20 days), medium was removed from the quadruplet wells to be analysed and the cells were washed 3 times with 0.9 % NaCl. The 96-well plate was incubated at 37 °C until the final time point was performed. Cells were then lysed by adding 50 µl/well of 0.1 M NaOH, and neutralysed with 20 µl/well of 2 M HCl. The plate was air-dried at RT overnight. The following day, 200 µl/well of assay buffer mix (containing 2-amino-2 methyl-propan-1-ol and the chromogen) was added to each well and incubated for 10 minutes at RT. To calibrate the assay, a calcium standard curve was also performed in a 96-well plate by applying a range of calcium standard (provided with the kit) concentrations varying from 0 to 100 µM. After addition of 200 µl/well of assay buffer mix to each well, the plate was incubated for 10 minutes at RT. Absorbance of each well and standard was then read against a reagent blank at a wavelength of 575 nm using a spectra MAX plus spectrophotometer and a sofMax software. Calcium concentrations using calibration were calculated as follows:

C_{sample} (mmol/l) = (A_{sample} / $A_{standard}$) x $C_{standard}$ (2.50)

C_{sample} (mg/dl) = (A_{sample} / $A_{standard}$) x $C_{standard}$ (10)

C= concentration

A=absorbance

2.4.9.7. Statistical Analysis

Statistics for human ES cell surface markers flow cytometry analysis were conducted based on three independent experiments. The mean value and the standard deviation were calculated using Microsoft Excel. Means were compared using the two-tailed student's T-test. Statistical analysis for the calcium assay was performed by using a F-test with Genstat (version 8.1, Lawes Agricultural Trust, 2005) software. Statistics for mitotic spreads were carried out using a Chi-squared test with 1 degree of freedom (χ_1^2). Values for $P < 0.05$ were considered statistically significant.

CHAPTER III

Engineering the *β-Casein* Gene in H9 Cells by Homologous Recombination

3.1. Introduction

3.2. Objectives

3.3. Results

3.4. Discussion

3.1. Introduction

The potential applications of hES cells include a source of tissue for regenerative medicine, models to study human diseases and early human embryonic development *in vitro*, and drug discovery and cytotoxicity testing. However, to bring nonautologous hES stem cell therapies to the clinic, rejection of hES-derived grafts by the patient's immune system must be avoided, directed differentiation to provide therapeutically relevant cell types must be established and strategies to purify differentiated cells from heterogenous populations, including potentially tumorigenic undifferentiated hES cells (Thomson *et al.,* 1998, Reubinoff *et al.*, 2000), must be developed. Genetic modification of hES cells could be the answer for many of these issues. Gene targeting will be probably required for some applications such as the modelling of human disease, the reduction of immunogenicity of hES cell-derived transplants and the dissection of pathways in developmental biology (for more details see section 1.4.2, chapter I). Whilst random integration of transgenes can be sufficient to provide stable transgene expression, for example to increase commitment to a specific lineage, it may lead to insertional inactivation of coding sequences, and/or reduced and aberrant expression of the transgene due to position effects (Clark *et al.*, 1994; Martin and Whitelaw, 1996). These problems can in principle be overcome by the use of gene targeting to introduce a transgene to a site known to be permissive for transgene expression (Wallace *et al.*, 2000). Thus gene targeting provides a safer and more effective route to stable transgene expression.

The mammary epithelium-specific *β-casein* gene was chosen as a candidate target site for reliable transgene expression in the H9 female hES cell line. The

primary advantage of targeting this gene is that it has no essential function out-with the mammary gland (Kumar *et al.,* 1994). Thus the possibility of targeting both alleles without risks of insertional mutagenesis of essential genes such as house keeping genes, also provides a low frequency of losing both copies of a transgene. This is an important consideration for strategies in which stem cells are manipulated to express "suicide genes" such as HSV-*tk* to provide protection against tumorigenesis (Bonini *et al.*, 1997; Schuldiner *et al.*, 2003). This gene was also chosen because targeting frequencies of up to 8 %, using a positive and negative selection (Capecchi, 1989), were achieved at this locus in mouse ES cells (Kolb, 2001).

Kolb (2001) introduced, by HR, a positive/negative *hprt*-expression cassette (Albertini, 2001) flanked by the *loxP* and *lox2272* heterospecific *lox* sites (Lee and Saito, 1998) at the *β-casein* gene in the HPRT-deficient murine ES cell line HM1 (Magin *et al.*, 1992). Because HM1 cells are HPRT-deficient, they are sensitive to HAT selection (Szybalski & Szybalska, 1962) and resistant to 6-TG selection (Stutts & Brockman, 1963) (see section 1.4.4.2, chapter 1, for details of the HPRT selection system). As a result, HM1 cell clones targeted with the *hprt* cassette at the *β-casein* gene could be selected in HAT selection. Using Cre-recombinase-mediated cassette exchange (RMCE) (Bethke and Sauer, 1997; Bouhassira *et al.*, 1997; Seibler *et al.*, 1998; Feng *et al.*, 1999; Trinh and Morrison, 2000) in a subsequent step, the *hprt*-expression cassette from a targeted cell clone was then exchanged for a promoterless *β-galactosidase* gene and a *PGK-hytk* expression cassette flanked by a *loxP* and a *lox2272* site. Site-specific recombinants carrying the modification could be recovered in hygromycin (Hy) and 6-TG (since these cells no longer expressed

hprt). Using this positive/negative selection strategy, 100 % of the 6-TG/Hy-resistant colonies carried the desired modification.

We sought to characterise the *β-casein* gene for its ability to provide reproducible transgene expression in H9 cells. Based on the work conducted by Kolb (2001), our objective was to tag, by HR, the same heterospecific *lox* sites described by Kolb such that RMCE can be subsequently applied.

At the start of this project, there was no report of gene targeting in hES cells and thus the establishment of a gene targeting protocol was a priority. Since then, Zwaka and Thomson as well as Urbach *et al.* have reported gene targeting at the *hprt* gene (Zwaka and Thomson, 2003; Urbach *et al.*, 2004), and *oct-4* gene (Zwaka and Thomson, 2003) in hES cells. The first part of this chapter describes the targeting strategy, the establishment of an efficient transfection protocol and the *β-casein* targeting experiments conducted in H9 cells. The second part of this chapter describes a preliminary experiment in which the chicken HS4 insulator elements (Chung *et al.*, 1993) were tested in hES cells as an alternative approach to overcome the problems associated with transgene expression, mentioned above. Insulators are boundary elements that exert their function by shielding genes from either the action of a distal enhancer and/or from the silencing effects of condensed chromatin. These elements could be useful in protecting transgenes randomly integrated in the genome for strategies that do not require gene targeting of hES cells. Additionally, insulators could prevent transgene silencing at loci that are either expressed at low levels or silent (such as the *β-casein* gene), against a potential surrounding condensed chromatin structure often observed at unexpressed loci (reviewed by Felsenfeld and Groudine, 2003).

3.2. Objectives

(i) To optimise the experimental conditions required for gene targeting experiments in H9 cells. (ii) To introduce incompatible *lox* sites by HR into the *β-casein* gene in H9 cells. (iii) To test the ability of HS4 insulators to protect transgenes from silencing in H9 cells.

3.3. Results

3.3.1. Development of a Targeting Protocol

3.3.2. *β-Casein* Targeting Experiments

3.3.3. Protective Effect of Insulators on Transgene Expression in H9 cells

---

3.3.1. Development of a Targeting Protocol

3.3.1.1. *β-Casein* Targeting Vector and Targeting Scheme

To introduce heterospecific *lox* sites into the *β-casein* gene, a human *β-casein* replacement vector pCas-EfN was generated (figure 3.1). This vector consists of a 2.3 kb 5' arm and a 5.4 kb 3' arm of homology which were amplified separately by "Expand™ High-fidelity PCR" (Roche) from H9 genomic (g) DNA. The 5' element includes the 3' end of the promoter region (between –2602 and –231 bp) and the 3' element includes most of the open reading frame of the gene (from exon 2 to the beginning of exon 8). An expression cassette, composed of the

human *EF1-α* (Elongation Factor 1-α) promoter (Mizushima and Nagata, 1990) driving a *neomycin (neo)* gene flanked by heterospecific *lox* sites *(loxP* and *lox 2272*), was inserted between the two arms of homology. The human *EF1-α* promoter (from pEF1/V5HisA, Invitrogen) was chosen to drive the *neo* cassette because of its strong and ubiquitous transcriptional activity in human cells (Mizushima and Nagata, 1990). Targeting one of the *β-casein* alleles with pCas-EfN would result in the replacement of a 4.9 kb *β-casein* sequence (spanning from the end of the promoter to the end of intron 1) with a 3.4 kb *EF1-α-loxP-neo-lox2272* cassette in the opposite transcriptional orientation to the gene (figure 3.1). This transgene would allow selection for both random and targeted events with the drug G418. Discrimination between random and targeted clones was achieved by Southern analysis.

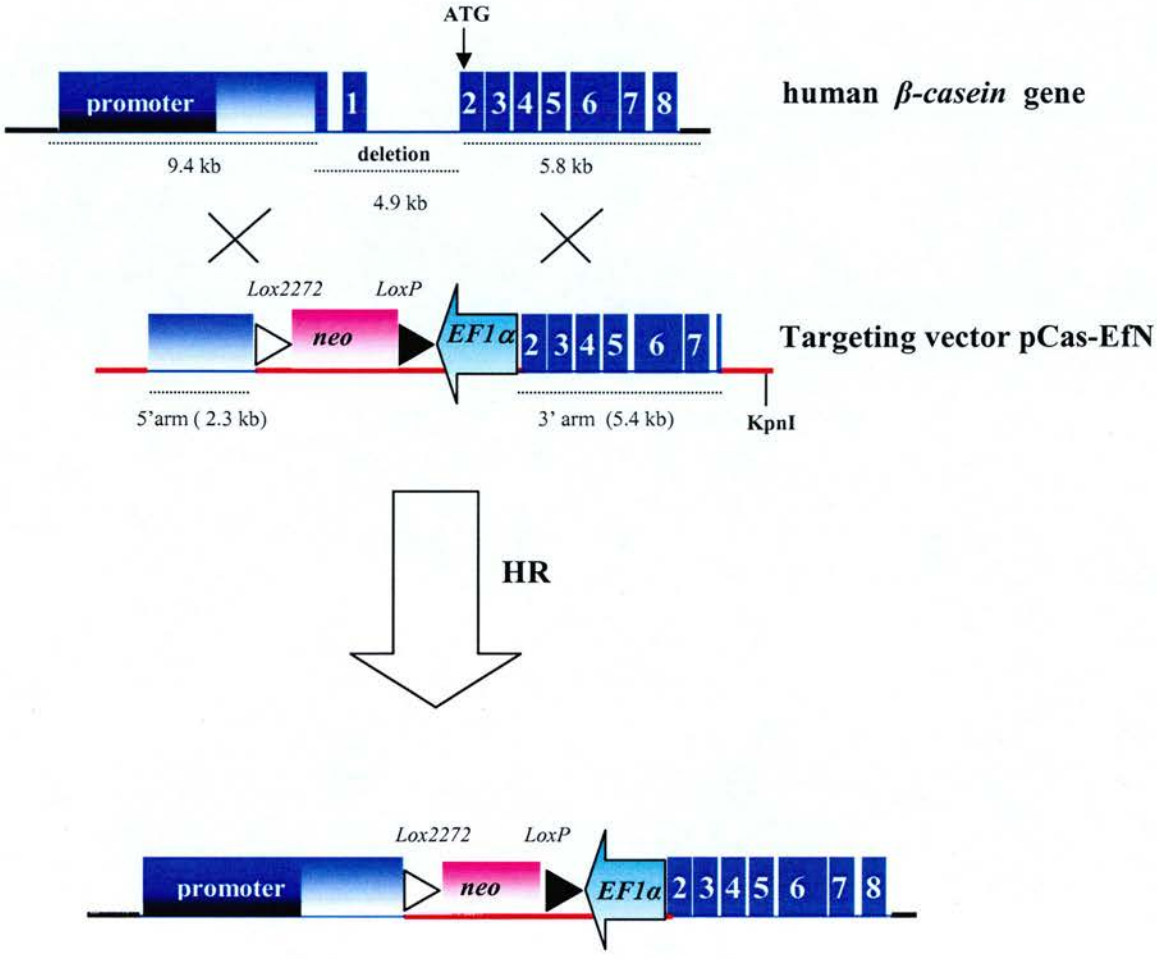


Figure 3.1. Schematic representation of the *β-casein* gene targeting in H9 cells. The *β-casein* promoter (9.4 kb) is represented by a box in a dark and light blue graduated shading. The light graduated shading represents the 5'arm of homology included in pCas-EfN. The 8 exons of the *β-casein* gene are presented as blue boxes with exon numbers indicated. The translational initiation start (ATG) is shown above the second exon. The *EF1-α* promoter and the *neo* expression cassette are indicated as a solid green arrow and solid pink box, respectively. The positions and orientations of the *loxP* and *lox2272* recombination sites are shown as a black and white arrow, respectively. The KpnI restriction site used to linearise pCas-EfN, prior to a targeting experiment is also indicated. The black and red thick lines represent the sequence outside the *β-casein* locus, and plasmid backbone, respectively. Cross-over recombination points between the targeting vector and the *β-casein* chromosomal sequence are indicated by an 'X'. Following homologous recombination (HR) a 4.9 kb *β-casein* sequence (spanning from the end of the promoter to the end of intron 1) is replaced with a 3.4 kb *EF1-α-loxP-neo-Lox2272* in the opposite transcriptional orientation of the gene, disrupting the *β-casein* gene.

3.3.1.2. Functionality of the *Neomycin* Expression Cassette

In order to verify the integrity of the *neomycin* expression cassette in pCas-EfN, 10^7 HM1 mouse ES cells (Magin *et al.*, 1992) were stably transfected by electroporation with equimolar amounts of the KpnI-linearised pCas-EfN targeting vector or KpnI-linearised pBS-PGKNeo using our standard mouse ES cell electroporation conditions (800 V, 3 μF). A plasmid (pBS-PGKNeo) containing a *PGKneo* expression cassette, was used as a positive transfection control in this study because it was shown, in our hands, to be efficient in generating G418-resistant colonies in HM1 cells. G418 selection (300 μg/ml) was applied the following day, and 12 days after electroporation the resulting G418-resistant colonies were stained with 10% Giemsa's stain solution (figure 3.2). HM1 cells transfected with pCas-EfN, produced a very large number of G418-resistant colonies (>500), rendering colony counting difficult. In contrast, HM1 cells transfected with pBS-PGKNeo produced ~135 colonies. Although this study suggests that the transcriptional activity of the *EF1-α* promoter is stronger than the *PGK* promoter, they could not be directly compared since different plasmids with different structures were used in the transfection. A mock control, in which cells were electroporated without DNA, did not yield any colonies, indicating that all the colonies obtained in the experimental plates were actually G418-resistant.

These results demonstrate that the *neomycin* expression cassette in pCas-EfN is functional and that the transcriptional activity of the *EF1-α* promoter is very strong leading to a large number of G418-resistant colonies in mouse ES cells.

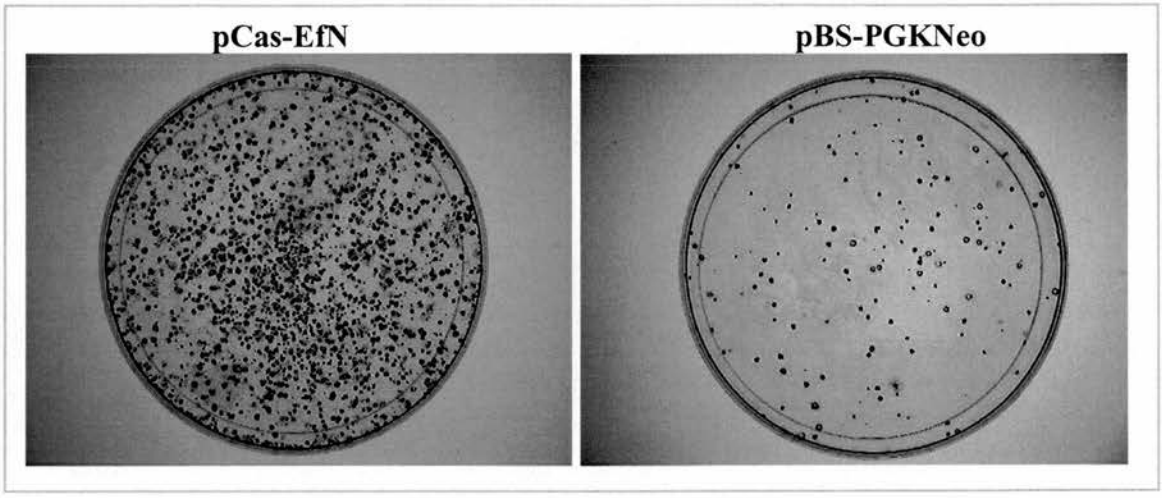


Figure 3.2. Functionality of the *neo* expression cassette from pCas-EfN in HM1 cells. Giemsa staining of G418-resistant HM1 colonies 12 days after transfection with pCas-EfN (left) or pBS-PGKNeo positive control plasmid (right).

3.3.1.3. Functionality of the *Lox* Sites in pCas-EfN

The sequencing of *loxP* and *lox2272* sites in the plasmid pCas-EfN showed that both *lox* sites were intact (Appendix 1, section 1.1). However, it was also necessary to test their functionality prior to performing a large scale targeting experiment, as the subsequent RMCE step relies on their function. The plasmid pCas-EfN was therefore assessed in an *in vitro* recombination assay (Kolb, 2001) for its ability to recombine with pB-RMCE2272 (gift of A. Kolb, appendix 1, section 2.4), a plasmid containing the same pair of heterospecific *lox* sites. Site-specific recombination between both vectors leads to reciprocal exchange of sequences within the *lox* sites. The plasmids pCas-EfN and pB-RMCE2272 were linearised with KpnI and ScaI, respectively, incubated with Cre recombinase (New England Biolabs), and the reactions were used as the template in a PCR reaction with T7 and T3 primers. These primers were selected because they should anneal specifically to pB-RMCE2272 and one of the two recombination products only, resulting in the amplification of 227 bp (from the pB-RMCE2272 plasmid) and 1850 bp if site-specific recombination had occurred (figure 3.3 A & B). PCR results showed the presence of the predicted bands when the two constructs were incubated along with Cre recombinase (figure 3.3 B). In contrast, when the two plasmids were used as template without prior treatment with Cre, only the 227 bp PCR product, arising from the pB-RMCE2272 plasmid, was detected (figure 3.3 B). These results indicate that cassette exchange was successful between the two plasmids in an *in vitro* assay, thus confirming the functionality of the *lox* sites.

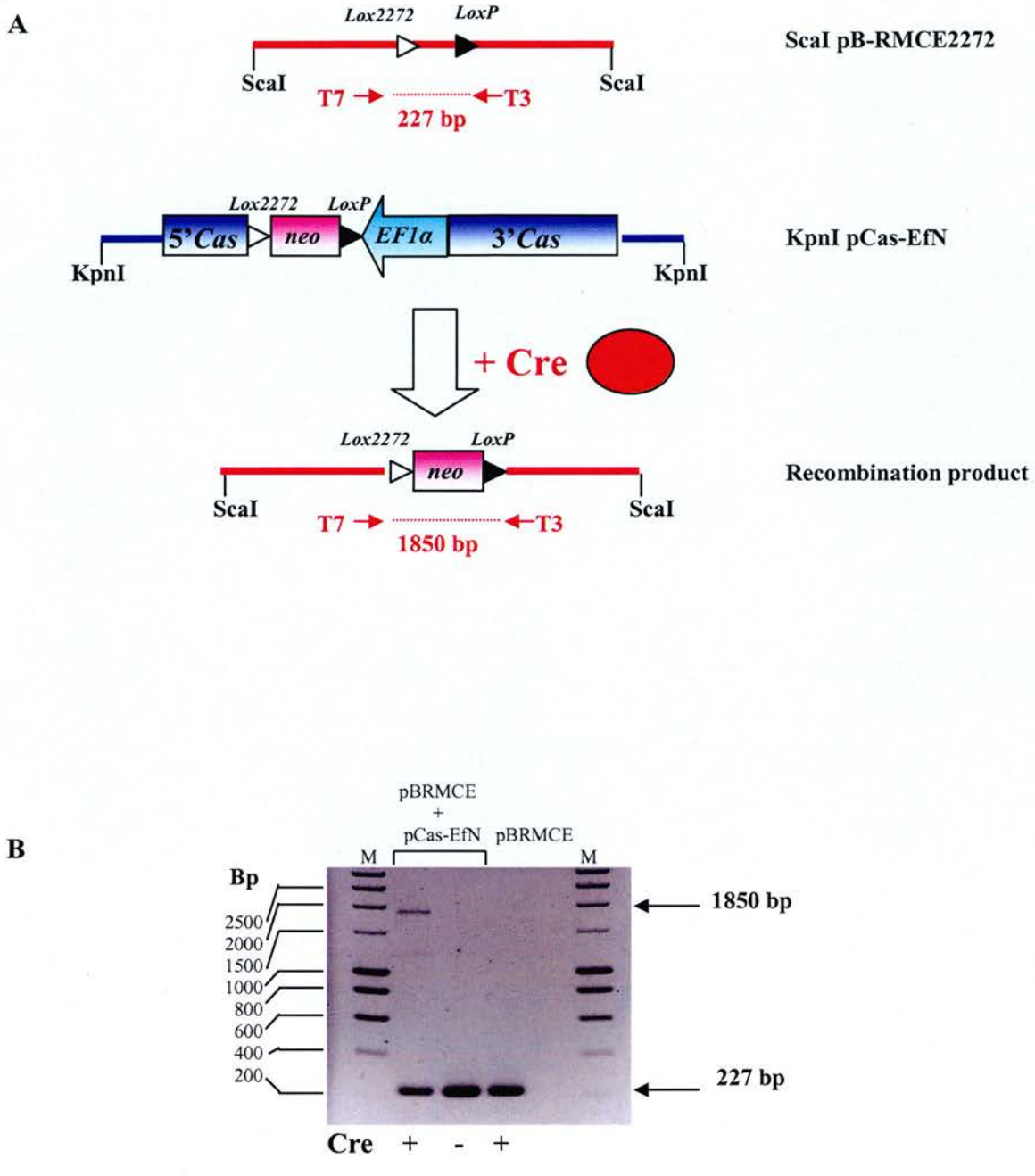


Figure 3.3. *In vitro* recombination assay between pCas-EfN and pB-RMCE2272. **(A)** Schematic representation of relevant segments of the plasmids pB-RMCE2272, pCas-EfN, and the potential recombination product arising from the recombination of each *lox* site with its counterpart. Positions of the primers and PCR product sizes are indicated in red. **(B)** PCR analysis of Cre reaction between plasmids pCas-EfN and pB-RMCE2272 treated with (+) or without (-) Cre recombinase using T7 and T3 primers. The 1850 bp band, indicative of site-specific recombination, can only be detected after Cre recombinase treatment of the two plasmids. M: Bioline hyperladder I.

3.3.1.4. G418 sensitivity of H9 cells

G418 (also known as geneticin) is an aminoglycosidic antibiotic commonly used for the selection of mammalian, plant or yeast cells transfected with the *neo* gene. Geneticin is an analogue of neomycin sulfate that interferes with the function of 80S ribosomes and protein synthesis (Daniels *et al.,* 1973). The product of the Tn*5* or Tn*601* aminoglycoside phosphotransferase gene (*neo*) confers resistance to G418 by phosphorylating the antibiotic, thus blocking its action (Southern & Berg, 1982). Prior to conducting a targeting experiment in H9 cells, it was important to determine the minimum concentration of G418 required to kill all the cells in a reasonable time (10 days). The working concentrations of G418 for mouse ES cells are usually around 300 μg/ml. However, because of genetic background variation and species differences, the H9 cell line might require a different concentration of G418 in order to recover G418-resistant colonies. The sensitivity of H9 cells to G418 was established by applying a range of G418 concentrations varying from 0 to 800 μg/ml to untransfected H9 cells. Ten days after selection, the cells were fixed and stained with 10% Giemsa R-66 solution. Figure 3.4 shows that 100 μg/ml of G418 is the minimal concentration that efficiently kills all untransfected H9 cells in 10 days and therefore was used for a subsequent *β-casein* targeting experiment in H9 cells.

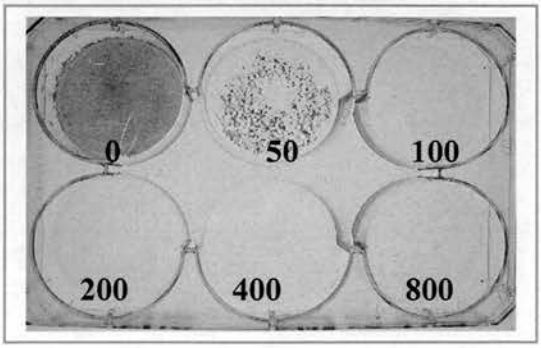


Figure 3.4. G418 sensitivity of H9 cells. H9 cells were exposed to a range of G418 concentrations (indicated in μg/ml) and after 10 days, were fixed and stained with 10 % Giemsa solution.

3.3.1.5. Establishment of a Transfection Protocol

Gene targeting by HR in murine ES cells is generally characterised by low frequencies (as the ratio between the number of homologous recombinants and the number of stably transfected clones) which vary between different target loci, from as little as 1 in 20 000 (Johnson *et al.*, 1989) to 1 in 3 (e.g. Te Riele *et al.*, 1992). Thus when an early report (Eiges *et al.*, 2001) showed that high stable transfection efficiencies (ratio between the total number of stable transfectants and the total number of electroporated cells) were difficult to achieve in hES cells, it became clear this would be a major obstacle to the routine isolation of gene-targeted hES cells. Eiges *et al.* (2001) assessed a number of transfection techniques and found the cationic polymer, ExGen 500 (Fermentas), to be most effective. This was used to produce the first reported transfected hES clones (expressing GFP), with a transfection efficiency of 3.3×10^{-5}. Whilst chemical transfection efficiencies are sufficient to generate transgenic hES cell lines, they would not provide sufficient clone numbers to optimise the chance of identifying gene-targeted clones from a transfection. Lipofection has also been shown to be less effective than electroporation for the production of gene-targeted human somatic cells (Yanez and Porter, 1999). Hence the focus of improvements for gene targeting in hES cells has been on electroporation methods (Zwaka and Thomson, 2003).

To optimise transfection efficiencies in H9 cells, a protocol based on standard mouse ES cell methods was first tested, with parameters adjusted to suit H9 cells (H. Priddle, unpublished data). In collaboration with A. Thomson (Roslin Institute), a pilot *β-casein* targeting experiment was conducted with the targeting vector pCas-EfN. 10^6 H9 cells were stably transfected in HBS (10 mM HEPES,

150 mM NaCl, pH 7.0) by electroporation (BioRad Gene Pulser II) at 200 V / 950 μF, with 50 μg of the KpnI-linearised targeting vector pCas-EfN (table 3.1). G418 selection (50 or 100 μg/ml) was applied 48 hours later and 11 days post-selection, 35 G418-resistant colonies were obtained with 50 μg/ml of selection, and 7 with 100 μg/ml of G418 (table 3.1). Therefore using this protocol, we obtained transfection efficiencies (between $1.4x10^{-5}$ and $7x10^{-5}$) similar to those reported by Eiges *et al.* (2001).

Cell line	Electroporation conditions	Number of cells treated	pCas-EfN (μg)	G418 (μg/ml)	G418^r colonies	Transfection efficiency
H9	200 V / 950 μF	10^6	50	50	35	$7x10^{-5}$
				100	7	$1.4x10^{-5}$

Table 3.1. Pilot *β-casein* targeting experiment with pCas-EfN in H9 cells.
10^6 H9 cells were electroporated with 50 μg of pCas-EfN using the conditions indicated and were maintained in G418 selection for 11 days.

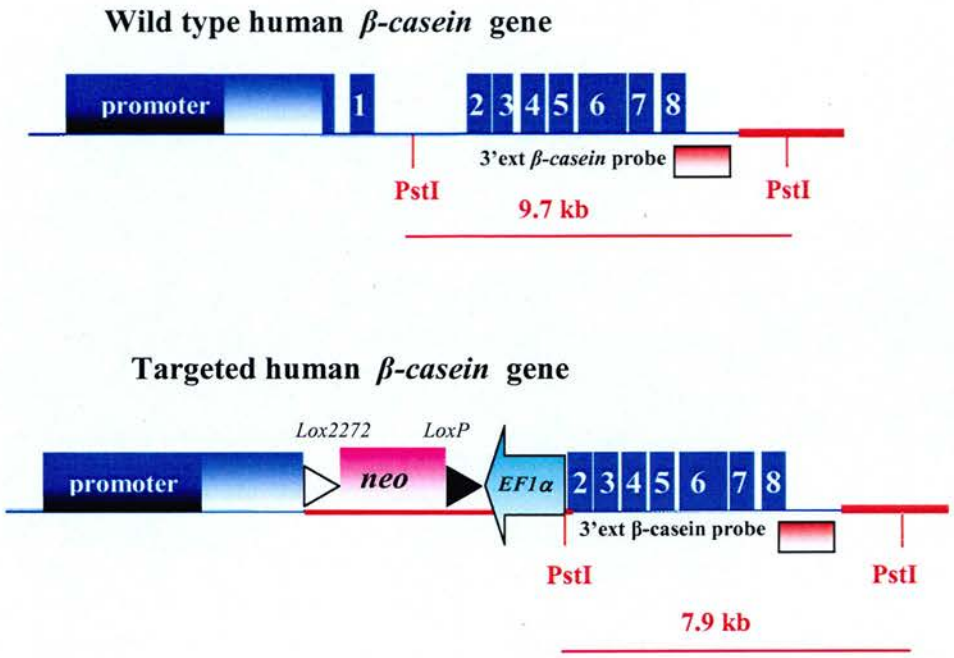
Because these transfection efficiencies were too low to be practical in the context of a targeting experiment, electroporation conditions were re-evaluated for H9 cells using the Eppendorf Multiporator system. This protocol utilises a hypo-osmolar buffer to expand the cell membrane and loosen its association with the cytoskeleton, reducing the required voltage. The Multiporator also applies the voltage in a shorter pulse, in the micro-second rather than milli-second range which is claimed to be much "gentler" for the cells (see the Eppendorf Multiporator Basic Applications Guide). By using this approach, stable transfection efficiencies of up to $2.3x10^{-4}$ were obtained in H9 cells with a *PGK-neo* plasmid (231 G418-resistant

clones were obtained after transfection of 10^6 H9 cells) using electroporation conditions of 300 V / 100 μs (H. Priddle, unpublished data). Therefore these transfection conditions were adopted for a subsequent *β-casein* targeting experiment.

3.3.1.6. Validation of the *β-Casein* Southern Probe

Southern analysis allows DNA fragments corresponding to a specific probe to be identified directly from a restriction digest of gDNA (Southern, 1975). In order to identify a targeting event at the *β-casein* gene in H9 cells by Southern blotting, a 365 bp *β-casein* probe (amplified by PCR from H9 gDNA) was generated and tested in a Southern screen with gDNA derived from wild-type H9 cells and G418-resistant clones from the *β-casein* pilot experiment. Figure 3.5 A shows the Southern PstI-diagnostic obtained with this probe with the wild-type allele (9.7 kb) and the targeted allele (7.9 kb). The *β-casein* probe was designed to recognise chromosomal sequences just beyond the 3' arm of homology of the *β-casein* gene used in pCas-EfN, so that random integrants would not be detected. To validate the probe, 10 μg of PstI-digested gDNA from wild-type H9 and the 42 H9 G418-resistant clones was analysed by Southern blotting. The presence of the predicted 9.7 kb PstI-fragment from the wild-type allele was detected in all the samples indicating that the probe was suitable for a *β-casein* targeting experiment (figure 3.5 B). The absence of the predicted 7.9 kb PstI fragment from the targeted allele, however indicated that none of the clones were the result of HR. Since targeting frequencies usually vary between 1:400 to 1: 2000 in mammalian cells, these results were expected as only 42 clones were analysed.

A



B

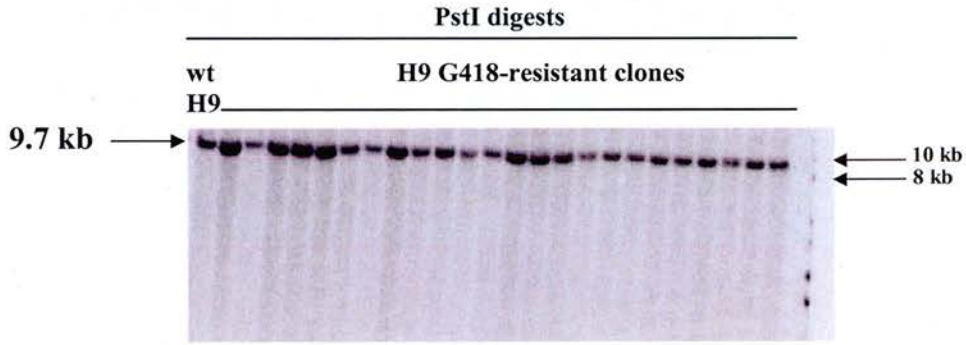


Figure 3.5. Validation of the *β-casein* probe. **(A)** Map of the relevant segments of the wild-type and targeted *β-casein* alleles showing the predicted PstI fragment diagnostics obtained with a *β-casein* probe external to the homology included in pCas-EfN. The *β-casein* sequence is represented as a thin blue line. The red line represents the sequence outside the *β-casein* locus. The position of the probe is indicated as a red box with the size of the predicted PstI fragment shown below. **(B)** Southern analysis of wild-type (wt) H9 and a representative subset of the 42 G418-resistant H9 subclones. The unmodified *β-casein* allele yielded the predicted 9.7 Kb fragment but the predicted 7.8 kb from the modified allele was not detected in any of the clones. The molecular ladder (Hyperladder, Bioline) is indicated on the right.

3.3.2. *β-Casein* Targeting Experiments

3.3.2.1. β-Casein Targeting Experiment I

Following the optimisation of experimental conditions in H9 cells, a large-scale *β-casein* targeting experiment was conducted. 10^6 H9 cells were electroporated with the Eppendorf Multiporator with either 50 or 100 μg of the KpnI-linearised targeting vector pCas-EfN (table 3.2). G418 selection (100 μg/ml) was applied 48 h later and fourteen days after selection, 137 and 177 G418-resistant colonies were produced with 50 and 100 μg of pCas-EfN, respectively (table 3.2). A control plate in which H9 cells were transfected in the absence of the targeting construct did not produce any G418-resistant colonies, indicating that those generated in the experimental plates were actually G418-resistant. Therefore the transfection efficiency obtained with this protocol was up to 13-fold greater than that obtained in the pilot targeting experiment (table 3.2). As a result of bacterial contamination, 48 clones were lost, therefore 266 G418-resistant clones were expanded individually to two wells of a 12-well plate for freezing and gDNA extraction.

Cell line	Electroporation Conditions	Number of cells treated	pCas-EfN (μg)	G418 (μg/ml)	G418^r colonies	Transfection efficiency
H9	300 V/100 μs	10^6	50	100	137	1.4×10^{-4}
		10^6	100	100	177*	1.8×10^{-4}

Table 3.2. β-casein targeting experiment I with pCas-EfN in H9 cells. * 48 clones were discarded as a result of bacterial contamination.

To screen for targeting events at the *β-casein* gene, PstI-digested gDNA from each clone was analysed by Southern blotting using the external *β-casein* probe (figure 3.5.A). The probe was found to hybridise to the predicted 9.7 kb PstI fragment from the wild-type allele in all the clones. However, the predicted 7.9 kb PstI fragment from the targeted allele was not detected in any of the clones analysed indicating that none of them were the result of HR (figure 3.6). These data show that the targeting frequency is below 0.38% since no targeting event was detected in 266 H9 G418-resistant clones.

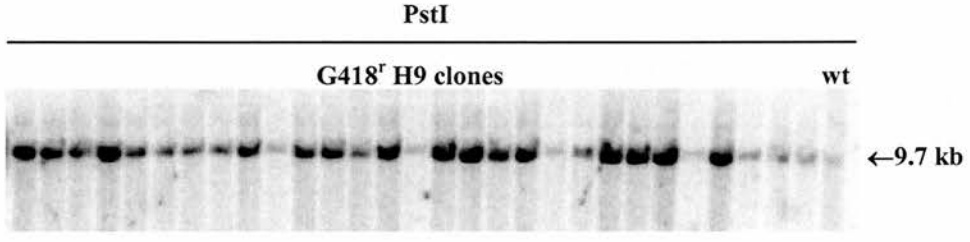


Figure 3.6. Southern blot analysis of a representative subset of the 266 G418-resistant H9 clones. The presence of the 9.7 kb PstI fragment from the wild-type allele can be detected but the absence of the 7.8 kb PstI fragment from the targeted allele indicates that none of the clones are the result of gene targeting at the *β-casein* gene.

3.3.2.2. *β-Casein* Targeting Experiment II

Although no targeting event was detected in H9 cells in the previous *β-casein* targeting experiment, 266 clones might not have been sufficient to detect a targeting event at the *β-casein* gene. Our strategy was therefore revised to increase the number of G418-resistant colonies. Data from the pilot experiment showed that using 50 rather than 100 μg/ml of G418 produced more colonies (35 versus 7) suggesting that colonies expressing low levels of *neo* could survive at 50 μg/ml but not at 100 μg/ml of G418. Because the *β-casein* gene is not expressed in hES cells, it is possible that the *neo* gene would be expressed at a low level at this locus due to a potential surrounding condensed chromatin structure, often observed at a transcriptionally inactive gene (Felsenfeld and Groudine, 2003). Based on this hypothesis, selection with 100 μg/ml of G418 may have killed targeted cells, whereas a primary G418 selection of 50 μg/ml might allow the survival of targeted cells. It has been suggested that transcription itself leads to distinct changes (i.e. "opening") in the structure of surrounding chromatin (Lee and Garrard, 1991). Thus it is possible that initial transcriptional activation of the *neo* gene at 50 μg/ml may lead to chromatin remodelling, thus rendering the *β-casein* locus more permissive to transgene expression. Allowing the transcription machinery to operate with a primary lower selection regime, may therefore allow the survival of targeted cells when G148 selection is subsequently raised to 100 μg/ml.

Based on this reasoning, another targeting experiment was conducted in which 10^6 H9 cells were electroporated with 50 μg of the KpnI linearised targeting vector pCas-EfN using the Eppendorf Multiporator system (table 3.3). G418 selection (50 μg/ml) was applied 48 h later and seven days post-electroporation, the

G418 concentration was doubled. This new selection regime produced 354 G418-resistant colonies, an increase of transfection efficiency of up to 2.5-fold (table 3.3), over the previous targeting experiment in which 50 μg of targeting vector was used (table 3.2). All the colonies obtained in this experiment were the result of resistance to the drug G418, as a mock control, in which cells were electroporated in the absence of DNA, did not produce any G418-resistant colonies following the same regime of G418 selection. The increase in G418-resistant colonies suggests that either more colonies expressing low levels of *neomycin* survived selection, or that this increase was the result of transfection efficiency variation between the two targeting experiments. Fourteen days after G418 selection, the 354 individual clones were picked and expanded up to 2 wells of a 12-well plate for freezing and gDNA extraction for Southern analysis.

Cell line	Electroporation conditions	Number of cells treated	pCas-EfN (μg)	G418 (μg/ml)	$G418^r$ colonies	Transfection efficiency	ETF (%)
H9	300 V/100 μs	10^6	50	50 (5 days) 100 (9 days)	354	$3.5x10^{-4}$	<0.3

Table 3.3. *β-casein* targeting experiment II with pCas-EfN in H9 cells. Effective targeting frequency (E.T.F): ratio between the number of homologous recombinants and the numbers of stable transfectants (see figure 3.7 for Southern analysis).

PstI-digested gDNA from each clone was analysed by Southern blotting to screen for targeted events with the external *β-casein* probe. The 9.7 kb PstI fragment from the wild-type allele was detected but not the 7.9 kb PstI fragment indicative of a targeted allele, showing that none of these clones were the result of HR at the *β-casein* gene (figure 3.7). These data indicate that targeting frequency at

the *β-casein* gene in H9 cells is below 0.3 % (< 1/354) using these experimental conditions, thus much lower than the 8 % reported by Kolb at the *β-casein* gene in mouse ES cells (Kolb, 2001). Kolb however, applied an enrichment strategy with the use of a positive/negative selection vector (Capecchi, 1989).

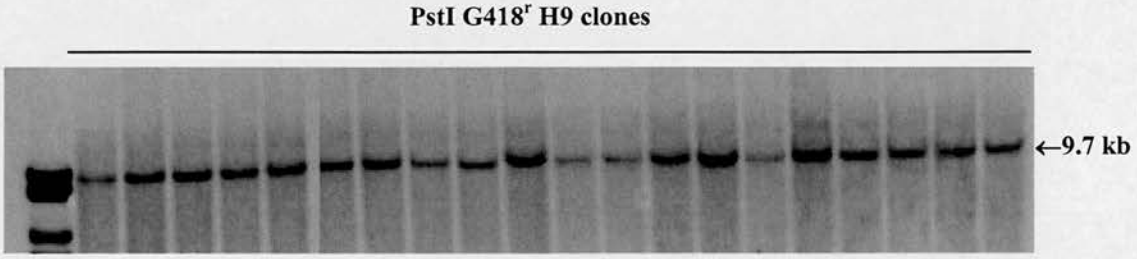


Figure 3.7. Southern blot analysis of a representative subset of the 354 G418-resistant H9 clones. The unmodified *β-casein* allele yielded the predicted 9.7 Kb fragment but the predicted 7.8 kb from the modified allele was not detected in any of the clones.

Failure to target the *β-casein* gene in H9 cells might be due to two possibilities; (i) The *β-casein* gene is refractory to gene targeting in H9 cells and therefore no targeting event has occurred; (ii) A targeting event has occurred but the targeted cell was killed because the threshold of resistance to G418 provided by *neo* expression was below 100 μg/ml. Because unexpressed genes (including the *β-casein* gene) have previously been targeted in mouse ES cells (Koller and Smithies, 1989; Johnson *et al.*, 1989; Kumar *et al.*, 1994; Kolb *et al.*, 1999; Kolb, 2001), it seems unlikely that hES cells are intrinsically refractory to gene targeting at the *β-casein* gene. If the *neo* gene was silenced at the *β-casein* gene, one solution would be to protect *neo* expression from the silencing effect of neighbouring chromatin. The use of insulator elements such as the chicken HS4 (Chung *et al.*, 1993) may be a solution.

3.3.3. Protective Effect of Insulators on Transgene Expression in H9 cells

3.3.3.1. Strategy

Chromosomal position effects on transgene expression have been a major problem in gene regulation studies in transgenic mice as well as in the development of successful gene therapy strategies (Clark *et al.*, 1994; Martin and Whitelaw, 1996). A considerable effort has therefore been made to identify the cis-regulatory requirements for tissue-specific, high-level and position-independent expression of transgenes. The existence of insulators next to some genes was shown to provide a barrier to shield the coding region from its surroundings, and as a result has led to the development of strategies to overcome chomosomal position effects (reviewed by Recillas-Targa *et al.*, 2004). Insulators are DNA elements located at the boundaries of a transcriptional domain. Their function is to shield a gene from either the action of a distal enhancer or silencer, and/or from the silencing effects of condensed chromatin (for more details, see section 1.2.2.1., chapter 1). As an example, a 1.2 kb 5' constitutive DNaseI-hypersensitive (HS) region with insulator properties from the chicken *β-globin* cluster was shown to protect endogenous gene or transgene expression from chromosomal position effects in mammalian cells (Chung *et al.*, 1993). In a later report, Chung *et al.* (1997) showed that the activity of the 1.2 kb HS4 fragment resides in a 250 bp core element which contains the constitutive DNase I-hypersensitive site (5'HS4) and that the insulating activity of the core element is multiplied when tandem copies are used on both sides of a transgene. Ma *et al.* (2003) showed that the inclusion of the 1.2 kb HS4 fragment in a *EF1α*-GFP lentiviral vector, provides significantly reduced variability of GFP

expression in the H9 cell line, compared with an *EF1α*-GFP lentiviral vector lacking the HS4 element.

Based on these data, we tested the ability of tandem copies of the 250 bp HS4 core elements to protect transgene expression in H9 cells. The insulating properties of HS4 were tested by comparing GFP expression between a plasmid containing a *PGK*GFP expression cassette (pPGKGFP-SV40Neo, given by A.Thomson) and a similar plasmid with a *PGK*GFP flanked on both sides by the HS4 insulators (pHS4PGKGFP-SV40Neo, given by T. Lacroix) (figure 3.8). 10^6 H9 cells were stably transfected with 50 µg of either DraIII-linearised pHS4PGKGFP-SV40Neo or DraIII-linearised pPGKGFP-SV40Neo using the Multiporator protocol. An *SV40neo* expression cassette was also included in the plasmids (figure 3.8) to allow the generation of G148-resistant colonies 10 days later. Examination by fluorescence microscopy 12 days post-electroporation, showed that few colonies expressed GFP in both transfection experiments (data not shown). Additionally, position effect variegation (PEV) was observed in most colonies. However, PEV was minimal in clones derived from the transfection with pHS4PGKGFP-SV40Neo (+ insulators) suggesting that insulators were able to protect GFP expression from PEV. Colonies from both transfections were pooled and expanded to a T25 flask for flow cytometric analysis to compare GFP expression levels.

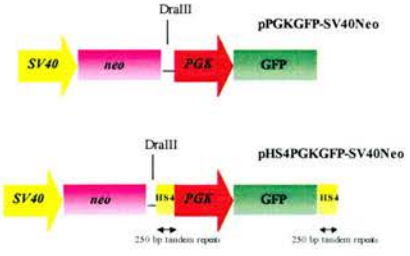
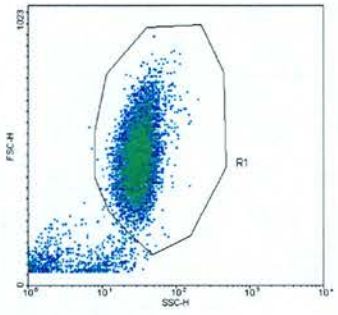


Figure 3.8. Effect of HS4 insulators on GFP expression in H9 cells. Cells were either transfected with DraIII-linearised pPGKGFP-SV40Neo (no insulators) or DraIII-linearised pHS4PGKGFP-SV40Neo (+insulators). The SV40-*neo* cassette provides resistance to the drug G418.

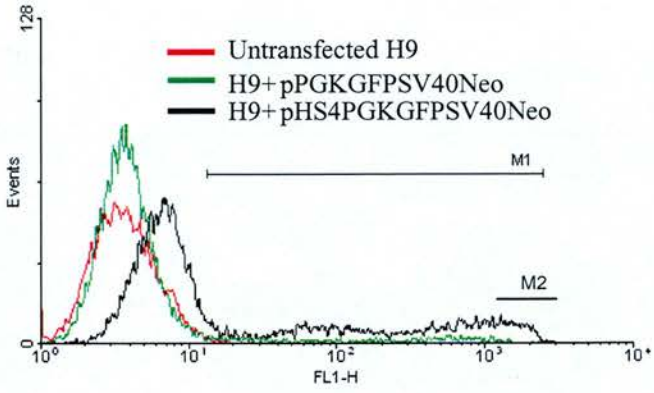
3.3.3.2. FACS Analysis

To compare GFP expression between the two transfections, cells were collected and examined by flow cytometry. Flow cytometry offers a quantitative way of assessing GFP expression levels in the cells. Untransfected H9 cells were used as a negative control for GFP expression. In this analysis, only the live cells were analysed. The dead cells were excluded based on their forward size scatter (FSC-H) and side scatter (SSC-H) profile. A representative density plot showing the region (R1) which includes only the live cells analysed is illustrated in figure 3.9 A. Quantification of GFP-positive cells showed that only 8.5 % of H9 cells transfected with pPGKGFP-SV40Neo (no insulators) expressed GFP. In contrast, ~ 36 % of H9 cells transfected with pHS4PGKGFP-SV40Neo (+ insulators) expressed GFP (figure 3.9 B & C). Thus a 4-fold increase of GFP-expressing cells was observed when the GFP expression cassette was flanked by HS4 elements consistent with an increase of protection of transgene expression. Among the GFP-positive population, ~ 3% (M2) of these cells were found to express GFP at higher levels than cells transfected with pPGKGFP-SV40Neo (figure 3.9 B). These preliminary data therefore suggest that HS4 elements insulate a transgene in H9 cells and could be of use in preventing the silencing of transgenes inserted at unexpressed or any other loci.

A



B



C

Sample		% GFP + cells
Untransfected H9	—	0.77
H9 + pPGKGFP-SV40Neo	—	8.51
H9 + pHS4PGKGFP-SV40Neo	—	36.1

Figure 3.9. Flow cytometric analysis of pooled H9-GFP+ cells transfected with pHS4PGKGFPSV40Neo or pPGKGFPSV40Neo. **(A)** Representative density plot showing the region R1 containing the live cells analysed. **(B)** Histogram representation of untransfected H9 cells (negative control) (red) and H9 cells transfected with either pHS4PGKGFPSV40Neo (black), or pPGKGFPSV40Neo (green). Data were generated from the analysis of at least 10000 viable cells. The gate M1 indicates the total number of GFP-positive cells and the gate M2 indicates the 3% of pHS4PGKGFPSV40Neo transfected-H9 cells expressing GFP at high levels. **(C)** Percentages of GFP+ cells for untransfected, pPGKGFPSV40Neo and pHS4PGKGFPSV40Neo transfected cells.

3.4. Discussion

We sought to engineer the *β-casein* gene in H9 cells by introducing, by gene targeting, a selectable *neo* marker flanked by incompatible *lox* sites and driven by the *EF1-α* promoter. Following gene targeting by HR, a promoterless selectable marker, also flanked by incompatible *lox* sites, would be exchanged against the *neo* marker by RMCE using a promoter trap strategy. Using an electroporation protocol, transfection efficiencies obtained in these targeting experiments reached up to $3.5x10^{-4}$ (table 3.3). This was 10-fold greater than the best transfection efficiency achieved (using a chemical transfection protocol) obtained so far in H9 cells (Eiges *et al.*, 2001). Southern analysis however, showed that none of the 668 G418-resistant colonies generated were the result of gene targeting at the *β-casein* gene.

Failure to target the *β-casein* gene in H9 cells could have been due to several reasons; One possibility is that the *β-casein* gene is less permissive to gene targeting because of its transcriptionally-inactive status. It has been suggested that transcription stimulates gene targeting frequency by creating a more favourable target for recombination enzymes, through the introduction of topological changes within the target (Thyagarajan *et al.*, 1995). While there is evidence that transcription can stimulate intrachromosomal HR (Nickoloff, 1992), extrachromosomal HR (Nickoloff and Reynolds, 1990) and gene targeting in mammalian cells (Thyagarajan *et al.*, 1995), it is also known that untrancribed genes can be targeted with frequencies similar to transcribed genes in murine ES cells (Johnson *et al.,* 1989; Koller and Smithies, 1989; Kumar *et al.*, 1994; Kolb *et al.,*

1999; Kolb, 2001). Thus these studies show that a low or nil transcription rate is probably not a limiting factor in gene targeting frequency.

Another possibility is that gene targeting has occurred at the *β-casein* locus but the *neo* was silenced or otherwise down-regulated, perhaps by heterochromatinisation of the locus, resulting in death of the targeted cells in the presence of G418. Marques *et al.* (manuscript in preparation) have previously attempted to target the *β-casein* gene in sheep foetal fibroblasts, and although PCR screening and restriction digestions of the PCR product indicated successful gene targeting for some of the clones, Southern screening failed to provide evidence of gene targeting in the candidate clones analysed. They claimed that targeted cell clones were originally mixed with random integrants which eventually took over the culture during expansion due to silencing of the selectable marker at the *β-casein* gene. Thus targeted clones could be identified by PCR but not by Southern analysis. Their result is consistent with a previous report in which failure to detect a targeting event in a human bladder carcinoma cell line was the result of cells not surviving drug selection when the transgene was inserted into an inactive chromosomal locus (Smithies *et al.*, 1985). Another group recently reported the lack of expression of a transgene targeted into the non-expressed *immunoglobulin-μ* and prion protein (*PrP*) genes in bovine foetal fibroblasts, possibly due to silencing of the integrated transgene as part of the silent locus (Kuroiwa *et al.,* 2004).

Similarly, if *neo* silencing occurs at the *β-casein* locus in H9 cells, targeted clones would have eventually died and all of the G418-resistant clones would be random integrants. It has been shown that targeted cells can be reliably identified by PCR in pools in which non-homologous recombinants cells outnumber homologous

recombinants by 100 to 1 (Kim *et al.*, 1991). Hence screening for a targeting event at the *β-casein* gene in H9 cells could be facilitated by combining individual early G418-resistant clones (obtained at 50 μg/ml G418) into small pools (<100 colonies) for analysis by PCR (Hanson and Sedivy, 1995). If gene targeting was confirmed, the targeting protocol could be modified by relaxing the selection as soon as colonies become evident and to re-clone surviving colonies in the absence of selection in order to identify targeted clones by Southern analysis.

The preliminary data presented in section 3.3.3 showed a 4-fold increase of GFP-expressing H9 cells when a GFP transgene was flanked on both sides by HS4 insulators, indicating that these elements protect transgene expression from silencing. Thus the addition of HS4 elements on both sides of the *neo* cassette in the *β-casein* targeting vector may be a solution in the future to improve targeting frequencies at the *β-casein* locus. A positive and negative selection could be employed (Mansour *et al.*, 1988; Capecchi, 1989) by including an HSV*tk* (herpes simplex virus thymidine kinase) (McKnight, 1980) expression cassette outside the regions of homology of the targeting construct in order to reduce the number of clones with random integration events in a cell population. In this enrichment scheme, G418 selection would be first applied to select for cells that have integrated the targeting vector in their genome. In a second step, the expression of the HSV*tk* expression cassette, which converts nucleoside analogues into toxic metabolites, would provide sensitivity to the drug gancyclovir (St Clair *et al.*, 1987). This selects against cells in which the entire targeting vector has integrated randomly. On the other hand, cells carrying a targeted integration should lose the HSV*tk* cassette during HR and therefore survive selection. Although a positive-negative selection strategy has been reported to produce modest

enrichments, typically ranging from 2-to 10-fold (Hanson and Sedivy, 1995), the addition of a second negative selection marker in the *β-casein* targeting vector could however increase this enrichment and should significantly ease the identification of targeted clones by reducing the number of random integrants (Hanson and Sedivy, 1995).

CHAPTER IV

Engineering the *Hprt* Gene in H1 Cells by Homologous Recombination

4.1. Introduction

The X-linked hypoxanthine-guanine phosphoribosyl transferase (*hprt)* gene is a housekeeping gene which is involved in the purine salvage pathway (Stout and Caskey, 1988). We sought to engineer the *hprt* gene because it presents several advantages: (i) The *hprt* gene is hemizygous in male cells and therefore a genetic modification of a single allele leads to complete loss of function; (ii) HPRT activity can be selected for or against (Albertini, 2001). HPRT catalyses the addition of purines to the sugar intermediate 5' phosphoribosyl-1-pyrophosphate (PRPP) in wild-type cells (for details of the HPRT selection system, see section 1.4.4.2). Either in the absence of purines or when *hprt* is not functional, the cells are forced to follow an alternative *de novo* synthesis pathway. Aminopterin selection inhibits the enzyme dihydrofolate reductase, which is essential for *de novo* synthesis. Hence cells with functional HPRT will survive Aminopterin selection if provided with Hypoxanthine and Thymidine in the medium (HAT) (Szybalski and Szybalska, 1962). Because wild-type cells are able to use the salvage pathway, they are not affected by the presence of HAT in the medium, unlike HPRT-deficient cells that have both purine ribonucleotide biosynthesis routes blocked and die in HAT. HPRT can also catalyse the addition of a toxic purine analogue, 6-thioguanine (Stutts and Brockman, 1963; Sharp *et al.*, 1973, Wahl *et al*, 1975) to PRPP in the salvage pathway, resulting in cell death. In contrast, HPRT-deficient cells do not show any alteration in their metabolism in the presence of 6-TG, since these cells lack an active salvage pathway. This ability to select for or against HPRT activity, therefore renders targeting frequencies very easy to estimate at a phenotypic level (i.e. colony counting) in a

male cell line (Albertini, 2001); (iii) The *hprt* locus is well known for its ability to provide tissue-specific reproducible transgene expression in several cell lines (Vivian *et al.*,1999; Evans *et al.*, 2000; Guillot *et al.,* 2000; Cvetkovic *et al.,* 2000; Minami *et al.*, 2002; Heany *et al.*, 2004) and therefore should be a good candidate gene for reliable transgene expression in hES cells; (iv) The mutation of this gene in hES cells would provide a model for Lesch-Nyhan syndrome (LNS). LNS is a severe human genetic disease that is caused by the malfunction of the *hprt* gene (Stout and Caskey, 1988). Although the mouse model of LNS exhibits the same biochemical defect as LNS patients, it does not mimic any of the behavioral symptoms associated with the disease (Bedell *et al.,* 1997; Elsea and Lucas, 2002).

Gene targeting has recently been described at the *hprt* gene in hES cells (Zwaka and Thomson, 2003; Urbach *et al.*, 2004). Zwaka and Thomson deleted regions of the last 3 exons of the gene in a subclone of the H1 male ES cell line (Thomson *et al.*, 1998) with a replacement construct using an electroporation protocol. Their targeting protocol resulted in a targeting frequency of up to 2 % in hES cells. Urbach *et al.*, (2004) inserted a transgene in the exon 2 of the gene in the H13 male ES cell line (Itskovitz-Eldor *et al.*, 2000) with a replacement construct using a chemical protocol which resulted in a targeting frequency of up to 1 %.

We sought to engineer the H1 male ES cell line by introducing, by homologous recombination (HR), a selectable marker flanked by heterospecific *lox* sites (Lee and Saito, 1998) into exon 3 of the *hprt* gene. The insertion of *lox* sites will permit the use of recombinase-mediated cassette exchange (RMCE) (Feng *et al.*, 1999; Trinh and Morrison, 2000; Kolb, 2001) to examine transgene expression at the *hprt* gene (see chapter 7).

4.2. Objectives

To engineer the *hprt* gene by HR in H1 cells by tagging the locus with heterospecific *lox* sites

4.3. Results

4.3.1. Targeting Scheme and *Hprt* Targeting Vector

To target the *hprt* gene in H1 cells, we used a replacement targeting vector, pHprt-EfN. Figure 4.1 shows the targeting scheme and the structure of the vector. The *hprt* homology used in this construct is a 6.9 kb HindIII fragment (Edwards *et al.,* 1990) comprising exons 2 and 3, intron 2 and fragments of introns 1 and 3. A *neomycin* expression cassette was inserted at the XhoI site in exon 3, in the opposite transcriptional orientation, separating the *hprt* sequence into 4.8 kb 5' and 2.1 kb 3' elements. Following double reciprocal recombination between the targeting vector and the *hprt* chromosomal sequences, a 3.4 kb *EF1-α-loxP-neopA-lox2272* expression cassette is inserted in exon 3 of the *hprt* gene (see paragraph below for the structure of pHprt-EfN). Thus this insertion permits selection for random and homologous recombinants with the drug G418, and for homologous recombinants alone with the drug 6-TG.

The pHprt-EfN targeting vector was used to introduce the *loxP* and *lox2272* heterospecific *lox* sites (Lee and Saito, 1998) into the *hprt* gene such that RMCE (Feng *et al.*, 1999; Trinh and Morrison, 2000; Kolb, 2001) can be subsequently applied. This construct is derived from the *p*E3Δ*neo* vector (Zheng *et al.,* 1991) which contains the 6.9 kb HindIII fragment (Edwards *et al.,* 1990) with an

HSV*tkneopA* expression cassette (Thomas & Capecchi, 1987) inserted at the XhoI site in exon 3 of the *hprt* gene. The plasmid pHprt-EfN was generated by replacing the HSV*tkneopA* with an *EF1-α-loxP-neopA-lox2272* expression cassette such that following HR, the *hprt* gene is tagged with heterospecific *lox* sites in the exon 3 (figure 4.1). Because the *EF1-α* promoter is not included between the *lox* sites, a promoter trap strategy (Jasin and Berg, 1988; Sedivy and Sharp, 1989) can subsequently be applied in an RMCE experiment. Following site-specific recombination, the *neo* selectable marker between the *lox* sites is exchanged with a promoterless selectable marker flanked by the same set of heterospecific *lox* sites which will then be driven by the *EF1-α* promoter (see chapter 7). This enrichment strategy therefore permits the direct selection of site-specific recombinants since the majority of random integrations will be transcriptionally silent and not survive selection. The ubiquitously expressed *EF1-α* promoter was chosen because it is a strong promoter (Mizushima & Nagata, 1990) and because preliminary transfection experiments in our laboratory showed that the transcriptional activity of this promoter was stronger than the HSV*tk* promoter in hES cells (H. Priddle, unpublished data).

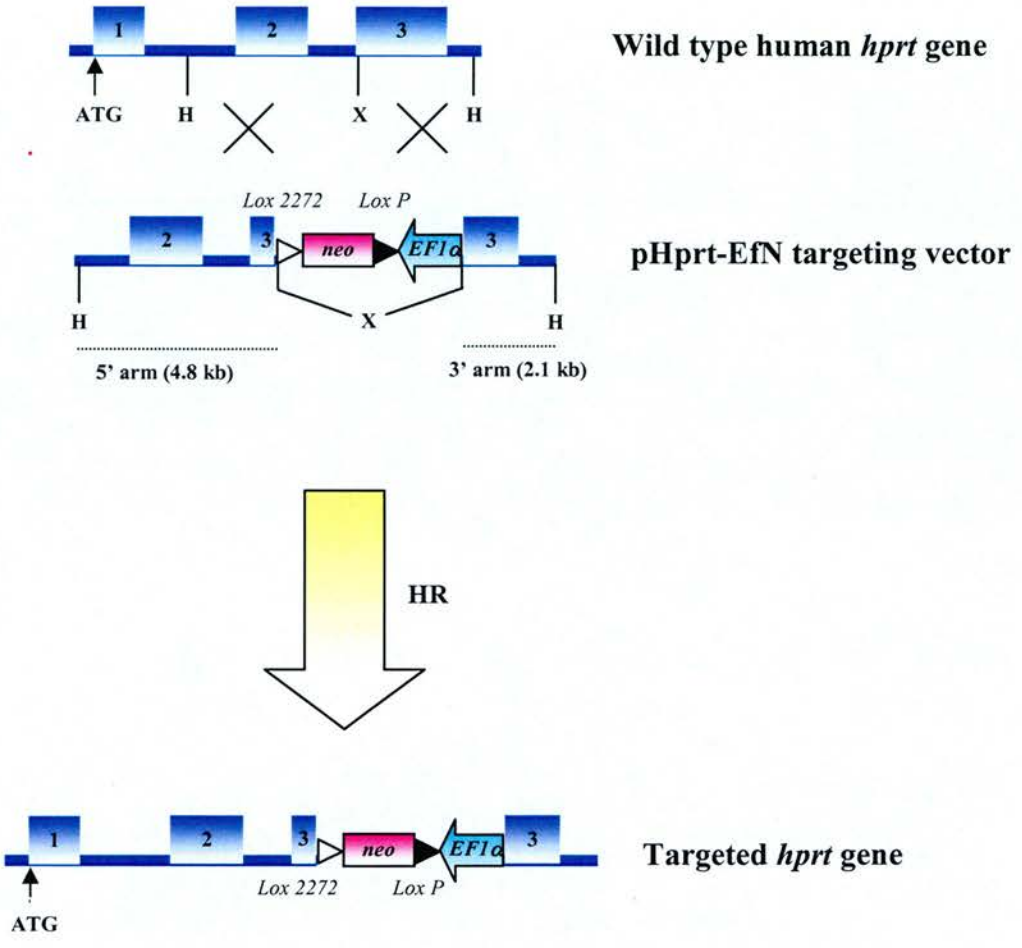


Figure 4.1. Schematic representation of the human *hprt* gene targeting in H1 cells. The structure of the relevant *hprt* sequence is represented by a thick blue line with boxes above marking the exons with numbers indicated. The translational initiation start (ATG) is shown below the first exon. The restriction sites HindIII (H) and XhoI (X), divide the *hprt* homology sequence into 4.8 kb 5' and 2.1 kb 3' arms. The *EF1-α* promoter is shown as a green solid arrow. The *neo* expression cassette is shown as a solid pink box. The positions and orientations of the *loxP* and *lox2272* recombination sites are indicated as solid black and white arrows, respectively. The two cross-over recombination points which take place between the targeting vector and the *hprt* chromosomal sequence are depicted by an 'X'. Following HR with the HindIII-digested targeting vector, targeted recombinants can be selected for, based on their resistance to 6-TG.

4.3.2. G418 Sensitivity in H1 Cells

As previously described in section 3.3.1.4 of chapter 3, G418 is an antibiotic which interferes with protein synthesis (Daniels *et al.,* 1973) and whose action is blocked by the expression of the *neomycin* gene (Southern & Berg, 1982). Targeting experiments at the *hprt* gene require a rapid primary G418 selection to ensure that targeted and non-targeted colonies are separated before adding 6-TG selection. The proximity of non-targeted cells to targeted cells could favour the exchange of the 5' phosphoribosyl-1-pyrophosphate (PRPP) and 6-TG metabolites which may kill targeted cells by a bystander effect or "kiss of death" (Andrade-Rozental *et al.*, 2000).

Although the G418 sensitivity threshold was previously determined in the female H9 cell line, it was necessary to establish a sensitivity threshold in the male H1 cell line, because a different genetic background may have different drug requirements to achieve the same effect. As a result, H1 cells were seeded at low density in a 6-well plate and exposed to a range of G418 concentrations varying from 0 to 800 μg/ml (figure 4.2). Ten days after selection, the cells were fixed and stained with 10 % Giemsa R-66 solution. As in H9 cells, this study showed that the antibiotic killed all H1 cells within 10 days at a minimum concentration of 100 μg/ml (Figure 4.2). For the subsequent targeting experiments in H1 cells, we chose to apply G418 selection at 50 μg/ml for 5 days before increasing the concentration to 100 μg/ml because we wanted to follow a similar selection regime to that used by Zwaka and Thomson in their *hprt* targeting experiment (Zwaka and Thomson, 2003).

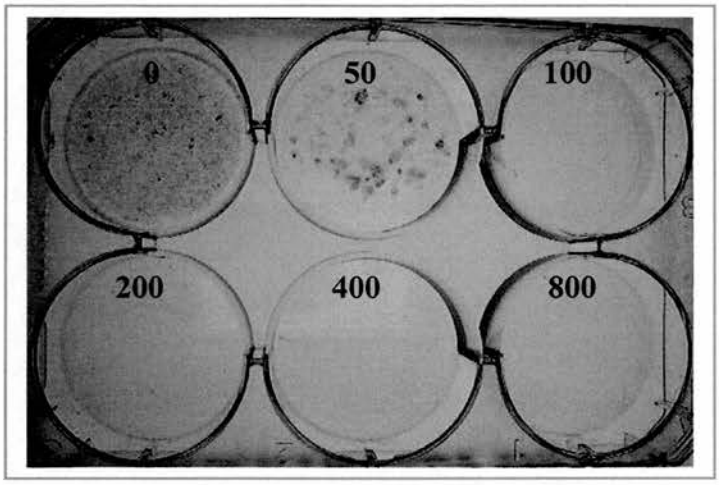


Figure 4.2. G418 sensitivity of H1 cells. H1 cells were subjected to varying concentrations of G418 (indicated in μg/ml) for ten days and were then fixed and stained with 10% Giemsa solution.

4.3.3. *Hprt* Targeting Experiments in H1 Cells

To target the *hprt* gene with the pHprt-EfN replacement vector, three different sets of electroporation conditions were compared in H1 cells. The first electroporation protocol (condition 1) was that established by Zwaka and Thomson who successfully target the *hprt* gene in a subline of H1 cells (Zwaka & Thomson, 2003). Using electroporation conditions of 320 V and 200 μF with the BioRad Gene Pulser II, they obtained 350 G418-resistant colonies. Of these, 7 were the result of gene targeting at the *hprt* gene (G418/6-TG-resistant). The two remaining electroporation protocols were established in our laboratory for H9 cells (described in section 3.3.1.5 of chapter 3) using the Eppendorf Multiporator (300 V / 100 μs) (condition 2) and the BioRad Gene Pulser II (200 V / 950 μF) (condition 3).

10^6 H1 cells were transfected with 50 μg of HindIII-digested pHprt-EfN using the three different electroporation conditions described above (table 4.1). The cells were exposed to 50 μg/ml of G418 selection 48 hours later, and 5 days post-transfection G418 concentration was increased to 100 μg/ml. 10 days post-transfection 600, 488 and 380 G418-resistant colonies were produced with the electroporation conditions 1, 2 and 3, respectively (table 4.1). A control experiment, in which H1 cells were electroporated without the vector pHprt-EfN, did not produce any G418-resistant colonies, suggesting that those generated in the experimental plates were actually G418-resistant. It is however not unlikely that in the experimental plates some G418-sensitive cells which survived at 50 μg/ml of G418, as shown on figure 4.2, would subsequently survive at 100 μg/ml by metabolic cooperation if these were in the proximity of G418-resistant cells. As a result, G418-sensitive cells surviving by metabolic cooperation may have killed some targeted

cells by a bystander effect (see section 4.3.2) once 6-TG selection was applied. Thus this selection regime may have affected targeting frequency.

6-TG (2.5 μg/ml) was applied 10 days post-transfection to select for homologous recombinants. It is critical to maintain the cells for at least a week in G418 selection before adding 6-TG selection, to clear residual *hprt* message and protein in targeted cells that can no longer synthesise the HPRT enzyme (Thomas and Capecchi, 1987). If 6-TG is applied too soon, residual HPRT enzyme will phosphoribosylate the toxic 6-TG producing lethal levels of 6-ThioGMP which will be fatal for targeted cells (described in section 1.4.4.2, chapter I). After 6 days of selection in 6-TG, 1 G418/6-TG-resistant colony was recovered with each of the electroporation conditions 1 (320 V / 200 μF) and 2 (300 V/ 100 μs) suggesting that these clones were targeted at the *hprt* gene (table 4.1). The third electroporation protocol (200 V / 950 μF) however, did not result in any 6-TG-resistant colony. Although the two 6-TG resistant clones obtained in this experiment were likely to be the result of a gene targeting event in the *hprt* gene, the possibility that a spontaneous *hprt* mutation could have arisen during routine culture of the H1 cells prior to the targeting experiments needed to be ruled out. Since these cells were never in HAT selection which would select for wild-type cells only, there might have already been a sub-population of spontaneous *hprt* mutants in the culture. However, if spontaneous mutation rates of *hprt* in hES cells are similar to those reported in mouse ES cells ($< 10^{-8}$ /cell/generation) (Chen *et al.,* 1998; Chen *et al.*, 2000; Munroe *et al.*, 2000; Cervantes, 2002), it was unlikely that this was the case. In order to distinguish between targeted and spontaneous mutation, these clones were expanded and analysed by PCR and Southern to confirm targeting (see section 4.3.4).

To confirm that targeting the *hprt* gene was reproducible in H1 cells, another targeting experiment was conducted with pHprt-EfN using the Eppendorf Multiporator (300 V / 100 μs). The same experimental conditions described in the first targeting experiment with pHprt-EfN were applied in this experiment. 233 G418-resistant colonies were obtained and of these, 1 was resistant to 6-TG selection suggesting that the *hprt* gene was targeted (table 4.1). These results show that even with decreased transfection efficiency ($2.3x10^{-4}$ versus $4.9x10^{-4}$ in the first targeting experiment), a 6-TG-resistant colony indicative of gene targeting at the *hprt* gene, can be obtained. This clone was expanded for PCR and Southern analysis to confirm a targeting event.

pHprt-EfN (μg)	cell line/ Passage No	Electroporation conditions	Number of cells	$G418^r$	$G418^r$ $6\text{-}TG^r$	E.T.F (%)	A.T.F	Transfection efficiency
	H1/73	(1) 320V/200 μF	$1x10^6$	600	1	0.16	10^{-6}	$6.0x10^{-4}$
50	H1/73	(2) 300 V/100 μs	$1x10^6$	488	1	0.20	10^{-6}	$4.9x10^{-4}$
	H1/73	(3) 200V/950 μF	$1x10^6$	380	0	0	0	$3.8x10^{-4}$
50	H1/76	(2) 300 V/100 μs	$1x10^6$	233	1	0.43	10^{-6}	$2.3x10^{-4}$

Table 4.1. *Hprt* targeting experiments with pHprt-EfN in H1 cells. E.T.F (Effective targeting frequency): ratio between number of homologous recombinants and number of integrants surviving selection (homologous and non-homologous). A.T.F (Absolute targeting frequency): ratio between number of homologous recombinants and number of cells electroporated. Transfection efficiency: ratio between number of integrants and number of cells electroporated.

4.3.4. Confirmation of a Targeting Event by PCR and Southern Analysis

To check that the G418/6-TG-resistant colonies were the result of gene targeting events at the *hprt* gene and not spontaneous mutation, the 3 colonies, obtained from the transfection of H1 cells with pHprt-EfN were expanded and characterised by PCR and Southern Blot analyses. The *hprt* gene inactivated by the targeted replacement of endogenous sequence with the *EF1α-LoxP-neo-Lox2272* cassette, produces a 6.7 kb PCR fragment with the two primers shown in figure 4.3.A. One primer anneals to the *hprt* sequence upstream of the 5' arm of homology included in the targeting vector (external primer) and a second primer anneals to the *neo* sequence (internal primer) (figure 4.3 A). gDNA was extracted from each cell line and used as the template in a PCR reaction with the 5' ext *hprt* and *neo* primers. The parental H1 cell line was used as a negative control in the PCR analysis. All the clones analysed produced the predicted 6.7 kb fragment expected for a gene targeting event (figure 4.3 B & C).

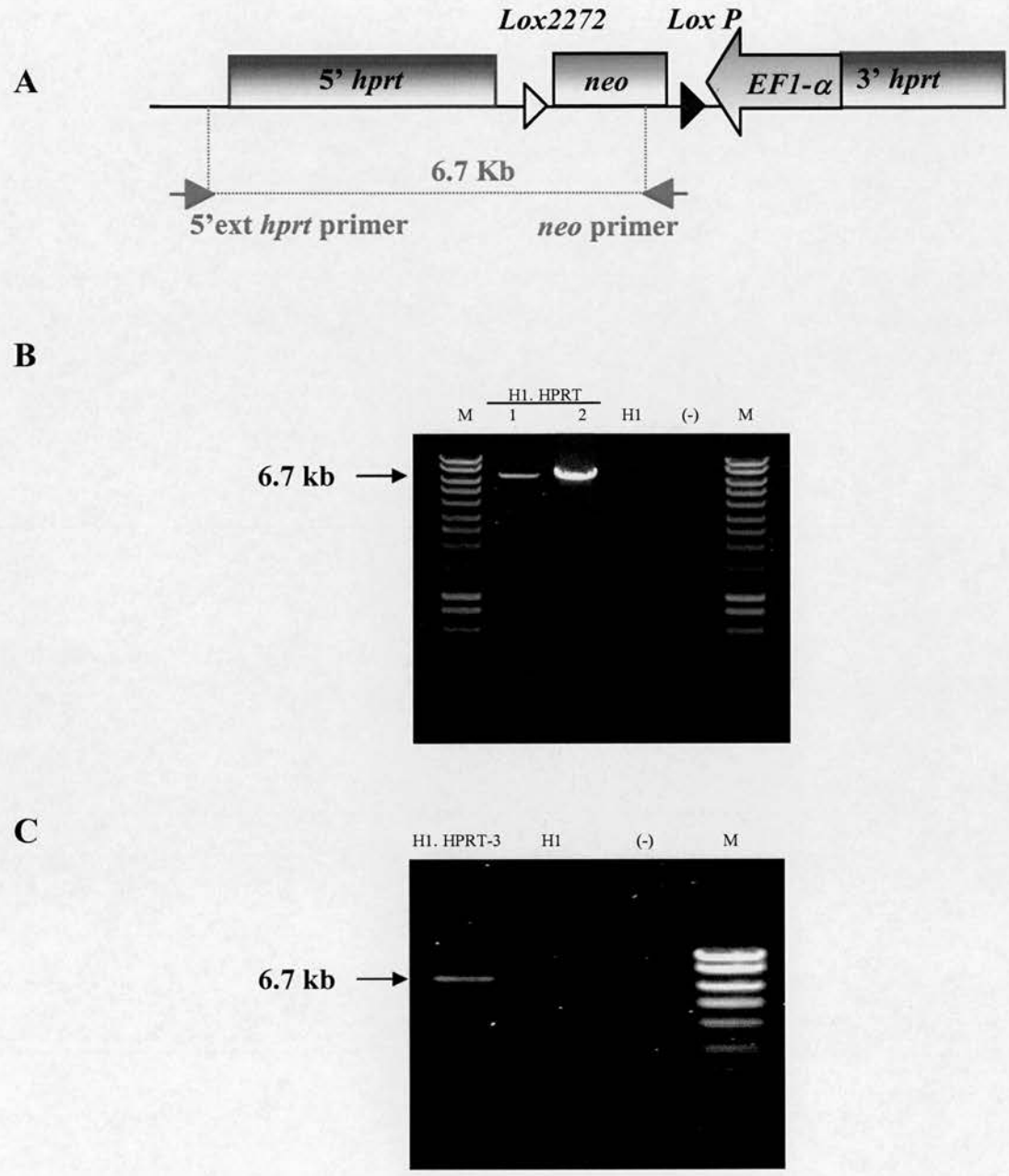


Figure 4.3. PCR screening for a targeting event at the *hprt* gene of the three G418/ 6-TG-resistant clones. **(A)** Map of the partial structure of the targeted *hprt* gene showing the predicted size of a PCR product generated with a reverse primer within the selectable marker (*neo* primer) and a forward primer 5' to the homology (5'ext *hprt* primer). **(B) (C)** PCR analysis of the three G418/6-TG-resistant clones (H1.HPRT-1, H1.HPRT-2 and H1.HPRT-3). Presence of a 6.7 kb band indicates a targeting event at the *hprt* gene. (-): no DNA template. M: 5 µl Hyperladder (Bioline).

The PCR results were confirmed by Southern analysis of BamHI- and EcoRI-digested gDNA from each clone. The probe (gift from David Russell) used for the screen consisted of a 1193 bp of *hprt* sequence internal to the 3' arm of homology included in the targeting construct. The wild-type allele is predicted to produce a 8.4 kb EcoRI and 20 kb BamHI fragment (figure 4.4 A). Because the *EF1α-loxP-neo-lox2272* cassette contains EcoRI and BamHI sites, its presence introduces new restriction sites in exon 3 of the *hprt* gene, resulting in the production of 6.9 kb EcoRI and 6.7 kb BamHI fragments (figure 4.4 A). As predicted from the restriction map of the modified *hprt* gene, Southern blot analysis shows the presence of 6.7 kb EcoRI and 6.9 kb BamHI fragments in the three clones recovered, confirming that they were the result of gene targeting by HR (figure 4.4 B). Random integration of the targeting vector should be detected as additional band(s) hybridising to the internal *hprt* probe. There is however always a possibility that random integration of the targeting vector may generate a band similar in size to that of the targeted integration. In this analysis, two digests were examined and in each case resulted in the detection of a single band. The probability that random integration of the targeting vector has produced a band similar in size to the targeted integration with both digests is therefore extremely low.

Together, these results show that transfection efficiencies of up to $6x10^{-4}$ and gene targeting frequencies of up to 0.43 % can be obtained at the *hprt* gene in H1 cells with the replacement vector pHprt-EfN using two different electroporation protocols.

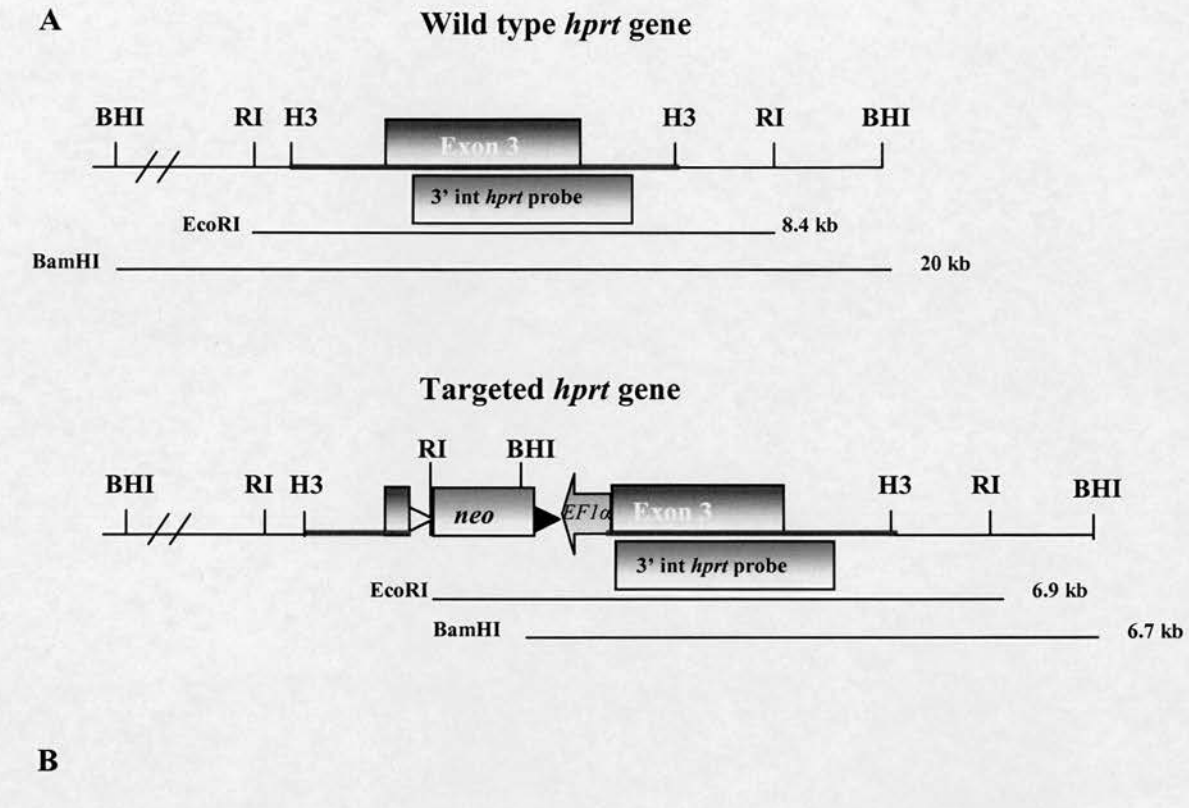


Figure 4.4. Southern screening for a targeting event at the *hprt* gene of the three G418/6-TG-resistant clones. **(A)** Map of the relevant segment of the wild-type and targeted *hprt* gene showing the predicted restriction fragments with an internal *hprt* probe. The *hprt* homology is represented as a thin blue line. The exon 3 in which the transgene has integrated is indicated as a blue box. The position of the *hprt* probe is shown as a red box with the size of the EcoRI and BamHI fragments indicated below. RI: EcoRI, BHI:BamHI. **(B)** Southern analysis of the 3 G418/6-TG-resistant clones (H1.HPRT-1, -2 and -3) obtained in the 2 targeting experiments with pHprt-EfN. The presence of a 6.9 kb EcoRI and 6.7 kb BamHI fragment indicates a targeting event at the *hprt* gene.

4.4. Discussion

We sought to engineer, by gene targeting, the *hprt* gene in H1 cells by the introduction of a *neo* marker driven by the *EF1-α* promoter and flanked by heterospecific *lox* sites such that RMCE could be subsequently applied. We generated three *hprt*-targeted cell lines using two different electroporation protocols. The first, established in our lab for H9 cells, used the Eppendorf Multiporator system, which produced transfection efficiencies of up to $4.9x10^{-4}$ and targeting frequencies of up to 0.43 % in our targeting experiments. The second protocol, established by Zwaka and Thomson (2003) for hES cells, used the BioRad Gene Pulser II, which gave a transfection efficiency of $6x10^{-4}$ and a targeting frequency of 0.16 % in our targeting experiment.

Using the electroporation conditions established by Zwaka and Thomson (320 V / 200 μF), we obtained a transfection efficiency 26-fold greater ($6x10^{-4}$) than that reported by Zwaka and Thomson ($2.3x10^{-5}$). Our transfection efficiencies were up to 600-fold greater than those obtained by Urbach *et al.*, (2004) (10^{-6}). However Urbach *et al.* used a different male cell line (H13) and a cationic reagent (ExGen 500) to transfect their cells both of which could have contributed to a low transfection efficiency. The difference in transfection efficiencies between our work and Zwaka and Thomson may reflect the use of a different culture system. They use a collagenase passaging regime in their culture system, which permits the disaggregation of hES cells into clumps of cells (Thomson *et al.*, 1998). In contrast, our culture system involves the use of trypsin-EGTA (TEG) which allows the disaggregation of hES cells into a single cell suspension. Because Zwaka and

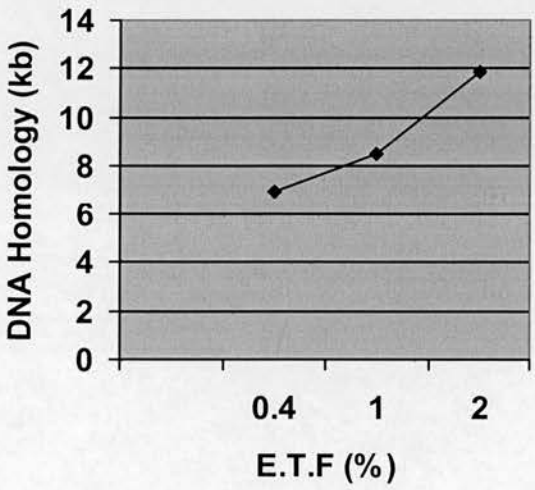
Thomson electroporated their cells as clumps, it is possible that the uptake of DNA has occurred in several cells in the clumps (i.e. multiple G418-resistant colonies may have occurred within one clump) and as a result, fewer G418-resistant colonies were detected. Alternatively, the difference of transfection efficiencies in both studies could have been the result of a difference of cell survival following electroporation due to experimental variations (e.g. passaging regime, number of cells electroporated, and medium composition). It is also possible that the transfection efficiency obtained in our work was much greater than Zwaka and Thomson because we used a targeting vector with a stronger promoter (*EF1-α*) to drive the expression of *neo*. Zwaka and Thomson used a targeting vector with the synthetic mutant polyoma enhanced HSV*tk* promoter (Thomas and Capecchi, 1987) driving *neo*, which has been shown, in our hands, to work poorly in H9 cells (H. Priddle, unpublished data).

Effective targeting frequencies (ratio between homologous recombinants and total (homologous and non-homologous) integrants surviving selection) at the *hprt* gene have been reported to reach 2 % using an electroporation approach in the male H1.1 subline (Zwaka and Thomson, 2003) and 1 % using a cationic reagent in the male H13 cell line (Urbach *et al.*, 2004). Our results showed that effective targeting frequencies of up to 0.4 % could be achieved by electroporation at the *hprt* gene in the H1 cell line (table 4.1). Several experimental differences between these studies could have accounted for this discrepancy. A major difference between our targeting experiment and the data presented by Zwaka and Thomson, or Urbach *et al.*, is that a different hES cell line was used in each study. Zwaka and Thomson used a clonal derivative of the H1 cell line (subclone H1.1). Thus the cells used in their study

consisted of homogenous population which may have been more permissive to gene targeting. In contrast, the parental H1 cell line was used in our targeting experiments, and since this cell line was not originally cloned from a single cell, the possibility that there is some variation among the undifferentiated cells in spite of their homogeneous appearance, cannot be ruled out (Thomson *et al.*, 1998). As a consequence, not all the H1 cells may have been permissive to the same extent to gene targeting. The difference of targeting frequencies between our work and Urbach's *et al.*, could also be the result of cell line difference. However, because there were so many experimental variations (e.g transfection conditions, *hprt* homologies, selectable markers, and *hprt* mutations), between the two studies, a single cause cannot be attributed. With only two published reports of gene targeting in two hES cell lines, it is presently unclear whether similar targeting frequencies will be obtained for the same gene between different hES cell lines.

The most likely explanation for the difference of targeting frequencies between these three studies is that both groups used an *hprt* targeting vector with a larger homology than we used to target the *hprt* gene. We used 6.9 kb of homology, whereas 11.9 kb and 8.5 kb of homology was used by Zwaka, Thomson, and Urbach *et al.*, respectively. Several studies have shown that targeting frequency in mammalian cells is dependent on homology length to some extent (Thomas and Capecchi, 1987; Shulman *et al.*, 1990; Hasty *et al.,* 1991a; Deng and Capecchi, 1992). Thomas and Capecchi showed in the first *hprt* targeting study conducted in mouse ES cells, a positive correlation between targeting frequencies and the length of *hprt* homology in the replacement targeting vector employed (Thomas and Capecchi, 1987). They observed a 16-fold increase of targeting frequencies with

only a 2.25-fold increase in homology. Similarly, Zwaka and Thomson (2003) achieved a 1.5-fold increase in targeting frequency at the *oct-4* gene in hES cells, when the total *oct-4* homology in the targeting vector was increased from 7.9 kb to 12.8 kb. A positive relationship between targeting frequencies and the length of *hprt* homology in the targeting vector was also observed in hES cells between our work and the work conducted by Zwaka, Thomson, and Urbach *et al.* (figure 4.5). However, the experimental conditions (i.e. cell line, transfection condition, and *hprt* mutation) were different between studies and therefore, it can only be suggested that an increase of homology in the targeting vector results in an increase of targeting frequency at the *hprt* locus in hES cells.



Group	Homology length (kb)	E.T.F (%)
Zwaka and Thomson, 2003	11.9	2
Urbach *et al.*, 2004	8.5	1
Di Domenico *et al.*	6.9	0.4

Figure 4.5. Gene targeting frequencies in hES cells at the *hprt* gene. Zwaka and Thomson deleted regions of the last 3 exons of the *hprt* gene in the H1.1 subline using 40 μg of a targeting construct containing 11.9 kb of isogenic *hprt* homology (by electroporation), resulting in an effective targeting frequency (E.T.F) of 2 %. Urbach *et al.*, introduced a mutation in exon 2 in the H13 cell line using 3-4 μg of a targeting construct with 8.5 kb of non-isogenic *hprt* homology (with a cationic reagent), resulting in an E.T.F of 1 %. We introduced a transgene in exon 3 in the H1 cell line using 50 μg of a targeting construct containing 6.9 kb of non-isogenic *hprt* homology (by electroporation,), resulting in an E.T.F of 0.4 %.

The use of isogenic DNA in Zwaka and Thomson's work could also have resulted in an increase of targeting frequency. Several lines of evidence indicate that the frequency of HR is strongly affected by the presence of vector-target mismatches in mouse ES cells. Te Riele *et al.* (1992) compared the frequencies of HR at the *retinoblastoma* (*Rb*) susceptibility gene in a mouse 129-derived ES cell line with isogenic (129-derived) and non-isogenic (BALB/c-derived) targeting constructs, and observed a 20- to 50-fold increase in targeting frequency with the isogenic DNA construct. The use of isogenic DNA was so efficient in recovering homologous recombinants (1 in 3 resistant colonies) that an enrichment protocol was not necessary. Similarly, Van Deursen and Wieringa (1992) showed that the mouse *creatine kinase M* (*ckm*) gene, characterised by a 2% divergence for the 129 and BALB/c alleles, can be targeted 25-fold more efficiently with isogenic DNA (129-derived) than non-isogenic DNA (BALB-c-derived). In contrast, other studies have shown high targeting frequencies without the use of isogenic DNA (Zijlstra *et al.,* 1989; Te Riele *et al.,* 1990) showing that parameters other than isogenecity greatly influence targeting frequency, or that some genes are more similar between mouse strains than others. Unlike mouse ES cells, it has been suggested that isogenicity is not important in human gene targeting (Sedivy *et al.*, 1999). Sedivy *et al.* have compiled gene targeting data in human cells including one study that featured the targeting of the *P53* gene in four different human cell lines using the same targeting vector. The homologous sequence used in this vector was amplified from the LL1 human embryonic skin fibroblast cell line. The targeting frequencies obtained with this vector in the LL1, LF1 (human embryonic lung fibroblast), TK6 (human lymphoblast) and HCT116 (human colorectal adenocarcinoma) cell lines were 1.1 %,

12 %, 0.8 % and 0.2 %, respectively. Thus targeting frequency was increased 11-fold in the LF1 cell line compared to the LL1 cell line (to which the vector was actually isogenic), suggesting that isogenicity was not a major issue in gene targeting frequencies in human cells (Sedivy *et al.*, 1999). Recently, Urbach *et al.* (2004) have also demonstrated that the *hprt* gene can be targeted with non-isogenic DNA in the H13 cell line at a similar frequency to that reported by Zwaka and Thomson in 2003 (figure 4.7), supporting the claims of Sedivy and collaborators. One possible explanation for the differences between human and mice, in respect to the efficiency of non-isogenic homologous recombination, is that the sequence divergence between two hES cell lines is much smaller than the difference between two mouse strains. Single-nucleotide polymorphisms (SNPs) in the human genome occur at a frequency of one per 500 to 1000 bp (Wang *et al.*, 1998; Sachidanandam, 2001; Zhao *et al.*, 2003). In contrast, SNPs are much more frequent in the sequences of different mice strains. For example, in the *ckm* gene, as many as 17 SNPs occur in a 150 bp stretch of sequence between the 129 and BALB/C alleles (Van Deursen and Wieringa, 1992). According to the National Centre for Biotechnology information (NCBI) database, the *hprt* sequence used in the pHprt-EfN targeting construct contains only 10 SNPs distributed along the 6.9 kb sequence (~ 1 SNP every 700 bp). Based on this information, it is less likely that the use of non-isogenic DNA in the context of our experiments might have had a negative effect on targeting frequency.

To conclude, we have generated three H1-derived cell lines that carry a positive selectable marker gene flanked by heterospecific *lox* sites at the *hprt* gene,

providing an ideal basis for the integration of any "floxed" marker genes by site-specific recombination.

CHAPTER V

Characterisation of an H1-Derived *Hprt*-Targeted Cell Line

5.1. Introduction

Like their murine counterparts, hES cells are characterised by a stable diploid karyotype and are capable of both self-renewal and differentiation into derivatives of the three germ layers (Thomson *et al.*, 1998; Schuldiner *et al.*, 2000; Reubinoff *et al.*, 2000; Itskovitz-Eldor *et al.*, 2000; Zeng *et al.,* 2004). To fulfil their potential in cell-based therapies, it is necessary that genetically-engineered hES cells retain these properties after manipulation. There are currently two reports of gene targeting in hES cells (Zwaka and Thomson, 2003; Urbach *et al.*, 2004) but the ability of gene targeted hES cells to retain a stable karyotype and to give rise to the three germ layers has not been investigated.

This chapter examines whether a clonally-derived H1.HPRT targeted cell line retains (i) euploidy, (ii) the expression of stem cell markers, and (iii) multipotency. A comparative study between the parental H1 hES cell line and the H1.HPRT-1 targeted cell line was conducted to address these questions.

5.2. Objectives

To compare ploidy, the expression of stem cell markers and multipotency between the H1 parental cell line and the H1.HPRT-1 targeted cell line.

5.3. Results

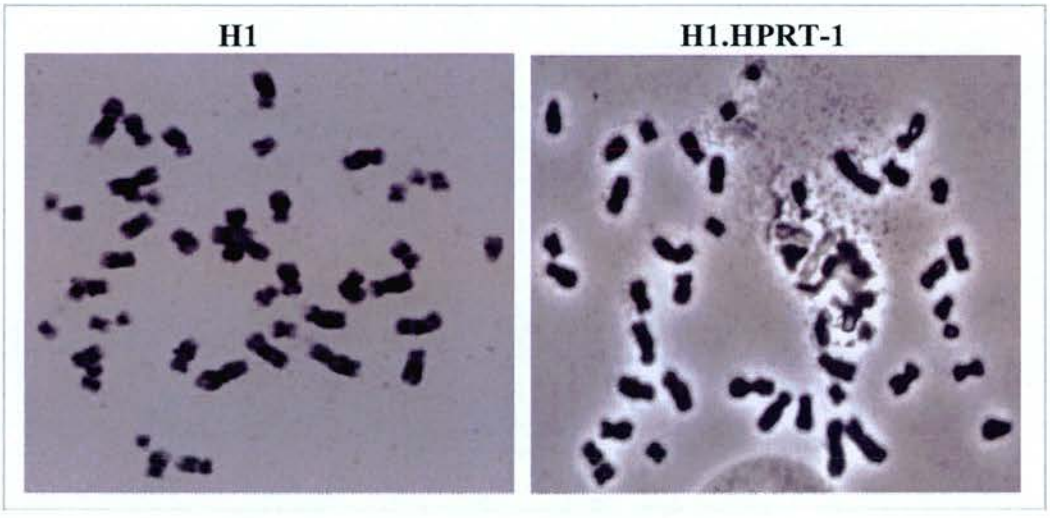
5.3.1. Karyotypic Analysis

There is increasing evidence that hES cells propagated with enzymatic methods (e.g. trypsin, collagenase) accumulate karyotypic abnormalities such as translocations, deletions and duplications during long-term culture (Inzunza *et al.*, 2004; Draper *et al.*, 2004; Rosler *et al.*, 2004; Mitalipova *et al.*, 2005; Maitra *et al.*, 2005; McWhir *et al.*, unpublished data). The occurrence of chromosomal abnormalities in hES cell lines destined for therapies is clearly a concern, given that *in vivo* karyotypic changes are often associated with malignancy (reviewed by Albertson *et al.,* 2003). For the potential of hES cell-based therapy to be realised, it is essential that these cells, including genetically modified cells, are shown to be safe. This will require that they be karyotypically stable over long-term culture.

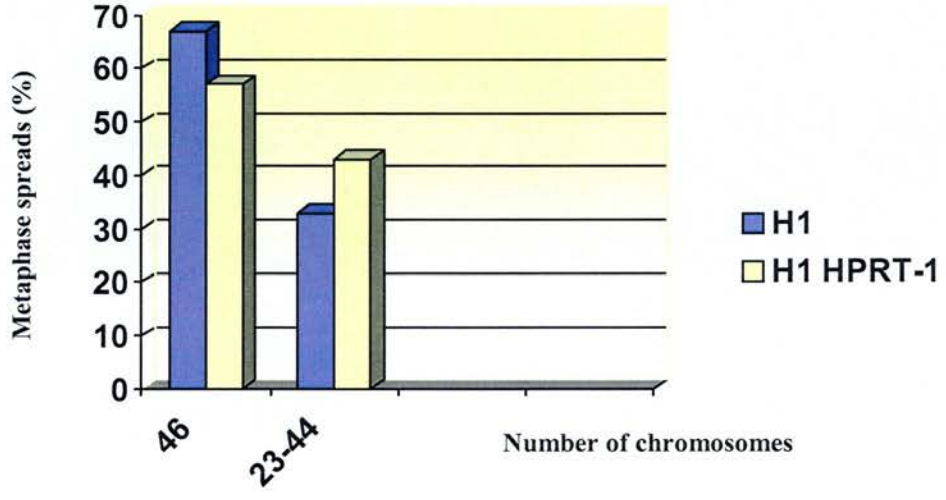
To investigate whether H1 cells retain a diploid karyotype after a gene targeting event and after propagation in culture, thirty chromosome spreads were analysed from both the targeted H1.HPRT-1 cell line (passage 85) and the parental

H1 cell line (passage 74), from which these cells were derived at the time of the gene targeting experiment. Metaphase spread analysis showed that 67 % and 57 % of the H1 and H1.HPRT-1 spreads, respectively, maintained a diploid male 46 XY karyotype (figure 5.1 B). 33 % and 43 % of the H1 and H1.HPRT-1 spreads, respectively, varied between 23 and 44 chromosomes (figure 5.1.B). The difference between the two cell lines was however not statistically significant ($p=0.6$, $\chi_1^2 = 0.28$). Thus these data show that H1 cells that have been subjected to gene targeting procedures and that have been propagated in culture for 11 passages, retain a predominantly diploid karyotype similar to the parental cells. Representative 46 XY chromosome spreads from H1 and H1.HPRT-1 cell lines is shown in figure 5.1 A.

A



B



Cell line	H1	H1.HPRT-1
Chromosomes numbers (% of metaphase spreads)	46 (67%) 23-44 (33%)	46 (57%) 23-44 (43%)

Figure 5.1. Analysis of 30 metaphase spreads of the parental H1 and the H1.HPRT-1 cell lines. **(A)** High power magnification of a male diploid 46 XY karoytype from H1 and H1 HPRT-1 cell lines. **(B)** Histogram representation of chromosome number for 30 metaphase spreads from H1 and H1.HPRT-1 cell lines.

5.3.2. Analysis of Stem Cell Markers

Two properties which characterise cells as embryonic stem cells are their ability to self-renew and their capacity to remain in an undifferentiated state until exposed to differentiation factors. Unless these criteria are fulfilled, cells cannot be classified as ES cells. An essential indication of the undifferentiated state of the cells is the expression of the transcription factor POU5F1 (also known as OCT-4), which is associated with pluripotency (Nichols *et al.*, 1998; Pesce *et al.*, 1999). Undifferentiated hES cells are also characterised by the expression of the cell-surface Stage Specific Embryonic Antigens SSEA-4 and SSEA-3 and the keratan sulphate-associated antigens TRA-1-81 and TRA-1-60 (Thomson *et al.*, 1998; Amit *et al.*, 2000; Xu *et al.*, 2001; Draper *et al.*, 2002; Zeng *et al.*, 2004). Upon differentiation, hES cells lose expression of all these markers. The undifferentiated state of the H1 cell line has already been confirmed by us and other groups by examination of the expression of these markers (Thomson *et al.,* 1998; Amit *et al.*, 2000; Xu *et al.*, 2001; Draper *et al.*, 2002; McWhir *et al.,* 2005, unpublished data). However, the expression of these markers has not been investigated in H1-derived targeted cell lines. We therefore sought to investigate whether the H1.HPRT-1 targeted cell line maintains the expression of stem cell markers.

5.3.2.1. *Oct-4* Expression Analysis

The POU-domain transcription factor OCT-4 is highly expressed in ES cells and has been shown to be essential for maintaining the undifferentiated state of ES cells (for a review see Pesce *et al.*, 1999). To assess the expression of *Oct-4* in both the parental H1 cell line and the H1.HPRT-1 targeted cell line, immunofluorescence microscopy directed at OCT-4 was performed (staining

performed by J. Fletcher, Roslin Institute). A clone derived from H1.HPRT-1 cells, H1.HPRT-1RMCE3, was used as a negative control as it exhibited a differentiated morphology and was characterised by a very slow proliferation rate (chapter 7, figure 7.8). Results showed that most H1 and H1.HPRT-1 cells expressed the OCT-4 protein in their nuclei (green staining) consistent with maintenance of an undifferentiated state (figure 5.2). A minority of cells did not detectably express the protein suggesting the presence of differentiated cells among the undifferentiated population. As predicted, *oct-4* expression was down regulated in the H1.HPRT-1.RMCE-3 cell line as shown by the lack of green staining in the merged picture. We conclude that the expression of the POU-domain transcription factor is maintained in the H1.HPRT-1 targeted cell line.

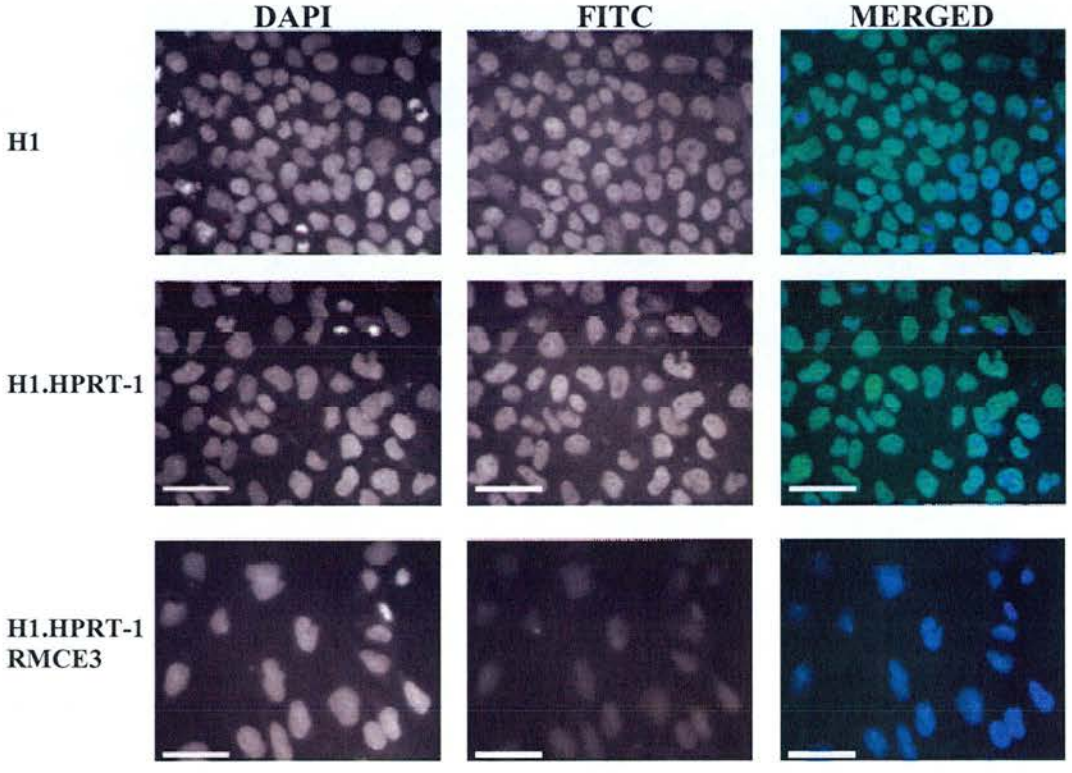


Figure 5.2. Oct-4 immunocytochemistry of H1 and H1.HPRT-1 cell lines. *Oct-4* immnuohistochemistry analysis showing positive expression of the *Oct-4* gene in the H1 and H1.HPRT-1 cell lines (middle panels). Dapi (4'-6-Diamidino-2-phenylindole), which is a DNA intercalating dye, can be observed in all nuclei (left panels). A merged picture of FITC (green) and Dapi (blue) channels is presented on the right panels. Scale bars = 50 μm.

5.3.2.2. Cell Surface Expression Analysis

Expression of the embryonic cell surface markers SSEA-4, TRA-1-60 and TRA-1-81 were analysed by flow cytometry in both the parental H1 and H1.HPRT-1 targeted cell line. The expression of SSEA-1, which is downregulated in ES cells and upregulated upon differentiation (Andrews *et al.*, 1996; Thomson and Marshall 1998; Thomson *et al*, 1998), was also quantified in both cell lines. Expression of a tetraspanin transmembrane protein, CD9, which is expressed in both mouse (Oka *et al.*, 2002) and hES cells (Carpenter *et al.*, 2004) was also evaluated. The H1 and H1.HPRT-1 cells were collected from a T25 flask and $5x10^5$ cells were stained with the appropriate antibodies. Isotype-matched controls (IgG3-FITC (SSEA-4), IgM-PE (SSEA-1, Tra-1-60 & Tra-1-81) and IgG1-PE (CD9)) were included to calibrate the instrument. In this analysis, only the live cells were taken into account. A representative density plot showing the gate (R1) which includes only the live cells is illustrated in figure 5.3 B. Flow cytometry showed, over 3 independent experiments, that more than 98% of the cells for both cell lines expressed the embryonic cell surface markers with no statistically-significant difference between cell lines ($P>0.05$, two-tailed t-test) confirming the undifferentiated phenotype for the majority of the cells (figure 5.3 A & C). Flow cytometry also showed that 5.5% (H1.HPRT-1) and 12 % (H1) of the cells examined expressed SSEA-1, consistent with the presence of a small proportion of differentiated cells among a largely undifferentiated population (figure 5.3 A & C). The proportion of SSEA-1 positive cells was not significantly different between the two cell lines ($P>0.05$, two-tailed t-test). These results show that the H1.HPRT-1 cell line, which was subjected to gene targeting procedures, expressed markers of an

undifferentiated ES phenotype (SSEA-4, TRA-1-60 and TRA-1-81) and differentiation phenotype (SSEA-1) at similar levels to the parental H1 cell line.

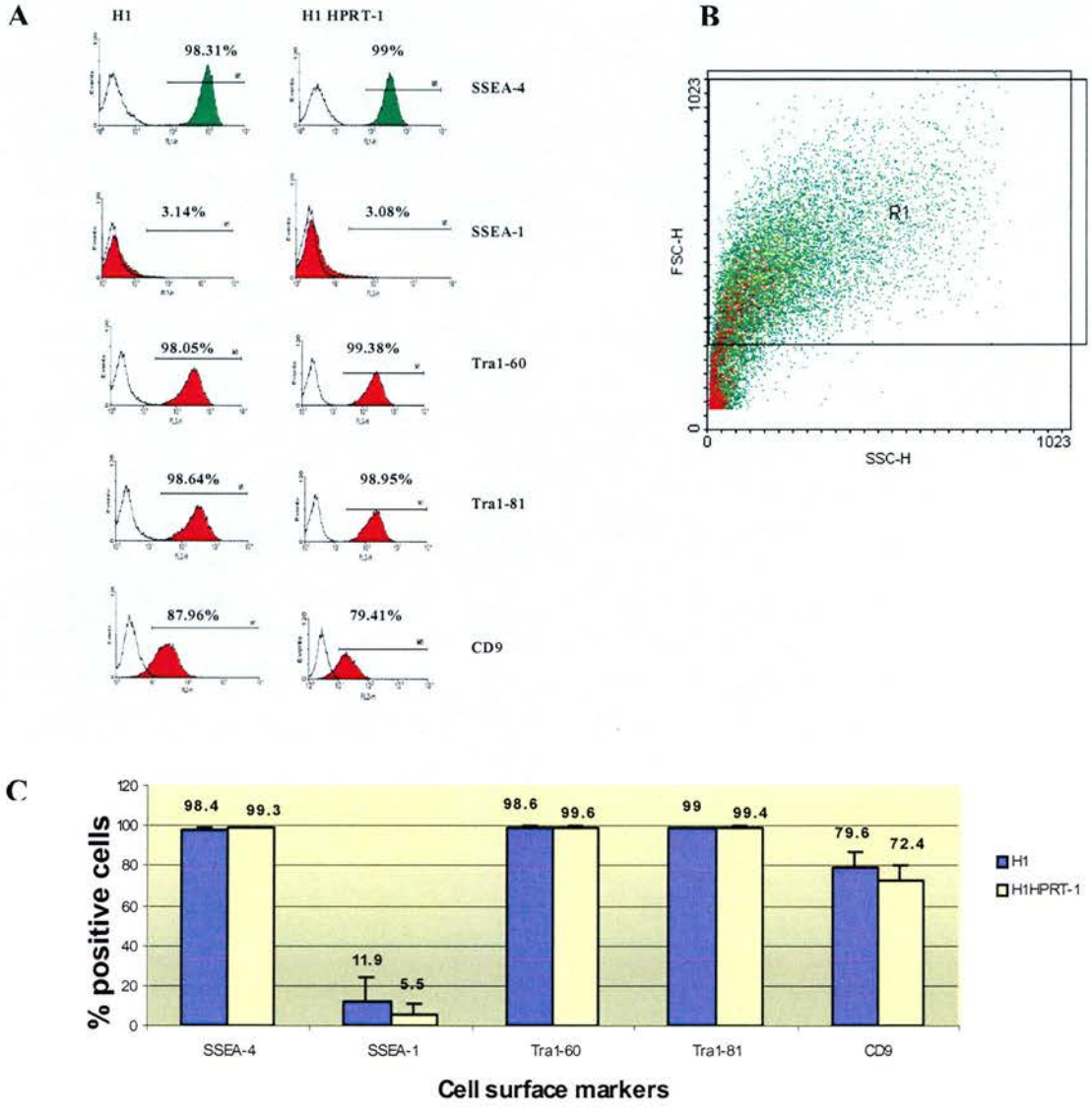


Figure 5.3. Flow cytometry analysis of cell surface marker expression for both the H1 and H1.HPRT-1 cell lines. **(A)** Representative expression analysis of SSEA-4, TRA-1-60, TRA-1-81, SSEA-1 and CD9 markers demonstrating that targeted H1.HPRT-1 cells are positive for the ES-specific markers SSEA-4, Tra-1-60 and Tra-1-81 and largely negative for SSEA-1 expression. The high-fluorescent intensity peak (FL1 or FL2 channel) represents positive cells, while the low-intensity peak represents background levels due to autofluorescence. Isotype controls: IgG3-FITC (SSEA-4), IgM-PE (SSEA-1, Tra-1-60 & Tra-1-81) and IgG1-PE (CD9). Data were generated from the analysis of at least 10,000 viable cells. **(B)** Representative density plot showing the region R1 containing the live cells analysed in all fluorescence plots. **(C)** Histogram statistics of cell surface marker expression in H1 and H1.HPRT-1 cells. Data are means ± standard deviation for three independent experiments.

5.3.3. Multipotency Analysis

5.3.3.1. Analysis of *In Vitro* Differentiation

Like their murine counterparts, hES cells spontaneously form three dimensional aggregates of cells known as embryoid bodies (EBs) when grown in suspension (Itskovitz-Eldor *et al.*, 2000). After 4 to 5 days in suspension in the absence of bFGF (required to keep the cells in an undifferentiated state), EBs may be plated onto adherent culture surfaces to form a variety of ectodermal, endodermal and mesodermal derivatives such as hematopoietic (Schuldiner *et al.*, 2000), endothelial (Levenberg *et al.*, 2002), cardiac muscle (Itskovitz-Eldor *et al.*, 2000; Kehat *et al.*, 2001, 2003; Xu *et al.*, 2002) and neuronal cell lineages (Reubinoff *et al.*, 2000; Itskovitz-Eldor *et al.*, 2000; Schuldiner *et al.*, 2001; Carpenter *et al.*, 2001). Both the female H9 and the male H1 cell lines have been shown to exhibit these properties. Currently, there are no published reports regarding the potential of a targeted hES cell line to differentiate into the 3 germ layers *in vitro*. This question was therefore addressed with the H1.HPRT-1 targeted cell line. Both the parental H1 and H1.HPRT-1 cell lines were subjected to *in vitro* differentiation and immunocytochemical analysis was performed to evaluate the expression of 3 lineage-specific markers: β-III-tubulin, α-feto protein (AFP) and troponin-T, representative of the ectoderm, endoderm and mesoderm respectively. α-β-Tubulin is an integral component of microtubules which are involved in several cytoskeletal functions (Oakley, 1992). The isotype β-III-tubulin, is found in the brain and dorsal root ganglia and appears to be localised to neurons of the central and peripheral nervous system, where its expression seems to increase during axonal outgrowth. AFP is a single-chain glycoprotein that is normally expressed in the fetal liver,

gastrointestinal tract and yolk sac (Abelev, 1971). In early fetal life it is one of the main plasma proteins. Anti-AFP was used to visualise endodermal cells in the derivatives of H1 cells. Troponin T is a microfilament protein closely associated with actin and its microfilament accessory proteins that appear in skeletal muscle myofibrils (Lin *et al.,* 1984). Anti-troponin T was used to detect skeletal muscle (mesoderm) in the differentiation cultures of H1 and H1.HPRT-1 cell lines.

H9, H1 and H1.HPRT-1 cells were disaggregated from a T75 flask with TEG into single cell suspensions. The H9 cell line was used as a positive control in this study as it had already been shown, in our hands, to express these lineage markers following *in vitro* differentiation. After 2 days of culture in suspension in conditioned medium supplemented with bFGF, the EBs were transferred to EB medium which contains serum but lacks bFGF. After 7 days in suspension, the H9 cells had grown into robust spherical EB-like structures (figure 5.4). In contrast, H1 and H1.HPRT-1 cells rarely aggregated and the low number of EBs generated (~5-fold fewer than H9) were very small and exhibited a less spherical shape compared with H9 cells (figure 5.4).

To determine the *in vitro* differentiation potential of each cell line, 7-day old EBs were plated intact onto gelatin-coated chamber slides (~5 to 10 EBs per well) and cultured for 3 weeks in EB medium before immunohistochemical analysis. The plating efficiency of the H1- and H1.HPRT-1-derived EBs was very low relative to H9-derived EBs. Most H9-derived EBs adhered to gelatin very efficiently. In contrast, most H1- and H1.HPRT-1-derived EBs did not adhere to the gelatin substrate in the chambers, resulting in a considerable loss of EBs following medium replacement 24 hours later. After 21 days of differentiation, several cell

types derived from H9 EBs could be observed, including areas of beating cardiomyocytes (figure 5.4). In contrast the few remaining H1- and H1.HPRT-1-derived small EBs mostly differentiated into cells with an epithelial-like morphology (figure 5.4).

Immunocytochemical staining for the expression of the lineage specific markers β-III-tubulin (ectoderm), AFP protein (endoderm) and troponin-T (mesoderm) revealed derivatives of the three germ layers in H9-derived EBs (figure 5.5). Unlike H9-derived cells, AFP and troponin-T expression could not be detected in H1- and H1.HPRT-1-derived cells indicating that these cell lines did not differentiate into these specific endodermal and mesodermal lineages. All the cells present in the H1-derived and H1.HPRT-1-derived cultures stained positively for β-III-tubulin. A staining control in which the primary antibody was omitted did not result in β-III tubulin staining indicating that the binding of the secondary antibody was specific (data not shown). Although these results first suggested that the H1-derived and H1.HPRT-1-derived cultures were capable of differentiating towards a neuronal lineage, neither cell line exhibited a neuronal cell type morphology (as determined by neuronal axon formation), and in this respect differed from the H9-derived cells. Additionally, β-III tubulin expression was also found in both H1 and H1.HPRT-1 undifferentiated cultures. Although β-III tubulin is considered a specific neuronal marker, its expression has been previously reported in undifferentiated hES cells (Zeng *et al.*, 2004). Taken together, these results indicate a lack of neuronal differentiation in both the H1 and H1.HPRT1 cell lines (figure 5.5).

To conclude, the H1 and H1.HPRT-1 cell lines failed to express AFP, troponin-T and neuronal β-III-tubulin markers. These results however, do not

necessarily reflect the lack of multipotentiality of these two cell lines but possibly the poor efficiency of robust EB generation.

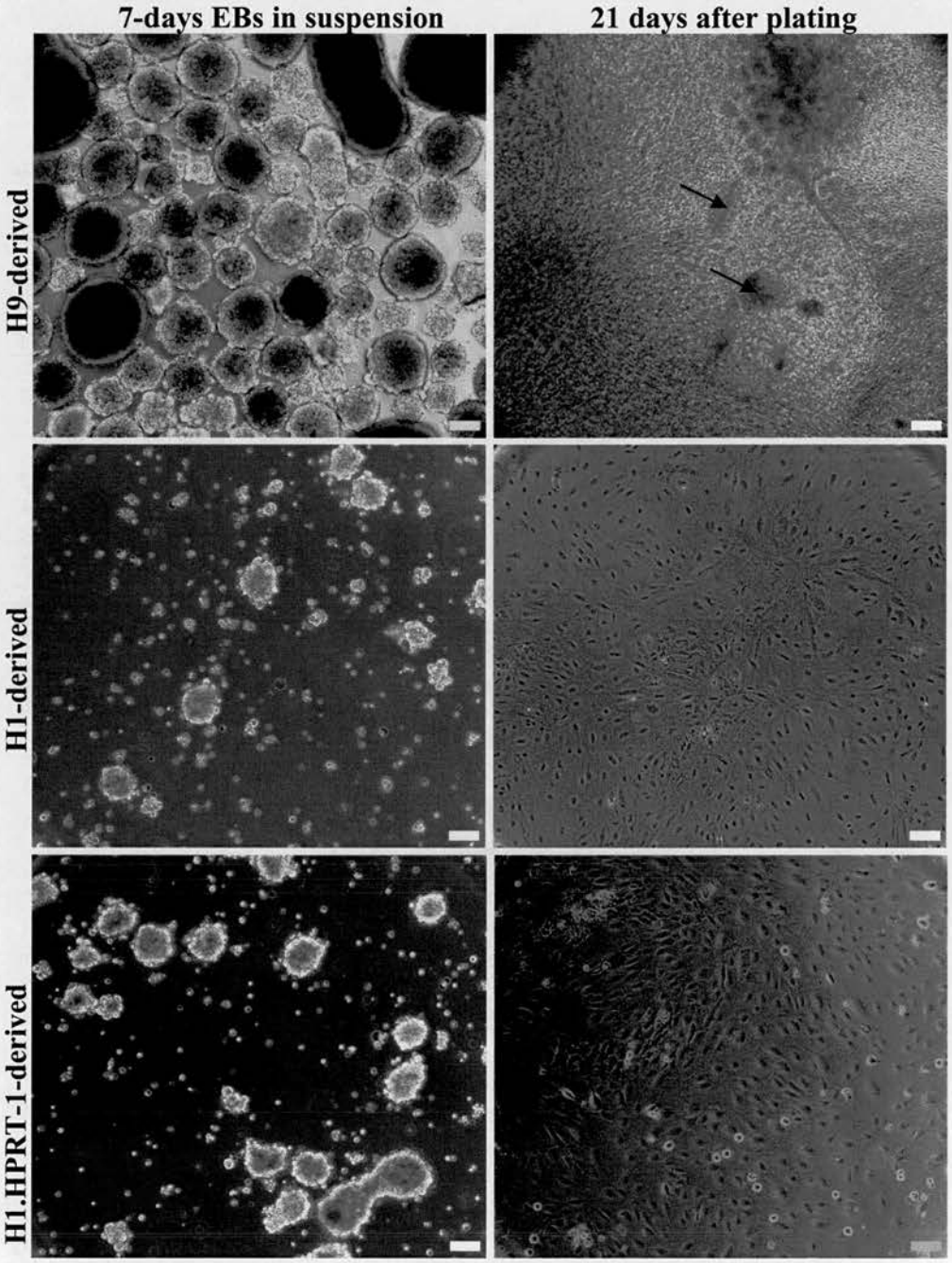


Figure 5.4. *In vitro* differentiation of H9, H1 and H1.HPRT-1 cell lines. Microscopic analysis of H9, H1 and H1.HPRT-1 derived EBs after 7 days in suspension culture (left panels) and 21 days after plating onto gelatin-coated chamber slides (right panels). Black arrows indicate areas of beating cardiomyocytes. Scale bars = 100 μm.

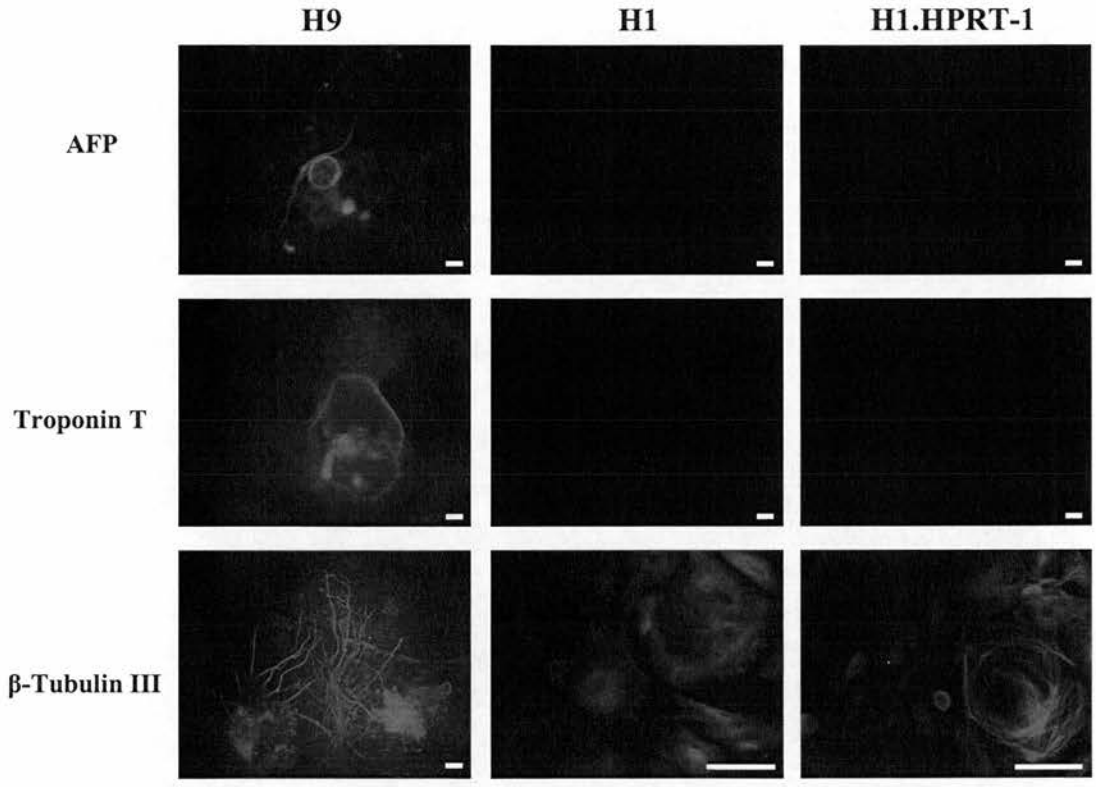


Figure 5.5. *In vitro* differentiation analysis. Immunohistochemical staining with anti-AFP (endoderm) and Troponin-T (mesoderm) expression markers shows positive staining for these 2 germ layers in H9-derived cells but not in H1 and H1.HPRT-1-derived cells. Unlike the H9-derived culture, neuronal axon staining with β-Tubulin III (ectoderm) is absent in the H1-derived and H1.HPRT-1-derived cultures indicating a lack of neuronal differentiation for these two cell lines. Scale bars = 100 μm.

5.3.3.2. Osteogenic potential of the H1.HPRT-1 cell line

To demonstrate whether H1.HPRT-1 cells could be induced to follow a specific fate, a differentiation protocol, which does not require the EB stage, was applied (work performed by D. Wojtacha, Roslin Institute). This protocol, recently developed in our lab for hES cells, induces hES cells to differentiate towards bone (Sottile *et al.*, 2003). It consists of EB medium supplemented with the osteogenic factors glucocorticoid dexamethasone, the extracellular matrix-promoting agent ascorbic acid, and β-glycerophosphate, which induce nodule formation and mineral deposition. The osteogenic differentiation potential of the H1.HPRT-1 cell line was compared to the H1 parental cell line. *In vitro* osteogenic differentiation was assessed at different time points over a period of 21 days by examination of mineralised nodule formation staining with the alizarin Red S dye (Sigma). Mineral deposition was measured using a quantitative calcium assay which had been previously optimised and adapted for H9 and H1 cells (Sottile *et al.*, 2003). This assay measures the calcium content of monolayer cultures using colorimetric methods. Alizarin Red S staining, calcium assay and analysis were performed by D. Wojtacha.

Undifferentiated H1 and H1.HPRT-1 cells were seeded at a density of 10^5 cells per well of a 6-well plate for the Alizarin Red staining, and 10^4 cells per well of a 96-well plate for the calcium assay. Conditioned medium was replaced 48 hours later (day 0) with EB medium supplemented or not with osteogenic factors (OS). The HEK 293 (human embryonic kidney) cell line (Graham *et al.*, 1977), which lacks osteogenic potential was used as a negative control in this experiment. The HEK 293 cells are characterised by a high proliferation rate compared with hES

cells. The HEK 293 cells were seeded at a density of 10^4 cells per well of 6-well plate and 10^3 cells per well of a 96-well plate, in order to prevent cultures becoming overconfluent and detaching over a period of 21 days (figure 5.6 A). Each cell line was analysed by Alizarin Red S staining at days 9, 13, 16 and 20. Both the H1 and H1.HPRT-1 cells formed nodules in response to OS treatment, which stained positively with Alizarin Red at day 9 (figure 5.6. A & B), showing calcium deposition. In contrast, no nodule formation was observed in cells cultured in the absence of OS supplements. Unlike H1 and H1.HPRT.1 cells, the HEK 293 cells did not stain with Alizarin Red in the presence of OS treatment and nodule formation was never observed (figure 5.6. A). The H1.HPRT-1 cells showed an increased level of nodule formation at each time point compared to the H1 cells suggesting that these cells had a greater osteogenic differentiation potential (figure 5.6. A & B). These results show that both the H1 and the H1.HPRT-1 cell lines form nodules and calcium deposits *in vitro,* consistent with osteogenic differentiation.

To quantify the mineralisation, cell matrix-associated calcium deposition analysis was performed in 96-well plates using a calcium assay kit (Randox). In this analysis, each cell line was seeded in 4 replicates and calcium deposition was quantified at days 9, 13, 16 and 20 for each replicate. This assay could not be performed on the HEK 293 cell line as they reached confluence very quickly, resulting in the detachment of the cells from the tissue culture plastic during the wash procedure. The results showed a significant ($P<0.001$, F-test) increase in calcium deposition over time in both the H1 and H1.HPRT-1 cells treated with osteogenic factors (figure 5.7). In contrast, calcium deposition was minimal in the absence of OS factors, and not significantly different between the two cell lines

($P>0.05$, F-test). The level of mineralisation in the OS-treated H1 cells was significantly lower ($P<0.01$, F-test) than that in the H1.HPRT-1 cells. Additionally, a delay in the onset of mineralisation in the OS-treated H1 cells was observed with respect to the H1.HPRT-1 cells (day 6) (figure 5.7). The difference of level and delay of mineralisation however, may reflect a difference in cell density rather than a difference of mineralisation potential. Both cell lines may have responded differently, in terms of proliferation, to the OS factors but because cell densities were not examined at each time point prior to performing the calcium assay, this possibility could not be excluded.

These data show that both the H1 hES cell line and the targeted H1.HPRT-1 cell line, when exposed to osteogenic factors, produce mineralised nodules *in vitro*, consistent with the ability of these cells to differentiate into bone-forming osteoblasts.

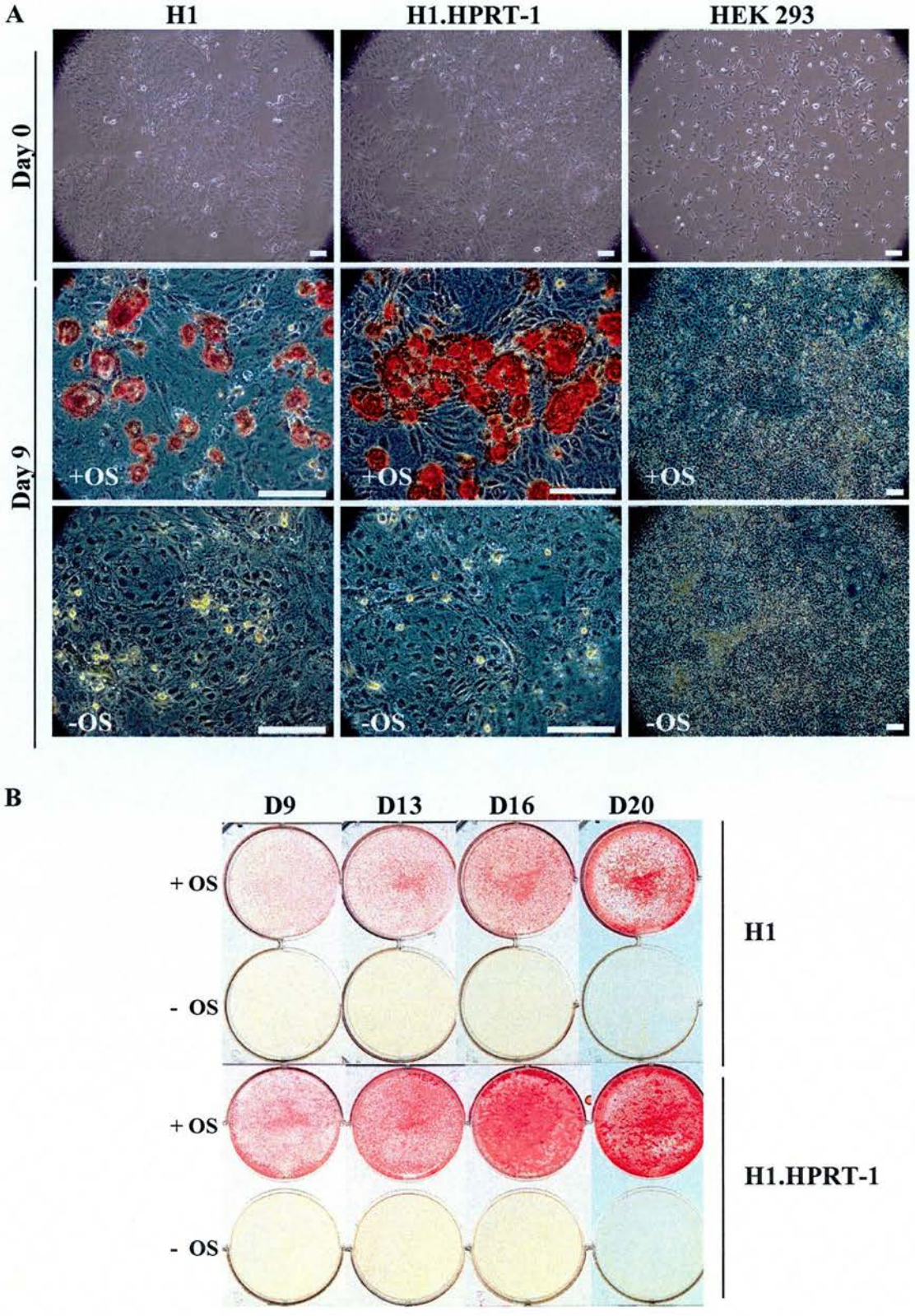


Figure 5.6. Osteogenic differentiation of H1 and H1.HPRT-1 cells *in vitro*. **(A)** Alizarin Red S staining of H1, H1.HPRT-1 and HEK 293 cells, nine days post-plating with or without OS factors. Red areas indicate staining of calcium deposition. Scale bars = 100 μm. **(B)** H1 and H1.HPRT-1 cell culture plates stained with Alizarin Red S at different points over a period of 20 days.

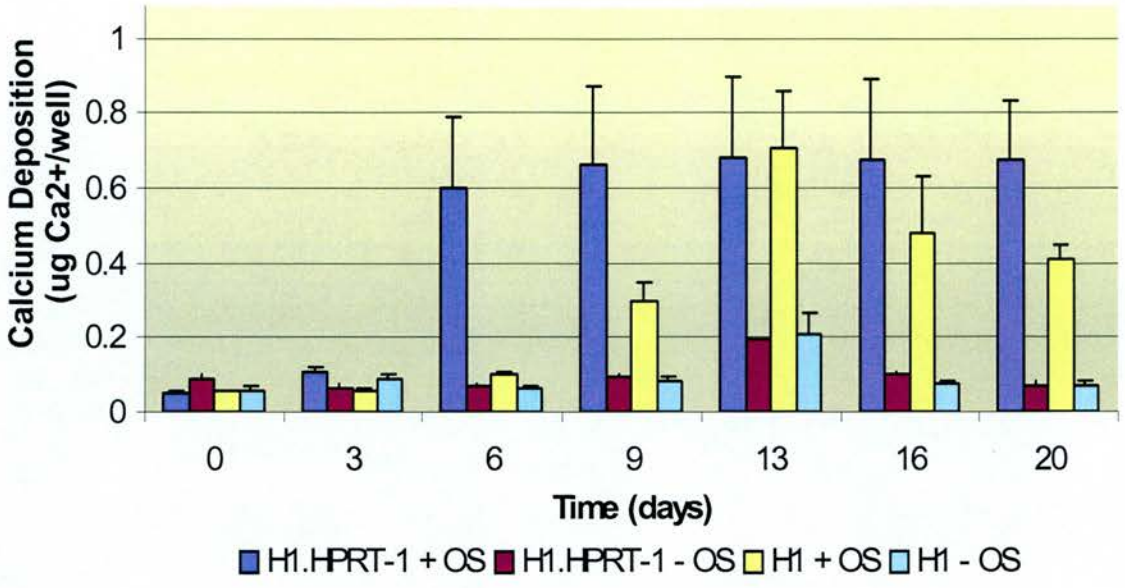


Figure 5.7. Quantitative mineralisation analysis. Time-course of calcium deposition of H1 and H1.HPRT-1 cells treated ± OS supplements. A significant increase of calcium deposition is observed in both H1 and H1.HPRT-1 cell lines treated with OS factors. However, the level of mineralisation in the OS-treated H1 cells is significantly lower than that in the OS-treated H1.HPRT-1 cells. Data are presented as means ± SEM (N=4).

5.4. Discussion

We investigated whether a hES cell line subjected to gene targeting procedures would remain euploid and retain the stem cell markers expression and multipotency of the parental H1 cell line. The data presented in this chapter show that, like the H1 cell line, the H1.HPRT-1 targeted cell line retains a diploid karyotype and expresses markers of an undifferentiated ES phenotype (OCT-4, SSEA-4, TRA-1-60 and TRA-1-81) and differentiated phenotype (SSEA-1) at similar levels. Thus gene targeting and/or the extra culture required for the gene targeting experiment did not affect the above properties.

In vitro differentiation analysis, failed to show multipotency of both the parental H1 and the H1.HPRT-1 cell lines, when using a protocol that involved the EB stage. It is likely that the problem resides in the poor efficiency of H1 cells in producing robust EBs under the conditions tested, and in the inability of these EBs to adhere to the gelatin substrate. This would be consistent with a previous report by Reubinoff *et al.* (2000), in which hES-derived EBs failed to differentiate *in vitro* due to extensive cell death and poor efficiency of good quality EB production. Since this report, several laboratories have shown that the production of robust EBs could be achieved, and that hES-derived EBs could differentiate into the three germ layers (e.g. Itskovitz-Eldor *et al.*, 2000). The inability of H1 and H1-derived cells to produce robust EBs raises the possibility that the trypsin-EGTA (TEG) passaging regime used in our culture system does not allow the generation of three-dimensional aggregates of H1 cells, since in our hands (data not shown) and in other labs (e.g. Xu *et al.*, 2002), good quality EBs can be generated from H1 cells treated with

collagenase. The TEG regime allows the disaggregation of hES cells into a single cell suspension, unlike the collagenase treatment which permits the disaggregation of hES cells into clumps of cells. Clumps of cells might therefore aggregate more efficiently into robust EBs than single cells. Arguing against this hypothesis however, is the ability of TEG-treated H9 cells to form robust EBs from single cell suspensions, as shown in figure 5.4. This might however reflect the genetic difference between the H9 and H1 cell lines. It is possible that TEG treatment causes the H1 and H1.HPRT-1 cells to lose their potential to aggregate and thus to differentiate *in vitro* towards lineages of the 3 germ layers. Consistent with *in vitro* differentiation data in H1 cells treated with TEG, *in vivo* differentiation studies in severe combined immunodeficiency (SCID) mice with TEG-treated H1 cells do not produce teratomas, unlike H9 cells treated with TEG (McWhir *et al.*, unpublished data). In contrast H9 and H1 cells treated with collagenase produce teratomas when injected into SCID mice (Thomson *et al.,* 1998). Taken together, these results suggest that TEG may affect the multipotentiality of H1 and H1-derived cells by a yet unknown mechanism.

The osteogenic differentiation experiment, which omits the EB stage, showed that both the H1 and the H1.HPRT-1 cell lines, when exposed to osteogenic factors, formed mineralised nodules *in vitro,* consistent with the ability of these cells to differentiate into bone-forming osteoblasts (mesoderm). Thus these data showed that gene targeting procedures did not affect the ability of H1-derived cells to form mineralised nodules *in vitro*. Although the degree of mineralisation was significantly greater in the H1.HPRT-1 targeted cell line than in the H1 cell line, it is unclear whether this was the result of cell line difference or a greater cell density of the

H1.HPRT-1 cells. The greater level of mineralisation of the H1.HPRT-1 cell line may be due to the fact that this cell line is a subclone of the H1 cells. Thomson *et al.* (1998) suggested that there may be some variation in the developmental potential among the undifferentiated H1 cells, in spite of their homogeneous appearance, because these cells were not cloned from a single cell. If this was true, a difference of developmental potential may have accounted for a reduced level of mineralisation in the H1 cells. In contrast, because the H1.HPRT-1 cell line is clonally-derived from H1 cells and should be therefore homogeneous, all the cells in the population would probably have a similar level of mineralisation.

Regardless of the difference between the two cell lines, further analysis would be required in order to formally confirm the osteogenic differentiation of the H1.HPRT-1 cell line. X-ray diffraction analysis of hES-derived nodules would show whether the diffraction pattern corresponds to that of hydroxyapatite crystals, the principal mineral component of bone matrix (Heaney, 2002). Immunocytochemistry and RT-PCR analysis would be required to show whether osteoblast-specific markers such as osteocalcin, collagen I α1 and bone sialoprotein (Sottile *et al.,* 2003) are expressed in the H1.HPRT-1 cell line.

To summarise, we have found no evidence that gene targeting procedures affect the properties of hES cells, since both H1 and H1-targeted cells retain predominantly a diploid karyotype, express stem cell markers at similar levels, and behave in a similar way in *in vitro* differentiation experiments. The difference observed between H9 and H1 cells in the *in vitro* differentiation studies indicates that the genetic variation between hES cell lines may require the establishment of

protocols specific to an individual hES cell line. These data therefore reflect the challenges of optimisation that we face presently with human ES research.

CHAPTER VI

Correction of an H1-Derived *HPRT*-Targeted Cell Line by Homologous Recombination

6.1. Introduction

6.2. Objective

6.3. Results

6.4. Discussion

6.1. Introduction

There are presently two published reports of targeted mutation by HR in hES cells (Zwaka and Thomson, 2003; Urbach *et al.*, 2004), but the correction of a mutant gene by targeted HR has as yet not been accomplished in hES cells. While genes can in principle be inactivated to create cellular models of genetic diseases in hES cells (Zwaka and Thomson, 2003; Urbach *et al.*, 2004), the ability to repair mutated genes in hES cells could also have great implications for gene therapy.

Most gene therapy strategies are based on gene addition by virus-driven random integration, in which a normal copy of a gene is transferred into the cell type of interest, where it coexist with the mutant copy of the same gene. However, gene addition remains a challenge owing to transgene silencing, inappropriate expression or misintegration (Clark *et al.*, 1994). An alternative approach has been the use of targeted gene correction via HR, which enables the production of a normal protein under the control of the cell's intrinsic regulatory signals. Some recent advances have been made in this area of research in mammalian cells *in vitro* as well *as ex vivo* (reviewed by Kren and Steer, 2002). A strategy based on the use of small DNA fragments (SDFs) (few hundred nucleotides) homologous to the targeted genomic sequence has been used to correct a deletion in the cystic fibrosis transmembrane conductance regulator gene (*cftr*) in primary and transformed human lung epithelial cells, reaching up to 1 % repair in the transfected cells (Goncz *et al.*, 1998). RNA-DNA oligonucleotide chimeratherapy is another technique that was developed to target site-specific nucleotide base changes to homologous genomic sequences. This technique was used to correct the sickle cell mutation in human

lymphoblasts (Cole-Strauss *et al.*, 1996). It has also been used to correct a dystrophin point mutation in Duchenne muscular dystrophy skeletal muscles in mice (Bertoni and Rando, 2002). Short single-strand oligonucleotide is an approach that has also been explored to correct point mutations. For instance, Igoucheva *et al.* (2001) showed that single-strand oligonucleotides (25-61 bases homologous to the target sequence) were capable of correcting a single point mutation in the mutant β-galactosidase gene in mammalian cells, with correction rate of 0.1 %. Another approach that holds great promise in gene therapy is the use of zinc-finger nucleases (ZFNs) (Kim *et al.*, 1996), which have been shown to stimulate HR in human cells (Porteus and Baltimore, 2003). These synthetic proteins are composed of a DNA-binding domain, which recognises a specific site near a mutation, and a DNA-cleaving domain, which introduces a double-strand break into the DNA, thus promoting HR. Urnov *et al.* (2005) have recently exploited this technology to correct a mutation in the *IL2Rγ* gene (which mutations cause X-linked severe combined immune deficiency) (Cavazzana-Calvo *et al.,* 2000) in transformed human cells. The co-introduction of the ZFNs and the wild-type template resulted in the correction of 15 to 20 % of the cells without the use of selection. Among these cells, 7 % acquired the desired correction on both X chromosomes. Additionally they reported around 5 % correction in human T cells, where correction frequencies are considerably lower without the introduction of double-strand breaks (~0.001%).

Lesch-Nyhan syndrome is a severe neurological disorder that is caused by the malfunction of the *hprt* gene (reviewed by Stout and Caskey, 1988). The seriousness of the disease makes it a candidate for eventual gene therapy. Several strategies have

been developed to rescue an HPRT-deficient phenotype in a broad range of cell lines, including the use of *hprt* minigenes (Selfridge *et al.*, 1992), *hprt*-expressing adenoviruses (Southgate *et al.*, 1999), and the use of an episomal transgene (Wade-Martins *et al.*, 2000). Gene targeting strategies have also been used to correct a mutated *hprt* gene in several cell lines, including mouse ES cells (Doetschman *et al.*, 1987; Thompson *et al.*, 1989; Koller *et al.*, 1989; Reid *et al.*, 1991; Magin *et al.*, 1992; Hatada *et al.*, 2000; Heaney *et al.*, 2004), and murine hematopoietic stem cells (Hatada *et al.*, 2000).

In this study, we sought to exemplify the correction of a mutated gene in hES cells by gene targeting. This chapter describes a preliminary targeting experiment in which a replacement vector was used to correct the targeted mutation in the HPRT-deficient H1.HPRT-1 cell line (obtained by targeting H1 cells with the pHPRT-EfN targeting vector).

6.2. Objective

To correct by HR a previously generated HPRT-targeted hES cell line using a replacement vector.

6.3. Results

6.3.1. Correction Scheme and Targeting Vector

The H1.HPRT-1 cell line obtained by targeting the *hprt* gene in H1 cells with the targeting vector pHPRT-EfN contains a 3.4 kb *EF1-α-loxP-neopA-lox2272* cassette inserted in exon 3 in the opposite transcriptional orientation to the *hprt* gene (figure 6.1). This cell line was used in a "correction" targeting experiment using the targeting vector, pTopo-HPRT, which contains 7.1 kb of *hprt* homology comprising exon 2 and 3, intron 2 and fragments of introns 1 and 3 (figure 6.1). Following double reciprocal recombination between pTopo-HPRT and the *hprt* chromosomal sequences, the chromosomal *hprt* sequence containing the 3.4 kb cassette is replaced with native *hprt* sequence, thereby correcting the *hprt* gene structure and restoring function. Because the *hprt* fragment only incorporates exons 2 and 3, it cannot behave like an *hprt* minigene (Selfridge *et al.*, 1992) to rescue HPRT-deficient cells by random integration on its own. Therefore, the HPRT-deficiency phenotype can only be rescued if HR has taken place at the mutated *hprt* gene.

The pTopo-HPRT replacement targeting vector consists of 7.1 kb of *hprt* homology comprising the 6.9 kb HindIII *hprt* fragment included in the pHPRT-EfN targeting construct (figure 6.1). The 7.1 kb *hprt* homology was obtained by using long range PCR of gDNA from H1 cells and subcloned into the Invitrogen pCR2.1-Topo vector (appendix 1, section 2.1). The inclusion of a selectable marker

in this vector is unnecessary because HPRT activity allows direct selection in HAT medium (Albertini, 2001). HAT medium inhibits the *de novo* pathway of purine synthesis and prevents growth of cells that lack an active salvage pathway (Szybalski & Szybalska, 1962). The mutant cells proliferate in the presence of 6-TG (Stutts & Brockman, 1963) and die in HAT medium while wild-type cells behave in the opposite fashion. Therefore corrected H1.HPRT-1 cells can be directly selected in HAT, rendering targeting frequencies easy to calculate at the phenotypic level.

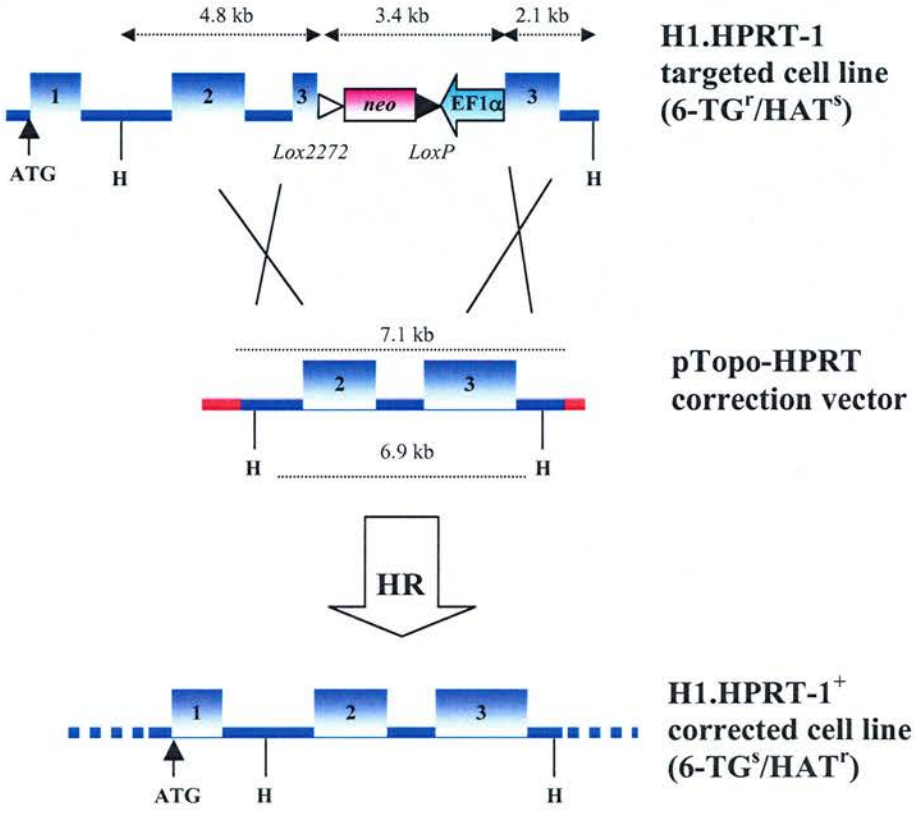


Figure 6.1. Schematic representation of the *hprt* gene correction by targeting in the H1.HPRT-1 mutant cell line. Partial structure of the *hprt* knock-out and corrected genes, and the correction vector used to restore *hprt* function and structure in the H1.HPRT-1 cell line. The *hprt* homology (7.1 kb) in the pTopo-HPRT targeting construct is represented by a thick blue line with blue boxes marking the exons. The pCR2.1 Topo vector backbone is represented as a thick red line. The two cross-over recombination points which take place between the correction vector and the *hprt* chromosomal sequence are depicted by an 'X'. HindIII (H) digestion of pTopo-HPRT generates a 6.9 kb *hprt* fragment and following HR, targeted recombinants can be selected for based on their newly-acquired resistance to HAT. The translational initiation start site (ATG) is indicated below exon 1 in the *hprt* knock-out and corrected genes.

6.3.2. *Hprt* Correction Experiment in H1.HPRT-1 Cells

To correct the *hprt* gene in the H1.HPRT-1 cell line, 10^6 H1.HPRT-1 cells were stably transfected using the Eppendorf Multiporator (300 V / 100 μs) with 50 μg of HindIII-digested pTopo-HPRT. HAT selection (0.1 mM hypoxanthine, 1 μM aminopterin, 10 μM thymidine) was applied 48 hours later to select for homologous recombinants carrying a corrected *hprt* gene. After 8 days in selection, 3 HAT-resistant colonies were produced suggesting that *hprt* function was restored in these clones by HR (table 6.1). A mock control transfection in which cells were electroporated in the absence of DNA did not produce any HAT-resistant colonies indicating that the H1.HPRT-1 cells were not contaminated with wild-type cells prior to electroporation. The 3 clones were expanded for subsequent analysis to confirm a targeting event.

cell line/ passage No	Electroporation conditions	Number of cells	pTopo-HPRT (μg)	HAT^r colonies	A.T.F
H1.HPRT-1/96	300 V/100 μs	$1x10^6$	50	3*	$3x10^{-6}$

Table 6.1. Targeted correction of the mutated *hprt* gene in H1.HPRT-1 cells by HR. ATF: Absolute targeting frequency: Ratio between the total number of homologous recombinants and the total number of cells electroporated. * 1 colony was lost during expansion.

6.3.3. Detection of a Targeting Event by PCR

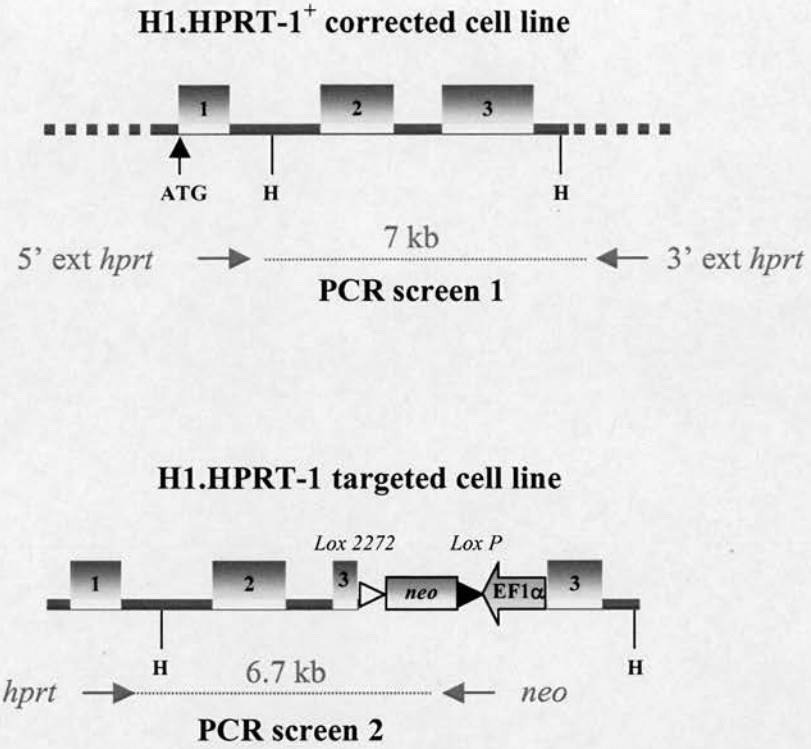
To confirm that the two HAT-resistant colonies (one clone was lost during expansion) were the result of a correction event at the *hprt* gene, gDNA was extracted from each clone and used as the template in a PCR screen (PCR screen 1). The corrected *hprt* or wild-type alleles should produce a 7 kb PCR fragment (since the screen does not distinguish between wild-type and corrected alleles) with two primers annealing to the *hprt* sequence 5' (5' ext *hprt*) and 3' (3' ext *hprt*) to the homology included in the pTopo-HPRT targeting vector (figure 6.2 A). The H1 parental cell line and the H1.HPRT-1 cell line, from which the HAT-resistant clones were derived, were used as controls in this analysis. PCR analysis showed that the two HAT-resistant clones and the wild-type H1 cells produced the 7 kb fragment (diagnostic for a wild-type or corrected allele). In contrast, this band was not detected with gDNA derived from the H1.HPRT-1 cell line (figure 6.2 B). Hence these results show that the two HAT-resistant clones are the result of a targeting correction event at the *hprt* gene.

A second PCR screen (PCR screen 2) was performed to check for any possibility of contamination with H1.HPRT-1 cells (which may have survived by metabolic cooperation) in the two *hprt* correctant cultures. In this screen, the H1.HPRT-1 cells should produce a 6.7 kb fragment with the 5'ext *hprt* primer and an internal *neo* primer (figure 6.2 A). In contrast, *hprt* correctant cells should not yield any product because of the absence of the *neo* expression cassette to which the *neo* primer anneals (figure 6.2 A). Results of the PCR showed the presence of the 6.7 kb fragment (diagnostic for the uncorrected allele) for the H1.HPRT-1 cells but also for the two *hprt* correctant clones (figure 6.2 B), indicating that both clones were

contaminated with H1.HPRT-1 cells. Survival of the H1.HPRT-1 cells in HAT selection may have been the result of positive metabolic cooperation ("kiss of life") (Hooper, 1982). This effect has previously been reported in other correction targeting experiments in the *hprt* gene in mouse ES cells (Thompson *et al.*, 1989).

Because homologous recombinants were directly selected in HAT, all the random integrants were eliminated, and as a result effective targeting frequency (ratio between the total number of homologous recombinants and the total number of integrants surviving selection) cannot be determined. The absolute targeting frequency (ATF) however, can be determined as the ratio between the total number of homologous recombinants (HAT-resistant clones) and the total number of cells electroporated. The ATF obtained in this experiment ($3x10^{-6}$) (table 6.1) was similar to that achieved for the targeted modification of the wild-type *hprt* gene (10^{-6}) (see chapter 4, table 4.1).

A



B

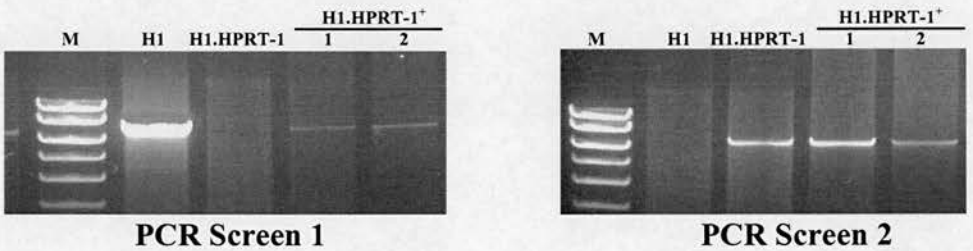


Figure 6.2. PCR screening for a targeted correction event at the *hprt* gene of the two HAT-resistant clones. **(A)** Partial structure of the *hprt* corrected and knock-out genes showing the predicted size of a PCR product generated using two external 5' and 3' *hprt* primers (7 kb) or the internal *neo* primer and the 5'external *hprt* primer (6.7 kb), respectively. **(B)** PCR analysis showing the presence of a 6.7 kb and 7 kb band for the 2 HAT-resistant clones. The 7 kb product shows correction event and the 6.7 kb product indicates that both clones are mixed with H1.HPRT-1 cells, which have survived HAT selection by metabolic cooperation. M: 5 µl Hyperladder (Bioline).

6.4. Discussion

In this study, we demonstrated that the targeted mutation in the H1.HPRT-1 targeted cell line could be functionally and structurally corrected by targeted replacement of the *EF1-α-loxP-neopA-lox2272* transgene with the endogenous *hprt* sequence. Using the pTopo-HPRT correction vector, we obtained three HAT-resistant cell lines derived from the H1.HPRT-1 deficient cell line (table 6.1) consistent with the restoration of HPRT activity in the H1.HPRT-1 cells. In contrast, untransfected H1.HPRT-1 cells did not result in the generation of HAT-resistant colonies, consistent with an absence of wild-type (wt) H1 cells in the H1.HPRT-1 culture. Subsequent PCR analysis confirmed the presence of *hprt* corrected structure for the 3 HAT-resistant clones. Thus the data presented in this chapter strongly support the idea that a correction event has occurred at the mutated *hprt* gene in H1.HPRT-1 cells.

The PCR analysis however showed that each HAT-resistant clone recovered had both mutant and corrected DNA sequences, suggesting that the correctant clones were contaminated with the H1.HPRT-1 cells from which these clones were derived. This is probably due to the persistence in the HAT-resistant culture of H1.HPRT-1 cells surviving by metabolic cooperation (Hooper, 1982). Ten days after isolating the HAT-resistant colonies, selection was relaxed and the clones were cultured in non-selective medium for over 30 days before gDNA extraction, which would allow uncorrected 6-TG-resistant/HAT-sensitive cells to expand along with the targeted/corrected ones. This has been previously reported in mouse ES cells

(Thompson *et al.*, 1989). Subjecting each clone to renewed HAT selection at clonal density should enable the subcloning of a pure corrected-HPRT cell line.

Although these preliminary data are very promising, Southern analysis will be necessary to confirm a correction event at the mutated *hprt* gene. Furthermore, our strategy could be refined to fully rule out the possibility that the generation of HAT-resistant colonies was the result of contamination with wt H1 cells. In this present study, the wt and corrected *hprt* structures cannot be distinguished because the homology used in pTopo-HPRT is isogenic. According to the National Centre for Biotechnology information (NCBI) database, the *hprt* homology used in the pTopo-HPRT targeting construct contains ~10 SNPs distributed along the 6.9 kb sequence (~ 1 SNP every 700 bp). Thus it may be possible to distinguish between a corrected and wt *hprt* gene by using non-isogenic gDNA to the H1 cell line, to correct the *hprt* gene (e.g. *hprt* homology amplified from H9-derived gDNA). Corrected clones would first be identified by PCR and Southern analysis. A correction event would then be confirmed by sequencing of PCR amplified gDNA flanking the polymorphic regions.

The targeted correction of the mutated *hprt* gene in H1-HPRT-1 cells was achieved at absolute targeting frequency (ATF) of $3x10^{-6}$ (table 6.1). Thus targeted correction of the mutated *hprt* gene can be achieved with similar frequencies in hES cells and murine ES cells (between $4x10^{-8}$ and $2x10^{-6}$) (Doetschman *et al.*, 1987; Thompson *et al.*, 1989; Magin *et al.*, 1992). These groups however used an insertion vector to correct a defective *hprt* locus in murine ES cells. Other groups have reported the use of a replacement vector to correct a defective *hprt* gene in the E14TG2a (Reid *et al.* 1991) and HM1 (Hatada *et al.,* 2000) mouse ES cell lines,

resulting in much greater absolute targeting frequencies (between 10^{-5} and $2x10^{-5}$) than those achieved in our study. The inactivity of the *hprt* gene in both the E14TG2a and HM1 cell lines is due to a spontaneous 55 kb deletion that includes the promoter and exons 1 and 2 of the *hprt* gene (Tsuda *et al.*, 1997). The pMP8 (Reid *et al.* 1991) and pMP8neo (Hatada *et al.,* 2000) targeting vectors used to correct both cell lines, contained 4 kb of mouse gDNA 5' to the deletion, 1.8 kb of human *hprt* gDNA including the promoter and exon 1, and 7 kb of mouse *hprt* gDNA including intron 1, exons 2 and 3. All together 7.6 kb homology to the locus was used to correct the *hprt* gene (Reid *et al.* 1991). We used a very similar length of homology (6.9 kb). Thus the length of homology cannot account for the difference in targeting frequency between our study and those reported by Reid *et al.*, and Hatada *et al.* The pTopo-HPRT correction vector used in our study to correct the mutated gene contains *hprt* gDNA homology derived from H1 cells. However, the homology used in the pHPRT-EfN targeting vector to modify the *hprt* gene, was not isogenic. As a result, the *hprt* homology used in pTopo-HPRT was non-isogenic to the mutated locus. As previously discussed in section 4.4. (chapter 4), there is no evidence that targeting frequency may be affected by polymorphic variation in human cells (Sedivy *et al.*, 1999). Additionally, the *hprt* gene has previously been targeted in hES cells at similar frequencies with either an isogenic (Zwaka and Thomson, 2003) or a non-isogenic targeting vector (Urbach *et al.*, 2004). Thus it seems unlikely that the use of isogenic DNA will increase targeting frequencies in hES cells.

Because there are so many experimental variations between our work and the work reported by Reid *et al.*, and Hatada *et al.* (e.g. cell lines, transfection conditions, *hprt* homologies, and *hprt* mutations), it is presently unclear why ATF

achieved in mouse ES cells is much greater than that achieved in our hands in hES cells. It is however probable that targeting frequencies in hES cells will be further improved with the establishment of optimal transfection conditions.

To conclude, these preliminary data support the exemplification of not only a targeted correction by HR in a chosen gene in hES cells, but also repeated targeting at the same locus. Thus this study demonstrates the feasibility of the gene targeting route to further manipulating hES cells, providing new means in gene therapy (Kren and Steer, 2002).

CHAPTER VII

Recombinase-Mediated Cassette Exchange at the *Hprt* Gene

7.1. Introduction

Gene targeting and site-specific recombination are complementary strategies that permit the alteration of specific genes in mammalian cells. The Cre site-specific DNA recombinase from the phage P1 specifically recognises a 34 bp recombination site called *loxP* (locus of X-over of P1) and catalyses reciprocal DNA strand exchange between two of these sites (Hoess and Abremski, 1984). By exploiting the properties of the Cre/*lox* system, the pre-positioning of a *loxP* site into a specific genetic locus, known for reproducible transgene expression, provides a means to introduce any transgenes flanked by two unidirectional *loxP* sites. In a first step, a *loxP* site is placed into a predetermined gene locus by HR. In a second step, a transgene of interest flanked by two unidirectional *loxP* sites can be integrated, in the presence of Cre, into the pre-existing *loxP* site (Kolb and Siddell, 1997; Kolb *et al.*, 1999). However, transgene integration is unstable because the plasmid can re-excise itself following recombination between two unidirectional *lox*P sites (Araki *et al.*, 1997; Kolb *et al.*, 1999). Since integrations are inter-molecular events while excisions are intra-molecular events, excisions are strongly favored. Recombinase-mediated cassette exchange (RMCE) is a method that aims at eliminating the re-excision problem (Bethke and Sauer, 1997; Bouhassira *et al.*, 1997; Seibler *et al.*, 1998). The RMCE approach is based on the observation that two recombination *lox* sites that differ in their central 8-bp-spacer region do not recombine with each other, whereas *lox* sites which possess the same spacer region recombine efficiently. By taking advantage of these mutated *lox* sites, a transgene flanked by heterospecific *lox*

sites (Lee and Saito, 1998) can be inserted by HR into a specific gene. This insertion allows in a subsequent step, the exchange, by a double reciprocal crossover replacement, of the transgene with a single copy of another transgene flanked by the same heterospecific *lox* sites (Bethke and Sauer, 1997; Bouhassira *et al.*, 1997; Seibler and Bode, 1998; Feng *et al.,* 1999; Trinh and Morrison, 2000; Kolb 2001; Li *et al.*, 2003; Cobellis *et al.*, 2005). In ES cells, effective RMCE frequencies (ratio between number of site-specific recombinants and number of integrants surviving selection (random and targeted)) between 54 and 100 % may be obtained using a negative and positive selection (Seibler *et al.*, 1998; Kolb, 2001). Effective targeting frequencies of 100% can be achieved by combining RMCE with a promoter trap strategy (Baubonis and Sauer, 1993 ; Verhoeyen *et al.,* 2001). As a result, the RMCE procedure has emerged as a powerful tool for manipulation of the eukaryotic genome and the production of multiple second-step modifications.

While the RMCE approach has now been used in a variety of cell lines including mouse ES cells (Seibler *et al.*, 1998; Feng *et al.*, 1999; Kolb, 2001), it has not yet been demonstrated in hES cells. The aim of the experiments described in this chapter was to exemplify RMCE at the previously tagged *hprt* gene in hES cells, and to test this locus for its ability to show reproducible transgene expression. Combined with the RMCE approach, we used a promoter trap strategy, in which following the desired recombination event, the reconstruction of a promoterless selectable marker in exon 3 of the *hprt* gene, ensures only the selection of single-copy site-specific recombinants.

The following study discusses the optimisation of the experimental procedures required to achieve site-specific recombination at the *hprt* gene, and the

assessment of transgene expression at the *hprt* gene in hES cells after site-specific recombination.

7.2. Objectives

(i) To optimise the RMCE protocol and exemplify RMCE in the previously tagged *hprt* gene in hES cells. (ii) To investigate whether site-specific recombinants at the *hprt* gene are characterised by reproducible transgene expression.

7.3. Results

7.3.1. Development of an RMCE Protocol

7.3.2. Transgene Expression Analysis of Site-Specific Recombinants at the *Hprt* gene

---

7.3.1. Development of an RMCE Protocol

7.3.1.1. RMCE Scheme and RMCE Vector

A promoter trap strategy was employed to enrich for site-specific recombinants at the previously tagged *hprt* gene. The general outline of the approach, detailed in figure 7.1, is as follows: Targeted clones (H1.HPRT cell clones) generated by HR in exon 3 of the *hprt* gene carry a *neomycin* (*neo*) gene flanked by heterospecific *lox* sites in the opposite transcriptional orientation of the *hprt* gene. The *neo* gene is under the transcriptional control of the *EF1-α* promoter, which is excluded from the "floxed" cassette. Cre-mediated site-specific recombination is performed by co-transfection of these clones with supercoiled RMCE plasmid (pfBSD-OctGFP) and pTurboCre plasmid (gift from T. Ley, Washington University, St. Louis, USA). The *cre* expression plasmid pTurboCre

was used because its excision activity between two *loxP* sites has been shown on several occasions in mouse ES cells (Westervelt *et al.*, 2003; Lin *et al.,* 2003; Engel *et al.*, 2004). Additionally, pTurboCre contains the strong ubiquitous *CAG* promoter (Niwa *et al.,* 1991), which produces high levels of Cre recombinase expression in mammalian cells. The *CAG* promoter combines the human cytomegalovirus immediate-early enhancer and a modified chicken β-actin promoter and first intron (Niwa *et al.,* 1991).

The RMCE construct (pfBSD-OctGFP), consists of a promoterless blasticidin (*BSD*) gene and the enhanced green fluorescent protein (EGFP) gene under transcriptional control of the human *oct-4* promoter (amplified by PCR from H9 cells). In common with the modified *hprt* gene, the RMCE vector is flanked by the same heterospecific *lox* sites. Because the *EF1-α* promoter is not included between the *lox* sites in the targeted locus, a promoter trap strategy (Jasin and Berg, 1988, Sedivy and Sharp, 1989) can be applied in a subsequent RMCE experiment. Upon site-specific recombination catalysed by Cre recombinase, the *neo* gene between the heterospecific *lox* sites is lost, conferring G418 sensitivity, and the promoterless *BSD* gene is now placed under the transcriptional control of the *EF1-α* promoter, conferring resistance to blasticidin. This enrichment strategy permits the direct selection of site-specific recombinants because the majority of random integrations will be transcriptionally silent and not survive blasticidin selection.

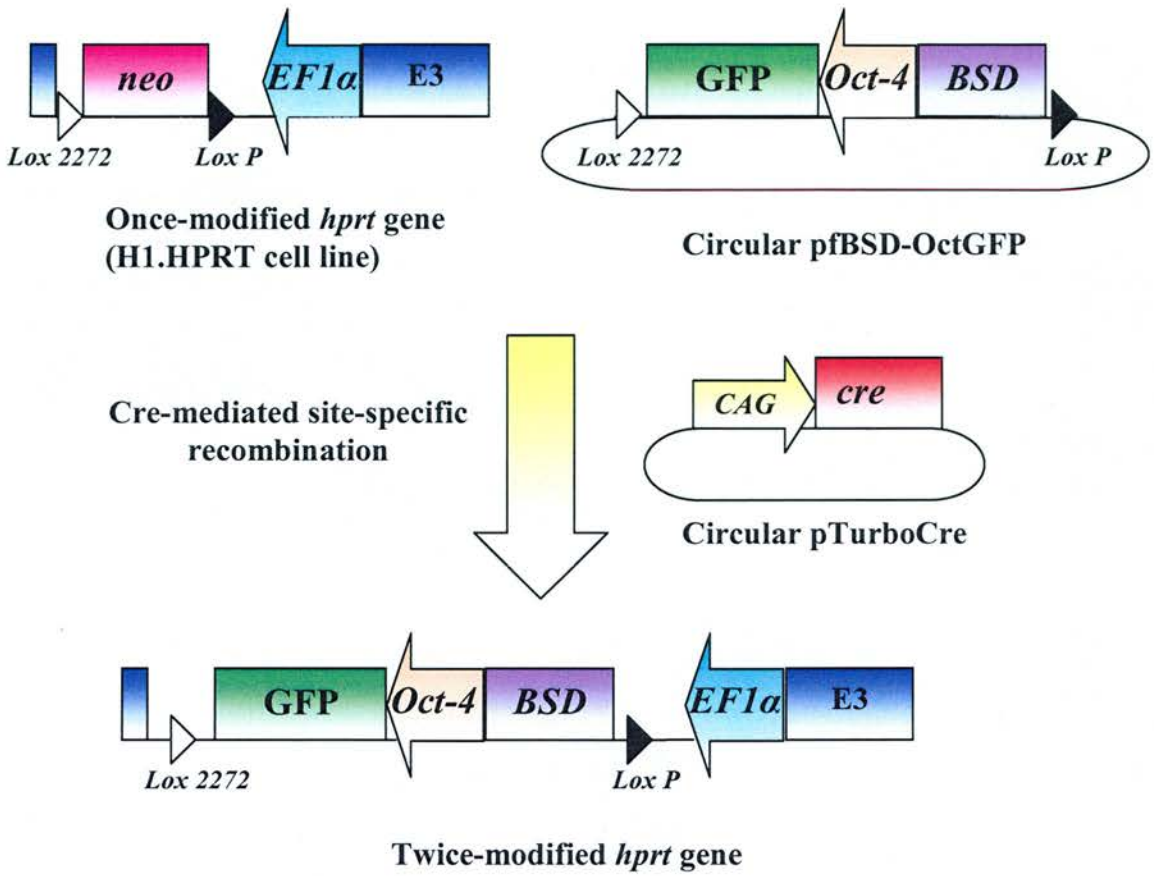


Figure 7.1. RMCE at the engineered *hprt* gene. Schematic representation of the *hprt* gene targeting by site-specific recombination in H1.HPRT targeted cells. The exon 3 of the *hprt* gene is depicted by a blue box. The positions of the *EF1-α*, *oct-4*, and *CAG* promoters are marked as green, orange and yellow arrows, respectively. The *neo*, *BSD*, GFP and *cre* genes are marked as a pink, purple, green and red boxes, respectively. Positions and orientations of the *loxP* and *lox2272* recombination sites are represented as solid and white arrows, respectively. The modified *hprt* gene, which incorporates a *neo* gene flanked by heterospecific *lox* sites under the transcriptional control of the *EF1-α* promoter, is used as the target for an RMCE experiment. Site-specific recombination at the *hprt* gene leads to the loss of the *neo* cassette (1.8 kb) and the insertion of a *BSD* gene (1.5 Kb) and an *Oct-4*GFP expression cassette (5 Kb). The *BSD* gene is inserted such that its expression is controlled by the *EF1-α* promoter and stable site-specific recombinants can be selected directly in blasticidin selection.

7.3.1.2. Blasticidin Kill Curve for H1 cells

Blasticidin S is a peptidyl nucleoside antibiotic isolated from the culture broth of *Streptomyces griseochromogenes*. It specifically inhibits protein synthesis in both prokaryotes and eukaryotes through inhibition of peptide bound formation in the ribosomal machinery (Yamaguchi *et al.,* 1965). Expression of the blasticidin S deaminase gene *BSD* from *Aspergillus terreus* confers resistance by converting the antibiotic to an inactive deaminohydroxy derivative compound (Kimura *et al.,* 1994). Blasticidin S is used to select for transfected cells carrying the *BSD* resistance gene.

An RMCE event with pfBSD-OctGFP in the H1.HPRT targeted cell lines results in the exchange of the *neo* gene with the *BSD* gene in exon 3 of the *hprt* gene. In order to determine the optimal concentration of antibiotic required to eliminate blasticidin-sensitive cells, a kill curve was calculated for untransfected H1 cells and for H1 cells transfected with a *BSD* expression vector pEfBSD-OCTGFP (a plasmid identical to pfBSD-OctGFP but in which the *BSD* cassette is under the control of the *EF1-α* promoter). Because 1μg/ml of blasticidin was sufficient in our hands to completely kill H9 cells (data not shown), drug selection ranging from 0 to 1 μg/ml was tested in untransfected and transfected H1 cells. Cells were cultured with blasticidin for 7 days and were subsequently stained with 10 % Giemsa's solution for colony counting.

As shown in table 7.1 and figure 7.2, concentrations between 0.4 and 0.8 μg/ml produced a similar number of resistant colonies in pEfBSD-OctGFP-transfected H1 cells, while fewer colonies were recovered at 1 μg/ml. Although blasticidin-resistant colonies were never produced in untransfected H1 cells (table

7.1), patches of cells surviving blasticidin selection were detected at 0.4 and 0.6 μg/ml but not at 0.8 μg/ml (figure 7.2 B). Below 0.4 μg/ml, selection had little effect on the cells. It was therefore decided that subsequent RMCE experiments would be conducted at a blasticidin concentration of 0.8 μg/ml to select for blasticidin-resistant site-specific recombinants.

Cell line	Number of cells	Plasmid DNA (μg)	Blasticidin concentration (μg/ml) / Number of colonies recovered					
			0	0.2	0.4	0.6	0.8	1
H1	$1x\ 10^6$	No DNA	*	*	0	0	0	0
H1	$1 x\ 10^6$	pEfBSDOctGFP (50)	*	*	23	27	24	11

Table 7.1. Blasticidin kill curve for H1 cells and for pEfBSD-OctGFP transfected H1 cells. (*) Selection level was too low to allow colony recovery.

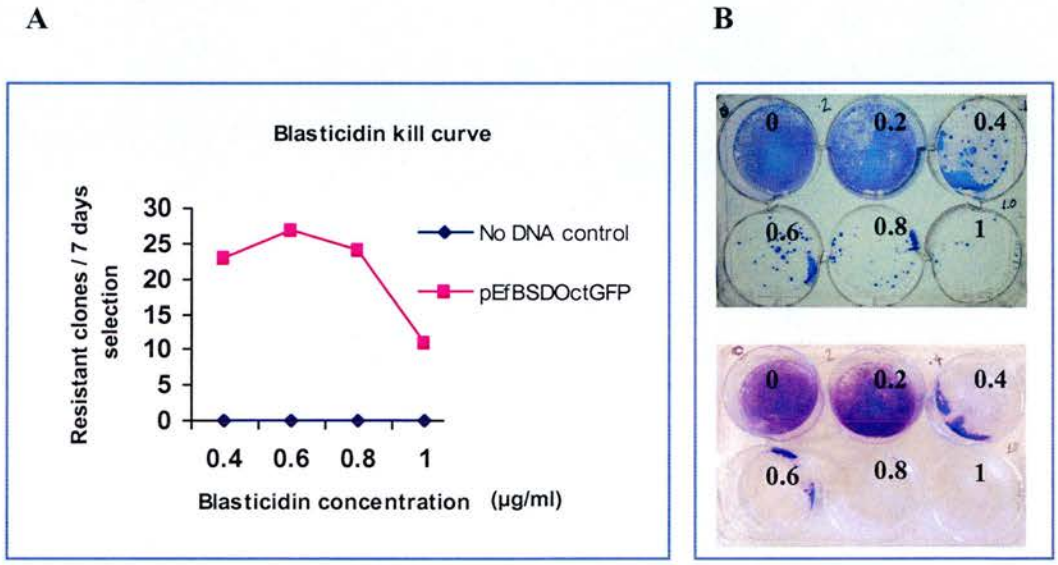
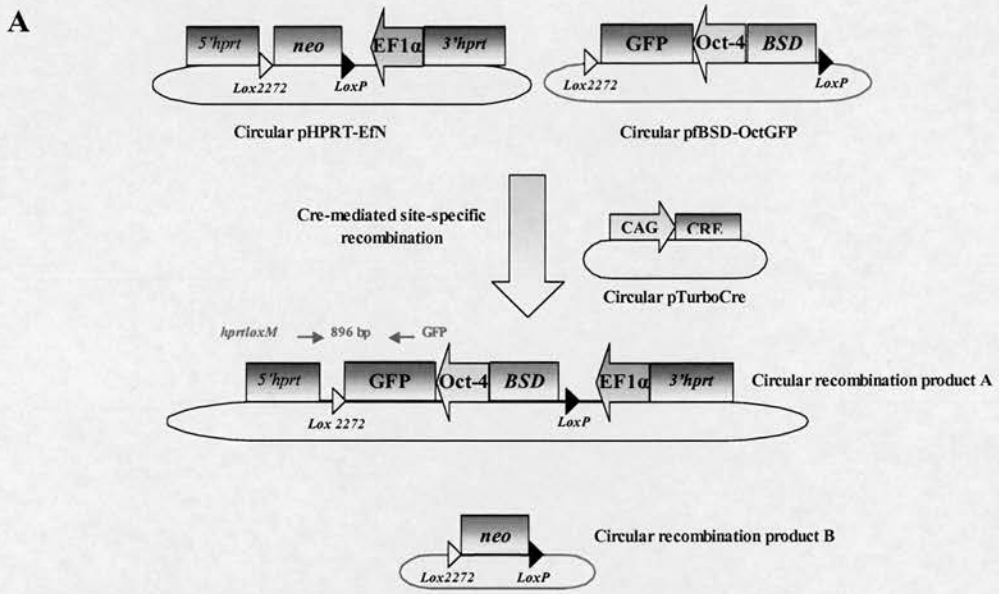


Figure 7.2. Blasticidin kill curve for untransfected H1 cells and for H1 cells transfected with pEfBSD-OctGFP. **(A)** Graph representation of blasticidin kill curve for untransfected H1 cells (blue) and for H1 cells transfected with pEfBSD-OctGFP (pink). Maximal colony recovery was achieved with 0.6 μg/ml of blasticidin in H1 cells transfected with pEfBSD-OctGFP. However all the untransfected H1 cells were killed at 0.8 μg/ml (see figure B). **(B)** pEfBSD-OtGFP transfected (top panel) and untransfected (lower panel) H1 cells were selected in blasticidin for seven days and were subsequently stained with 10 % Giemsa's solution. Blasticidin concentrations ranging from 0 to 1 μg/ml are indicated. Optimal colony recovery in H1 cells transfected with pEfBSD-OctGFP can be achieved with 0.8 μg/ml of blasticidin selection.

7.3.1.3. Transient Recombination Assay in H1 Cells

Sequencing of the *hprt* targeting vector (pHPRT-EfN) and the RMCE vector (pfBSD-OctGFP) showed that both heterospecific *lox* sites were intact (Appendix 1, sections 1.2 and 1.3). To confirm functionality of the *lox* sites, a recombination assay was developed in H1 cells using a lipofection protocol which had been optimised for H9 cells (H. Priddle, unpublished data). H1 cells were transiently co-transfected with the circular plasmids pHPRT-EfN (1.17 μg), pfBSD-OctGFP (0.9 μg), and pTurboCre (3.2 μg), using lipofectamine 2000 (Invitrogen) (figure 7.3 A). A 3:1 molar excess of pTurboCre was used to ensure that most of the cells transfected with pHPRT-EfN and pfBSD-OctGFP also incorporated the *cre* expression plasmid. As a control reaction, H1 cells were co-transfected with pHPRT-EfN and pfBSD-OctGFP in the absence of pTurboCre. Microscopic analysis performed 48 hours later showed that a large percentage of the cells expressed GFP (arising from transient expression of pfBSD-OctGFP), indicating DNA uptake (data not shown). gDNA was extracted for PCR analysis 48 hours post-lipofection to allow for expression of *cre* and cassette exchange between the targeting and RMCE constructs. Site-specific recombination between both vectors leads to reciprocal exchange of sequences within the *lox* sites, and produces an 896 bp PCR product with the *hprtloxM* and GFP primers (designed to anneal to one of the two recombination products only) (Figure 7.3 A and B). PCR analysis of gDNA derived from cells treated with all three plasmids showed the presence of the predicted 896 bp product arising from the recombination product A, whereas no signal was detected in the absence of the pTurboCre (figure 7.3 B). These results showed that cassette exchange had occurred between the *hprt* targeting and RMCE

constructs in H1 cells, thus confirming the functionality of the *lox P* and *lox2272* sites. In addition, these data show that pTurboCre is not only able, but is required, to direct site-specific cassette exchange of two plasmids flanked by heterospecific *l*ox sites in H1 cells.



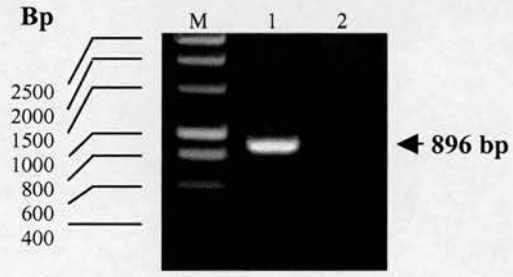


Figure 7.3. Transient recombination assay between pHPRT-EfN and pfBSD-OctGFP in H1 cells. **(A)** H1 cells transiently co-transfected with supercoiled plasmids pfBSD-OctGFP, pHPRT-EfN and pTurboCre. Transient expression of the pTurboCre plasmid results in the generation of two recombination products (A and B) arising from the recombination of each *lox* site with its counterpart. Position of the primers' binding sites are indicated by red arrows. **(B)** PCR analysis of gDNA extracted from cells co-transfected with pHPRT-EfN, pfBSD-OctGFP and pTurboCre (**lane 1**) or with pHPRT-EfN and pfBSD-OctGFP only (**lane 2**). The 896 bp band indicative of RMCE between pHPRT-EfN and pfBSD-OctGFP can be detected only in the presence of pTurboCre. The sizes of the DNA fragments in the marker lane (M) (Bioline hyperladder) are indicated.

7.3.1.4. RMCE Experiments by Electroporation

To achieve RMCE in the H1.HPRT targeted cell lines, we tested two electroporation protocols that resulted in high transfection efficiencies in the *hprt* targeting experiments previously conducted in H1 cells (up to $6x10^{-4}$) (see chapter 4, table 4.1).

An unknown parameter was the optimal molar ratio of pTurboCre:pfBSD-OctGFP plasmids required for efficient RMCE. Three different pTurboCre:pfBSD-OctGFP DNA molar ratios were tested, based on data obtained by A. Kolb (2001) who achieved high recombination frequencies using a 2:1 to a 4:1 pCre:pRMCE DNA molar ratio in mouse ES cells (Kolb, 2001). Kolb introduced, by HR, a *PGK-hprt* expression cassette flanked by heterospecific *lox* sites (*loxP* and *lox 2272*) into the *β-casein* gene in the HPRT-deficient mouse cell line HM1 (Magin *et al.*, 1992). In a subsequent step, a series of RMCE vectors were tested for cassette exchange by site-specific recombination at the modified *β-casein* gene. Following site-specific recombination, the *PGK-hprt* selection marker was exchanged for a new selection marker, providing both a positive selection (presence of the new marker gene) and negative selection (absence of the *hprt* expression cassette). Using this enrichment scheme, absolute RMCE frequencies (ratio between the number of site-specific recombinants and the number of cells electroporated) of approximately 3×10^{-6} were observed, with 100 % of the clones carrying the desired modification.

Based on Kolb's work, the first RMCE experiment was performed using a 4:1 molar ratio of pTurboCre:pfBSDOctGFP. The H1.HPRT-1 and H1.HPRT-2 cell lines were transiently co-transfected with 50 μg of supercoiled pTurboCre and 25 μg of supercoiled pfBSD-OctGFP plasmids using the Eppendorf

Multiporator (300 V/ 100 μs) (table 7.2, experiments 1). Both plasmids were expected to remain episomal and maintain their expression for a few days before being degraded, allowing sufficient accumulation of Cre for a site-specific recombination event at the *hprt* gene. Supercoiled DNA was used to reduce the possibility of random integration (Taniguchi *et al.*, 1998).

Illegitimate recombination was measured by transiently transfecting H1.HPRT-1 and H1.HPRT-2 cells with 25 μg of circular pfBSD-OctGFP in the absence of pTurboCre. Blasticidin selection was applied 48 hours post-transfection and after fourteen days in selection, none of the cells survived selection, suggesting that cassette exchange had not taken place in both the H1.HPRT-1 and H1.HPRT-2 targeted cell lines (table 7.2).

In a second experiment, electroporation conditions (320 V/ 200 μF) established in hES cells by Zwaka and Thomson (Zwaka and Thomson, 2003) were also tested. Additionally the amount of pfBSD-OctGFP was increased to 50 μg to ensure that there was sufficient RMCE plasmid for a site-specific recombination event. Finally, two new molar ratios of pTurboCre:pfBSD-OctGFP (2:1 and 3:1) were tested (table 7.2, experiment 2). An increase in cell death was observed 24 hours post-transfection with both sets of electroporation conditions, possibly as a consequence of DNA toxicity due to an increase of total DNA transfected (up to 138 μg). Following blasticidin selection for 14 days, none of the cells survived, suggesting that site-specific recombination had not occured at the *hprt* gene.

A third experiment was performed, in which a 3:1 pTurboCre:pfBSD-OctGFP molar ratio and a total DNA amount of 50 μg (to minimise DNA toxicity) were used. The plasmid pfBSD-OctGFP was also

transfected in the absence of pTurboCre to control for illegitimate recombination. Blasticidin selection was applied 48 hour later, and 9 days following electroporation 1 blasticidin-resistant colony was obtained. The control reaction in which cells were transfected with pfBSD-OctGFP did not produce any blasticidin-resistant colonies, suggesting successful cassette exchange in the experimental plate (table 7.2, experiment 3). To confirm that this clone was the result of RMCE rather than a clone which had acquired resistance by integration downstream of a random promoter, the colony was expanded for subsequent gDNA analysis (see section 7.3.1.6).

Experiment	Electroporation conditions	Cell line/ passage no	Number of cells	pTurboCre: pfBSDOctGFP molar ratio	pCre: pRMCE (μg)	BSD^r clones	A.R.F
1	300V/100μs	H1 HPRT-1/88	$1x10^6$	4:1	50:25	0	0
		H1 HPRT-2/86	$1x10^6$	4:1	50:25	0	0
2	300V/100 μs 320V/200 μF	H1 HPRT-2/86	$1x10^6$	2:1	59:50	0	0
				3:1	88:50	0	0
3	300V/100 μs	H1HPRT-1/97	$1x10^6$	3:1	31:19	1	10^{-6}

Table 7.2. Summary of RMCE experiments in both the H1.HPRT-1 and H1.HPRT-2 cell lines using electroporation. Absolute RMCE frequency (ARF): ratio between the number of site-specific recombinants and the number of cell electroporated.

7.3.1.5. RMCE Experiments by Lipofection

The recovery of a single blasticidin-resistant colony using an electroporation protocol required a large number of cells and high amount of DNA. Site-specific recombination has previously been demonstrated in mammalian cells using lipofection (e.g. Baubonis and Sauer, 1993; Masui *et al.*, 2005). It was therefore decided to establish a lipofection protocol in order to scale down the experiment such that more transfections could be performed simultaneously. In order to optimise experimental conditions in H1.HPRT targeted cell lines, three independent experiments were performed and the variables cell density, total DNA concentration and DNA molar ratio were investigated.

In the first experiment, H1.HPRT-2 targeted cells were plated at a density of $3x10^5$ cells per well of a 6-well plate and were transfected 48 hours later (~90 % confluent) with 2.5 μg of supercoiled pTurboCre:pfBSD-OctGFP DNA at a 3:1 molar ratio (determined as the effective DNA molar ratio to produce blasticidin-resistant colonies in the electroporation protocol) (table 7.3). 2.5 μg of total DNA was used because an increase of cell death had been observed with greater amounts of DNA (data not shown). Transient GFP expression (arising from the introduction of pfBSD-OctGFP in the H1.HPRT-2 cells) examined by fluorescence microscopy 48 hours post-lipofection, showed that only a small percentage of cells transiently expressed GFP, indicating poor DNA uptake (figure 7.4).

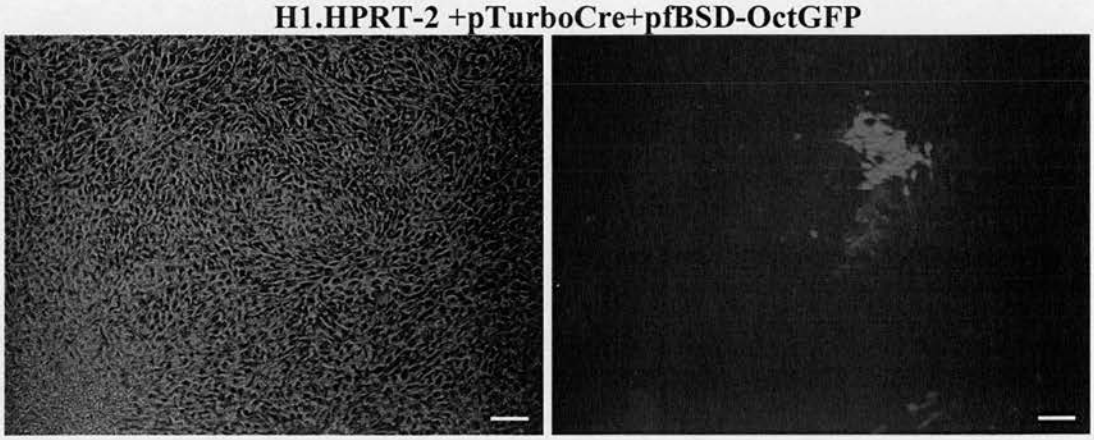


Figure 7.4. RMCE experiment 1 in the H1.HPRT-2 cell line using lipofection. Fluorescence microscopy (right panel) and phase-contrast microscopy (left panel) of H1.HPRT-2 cells transiently expressing GFP 48 hours following lipofection with supercoiled pfBSD-OCTGFP and pTurboCre. The low number of H1.HPRT-2 cells transiently expressing GFP show poor DNA uptake. Scale bars = 200 μm.

H1.HPRT-2 cells co-transfected with pTurboCre and pfBSD-OctGFP were subjected to blasticidin selection 48 hours post-lipofection, resulting in the recovery of 2 blasticidin-resistant colonies in the same well after 7 days in selection (table 7.3). These results show that even with poor DNA uptake, the recovery of blasticidin-resistant colonies is possible. A control experiment in which H1.HPRT-2 cells were transfected with pfBSD-OctGFP alone did not produce any blasticidin-resistant colonies (table 7.3). Together, these results suggest that the 2 blasticidin-resistant clones recovered in the experimental plate were the result of cassette exchange by site-specific recombination at the *hprt* gene rather than illegitimate recombination. During the subcloning procedure, these 2 blasticidin-resistant colonies were lost and thus cassette exchange could not be confirmed.

Cell line / passage number	Number of cells seeded / well	Lipofectamine:DNA ratio (µl:µg)	pTurboCre: pfBSD-OctGFP molar ratio	pTurboCre: pfBSD-octGFP (µg)	BSD^r clones
H1.HPRT-2/ 81	$3x10^5$	4.8:1	3:1	1.6:0.9	2
				0:2.5	0

Table 7.3. RMCE experiment 1 in the H1.HPRT-2 cell line using lipofection.

To determine whether the recovery of blasticidin-resistant colonies was dependent on cell density at the time of lipofection and total DNA amount, different cell densities and DNA amounts were tested in the H1.HPRT-1 cell line. 10^5, $2x10^5$ and $4x10^5$ H1.HPRT-1 cells were seeded per well of a 6-well plate, resulting in ~50%, 80% and 100% confluent cultures, respectively, at the time of lipofection (table 7.4). The H1.HPRT-1 targeted cells were transiently co-transfected with different amounts of the pTurboCre and pfBSD-OctGFP plasmids (3:1 molar ratio) varying from 0 to 8 µg, and blasticidin selection was applied 48 hours post-transfection. Following blasticidin selection for 9 days, 1 blasticidin-resistant colony was recovered in each well treated with 1, 2 and 4 µg of total DNA, at a plating cell density of $2X10^5$ cells per well (80% confluent at the time of lipofection) indicating that these experimental conditions were best suited for colony recovery (table 7.4). No blasticidin-resistant colonies were produced when cells were transfected with pfBSD-OctGFP alone, suggesting that the colonies recovered were the result of cassette exchange at the *hprt* gene. Colonies were expanded for subsequent analysis.

Cell line/ Passage number	pTurboCre: pfBSD-OctGFP molar ratio	Number of cells seeded per well	Total DNA (μg) Blasticidin-resistant colonies					
			0	0.5	1	2	4	8
H1.HPRT-1/ 91	3:1	$1x10^5$	0	0	0	0	0	0
		$2x10^5$	0	0	1	1	1	0
		$4x10^5$	0	0	0	0	0	0

Table 7.4. RMCE experiment 2 in the H1.HPRT-1 cell line using lipofection.

A third experiment was performed to investigate whether an increase of colony recovery was dependent on an increase of pfBSD-OctGFP DNA, while maintaining the total amount of DNA at 2.5 μg. To address this question, a 2:1 and 3:1 pTurboCre:pfBSD-OctGFP molar ratio were compared (table 7.5). Cells were seeded at densities of 10^5, $2x10^5$ or $4x10^5$ per well in duplicate such that prior to transfection, they could be counted to determine the absolute RMCE frequency (ARF) (ratio between the number of site-specific recombinants and the number of cells treated). H1.HPRT-1 cells were co-transfected with 2.5 μg of a 2:1 or 3:1 pTurboCre:pfBSDOctGFP molar ratio. Blasticidin-resistant colonies were recovered at both DNA molar ratios and at both $2x10^5$ and $4x10^5$ plating cell densities (table 7.5). However, the highest number of colonies recovered was again at a plating cell density of $2x10^5$ (~90% confluent at lipofection) with a 2:1 molar ratio of pTurboCre to pfBSD-OctGFP (2 colonies at 2:1 molar ratio versus 1 colony at 3:1 molar ratio). These results suggest that an excess molar ratio of the pTurboCre plasmid over the RMCE plasmid may not be necessary to achieve site-specific recombination in hES cells. These clones were expanded for characterisation.

Cell line/ Passage number	Total DNA (μg)	Number of cells seeded per well	Number of cells/well lipofected	BSD^r clones		ARF	
				2:1*	3:1*	2:1*	3:1*
H1.HPRT-1/ 99	2.5	$1x10^5$	$3x10^5$	0	0	0	0
		$2x10^5$	$5x10^5$	2	1	$4x10^{-6}$	$2x10^{-6}$
		$4x10^5$	$7x10^5$	1	0	$1.4x10^{-6}$	0

Table 7.5. RMCE experiment 3 in the H1.HPRT-1 cell line using lipofection. (*) 2:1 & 3:1 represent the pTurboCre:pfBSDOctGFP molar ratio. ARF (Absolute RMCE frequency): ratio between the number of site-specific recombinants and the number of cells treated per well (see section 7.3.1.6).

Given the data collected from three independent experiments, it was concluded that the most efficient plating cell density was $2x10^5$, such that cells would reach ~80 to 90 % confluency at the time of lipofection. A 2:1 or a 3:1 pTurboCre:pfBSD-OctGFP molar ratio using 2.5 μg of total DNA were both suitable for colony recovery. It is very likely that further optimisation of experimental conditions (e.g. optimal pTurboCre:pfBSD-OctGFP molar ratio, and transfection efficiency) in the H1.HPRT targeted cell lines will improve blasticidin-resistant colony recovery and therefore RMCE frequencies.

7.3.1.6. Detection of an RMCE Event by PCR and Southern Analysis

Five blasticidin-resistant clones isolated from the electroporation and lipofection experiments described in sections 7.3.1.4 and 7.3.1.5 (one from the electroporation, two from the lipofection experiment 2 and two from the lipofection experiment 3) were analysed by PCR amplification of gDNA. The *hprt* gene modified by HR and subsequent site-specific recombination produces a 197 bp PCR product amplified with the internal *BSD* primer and the external *EF1-α* primer (figures 7.5 and 7.6). PCR analysis showed that all the clones yielded the predicted 197 bp fragment consistent with a site-specific recombination event at the *hprt* gene (figure 7.6).

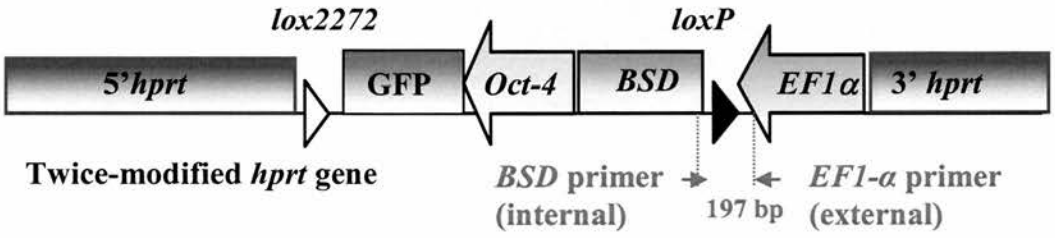


Figure 7.5. Map of the structure of the twice-modified *hprt* gene following an RMCE event. The internal (*BSD*) and external (*EF1-α*) primers and the predicted size of PCR product (197 bp) are indicated.

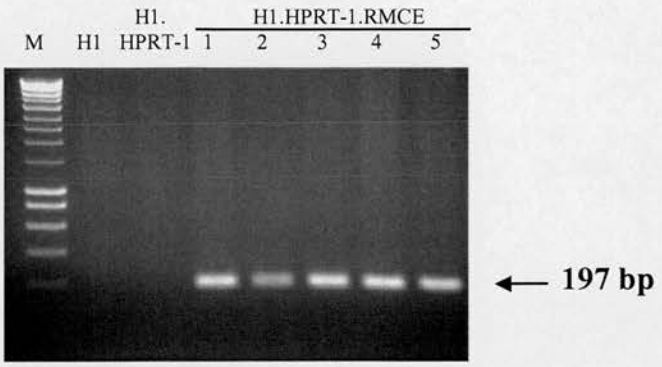


Figure 7.6. PCR screening for an RMCE event at the modified *hprt* gene. PCR analysis of five H1.HPRT-1.RMCE clones obtained by electroporation (H1.HPRT-1.RMCE-1) and lipofection (H1.HPRT-1.RMCE-2, -3, -4 and -5). A 197 bp band can be detected as expected for an RMCE event at the modified *hprt* gene. M: Bioline hyperladder.

To confirm that these clones were the result of site-specific recombination at the *hprt* gene, EcoRI-digested gDNA from two clones (H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE-3) was analysed by Southern blotting. The 1193 bp 3' *hprt* probe (used in the first targeted modification of the *hprt* gene) was predicted to hybridise to a 8.4 kb EcoRI fragment from the wild-type allele, a 6.9 kb fragment from the allele modified by HR and a 9.9 kb fragment from the allele modified by site-specific recombination (figure 7.7 A). Southern analysis showed the presence of the predicted 9.9 kb EcoRI fragment in both the H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE-3 clones, confirming that these clones were the result of site-specific recombination at the modified *hprt* gene (figure 7.7 B).

Taken together, these data show that the promoter trap strategy employed in this study allows the recovery of blasticidin-resistant colonies, of which 100 % carry the desired modification at the *hprt* gene. In addition, absolute RMCE frequencies of 10^{-6} and 4 x 10^{-6} (tables 7.2 and 7.5) can be achieved by

electroporation and lipofection, respectively. Thus the RMCE frequencies obtained in hES cells are in the range of those obtained in mouse ES cells using the same heterospecific *lox* sites and a promoter trap strategy (between 2.4 x 10^{-6} and 1.6 x 10^{-5}) (Kolb, 2001). Finally, absolute RMCE frequencies at the *hprt* gene can compare to absolute targeting frequencies obtained in the first modification of the *hprt* gene by HR in H1 cells (10^{-6}). It is probable that RMCE frequencies will be further enhanced with the development of optimised transfection protocols.

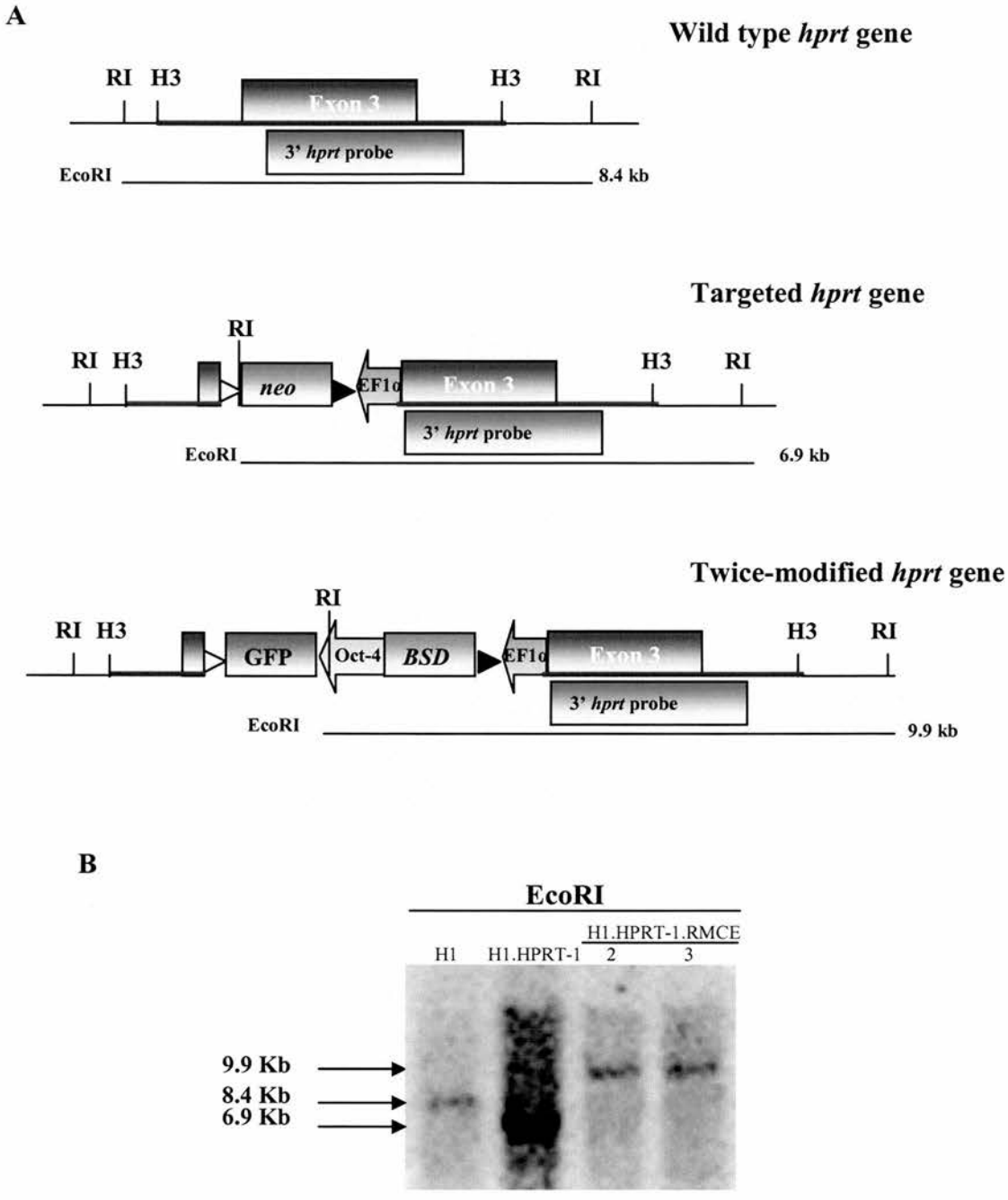


Figure 7.7. Southern screening for a site-specific recombination event at the *hprt* gene. **(A)** Map of the wild-type, once-modified (by HR) and twice-modified (by site-specific recombination) *hprt* gene showing fragments digests diagnostics using a 3' *hprt* probe. **(B)** Southern analysis of H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE–3 clones obtained by lipofection. The 9.9 kb EcoRI fragment confirms that these clones are the result of site-specific recombination at the *hprt* gene.

7.3.2. Transgene Expression Analysis of Site-Specific Recombinants at the *Hprt* Gene

7.3.2.1. GFP Expression Analysis

The aim of this project was to characterise transgene expression at the *hprt* gene by combining gene targeting and site-specific recombination in hES cells. All the site-specific recombinants contain an *EF1-α-BSD-oct-4*GFP transgene at the *hprt* gene (figure 7.1). Because *oct-4* is only expressed in undifferentiated ES cells (Pesce *et al.*, 1999), it was predicted that GFP would be expressed in undifferentiated cells and progressively down regulated upon differentiation. To confirm that GFP was expressed in the site-specific recombinants, GFP expression was examined in undifferentiated ES cells by fluorescence microscopy and flow cytometry analysis.

Fluorescence microscopy showed that all the blasticidin-resistant colonies generated by electroporation and lipofection expressed GFP 9 to 14 days post-transfection. However, following this analysis, progressive down-regulation of GFP expression was observed in all the blasticidin-resistant clones. By 4 to 6 days following the first fluorescence microscopy analysis, no detectable levels of GFP expression were observed. The fact that all the clones remained resistant to blasticidin, indicated that the *BSD-oct-4*-GFP transgene was stably integrated in the genome. Additionally, transient persistence of the pfBSD-OctGFP plasmid in the cells was only expected for 2 to 3 days following lipofection. Since EGFP has a half-life of 24 hours, it seems unlikely that progressive GFP silencing was due to elimination of transient persistence of pfBSD-OctGFP in the blasticidin-resistant clones. A representative fluorescence microscopy analysis is shown on figure 7.8 for one blasticidin-resistant colony obtained by electroporation (H1.HPRT-1.RMCE-1) and two blasticidin-resistant colonies obtained by lipofection (H1.HPRT-1.RMCE-2

and H1.HPRT-1.RMCE-3). These results suggest that either these clones have spontaneously differentiated, and therefore *oct-4* is down-regulated, or that GFP expression is progressively down-regulated possibly by transcriptional interference, a mechanism commonly observed in loci in which two independent transcriptional units are introduced (Kadesch and Berg, 1986; Proudfoot, 1986).

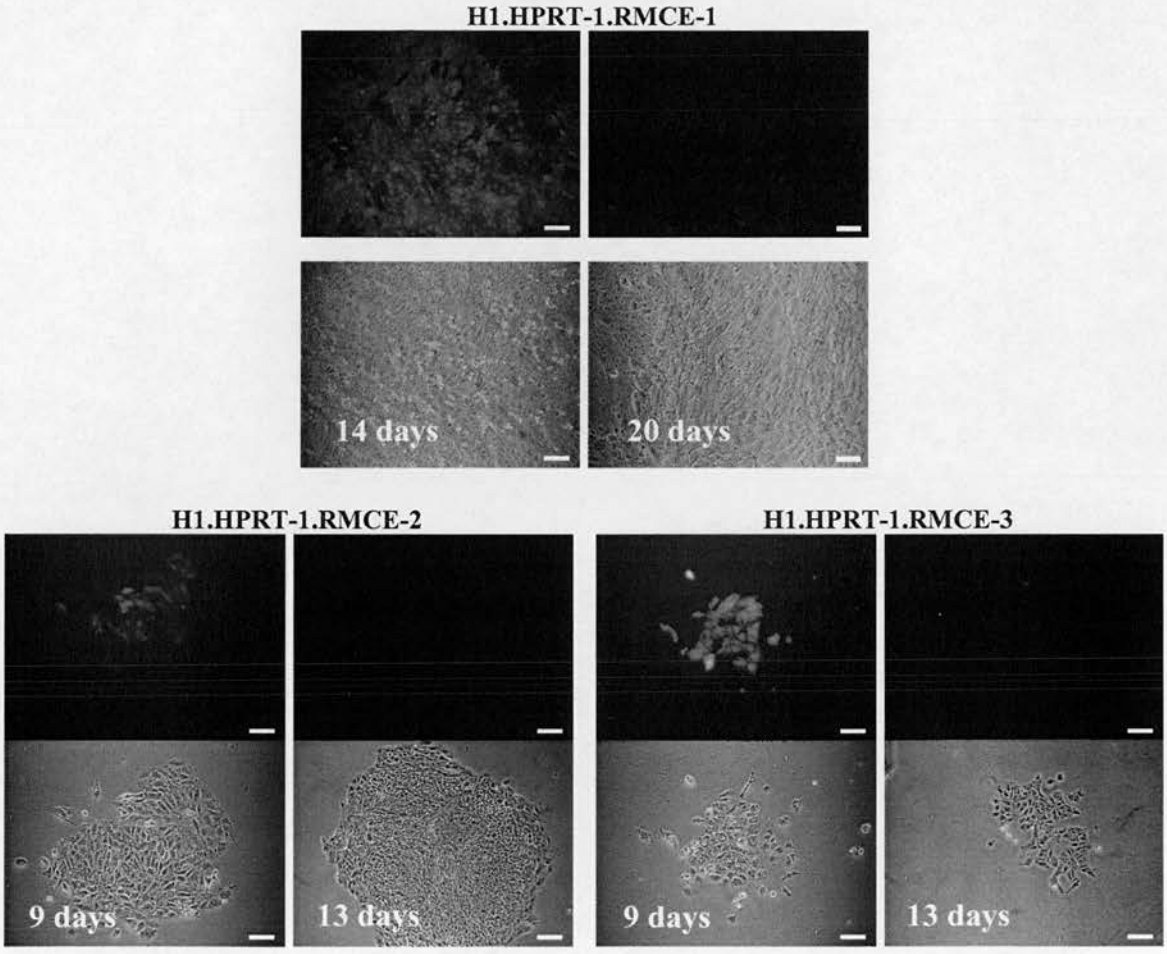


Figure 7.8. Analysis of GFP expression of the *hprt* site-specific recombinants. Fluorescence (top panels) and phase-contrast (bottom panels) microscopy showing down-regulation of GFP expression in site-specific recombinants obtained by electroporation (H1.HPRT-1.RMCE-1) and lipofection (H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE-3). The number of days following transfection is indicated. Scale bars = 100 μm

To confirm down-regulation of GFP expression, flow cytometry analysis was performed on the H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE-3 cell clones that had been expanded over a period of 5 weeks. Untransfected H1 cells were used as a negative control for GFP expression. As a positive control for GFP expression, pooled H1 cells expressing an *oct-4*-GFP transgene (H1.EfBSD-OctGFP cells), were also sorted. The H1.EfBSD-OctGFP cells were generated by stable transfection of H1 cells with the pEfBSD-OctGFP plasmid (a plasmid identical to pfBSD-OctGFP but in which the *BSD* cassette is under the control of the *EF1-α* promoter) (figure 7.9.A). Blasticidin-resistant colonies obtained with pEfBSD-OctGFP were pooled ten days following selection and expanded over a period of two weeks in blasticidin selection before FACS analysis. Pools rather than subclones were studied to minimise transgene expression variation due to site of incorporation. Fluorescence microscopy analysis of the pooled H1.EfBSDOctGFP cells showed variegated GFP expression as evident from the presence of cells not expressing GFP in the pooled population (figure 7.9 B).

While FACS analysis showed that 15 % of H1.EfBSD-OctGFP cells expressed GFP, similar background levels of fluorescence between the wild-type H1 cells and the two *hprt* site-specific recombinants were observed, confirming down-regulation of GFP expression in the *hprt* site-specific recombinants (figure 7.9 C).

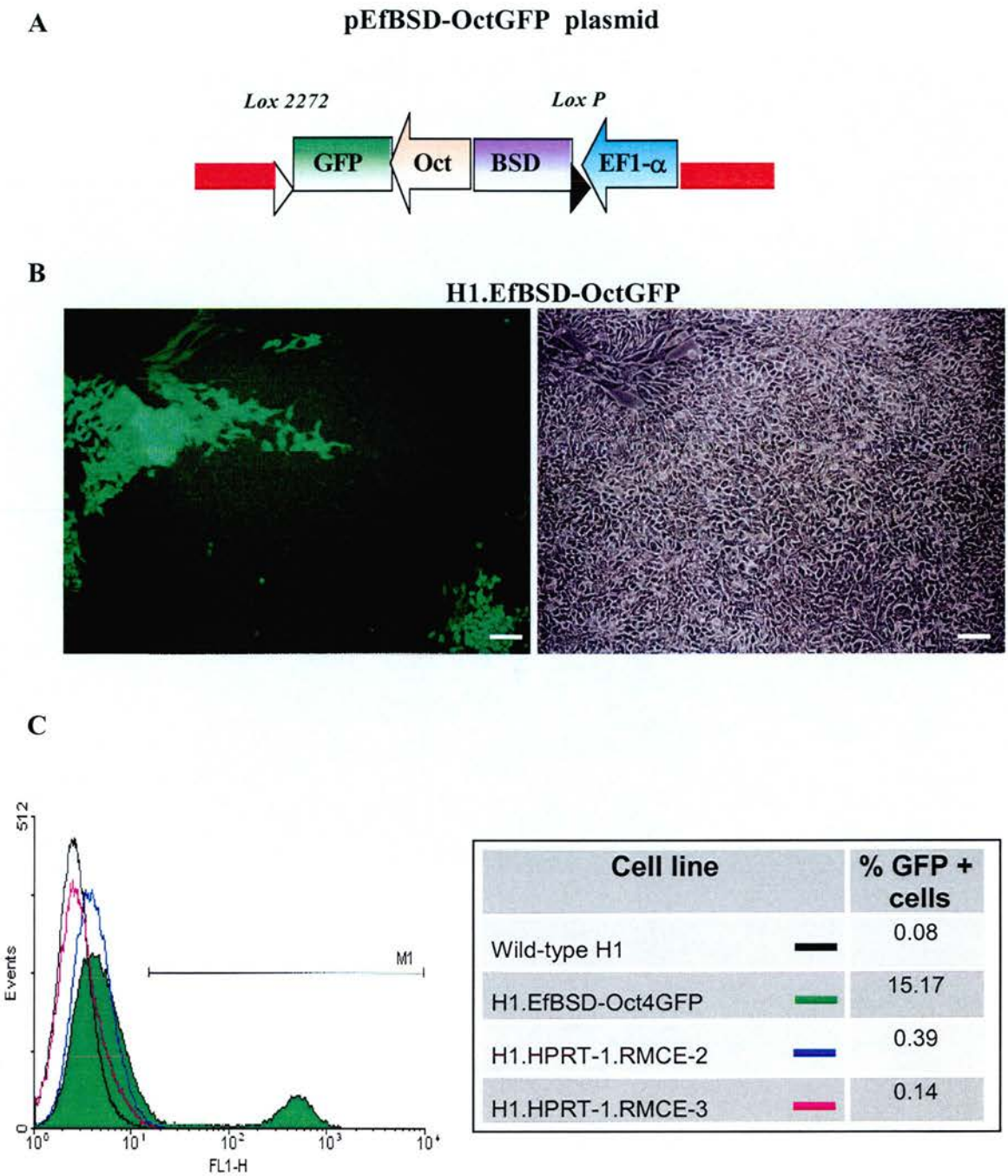


Figure 7.9. Flow cytometric analysis of GFP expression in the H1.HPRT-1.RMCE site-specific recombinants. **(A)** Schematic representation of the pEfBSDOctGFP plasmid used to generate the pooled H1.EfBSDOctGFP cells (positive control for GFP expression). **(B)** Fluorescence (left) and phase-contrast (right) microscopy showing GFP variegation in the H1.EfBSD-OctGFP cells. Scale bars = 200 μm. **(C)** Flow cytometric analysis confirming down-regulation of GFP expression for the two *hprt* site-specific integrants, H1.HPRT-1.RMCE-2 (blue) and H1.HPRT-1.RMCE-3 (pink). Data were generated from the analysis of at least 10, 000 viable cells. The gate M1 indicates the total number of GFP+ cells for each cell line in the graph and table.

7.3.2.2. *Oct-4* Expression Analysis

As described in section 7.3.2.1, the GFP gene is under the transcriptional control of the ES-specific human *oct-4* promoter in the *hprt* site-specific recombinants. GFP silencing observed in the site-specific recombinants may have arisen from spontaneous differentiation of the cell clones possibly as a result of poor quality conditioned medium or as a result of the transfection procedure. To rule out this possibility, endogenous *oct-4* expression was examined by immunocytochemistry (in collaboration with J. Fletcher, Roslin Institute) in the H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE-3 cell clones (figure 7.10). The H1.HPRT-1.RMCE-2 clone was chosen because it exhibits an ES-like phenotype, and is representative of the remaining site-specific recombinants obtained in the course of the RMCE experiments. In contrast, the H1.HPRT-1.RMCE-3 cell clone was selected because it exhibits a differentiated morphology (figure 7.8), and it was predicted that the endogenous *oct-4* would not be expressed. The H1.HPRT-1 cell line was used as a positive control for *oct-4* staining as it was previously shown to express *oct-4* (see chapter 5, section 5.3.2.1). OCT-4 staining showed that the H1.HPRT-1.RMCE-2 cell line expressed the endogenous *oct-4*, suggesting that down-regulation of the transgenic human *oct-4* promoter was not the result of spontaneous differentiation (figure 7.10). However, *oct-4* expression was down regulated in the H1.HPRT-1.RMCE-3 cell line, consistent with the differentiated cellular morphology previously observed by microscopy analysis (figure 7.8). It is therefore very likely that GFP down-regulation was the result of spontaneous differentiation in the H1.HPRT-1.RMCE-3 clone. Because the remaining *hprt* site-specific recombinants exhibited an undifferentiated phenotype throughout expansion

(data not shown), these results argue that spontaneous differentiation was not responsible for GFP silencing in the H1.HPRT-1.RMCE-2 clone and in the remaining *hprt* site-specific recombinants. *Oct-4* expression analysis will however be required to rule out the possibility of spontaneous differentiation in the remaining *hprt* site-specific integrants.

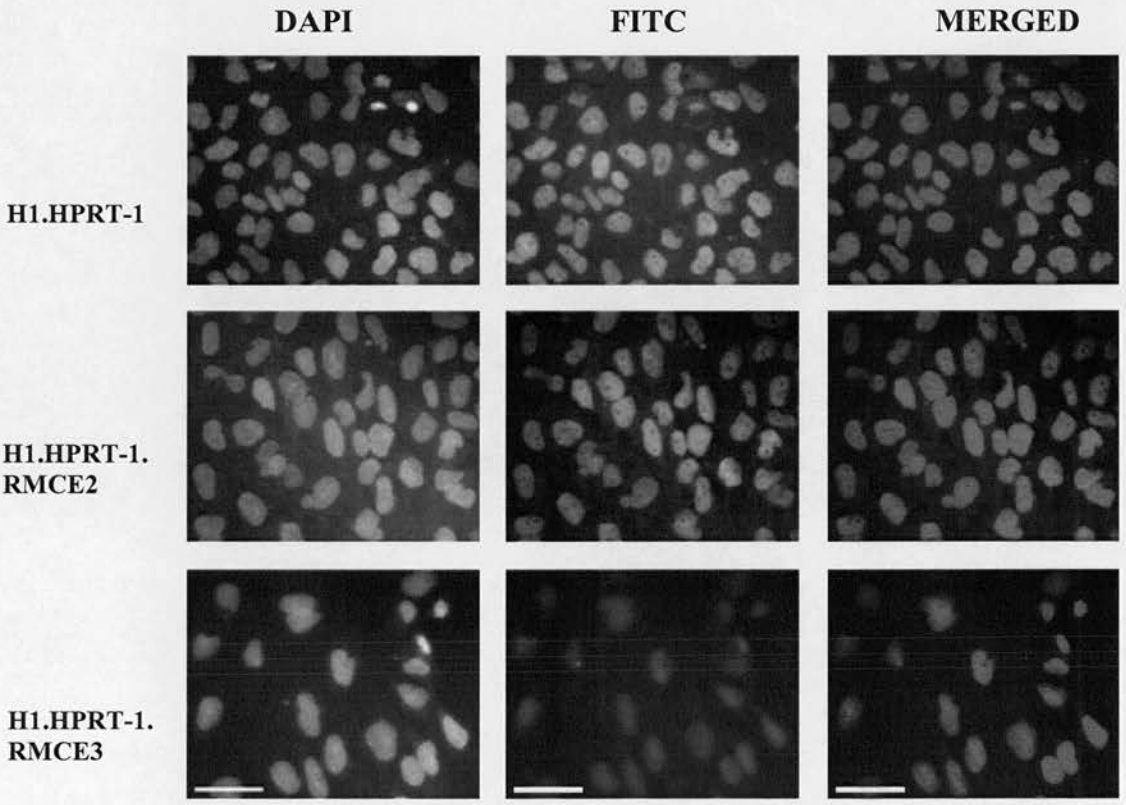


Figure 7.10. *Oct-4* immunocytochemistry of H1.HPRT-1, H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE-3 cell lines. *Oct-4* immnuohistochemistry analysis showing positive expression of the endogenous *Oct-4* gene in the H1.HPRT-1 and H1.HPRT-1.RMCE-2 cell lines, but not in the H1.HPRT-1.RMCE-3 cell line (middle panels). Dapi (4'-6-Diamidino-2-phenylindole), which is a DNA intercalating dye, can be observed in all nuclei (left panels). A merged picture of FITC (green) and Dapi (blue) channels is presented on the right panels. Scale bars = 50 μm.

7.4. Discussion

It was demonstrated for the first time that RMCE can be successfully achieved at the *hprt* gene in hES cells. Using a biphasic recombination strategy (Kolb *et al.*, 1999; Kolb, 2001), the *hprt* gene (exon 3) was first tagged by HR with a *neo* expression cassette flanked by incompatible *lox* sites. In a subsequent step, RMCE allowed the exchange of the *neo* gene with a *BSD-Oct-4*GFP transgene flanked by incompatible *lox* sites. An RMCE rate of 100 % was achieved by coupling site-specific recombination to a promoter trap strategy.

Site-specific recombinants were recovered at the *hprt* gene using either electroporation or lipofection protocols. Under the present experimental conditions, absolute RMCE frequencies of 10^{-6} by electroporation, and up to 4 x 10^{-6} by lipofection were observed at the *hprt* gene in the H1.HPRT-1 cell line. Using a positive and negative selection strategy in mouse ES cells, Kolb (2001) showed that absolute RMCE frequencies between 2.4 x 10^{-6} and 1.6 x 10^{-5} could be achieved at the *β-casein* gene tagged with a *loxP* and *lox2272* sites. Thus the RMCE frequencies observed in hES cells are in the range of those obtained in mouse ES cells.

EGFP expression was analysed following a site-specific recombination event at the *hprt* gene. Because EGFP is under the transcriptional control of the human *oct-4* promoter, the *hprt* site-specific recombinants should express GFP in their undifferentiated state and progressively down-regulate GFP upon differentiation. Fluorescence microscopy analysis performed on three clones 9 days post-lipofection (H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE–3) or 14 days post-electroporation (H1.HPRT-1.RMCE-1) showed that these blasticidin-resistant *hprt* site-specific

recombinants initially expressed GFP. However, progressive down-regulation of GFP was observed in these site-specific recombinants, and by 13 days post-lipofection or 20 days post-electroporation, no GFP-expressing cells could be detected by fluorescent microscopy analysis (figure 7.8). GFP silencing was further confirmed by flow cytometry analysis in the H1.HPRT-1.RMCE-2 and H1.HPRT-1.RMCE–3 site-specific integrants. Although GFP expression was not analysed by flow cytometry for the remaining *hprt* site-specific integrants, progressive GFP down-regulation was also observed in these clones by fluorescence microscopy analysis (data not shown).

Immunocytochemistry analysis of the ES-specific OCT-4 protein showed that the H1.HPRT-1.RMCE-2 site-specific integrant expressed the endogenous *oct-4*. However, endogenous *oct*-4 expression was found to be down-regulated in the H1.HPRT-1.RMCE-3 cell line. Thus while spontaneous differentiation was likely responsible for GFP down-regulation in the H1.HPRT-1.RMCE-3 cell line, GFP silencing was the result of a different mechanism in the H1.HPRT-1.RMCE-2 cell line.

These data may suggest that the *hprt* gene is not a suitable locus to obtain reproducible transgene expression in hES cells. Several studies have however demonstrated reproducible tissue-specific transgene expression at the *hprt* gene in mouse ES cells (Bronson *et al.*, 1996; Cvetkovic *et al.*, 2000; Evans *et al.*, 2000; Guillot *et al.*, 2000; Minami, *et al.*, 2002; Heaney *et al.*, 2004). Certain of these reports also described reproducible transgene expression with transgenes introduced in exon 3 in the opposite transcriptional orientation of the *hprt* gene (Shaw-White *et al.*, 1993; Le *et al.*, 2003).

Although the reason for progressive GFP silencing at the *hprt* locus in hES cells is presently unclear, one possible explanation is that GFP down-regulation was the result of transcriptional interference (TI) between the two transgenic promoters (*EF1-α* and *oct-4*) present at *hprt* locus in the site-specific integrants. TI often occurs when two transcriptionally active units lie adjacent in head-to-tail tandem on a chromosome (Kadesch and Berg, 1986; Proudfoot, 1986; Emerman and Termin, 1986; Shaw-White *et al.*, 1993; Clark *et al.*, 1997; Villemure *et al.*, 2001; Hasegawa and Nakatsuji, 2002; Eszterhas *et al.,* 2002). Interference between two promoters may affect the expression of two transgenes contained in a single construct, and is probably caused by competitive binding of transcription factors and/or modification of DNA structure (induced by transcription) at one site that affect the other site (Emerman and Temin, 1984). Eszterhas *et al.* (2002) developed an elegant system to study TI between two nearly identical transcriptional units (CMV-GFPpA and CMV-YFPpA) juxtaposed in various arrangements (tandem, divergent, and convergent) at a same locus. They used RMCE to insert the constructs into two previously-tagged genomic sites in mouse erythroleukemia (MEL) cells. To study TI, they compared by FACS analysis the mean expression of each gene in tandem constructs in one site to expression of an identical single gene integrated at that site in the same orientation. Their results showed that expression from each of two active transcription units arranged in tandem was less than expression from a single unit in the same integration site and orientation. The downstream gene was down regulated to 10 % (when compared to the expression of the single unit) in one locus, and to a basal transcription level in the other locus. The upstream gene was also down-regulated but maintained from 50 % to 80 % of expression depending on the

locus. Thus, the downstream (3') unit was more affected by TI than the upstream (5') unit. Removal of the CMV promoter from the upstream transcription unit fully restored expression of the down-stream transcription unit, showing that the suppression of the downstream gene is dependent on transcription of the upstream gene.

In our study, the two transcription units present in the *hprt* site-specific recombinants were also arranged in head-to-tail tandem manner (*EF1-α-loxP-BSD-oct-4*-GFP-*lox2272*), as shown on figure 7.1. While the upstream *BSD* gene was always expressed (as indicated by constant resistance to blasticidin selection), the downstream GFP unit was progressively down-regulated. Thus based on Eszterhas *et al.*'s data, it is possible that the expression of the upstream *BSD* gene has affected the expression of the downstream GFP gene at the *hprt* locus in hES cells.

TI from the *BSD* transcription unit could be tested by the elimination of the *EF1-α* promoter and comparison of GFP expression and transcripts production with or without expression of the *BSD* gene. A positive and negative selection system such as hygromycin-thymidine kinase (*hytk*) fusion gene would be required in order to recover *hprt* site-specific recombinants lacking the *EF1-α* promoter. In this strategy, the *hprt* gene would first be tagged with a *loxP-EF1-α-hytk-lox2272* transgene by HR, and targeted clones would be selected in hygromycin and 6-TG. Following site-specific recombination with either *loxP-EF1-α-BSD-Oct-4*-GFP-*lox2272,* or *loxP-BSD-Oct-4*-GFP-*lox2272,* site-specific recombinants would be directly selected in gancyclovir. TI by *BSD* expression would be confirmed if GFP

was expressed in the site-specific recombinants lacking the *EF1-α* promoter, but not in those containing the *EF1-α* promoter.

If GFP silencing observed in the *hprt* site-specific recombinants was the result of TI by *BSD* expression, this could possibly be avoided by introducing in the RMCE construct the core HS4 insulator elements between the two transcriptional units (Hasegawa and Nakatsuji, 2002; Villemure *et al.*, 2001). Hasegawa and Nakatsuji (2002) showed that two expression units in a single construct are expressed in an independent manner and are controlled by their respective regulatory element only if HS4 insulators are placed between the two transcription units. Similarly, Villemure *et al.* (2001) showed that insertion of single copy of the HS4 insulator between two strong ubiquitous transcriptional units in a double gene construct prevents the interference of the upstream gene in the stable transfected cells. Alternatively, a divergent arrangement of the two transcriptional units (*EF1-α-loxP-BSD*-GFP-*Oct-4-lox2272*) may be more suitable for the expression of both units, as demonstrated by Eszterhas *et al.* (2002).

To conclude, we showed that Cre-mediated cassette exchange can be achieved at the *hprt* gene in hES cells with a 100 % rate of site-specific recombinants recovery. Further investigations (discussed above) will however be required to determine the cause for GFP silencing in the site-specific recombinants.

CHAPTER VIII

Concluding Remarks

The primary goal of this project was to generate a hES cell line with reproducible transgene expression for future use in transgenic approaches. This issue was addressed in hES cells, by combining gene targeting and site-specific recombination at the *β-casein* and *hprt* genes. At the start of this project, established efficient transfection protocols adapted for hES cells were not available. Thus transfection efficiencies were too low (Eiges *et al.*, 2001) to be practical in the context of a targeting experiment because of the predominance of illegitimate recombination over homologous recombination. Therefore much of this work has focused on the establishment of targeting and RMCE protocols in hES cells.

The establishment of electroporation protocols with high transfection efficiencies (up to $4.9x10^{-4}$) has permitted the isolation of homologous-recombinants with targeting frequencies of up to 0.4 % at the *hprt* gene in the male hES cell line (H1). Although Southern blotting of the junctions at the 5' and 3' end (using a 5' and 3' external probe) will be required to confirm a correct replacement event for the three *hprt*-targeted hES cell lines generated, our data show that reasonable targeting frequencies can be achieved at an expressed locus in hES cells. In contrast, gene targeting at the unexpressed *β-casein* gene in the female hES cell line (H9) was unsuccessful, despite the high transfection efficiencies achieved (up to $3.5x10^{-4}$). These results may reflect a difference of targeting efficiency between cell lines (this question could easily be answered by conducting a *β-casein* targeting experiment in H1 cells), and/or that gene targeting efficiency in hES cells may be dependent on the locus and/or transcriptional status of the gene. It has been shown that transcription enhances HR in eukaryotes (Nickoloff and Reynolds, 1990). According to this report, failure to target the untranscribed *β-casein* gene would not be unexpected,

since only one targeting event was obtained at the ubiquitously expressed *hprt* gene in each targeting experiment (table 4.1).

Alternatively it is possible that a targeting event has occurred at the *β-casein* locus but targeted clones could not be detected as a result of transgene silencing in the *β-casein* gene, perhaps by heterochromatinisation of the locus. This scenario would be consistent with a previous report in which a transgene was silenced when targeted to a transcriptionally-inactive gene in bovine fetal fibroblast (Kuroiwa *et al.*, 2004). As discussed in chapter 3 (section 3.4), PCR screening may be sensitive enough to detect a targeting event at the *β-casein* locus among small pools of early G418-resistant colonies and therefore should be considered to confirm a targeting event at the *β-casein* locus. If a targeting event was confirmed, then selection could be relaxed as soon as colonies are formed in order to recover targeted clones. The data presented in section 3.3.3 (chapter 3) suggests that insulators HS4 can protect transgene expression from silencing. Insulators flanking the *neo* expression cassette in the targeting vector may therefore provide another means to recover targeted clones at the *β-casein* locus.

Given the low targeting frequencies achieved at the *hprt* locus in hES cells, strategies that have been designed to enhance targeting frequencies in other mammalian systems could be applied in hES cells. For instance, approaches aiming to favour the HR pathway over the illegitimate recombination pathway by controlling the expression of key enzymes involved in recombination and DNA repair, could be explored in the context of this work. It was recently shown in our lab that over-expression of the mammalian recombinase Rad51 (involved in double-strand break DNA repair by HR) or down-regulation of poly(ADP-ribose) polymerase (PARP)

(involved in double strand break DNA repair by illegitimate recombination) increases targeting frequency by 4- and 3.3- fold, respectively, at the *hprt* locus in mouse ES cells (Dominguez-Bendala *et al.*, 2006, Dominguez-Bendala *et al.*, 2003). It would therefore be interesting to determine whether similar enhancement of targeting frequencies could be observed at both the *β-casein* and *hprt* loci in hES cells.

This work has also demonstrated the feasibility of correcting, by HR, the previously modified *hprt* gene, thereby establishing the basic principle that therapeutic cloning combined with gene repair in hES cells could be used as a novel avenue for gene therapy strategies. As outlined in chapter 6 (section 6.4), future work would require the confirmation of a correction event at the modified *hprt* gene by Southern analysis.

We have also shown that the modified *hprt* gene tagged with heterospecific *lox* sites could be further modified by Cre-mediated site-specific recombination, thus exemplifying this alternative approach for the genetic manipulation of hES cells. Furthermore, when using a promoter trap strategy, a 100-fold enrichment for site-specific recombinants was achieved, demonstrating the efficiency of this method to enrich for site-specific recombinants in hES cells. Southern blotting of the junctions at the 5' and 3' end will however be required to confirm a correct site-specific recombination event for all the *hprt* site-specific recombinants. Absolute RMCE frequencies ranging from 10^{-6} (by electroporation) (table 7.2) to $4x10^{-6}$ (by lipofection) (table 7.5) were similar to absolute targeting frequencies obtained in the first modification (10^{-6}) (table 4.1) and correction ($3x10^{-6}$) (table 6.1) of the *hprt* gene. It is strongly believed that RMCE frequency was hampered by the low cell

survival and DNA uptake in the lipofection protocol and that RMCE frequencies could be further enhanced with the improvement of lipofection protocols.

Karyotypic stability, pluripotency, and stemness of an *hprt*-targeted cell line were also issues addressed in this work. Like the parental H1 cell line, an *hprt*-targeted H1 cell line predominantly maintains a normal diploid karyotype, and expresses common hES cell markers (OCT-4, SSEA-4, TRA-1-60 and TRA-1-81). These results showed that gene targeting did not affect the above properties. However, *in vitro* differentiation experiments failed to establish pluripotency of the parental H1 cell line and the *hprt* targeted cell line. Nevertheless, there was no evidence found that the gene targeting procedures affect the *in vitro* differentiation potential of the *hprt* targeted cell line, since pluripotency was not established in the parental H1 cell line. It is possible that the inability to show pluripotency in these cell lines was due to the lack of production of robust EBs, possibly related to the use of the TEG (trypsin/EGTA)-based disaggregation regime. This is based on the observation that collagenase-based disaggregation regime allows the production of robust EBs from H1 cells. Spontaneous *in vitro* differentiation of collagenase-treated H1 cells into the 3 germ layers has been demonstrated without aggregation into EBs under prolonged suboptimal culture conditions (Reubinoff *et al.*, 2000). This work also demonstrated that by adding osteogenic factors to the medium, TEG-treated cells can be induced to form mineralised nodules (consistent with bone formation) *in vitro* without the EB stage. Thus, spontaneous differentiation without aggregation into EBs may be an alternative approach to investigate pluripotency in the parental H1 cell line and the *hprt* targeted cell line passaged with TEG. An important question that remains to be answered is whether the cell lines generated after double

modification of the *hprt* gene either by HR in the case of the *hprt* correction experiment, or by site-specific recombination, retain the properties of hES cells.

The primary goal of this work was to obtain a hES cell line tagged with heterospecific *lox* sites at the *hprt* locus, such that reproducible transgene expression could be achieved with any transgenes flanked with the same heterospecific *lox* sites following a site-specific recombination event at this locus. The data presented in this work, however, showed that a GFP transgene under the control of the human *oct-4* promoter introduced at the *hprt* locus by site-specific recombination, was progressively down-regulated. It is presently unclear whether this was the result of transcriptional interference (TI) from the expression of the upstream *BSD* cassette, or position effect due to the site of integration. As discussed in section 7.4 (chapter 7), future work would require the removal of the promoter (i.e. *EF1-α*) driving the *BSD* gene to confirm TI at the *hprt* gene. Furthermore, a divergent arrangement between the two transcriptional units (*oct-4*-GFP and *EF1-α-BSD*) should also be examined. Indeed, it has been shown that such an arrangement is more suitable for appropriate expression of two transcription units (Eszterhas *et al.,* 2002). Alternatively, GFP silencing in the *hprt* site-specific integrants could possibly be overcome by flanking each transcription unit in the RMCE vector with HS4 insulators, allowing each unit to be expressed independently at the *hprt* locus.

In summary, it has been shown that the *hprt* gene can be modified in hES cells using conventional gene targeting with similar frequencies to those achieved in murine ES cells. Importantly, this work presents evidence that the modified *hprt* locus tagged with heterospecific *lox* sites can be repeatedly modified either by conventional gene targeting (*hprt* correction experiment) or site-specific

recombination (RMCE experiment) in hES cells. Thus, the basic principle that a hES cell line genetically modified at a specific locus can be re-used for further genetic modification has been established. Although there is still a long way to bring genetically modified hES-derived cells to the clinic, an immediate application would be to introduce engineered lineage-specific reporter genes at the *hprt* locus (dependent on whether reproducible transgene expression is eventually achieved at this locus). This would allow the purification of a particular cell type (using for instance FACS separation) and enable the study of gene expression and function in a human developmental context as well as the screening of drug toxicity and provide a source of a desired cell type for transplantation purposes. With respect to the latter, our lab is specifically interested in the purification of bone progenitors cells for transplantation therapy. Therefore, the generation of cell clones tagged with heterospecific *lox* sites at loci that provide reliable transgene expression will be particularly useful for the introduction of a reporter gene (e.g. GFP) under the transcriptional control of a bone-specific promoter (e.g. osteocalcin and *CBFA1*) for the enrichment of osteoblasts. Following induction towards the osteogenic lineage, purification of osteoblats would be achieved by sorting GFP-expressing cells.

References

Abelev, G. I. 1971, "Alpha-fetoprotein in ontogenesis and its association with malignant tumors", *Adv.Cancer Res.*, vol. 14, pp. 295-358.

Abremski, K. & Hoess, R. 1984, "Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein", *J.Biol.Chem.*, vol. 259, no. 3, pp. 1509-1514.

al Shawi, R., Kinnaird, J., Burke, J., & Bishop, J. O. 1990, "Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect", *Mol.Cell Biol*, vol. 10, no. 3, pp. 1192-1198.

Alami, R., Greally, J. M., Tanimoto, K., Hwang, S., Feng, Y. Q., Engel, J. D., Fiering, S., & Bouhassira, E. E. 2000, "Beta-globin YAC transgenes exhibit uniform expression levels but position effect variegation in mice", *Hum.Mol.Genet.*, vol. 9, no. 4, pp. 631-636.

Albertini, R. J. 2001, "HPRT mutations in humans: biomarkers for mechanistic studies", *Mutat.Res.*, vol. 489, no. 1, pp. 1-16.

Albertson, D. G., Collins, C., McCormick, F., & Gray, J. W. 2003, "Chromosome aberrations in solid tumors", *Nat.Genet.*, vol. 34, no. 4, pp. 369-376.

Amit, M., Carpenter, M. K., Inokuma, M. S., Chiu, C. P., Harris, C. P., Waknitz, M. A., Itskovitz-Eldor, J., & Thomson, J. A. 2000, "Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture", *Dev.Biol.*, vol. 227, no. 2, pp. 271-278.

Amit, M., Margulets, V., Segev, H., Shariki, K., Laevsky, I., Coleman, R., & Itskovitz-Eldor, J. 2003, "Human feeder layers for human embryonic stem cells", *Biol.Reprod.*, vol. 68, no. 6, pp. 2150-2156.

Andrade-Rozental, A. F., Rozental, R., Hopperstad, M. G., Wu, J. K., Vrionis, F. D., & Spray, D. C. 2000, "Gap junctions: the "kiss of death" and the "kiss of life"", *Brain Res.Brain Res.Rev.*, vol. 32, no. 1, pp. 308-315.

Andrews, P. W., Damjanov, I., Simon, D., Banting, G. S., Carlin, C., Dracopoli, N. C., & Fogh, J. 1984, "Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2. Differentiation in vivo and in vitro", *Lab Invest*, vol. 50, no. 2, pp. 147-162.

Andrews, P. W., Casper, J., Damjanov, I., Duggan-Keen, M., Giwercman, A., Hata, J., von, K. A., Looijenga, L. H., Millan, J. L., Oosterhuis, J. W., Pera, M., Sawada, M., Schmoll, H. J., Skakkebaek, N. E., van, P. W., & Stern, P. 1996, "Comparative analysis of cell surface antigens expressed by cell lines derived from human germ cell tumours", *Int.J.Cancer*, vol. 66, no. 6, pp. 806-816.

Araki, K., Araki, M., & Yamamura, K. 1997, "Targeted integration of DNA using mutant lox sites in embryonic stem cells", *Nucleic Acids Res.*, vol. 25, no. 4, pp. 868-872.

Arase, H., Mocarski, E. S., Campbell, A. E., Hill, A. B., & Lanier, L. L. 2002, "Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors", *Science*, vol. 296, no. 5571, pp. 1323-1326.

Asakura, A. & Rudnicki, M. A. 2002, "Side population cells from diverse adult tissues are capable of in vitro hematopoietic differentiation", *Exp.Hematol.*, vol. 30, no. 11, pp. 1339-1345.

Askew, G. R., Doetschman, T., & Lingrel, J. B. 1993, "Site-directed point mutations in embryonic stem cells: a gene-targeting tag-and-exchange strategy", *Mol.Cell Biol.*, vol. 13, no. 7, pp. 4115-4124.

Assady, S., Maor, G., Amit, M., Itskovitz-Eldor, J., Skorecki, K. L., & Tzukerman, M. 2001, "Insulin production by human embryonic stem cells", *Diabetes*, vol. 50, no. 8, pp. 1691-1697.

Babinet, C. & Cohen-Tannoudji, M. 2001, "Genome engineering via homologous recombination in mouse embryonic stem (ES) cells: an amazingly versatile tool for the study of mammalian biology", *An.Acad.Bras.Cienc.*, vol. 73, no. 3, pp. 365-383.

Baer, A. & Bode, J. 2001, "Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes", *Curr.Opin.Biotechnol.*, vol. 12, no. 5, pp. 473-480.

Barnes, W. M. 1994, "PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 6, pp. 2216-2220.

Baubonis, W. & Sauer, B. 1993, "Genomic targeting with purified Cre recombinase", *Nucleic Acids Res.*, vol. 21, no. 9, pp. 2025-2029.

Bedell, M. A., Largaespada, D. A., Jenkins, N. A., & Copeland, N. G. 1997, "Mouse models of human disease. Part II: recent progress and future directions", *Genes Dev.*, vol. 11, no. 1, pp. 11-43.

Bertoni, C. & Rando, T. A. 2002, "Dystrophin gene repair in mdx muscle precursor cells in vitro and in vivo mediated by RNA-DNA chimeric oligonucleotides", *Hum.Gene Ther.*, vol. 13, no. 6, pp. 707-718.

Bethke, B. & Sauer, B. 1997, "Segmental genomic replacement by Cre-mediated recombination: genotoxic stress activation of the p53 promoter in single-copy transformants", *Nucleic Acids Res.*, vol. 25, no. 14, pp. 2828-2834.

Bird, A. P. & Wolffe, A. P. 1999, "Methylation-induced repression--belts, braces, and chromatin", *Cell*, vol. 99, no. 5, pp. 451-454.

Blackwell, T. K., Moore, M. W., Yancopoulos, G. D., Suh, H., Lutzker, S., Selsing, E., & Alt, F. W. 1986, "Recombination between immunoglobulin variable region gene segments is enhanced by transcription", *Nature*, vol. 324, no. 6097, pp. 585-589.

Bode, J., Benham, C., Ernst, E., Knopp, A., Marschalek, R., Strick, R., & Strissel, P. 2000a, "Fatal connections: when DNA ends meet on the nuclear matrix", *J.Cell Biochem.Suppl*, vol. Suppl 35, pp. 3-22.

Bode, J., Schlake, T., Iber, M., Schubeler, D., Seibler, J., Snezhkov, E., & Nikolaev, L. 2000b, "The transgeneticist's toolbox: novel methods for the targeted modification of eukaryotic genomes", *Biol.Chem.*, vol. 381, no. 9-10, pp. 801-813.

Bode, J., Benham, C., Knopp, A., & Mielke, C. 2000c, "Transcriptional augmentation: modulation of gene expression by scaffold/matrix-attached regions (S/MAR elements)", *Crit Rev.Eukaryot.Gene Expr.*, vol. 10, no. 1, pp. 73-90.

Bollag, R. J., Waldman, A. S., & Liskay, R. M. 1989, "Homologous recombination in mammalian cells", *Annu.Rev.Genet.*, vol. 23, pp. 199-225.

Bonifer, C., Vidal, M., Grosveld, F., & Sippel, A. E. 1990, "Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice", *EMBO J.*, vol. 9, no. 9, pp. 2843-2848.

Bonifer, C., Yannoutsos, N., Kruger, G., Grosveld, F., & Sippel, A. E. 1994, "Dissection of the locus control function located on the chicken lysozyme gene domain in transgenic mice", *Nucleic Acids Res.*, vol. 22, no. 20, pp. 4202-4210.

Bonini, C., Ferrari, G., Verzeletti, S., Servida, P., Zappone, E., Ruggieri, L., Ponzoni, M., Rossini, S., Mavilio, F., Traversari, C., & Bordignon, C. 1997, "HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia", *Science*, vol. 276, no. 5319, pp. 1719-1724.

Bouhassira, E. E., Westerman, K., & Leboulch, P. 1997, "Transcriptional behavior of LCR enhancer elements integrated at the same chromosomal locus by recombinase-mediated cassette exchange", *Blood*, vol. 90, no. 9, pp. 3332-3344.

Bradley, J. A., Bolton, E. M., & Pedersen, R. A. 2002, "Stem cell medicine encounters the immune system", *Nat.Rev.Immunol.*, vol. 2, no. 11, pp. 859-871.

Brandon, E. P., Idzerda, R. L., & McKnight, G. S. 1995, "Targeting the mouse genome: a compendium of knockouts (Part II)", *Curr.Biol.*, vol. 5, no. 7, pp. 758-765.

Bronson, S. K., Plaehn, E. G., Kluckman, K. D., Hagaman, J. R., Maeda, N., & Smithies, O. 1996, "Single-copy transgenic mice with chosen-site integration", *Proc.Natl.Acad.Sci.U.S.A*, vol. 93, no. 17, pp. 9067-9072.

Buchholz, F., Refaeli, Y., Trumpp, A., & Bishop, J. M. 2000, "Inducible chromosomal translocation of AML1 and ETO genes through Cre/loxP-mediated recombination in the mouse", *EMBO Rep.*, vol. 1, no. 2, pp. 133-139.

Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A., & Felsenfeld, G. 2002, "The insulation of genes from external enhancers and silencing chromatin", *Proc.Natl.Acad.Sci.U.S.A*, vol. 99 Suppl 4, pp. 16433-16437.

Butner, K. & Lo, C. W. 1986, "Modulation of tk expression in mouse pericentromeric heterochromatin", *Mol.Cell Biol*, vol. 6, no. 12, pp. 4440-4449.

Campbell, K. H., J. McWhir, W. A. Ritchie, and I. Wilmut, 1996, "Sheep cloned by nuclear transfer from a cultured cell line", *Nature*, v. 380, no. 6569, p. 64-66.

Capecchi, M. R. 1989, "Altering the genome by homologous recombination", *Science*, vol. 244, no. 4910, pp. 1288-1292.

Carlberg, A. L., Pucci, B., Rallapalli, R., Tuan, R. S., & Hall, D. J. 2001, "Efficient chondrogenic differentiation of mesenchymal cells in micromass culture by retroviral gene transfer of BMP-2", *Differentiation*, vol. 67, no. 4-5, pp. 128-138.

Carpenter, A. T. 1982, "Mismatch repair, gene conversion, and crossing-over in two recombination-defective mutants of Drosophila melanogaster", *Proc.Natl.Acad.Sci.U.S.A*, vol. 79, no. 19, pp. 5961-5965.

Carpenter, M. K., Inokuma, M. S., Denham, J., Mujtaba, T., Chiu, C. P., & Rao, M. S. 2001, "Enrichment of neurons and neural precursors from human embryonic stem cells", *Exp.Neurol.*, vol. 172, no. 2, pp. 383-397.

Carpenter, M. K., Rosler, E. S., Fisk, G. J., Brandenberger, R., Ares, X., Miura, T., Lucero, M., & Rao, M. S. 2004, "Properties of four human embryonic stem cell lines maintained in a feeder-free culture system", *Dev.Dyn.*, vol. 229, no. 2, pp. 243-258.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint, B. G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J. L., Bousso, P., Deist, F. L., & Fischer, A. 2000, "Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease", *Science*, vol. 288, no. 5466, pp. 669-672.

Cearley, J. A. & Detloff, P. J. 2001, "Efficient repetitive alteration of the mouse Huntington's disease gene by management of background in the tag and exchange gene targeting strategy", *Transgenic Res.*, vol. 10, no. 6, pp. 479-488.

Cecconi, F. & Meyer, B. I. 2000, "Gene trap: a way to identify novel genes and unravel their biological function", *FEBS Lett.*, vol. 480, no. 1, pp. 63-71.

Cedar, H. 1988, "DNA methylation and gene activity", *Cell*, vol. 53, no. 1, pp. 3-4.

Cervantes, R. B., Stringer, J. R., Shao, C., Tischfield, J. A., & Stambrook, P. J. 2002, "Embryonic stem cells and somatic cells differ in mutation frequency and type", *Proc.Natl.Acad.Sci.U.S.A*, vol. 99, no. 6, pp. 3586-3590.

Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L., & Jaenisch, R. 1998, "DNA hypomethylation leads to elevated mutation rates", *Nature*, vol. 395, no. 6697, pp. 89-93.

Chen, Y., Yee, D., Dains, K., Chatterjee, A., Cavalcoli, J., Schneider, E., Om, J., Woychik, R. P., & Magnuson, T. 2000, "Genotype-based screen for ENU-induced mutations in mouse embryonic stem cells", *Nat.Genet.*, vol. 24, no. 3, pp. 314-317.

Cheng, L., Hammond, H., Ye, Z., Zhan, X., & Dravid, G. 2003, "Human adult marrow cells support prolonged expansion of human embryonic stem cells in culture", *Stem Cells*, vol. 21, no. 2, pp. 131-142.

Chung, J. H., Whiteley, M., & Felsenfeld, G. 1993, "A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila", *Cell*, vol. 74, no. 3, pp. 505-514.

Chung, J. H., Bell, A. C., & Felsenfeld, G. 1997, "Characterization of the chicken beta-globin insulator", *Proc.Natl.Acad.Sci.U.S.A*, vol. 94, no. 2, pp. 575-580.

Clark, A. J., Bissinger, P., Bullock, D. W., Damak, S., Wallace, R., Whitelaw, C. B., & Yull, F. 1994, "Chromosomal position effects and the modulation of transgene expression", *Reprod.Fertil.Dev.*, vol. 6, no. 5, pp. 589-598.

Clark, A. J., Harold, G., & Yull, F. E. 1997, "Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice", *Nucleic Acids Res.*, vol. 25, no. 5, pp. 1009-1014.

Clarke, A. R. 1994, "Murine genetic models of human disease", *Curr.Opin.Genet.Dev.*, vol. 4, no. 3, pp. 453-460.

Cobellis, G., Nicolaus, G., Iovino, M., Romito, A., Marra, E., Barbarisi, M., Sardiello, M., Di Giorgio, F. P., Iovino, N., Zollo, M., Ballabio, A., & Cortese, R. 2005, "Tagging genes with cassette-exchange sites", *Nucleic Acids Res.*, vol. 33, no. 4, p. e44.

Cogoni, C., Irelan, J. T., Schumacher, M., Schmidhauser, T. J., Selker, E. U., & Macino, G. 1996, "Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation", *EMBO J.*, vol. 15, no. 12, pp. 3153-3163.

Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Gryn, J., Holloman, W. K., & Kmiec, E. B. 1996, "Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide", *Science* , vol. 273, no. 5280, pp. 1386-1389.

Collas, P., Liang, M. R., Vincent, M., & Alestrom, P. 1999, "Active transgenes in zebrafish are enriched in acetylated histone H4 and dynamically associate with RNA Pol II and splicing complexes", *J.Cell Sci.*, vol. 112 (Pt 7), pp. 1045-1054.

Collins, E. C., Pannell, R., Simpson, E. M., Forster, A., & Rabbitts, T. H. 2000, "Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development", *EMBO Rep.*, vol. 1, no. 2, pp. 127-132.

Colman, A., and Kind, A, 2000, "Therapeutic cloning: concepts and practicalities", *Trends Biotechnol.*, v. 18, no. 5, p. 192-196.

Colter, D. C., Class, R., DiGirolamo, C. M., & Prockop, D. J. 2000, "Rapid expansion of recycling stem cells in cultures of plastic-adherent cells from human bone marrow", *Proc.Natl.Acad.Sci.U.S.A*, vol. 97, no. 7, pp. 3213-3218.

Cowan, C. A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J. P., Wang, S., Morton, C. C., McMahon, A. P., Powers, D., & Melton, D. A. 2004, "Derivation of embryonic stem-cell lines from human blastocysts", *N.Engl.J.Med.*, vol. 350, no. 13, pp. 1353-1356.

Cromme, F. V., Airey, J., Heemels, M. T., Ploegh, H. L., Keating, P. J., Stern, P. L., Meijer, C. J., & Walboomers, J. M. 1994, "Loss of transporter protein, encoded by the TAP-1 gene, is highly correlated with loss of HLA expression in cervical carcinomas", *J.Exp.Med.*, vol. 179, no. 1, pp. 335-340.

Cvetkovic, B., Yang, B., Williamson, R. A., & Sigmund, C. D. 2000, "Appropriate tissue- and cell-specific expression of a single copy human angiotensinogen transgene specifically targeted upstream of the HPRT locus by homologous recombination", *J.Biol.Chem.*, vol. 275, no. 2, pp. 1073-1078.

D'Amour, K. A. & Gage, F. H. 2002, "Are somatic stem cells pluripotent or lineage-restricted?", *Nat.Med.*, vol. 8, no. 3, pp. 213-214.

D'Ippolito, G., Diabira, S., Howard, G. A., Menei, P., Roos, B. A., & Schiller, P. C. 2004, "Marrow-isolated adult multilineage inducible (MIAMI) cells, a unique population of postnatal young and old human cells with extensive expansion and differentiation potential", *J.Cell Sci.*, vol. 117, no. Pt 14, pp. 2971-2981.

Dang, Q., Auten, J., & Plavec, I. 2000, "Human beta interferon scaffold attachment region inhibits de novo methylation and confers long-term, copy number-dependent expression to a retroviral vector", *J.Virol.*, vol. 74, no. 6, pp. 2671-2678.

Daniels, P. J. L., Yehaskel, A. S. & Morton, J. B. 1973, "*13th Inter Conf. Antimicrobial Agents & Chemotherapy*", Washington, Abstr. 137

Deng, C. & Capecchi, M. R. 1992, "Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus", *Mol.Cell Biol.*, vol. 12, no. 8, pp. 3365-3371.

Detloff, P. J., Lewis, J., John, S. W., Shehee, W. R., Langenbach, R., Maeda, N., & Smithies, O. 1994, "Deletion and replacement of the mouse adult beta-globin genes by a "plug and socket" repeated targeting strategy", *Mol.Cell Biol*, vol. 14, no. 10, pp. 6936-6943.

DeWitt, N. & Knight, J. 2002, "Biologists question adult stem-cell versatility", *Nature*, vol. 416, no. 6879, p. 354.

Dickinson, P., Dorin, J. R., & Porteous, D. J. 1995, "Modelling cystic fibrosis in the mouse", *Mol.Med.Today*, vol. 1, no. 3, pp. 140-148.

Dickinson, P., Kimber, W. L., Kilanowski, F. M., Webb, S., Stevenson, B. J., Porteous, D. J., & Dorin, J. R. 2000, "Enhancing the efficiency of introducing precise mutations into the mouse genome by hit and run gene targeting", *Transgenic Res.*, vol. 9, no. 1, pp. 55-66.

Dillon, N. & Grosveld, F. 1994, "Chromatin domains as potential units of eukaryotic gene function", *Curr.Opin.Genet.Dev.*, vol. 4, no. 2, pp. 260-264.

Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S., & Smithies, O. 1987, "Targetted correction of a mutant HPRT gene in mouse embryonic stem cells", *Nature*, vol. 330, no. 6148, pp. 576-578.

Doetschman, T., Maeda, N., & Smithies, O. 1988, "Targeted mutation of the Hprt gene in mouse embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 85, no. 22, pp. 8583-8587.

Dominguez-Bendala, J., H. Priddle, A. Clarke, and J. McWhir, 2003, "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.*, v. 286, no. 2, p. 298-307.

Dominguez-Bendala, J., and J. McWhir, 2004, "Enhanced gene targeting frequency in ES cells with low genomic methylation levels", *Transgenic Res.*, v. 13, no. 1, p. 69-74.

Dominguez-Bendala, J., M. Masutani, and J. McWhir, 2006, "Down-regulation of PARP-1, but not of Ku80 or DNA-PK(cs), results in higher gene targeting efficiency", *Cell Biol.Int.*.

Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C. A., Jr., Butel, J. S., & Bradley, A. 1992, "Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours", *Nature*, vol. 356, no. 6366, pp. 215-221.

Dorer, D. R., and S. Henikoff, 1997, "Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans", *Genetics*, v. 147, no. 3, p. 1181-1190.

Dorin, J. R., Inglis, J. D., & Porteous, D. J. 1989, "Selection for precise chromosomal targeting of a dominant marker by homologous recombination", *Science*, vol. 243, no. 4896, pp. 1357-1360.

Draper, J. S., Pigott, C., Thomson, J. A., & Andrews, P. W. 2002, "Surface antigens of human embryonic stem cells: changes upon differentiation in culture", *J.Anat.*, vol. 200, no. Pt 3, pp. 249-258.

Draper, J. S., Smith, K., Gokhale, P., Moore, H. D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T. P., Thomson, J. A., & Andrews, P. W. 2004, "Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells", *Nat.Biotechnol.*, vol. 22, no. 1, pp. 53-54.

Driscoll, M. C., Dobkin, C. S., & Alter, B. P. 1989, "Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites", *Proc.Natl.Acad.Sci.U.S.A*, vol. 86, no. 19, pp. 7470-7474.

Edwards, A., Voss, H., Rice, P., Civitello, A., Stegemann, J., Schwager, C., Zimmermann, J., Erfle, H., Caskey, C. T., & Ansorge, W. 1990, "Automated DNA sequencing of the human HPRT locus", *Genomics*, vol. 6, no. 4, pp. 593-608.

Eiges, R., Schuldiner, M., Drukker, M., Yanuka, O., Itskovitz-Eldor, J., & Benvenisty, N. 2001, "Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells", *Curr.Biol*, vol. 11, no. 7, pp. 514-518.

Ellis, J. & Pannell, D. 2001, "The beta-globin locus control region versus gene therapy vectors: a struggle for expression", *Clin.Genet.* , vol. 59, no. 1, pp. 17-24.

Elsea, S. H. & Lucas, R. E. 2002, "The mousetrap: what we can learn when the mouse model does not mimic the human disease", *ILAR.J.*, vol. 43, no. 2, pp. 66-79.

Emerman, M. & Temin, H. M. 1984, "Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism", *Cell*, vol. 39, no. 3 Pt 2, pp. 449-467.

Emerman, M. & Temin, H. M. 1986, "Comparison of promoter suppression in avian and murine retrovirus vectors", *Nucleic Acids Res.*, vol. 14, no. 23, pp. 9381-9396.

Emery, D. W., Yannaki, E., Tubb, J., & Stamatoyannopoulos, G. 2000, "A chromatin insulator protects retrovirus vectors from chromosomal position effects", *Proc.Natl.Acad.Sci.U.S.A*, vol. 97, no. 16, pp. 9150-9155.

Engebrecht, J., Hirsch, J., & Roeder, G. S. 1990, "Meiotic gene conversion and crossing over: their relationship to each other and to chromosome synapsis and segregation", *Cell*, vol. 62, no. 5, pp. 927-937.

Engel, N., West, A. G., Felsenfeld, G., & Bartolomei, M. S. 2004, "Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations", *Nat.Genet.*, vol. 36, no. 8, pp. 883-888.

Erices, A., Conget, P., & Minguell, J. J. 2000, "Mesenchymal progenitor cells in human umbilical cord blood", *Br.J.Haematol.*, vol. 109, no. 1, pp. 235-242.

Eszterhas, S. K., Bouhassira, E. E., Martin, D. I., & Fiering, S. 2002, "Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position", *Mol.Cell Biol.*, vol. 22, no. 2, pp. 469-479.

Evans, M. J. & Kaufman, M. H. 1981, "Establishment in culture of pluripotential cells from mouse embryos", *Nature*, vol. 292, no. 5819, pp. 154-156.

Evans, V., Hatzopoulos, A., Aird, W. C., Rayburn, H. B., Rosenberg, R. D., & Kuivenhoven, J. A. 2000, "Targeting the Hprt locus in mice reveals differential regulation of Tie2 gene expression in the endothelium", *Physiol Genomics*, vol. 2, no. 2, pp. 67-75.

Faugeron, G. 2000, "Diversity of homology-dependent gene silencing strategies in fungi", *Curr.Opin.Microbiol.*, vol. 3, no. 2, pp. 144-148.

Feinstein, S. C., Ross, S. R., & Yamamoto, K. R. 1982, "Chromosomal position effects determine transcriptional potential of integrated mammary tumor virus DNA", *J.Mol.Biol.*, vol. 156, no. 3, pp. 549-565.

Felsenfeld, G. & Groudine, M. 2003, "Controlling the double helix", *Nature*, vol. 421, no. 6921, pp. 448-453.

Feng, Y. Q., Seibler, J., Alami, R., Eisen, A., Westerman, K. A., Leboulch, P., Fiering, S., & Bouhassira, E. E. 1999, "Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange", *J.Mol.Biol.*, vol. 292, no. 4, pp. 779-785.

Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Miliou, A., Jones, M., & Kioussis, D. 1996, "Locus control region function and heterochromatin-induced position effect variegation", *Science*, vol. 271, no. 5252, pp. 1123-1125.

Flavell, R. B. 1994, "Inactivation of gene expression in plants as a consequence of specific sequence duplication", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 9, pp. 3490-3496.

Forrester, W. C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G., & Groudine, M. 1987, "Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids", *Nucleic Acids Res.*, vol. 15, no. 24, pp. 10159-10177.

Frankel, M. S. 2000, "In search of stem cell policy", *Science*, vol. 287, no. 5457, p. 1397.

Freytag, S. O., Paielli, D., Wing, M., Rogulski, K., Brown, S., Kolozsvary, A., Seely, J., Barton, K., Dragovic, A., & Kim, J. H. 2002, "Efficacy and toxicity of replication-competent adenovirus-mediated double suicide gene therapy in combination with radiation therapy in an orthotopic mouse prostate cancer model", *Int.J.Radiat.Oncol.Biol.Phys.*, vol. 54, no. 3, pp. 873-885.

Friedrich, G. & Soriano, P. 1991, "Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice", *Genes Dev.*, vol. 5, no. 9, pp. 1513-1523.

Fukushige, S. & Sauer, B. 1992, "Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 89, no. 17, pp. 7905-7909.

Gallie, D. R. 1998, "Controlling gene expression in transgenics", *Curr.Opin.Plant Biol.*, vol. 1, no. 2, pp. 166-172.

Garrick, D., Fiering, S., Martin, D. I., & Whitelaw, E. 1998, "Repeat-induced gene silencing in mammals", *Nat.Genet.*, vol. 18, no. 1, pp. 56-59.

Gerasimova, T. I., Byrd, K., & Corces, V. G. 2000, "A chromatin insulator determines the nuclear localization of DNA", *Mol.Cell*, vol. 6, no. 5, pp. 1025-1035.

Ginger, M. R. & Grigor, M. R. 1999, "Comparative aspects of milk caseins", *Comp Biochem.Physiol B Biochem.Mol.Biol.*, vol. 124, no. 2, pp. 133-145.

Goetze, S., Gluch, A., Benham, C., & Bode, J. 2003, "Computational and in vitro analysis of destabilized DNA regions in the interferon gene cluster: potential of predicting functional gene domains", *Biochemistry*, vol. 42, no. 1, pp. 154-166.

Goncz, K. K., Kunzelmann, K., Xu, Z., & Gruenert, D. C. 1998, "Targeted replacement of normal and mutant CFTR sequences in human airway epithelial cells using DNA fragments", *Hum.Mol.Genet.*, vol. 7, no. 12, pp. 1913-1919.

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., & Ruddle, F. H. 1980, "Genetic transformation of mouse embryos by microinjection of purified DNA", *Proc.Natl.Acad.Sci.U.S.A*, vol. 77, no. 12, pp. 7380-7384.

Gossler, A., Doetschman, T., Korn, R., Serfling, E., & Kemler, R. 1986, "Transgenesis by means of blastocyst-derived embryonic stem cell lines", *Proc.Natl.Acad.Sci.U.S.A*, vol. 83, no. 23, pp. 9065-9069.

Graham, F. L., Smiley, J., Russell, W. C., & Nairn, R. 1977, "Characteristics of a human cell line transformed by DNA from human adenovirus type 5", *J.Gen.Virol.*, vol. 36, no. 1, pp. 59-74.

Gronthos, S., Zannettino, A. C., Hay, S. J., Shi, S., Graves, S. E., Kortesidis, A., & Simmons, P. J. 2003, "Molecular and cellular characterisation of highly purified stromal stem cells derived from human bone marrow", *J.Cell Sci.*, vol. 116, no. Pt 9, pp. 1827-1835.

Gropp, M., Itsykson, P., Singer, O., Ben Hur, T., Reinhartz, E., Galun, E., & Reubinoff, B. E. 2003, "Stable genetic modification of human embryonic stem cells by lentiviral vectors", *Mol.Ther.*, vol. 7, no. 2, pp. 281-287.

Grosveld, F., van Assendelft, G. B., Greaves, D. R., & Kollias, G. 1987, "Position-independent, high-level expression of the human beta-globin gene in transgenic mice", *Cell*, vol. 51, no. 6, pp. 975-985.

Grusby, M. J., Auchincloss, H., Jr., Lee, R., Johnson, R. S., Spencer, J. P., Zijlstra, M., Jaenisch, R., Papaioannou, V. E., & Glimcher, L. H. 1993, "Mice lacking major histocompatibility complex class I and class II molecules", *Proc.Natl.Acad.Sci.U.S.A*, vol. 90, no. 9, pp. 3913-3917.

Gschwind, M. & Huber, G. 1998, "Introduction of hereditary disease-associated mutations into the beta-amyloid precursor protein gene of mouse embryonic stem cells: a comparison of homologous recombination methods", *Mol.Cell Biol.*, vol. 18, no. 8, pp. 4651-4658.

Gu, H., Marth, J. D., Orban, P. C., Mossmann, H., & Rajewsky, K. 1994, "Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting", *Science*, vol. 265, no. 5168, pp. 103-106.

Guillot, P. V., Liu, L., Kuivenhoven, J. A., Guan, J., Rosenberg, R. D., & Aird, W. C. 2000, "Targeting of human eNOS promoter to the Hprt locus of mice leads to tissue-restricted transgene expression", *Physiol Genomics*, vol. 2, no. 2, pp. 77-83.

Hanson, K. D. & Sedivy, J. M. 1995, "Analysis of biological selections for high-efficiency gene targeting", *Mol.Cell Biol.*, vol. 15, no. 1, pp. 45-51.

Hasegawa, K. & Nakatsuji, N. 2002, "Insulators prevent transcriptional interference between two promoters in a double gene construct for transgenesis", *FEBS Lett.*, vol. 520, no. 1-3, pp. 47-52.

Hasty, P., Rivera-Perez, J., & Bradley, A. 1991a, "The length of homology required for gene targeting in embryonic stem cells", *Mol.Cell Biol.*, vol. 11, no. 11, pp. 5586-5591.

Hasty, P., Ramirez-Solis, R., Krumlauf, R., & Bradley, A. 1991b, "Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells", *Nature*, vol. 350, no. 6315, pp. 243-246.

Hatada, S., Nikkuni, K., Bentley, S. A., Kirby, S., & Smithies, O. 2000, "Gene correction in hematopoietic progenitor cells by homologous recombination", *Proc.Natl.Acad.Sci.U.S.A*, vol. 97, no. 25, pp. 13807-13811.

Hay, D. C., Sutherland, L., Clark, J., & Burdon, T. 2004, "Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells", *Stem Cells*, vol. 22, no. 2, pp. 225-235.

Heaney, J. D., Rettew, A. N., & Bronson, S. K. 2004, "Tissue-specific expression of a BAC transgene targeted to the Hprt locus in mouse embryonic stem cells", *Genomics*, vol. 83, no. 6, pp. 1072-1082.

Heaney, RP. 2002, "Calcium, In *Principles of bone biology*", J.P. Bilezikian, LR Raisz, and GA Rodan, eds, academic press, san diego, pp. 1325-1337.

Henikoff, S. 1990, "Position-effect variegation after 60 years", *Trends Genet.*, vol. 6, no. 12, pp. 422-426.

Henikoff, S. 1998, "Conspiracy of silence among repeated transgenes", *Bioessays*, vol. 20, no. 7, pp. 532-535.

Hicklin, D. J., Wang, Z., Arienti, F., Rivoltini, L., Parmiani, G., & Ferrone, S. 1998, "beta2-Microglobulin mutations, HLA class I antigen loss, and tumor progression in melanoma", *J.Clin.Invest*, vol. 101, no. 12, pp. 2720-2729.

Hoess, R. H. & Abremski, K. 1984, "Interaction of the bacteriophage P1 recombinase Cre with the recombining site loxP", *Proc.Natl.Acad.Sci.U.S.A*, vol. 81, no. 4, pp. 1026-1029.

Holden, C. & Vogel, G. 2002, "Stem cells. Plasticity: time for a reappraisal?", *Science*, vol. 296, no. 5576, pp. 2126-2129.

Holliday, R. 1964, "A mechanism for gene conversion in fungi", Genet Res. 5:282-304

Hooper, M. L. 1982, "Metabolic co-operation between mammalian cells in culture", *Biochim.Biophys.Acta*, vol. 651, no. 2-3, pp. 85-103.

Hooper, M., Hardy, K., Handyside, A., Hunter, S., & Monk, M. 1987, "HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells", *Nature*, vol. 326, no. 6110, pp. 292-295.

Hooper, M. L. 1992, "Embryonal Stem Cells: Introducing Planned Changes into the Animal Germline", (Harwood, Switzerland).

Houdebine, L. M. 2000, "Transgenic animal bioreactors", *Transgenic Res.*, vol. 9, no. 4-5, pp. 305-320.

Hovatta, O., Mikkola, M., Gertow, K., Stromberg, A. M., Inzunza, J., Hreinsson, J., Rozell, B., Blennow, E., Andang, M., & hrlund-Richter, L. 2003, "A culture system using human foreskin fibroblasts as feeder cells allows production of human embryonic stem cells", *Hum.Reprod.*, vol. 18, no. 7, pp. 1404-1409.

Huber, M. C., Bosch, F. X., Sippel, A. E., & Bonifer, C. 1994, "Chromosomal position effects in chicken lysozyme gene transgenic mice are correlated with suppression of DNase I hypersensitive site formation", *Nucleic Acids Res.*, vol. 22, no. 20, pp. 4195-4201.

Hubner, K., Fuhrmann, G., Christenson, L. K., Kehler, J., Reinbold, R., De La, F. R., Wood, J., Strauss, J. F., III, Boiani, M., & Scholer, H. R. 2003, "Derivation of oocytes from mouse embryonic stem cells", *Science* , vol. 300, no. 5623, pp. 1251-1256.

Huss, R., Lange, C., Weissinger, E. M., Kolb, H. J., & Thalmeier, K. 2000, "Evidence of peripheral blood-derived, plastic-adherent CD34(-/low) hematopoietic stem cell clones with mesenchymal stem cell characteristics", *Stem Cells*, vol. 18, no. 4, pp. 252-260.

Igoucheva, O., Alexeev, V., & Yoon, K. 2001, "Targeted gene correction by small single-stranded oligonucleotides in mammalian cells", *Gene Ther.*, vol. 8, no. 5, pp. 391-399.

Ikeda, H. & Matsumoto, T. 1979, "Transcription promotes recA-independent recombination mediated by DNA-dependent RNA polymerase in Escherichia coli", *Proc.Natl.Acad.Sci.U.S.A*, vol. 76, no. 9, pp. 4571-4575.

Inoue, T., Yamaza, H., Sakai, Y., Mizuno, S., Ohno, M., Hamasaki, N., & Fukumaki, Y. 1999, "Position-independent human beta-globin gene expression mediated by a recombinant adeno-associated virus vector carrying the chicken beta-globin insulator", *J.Hum.Genet.*, vol. 44, no. 3, pp. 152-162.

Inzunza, J., Sahlen, S., Holmberg, K., Stromberg, A. M., Teerijoki, H., Blennow, E., Hovatta, O., & Malmgren, H. 2004, "Comparative genomic hybridization and karyotyping of human embryonic stem cells reveals the occurrence of an isodicentric X chromosome after long-term cultivation", *Mol.Hum.Reprod.*, vol. 10, no. 6, pp. 461-466.

Ishii, K., Arib, G., Lin, C., Van, H. G., & Laemmli, U. K. 2002, "Chromatin boundaries in budding yeast: the nuclear pore connection", *Cell*, vol. 109, no. 5, pp. 551-562.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., & Benvenisty, N. 2000, "Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers", *Mol.Med*, vol. 6, no. 2, pp. 88-95.

Jaenisch, R., Jahner, D., Nobis, P., Simon, I., Lohler, J., Harbers, K., & Grotkopp, D. 1981, "Chromosomal position and activation of retroviral genomes inserted into the germ line of mice", *Cell*, vol. 24, no. 2, pp. 519-529.

Jasin, M. & Berg, P. 1988, "Homologous integration in mammalian cells without target gene selection", *Genes Dev.*, vol. 2, no. 11, pp. 1353-1363.

Jasin, M., Moynahan, M. E., & Richardson, C. 1996, "Targeted transgenesis", *Proc.Natl.Acad.Sci.U.S.A*, vol. 93, no. 17, pp. 8804-8808.

Jeannotte, L., Ruiz, J. C., & Robertson, E. J. 1991, "Low level of Hox1.3 gene expression does not preclude the use of promoterless vectors to generate a targeted gene disruption. off", *Mol.Cell Biol.*, vol. 11, no. 11, pp. 5578-5585.

Jiang, Y., Jahagirdar, B. N., Reinhardt, R. L., Schwartz, R. E., Keene, C. D., Ortiz-Gonzalez, X. R., Reyes, M., Lenvik, T., Lund, T., Blackstad, M., Du, J., Aldrich, S., Lisberg, A., Low, W. C., Largaespada, D. A., & Verfaillie, C. M. 2002a, "Pluripotency of mesenchymal stem cells derived from adult marrow", *Nature*, vol. 418, no. 6893, pp. 41-49.

Jiang, Y., Vaessen, B., Lenvik, T., Blackstad, M., Reyes, M., & Verfaillie, C. M. 2002b, "Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain", *Exp.Hematol.*, vol. 30, no. 8, pp. 896-904.

Johnson, R. S., Sheng, M., Greenberg, M. E., Kolodner, R. D., Papaioannou, V. E., & Spiegelman, B. M. 1989, "Targeting of nonexpressed genes in embryonic stem cells via homologous recombination", *Science*, vol. 245, no. 4923, pp. 1234-1236.

Joyce, C. M. & Grindley, N. D. 1983, "Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I of Escherichia coli", *Proc.Natl.Acad.Sci.U.S.A*, vol. 80, no. 7, pp. 1830-1834.

Joyner, A. L., Skarnes, W. C., & Rossant, J. 1989, "Production of a mutation in mouse En-2 gene by homologous recombination in embryonic stem cells", *Nature*, vol. 338, no. 6211, pp. 153-156.

Kadesch, T. & Berg, P. 1986, "Effects of the position of the simian virus 40 enhancer on expression of multiple transcription units in a single plasmid", *Mol.Cell Biol.*, vol. 6, no. 7, pp. 2593-2601.

Kanda, S., Shiroi, A., Ouji, Y., Birumachi, J., Ueda, S., Fukui, H., Tatsumi, K., Ishizaka, S., Takahashi, Y., & Yoshikawa, M. 2003, "In vitro differentiation of hepatocyte-like cells from embryonic stem cells promoted by gene transfer of hepatocyte nuclear factor 3 beta", *Hepatol.Res.*, vol. 26, no. 3, pp. 225-231.

Karp, J. M., Ferreira, L. S., Khademhosseini, A., Kwon, A. H., Yeh, J., & Langer, R. 2005, "Cultivation of human embryonic stem cells without the embryoid body step enhances osteogenesis in vitro", *Stem Cells*.

Karpen, G. H. 1994, "Position-effect variegation and the new biology of heterochromatin", *Curr.Opin.Genet.Dev.*, vol. 4, no. 2, pp. 281-291.

Kaufman, D. S., Hanson, E. T., Lewis, R. L., Auerbach, R., & Thomson, J. A. 2001, "Hematopoietic colony-forming cells derived from human embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 98, no. 19, pp. 10716-10721.

Keefer, C. L. 2004, "Production of bioproducts through the use of transgenic animal models", *Anim Reprod.Sci.*, vol. 82-83, pp. 5-12.

Kehat, I., Kenyagin-Karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-Eldor, J., & Gepstein, L. 2001, "Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes", *J.Clin.Invest*, vol. 108, no. 3, pp. 407-414.

Kehat, I., Amit, M., Gepstein, A., Huber, I., Itskovitz-Eldor, J., & Gepstein, L. 2003, "Development of cardiomyocytes from human ES cells", *Methods Enzymol.*, vol. 365, pp. 461-473.

Kelly, W. G. & Fire, A. 1998, "Chromatin silencing and the maintenance of a functional germline in Caenorhabditis elegans", *Development*, vol. 125, no. 13, pp. 2451-2456.

Kilby, N. J., Snaith, M. R., & Murray, J. A. 1993, "Site-specific recombinases: tools for genome engineering", *Trends Genet.*, vol. 9, no. 12, pp. 413-421.

Kim, H. S., Popovich, B. W., Shehee, W. R., Shesely, E. G., & Smithies, O. 1991, "Problems encountered in detecting a targeted gene by the polymerase chain reaction", *Gene*, vol. 103, no. 2, pp. 227-233.

Kim, Y. G., Cha, J., & Chandrasegaran, S. 1996, "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain", *Proc.Natl.Acad.Sci.U.S.A*, vol. 93, no. 3, pp. 1156-1160.

Kimura, M., Takatsuki, A., & Yamaguchi, I. 1994, "Blasticidin S deaminase gene from Aspergillus terreus (BSD): a new drug resistance gene for transfection of mammalian cells", *Biochim.Biophys.Acta*, vol. 1219, no. 3, pp. 653-659.

Kioussis, D., Vanin, E., deLange, T., Flavell, R. A., & Grosveld, F. G. 1983, "Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia", *Nature*, vol. 306, no. 5944, pp. 662-666.

Kleinsmith, L. J. & Pierce, G. B., Jr. 1964, "Multipotentiality of single embryonal carcinoma cells", *Cancer Res.*, vol. 24, pp. 1544-1551.

Klug, M. G., Soonpaa, M. H., Koh, G. Y., & Field, L. J. 1996, "Genetically selected cardiomyocytes from differentiating embronic stem cells form stable intracardiac grafts", *J.Clin.Invest*, vol. 98, no. 1, pp. 216-224.

Kolb, A. F. & Siddell, S. G. 1996, "Genomic targeting with an MBP-Cre fusion protein", *Gene*, vol. 183, no. 1-2, pp. 53-60.

Kolb, A. F. & Siddell, S. G. 1997, "Genomic targeting of a bicistronic DNA fragment by Cre-mediated site-specific recombination", *Gene*, vol. 203, no. 2, pp. 209-216.

Kolb, A. F., Ansell, R., McWhir, J., & Siddell, S. G. 1999, "Insertion of a foreign gene into the beta-casein locus by Cre-mediated site-specific recombination", *Gene*, vol. 227, no. 1, pp. 21-31.

Kolb, A. F. 2001, "Selection-marker-free modification of the murine beta-casein gene using a lox2272 [correction of lox2722] site", *Anal.Biochem.*, vol. 290, no. 2, pp. 260-271.

Koller, B. H. & Smithies, O. 1989, "Inactivating the beta 2-microglobulin locus in mouse embryonic stem cells by homologous recombination", *Proc.Natl.Acad.Sci.U.S.A*, vol. 86, no. 22, pp. 8932-8935.

Koller, B. H., Hagemann, L. J., Doetschman, T., Hagaman, J. R., Huang, S., Williams, P. J., First, N. L., Maeda, N., & Smithies, O. 1989, "Germ-line transmission of a planned alteration made in a hypoxanthine phosphoribosyltransferase gene by homologous recombination in embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 86, no. 22, pp. 8927-8931.

Kren, B. T. & Steer, C. J. 2002, "The application of DNA repair vectors to gene therapy", *Curr.Opin.Biotechnol.*, vol. 13, no. 5, pp. 473-481.

Kuci, S., Wessels, J. T., Buhring, H. J., Schilbach, K., Schumm, M., Seitz, G., Loffler, J., Bader, P., Schlegel, P. G., Niethammer, D., & Handgretinger, R. 2003, "Identification of a novel class of human adherent", *Blood*, vol. 101, no. 3, pp. 869-876.

Kuehn, M. R., Bradley, A., Robertson, E. J., & Evans, M. J. 1987, "A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice", *Nature*, vol. 326, no. 6110, pp. 295-298.

Kumar, S., Clarke, A. R., Hooper, M. L., Horne, D. S., Law, A. J., Leaver, J., Springbett, A., Stevenson, E., & Simons, J. P. 1994, "Milk composition and lactation of beta-casein-deficient mice", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 13, pp. 6138-6142.

Kuroiwa, Y., Kasinathan, P., Matsushita, H., Sathiyaselan, J., Sullivan, E. J., Kakitani, M., Tomizuka, K., Ishida, I., & Robl, J. M. 2004, "Sequential targeting of the genes encoding immunoglobulin-mu and prion protein in cattle", *Nat.Genet.*, vol. 36, no. 7, pp. 775-780.

Kyba, M., Perlingeiro, R. C., & Daley, G. Q. 2002, "HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors", *Cell*, vol. 109, no. 1, pp. 29-37.

Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D., & Costantini, F. D. 1983, "A foreign beta-globin gene in transgenic mice: integration at abnormal chromosomal positions and expression in inappropriate tissues", *Cell*, vol. 34, no. 2, pp. 343-358.

Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R., & Berns, A. 1991, "Simplified mammalian DNA isolation procedure", *Nucleic Acids Res.*, vol. 19, no. 15, p. 4293.

Langer, S. J., Ghafoori, A. P., Byrd, M., & Leinwand, L. 2002, "A genetic screen identifies novel non-compatible loxP sites", *Nucleic Acids Res.*, vol. 30, no. 14, pp. 3067-3077.

Lanzendorf, S. E., Boyd, C. A., Wright, D. L., Muasher, S., Oehninger, S., & Hodgen, G. D. 2001, "Use of human gametes obtained from anonymous donors for the production of human embryonic stem cell lines", *Fertil.Steril.*, vol. 76, no. 1, pp. 132-137.

Le, M. H., Lallemand, Y., & Brulet, P. 1990, "Targeted replacement of the homeobox gene Hox-3.1 by the Escherichia coli lacZ in mouse chimeric embryos", *Proc.Natl.Acad.Sci.U.S.A*, vol. 87, no. 12, pp. 4712-4716.

Le, Y., Gagneten, S., Larson, T., Santha, E., Dobi, A., Agoston, D., & Sauer, B. 2003, "Far-upstream elements are dispensable for tissue-specific proenkephalin expression using a Cre-mediated knock-in strategy", *J.Neurochem.*, vol. 84, no. 4, pp. 689-697.

Lee, G. & Saito, I. 1998, "Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination", *Gene*, vol. 216, no. 1, pp. 55-65.

Lee, J. B., Lee, J. E., Park, J. H., Kim, S. J., Kim, M. K., Roh, S. I., & Yoon, H. S. 2005, "Establishment and maintenance of human embryonic stem cell lines on human feeder cells derived from uterine endometrium under serum-free condition", *Biol.Reprod.*, vol. 72, no. 1, pp. 42-49.

Lee, M. S. & Garrard, W. T. 1991, "Transcription-induced nucleosome 'splitting': an underlying structure for DNase I sensitive chromatin", *EMBO J.*, vol. 10, no. 3, pp. 607-615.

Lee, R. S., Grusby, M. J., Laufer, T. M., Colvin, R., Glimcher, L. H., & Auchincloss, H., Jr. 1997, "CD8+ effector cells responding to residual class I antigens, with help from CD4+ cells stimulated indirectly, cause rejection of "major histocompatibility complex-deficient" skin grafts", *Transplantation*, vol. 63, no. 8, pp. 1123-1133.

Levenberg, S., Golub, J. S., Amit, M., Itskovitz-Eldor, J., & Langer, R. 2002, "Endothelial cells derived from human embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 99, no. 7, pp. 4391-4396.

Lewandoski, M. 2001, "Conditional control of gene expression in the mouse", *Nat.Rev.Genet.*, vol. 2, no. 10, pp. 743-755.

Lewis, E. B. 1950, "The phenomenon of position effect", *Adv.Genet.*, vol. 3, pp. 73-115.

Li, M., Pevny, L., Lovell-Badge, R., & Smith, A. 1998, "Generation of purified neural precursors from embryonic stem cells by lineage selection", *Curr.Biol.*, vol. 8, no. 17, pp. 971-974.

Li, Q., Peterson, K. R., Fang, X., & Stamatoyannopoulos, G. 2002, "Locus control regions", *Blood*, vol. 100, no. 9, pp. 3077-3086.

Li, X. G., Yan, H. H., Liu, D. P., Hao, D. L., & Liang, C. C. 2003, "Cre-mediated site-specific cassette exchange in erythroid cell", *Sheng Wu Hua Xue.Yu Sheng Wu Wu Li Xue.Bao.(Shanghai)*, vol. 35, no. 10, pp. 947-951.

Li, X. J., Du, Z. W., Zarnowska, E. D., Pankratz, M., Hansen, L. O., Pearce, R. A., & Zhang, S. C. 2005, "Specification of motoneurons from human embryonic stem cells", *Nat.Biotechnol.*, vol. 23, no. 2, pp. 215-221.

Lien, L. L., Lee, Y., & Orkin, S. H. 1997, "Regulation of the myeloid-cell-expressed human gp91-phox gene as studied by transfer of yeast artificial chromosome clones into embryonic stem cells: suppression of a variegated cellular pattern of expression requires a full complement of distant cis elements", *Mol.Cell Biol.*, vol. 17, no. 4, pp. 2279-2290.

Lin, J. J., Matsumura, F., & Yamashiro-Matsumura, S. 1984, "Tropomyosin-enriched and alpha-actinin-enriched microfilaments isolated from chicken embryo fibroblasts by monoclonal antibodies", *J.Cell Biol.*, vol. 98, no. 1, pp. 116-127.

Lin, S. P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., & Ferguson-Smith, A. C. 2003, "Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12", *Nat.Genet.*, vol. 35, no. 1, pp. 97-102.

Lindsay, E. A., Botta, A., Jurecic, V., Carattini-Rivera, S., Cheah, Y. C., Rosenblatt, H. M., Bradley, A., & Baldini, A. 1999, "Congenital heart disease in mice deficient for the DiGeorge syndrome region", *Nature*, vol. 401, no. 6751, pp. 379-383.

Liu, Y. P., Dambaeva, S. V., Dovzhenko, O. V., Garthwaite, M. A., & Golos, T. G. 2005, "Stable Plasmid-based siRNA Silencing of Gene Expression in Human Embryonic Stem Cells", *Stem Cells Dev.*, vol. 14, no. 5, pp. 487-492.

Ma, Y., Ramezani, A., Lewis, R., Hawley, R. G., & Thomson, J. A. 2003, "High-level sustained transgene expression in human embryonic stem cells using lentiviral vectors", *Stem Cells*, vol. 21, no. 1, pp. 111-117.

Magin, T. M., McWhir, J., & Melton, D. W. 1992, "A new mouse embryonic stem cell line with good germ line contribution and gene targeting frequency", *Nucleic Acids Res.*, vol. 20, no. 14, pp. 3795-3796.

Maitra, A., Arking, D. E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., Sui, G., Cutler, D. J., Liu, Y., Brimble, S. N., Noaksson, K., Hyllner, J., Schulz, T. C., Zeng, X., Freed, W. J., Crook, J., Abraham, S., Colman, A., Sartipy, P., Matsui, S. I., Carpenter, M., Gazdar, A. F., Rao, M., & Chakravarti, A. 2005, "Genomic alterations in cultured human embryonic stem cells", *Nat.Genet.*, vol. 37, no. 10, pp. 1099-1103.

Mansour, S. L., Thomas, K. R., & Capecchi, M. R. 1988, "Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes", *Nature*, vol. 336, no. 6197, pp. 348-352.

Mansour, S. L., Thomas, K. R., Deng, C. X., & Capecchi, M. R. 1990, "Introduction of a lacZ reporter gene into the mouse int-2 locus by homologous recombination", *Proc.Natl.Acad.Sci.U.S.A*, vol. 87, no. 19, pp. 7688-7692.

Martin, D. I. & Whitelaw, E. 1996, "The vagaries of variegating transgenes", *Bioessays*, vol. 18, no. 11, pp. 919-923.

Martin, G. R. 1981, "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 78, no. 12, pp. 7634-7638.

Masui, S., Shimosato, D., Toyooka, Y., Yagi, R., Takahashi, K., & Niwa, H. 2005, "An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit", *Nucleic Acids Res.*, vol. 33, no. 4, p. e43.

Matsui, Y., Zsebo, K., & Hogan, B. L. 1992, "Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture", *Cell*, vol. 70, no. 5, pp. 841-847.

Matzke, M. A., Mette, M. F., Aufsatz, W., Jakowitsch, J., & Matzke, A. J. 1999, "Host defenses to parasitic sequences and the evolution of epigenetic control mechanisms", *Genetica*, vol. 107, no. 1-3, pp. 271-287.

Matzke, M. A., Mette, M. F., & Matzke, A. J. 2000, "Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates", *Plant Mol.Biol.*, vol. 43, no. 2-3, pp. 401-415.

Matzke, M., Mette, M. F., Jakowitsch, J., Kanno, T., Moscone, E. A., van der, W. J., & Matzke, A. J. 2001, "A test for transvection in plants: DNA pairing may lead to trans-activation or silencing of complex heteroalleles in tobacco", *Genetics*, vol. 158, no. 1, pp. 451-461.

McBurney, M. W., Mai, T., Yang, X., & Jardine, K. 2002, "Evidence for repeat-induced gene silencing in cultured Mammalian cells: inactivation of tandem repeats of transfected genes", *Exp.Cell Res.*, vol. 274, no. 1, pp. 1-8.

McKnight, R. A., Shamay, A., Sankaran, L., Wall, R. J., & Hennighausen, L. 1992, "Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice", *Proc.Natl.Acad.Sci.U.S.A*, vol. 89, no. 15, pp. 6943-6947.

McKnight, S. L. 1980, "The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene", *Nucleic Acids Res.* , vol. 8, no. 24, pp. 5949-5964.

Melton, D. W. 1994, "Gene targeting in the mouse", *Bioessays*, vol. 16, no. 9, pp. 633-638.

Meselson, M. S. & Radding, C. M. 1975, "A general model for genetic recombination", *Proc.Natl.Acad.Sci.U.S.A*, vol. 72, no. 1, pp. 358-361.

Metzger, D. & Feil, R. 1999, "Engineering the mouse genome by site-specific recombination", *Curr.Opin.Biotechnol.*, vol. 10, no. 5, pp. 470-476.

Michalowski, S. M., Allen, G. C., Hall, G. E., Jr., Thompson, W. F., & Spiker, S. 1999, "Characterization of randomly-obtained matrix attachment regions (MARs) from higher plants", *Biochemistry*, vol. 38, no. 39, pp. 12795-12804.

Minami, T., Donovan, D. J., Tsai, J. C., Rosenberg, R. D., & Aird, W. C. 2002, "Differential regulation of the von Willebrand factor and Flt-1 promoters in the endothelium of hypoxanthine phosphoribosyltransferase-targeted mice", *Blood*, vol. 100, no. 12, pp. 4019-4025.

Mitalipova, M., Calhoun, J., Shin, S., Wininger, D., Schulz, T., Noggle, S., Venable, A., Lyons, I., Robins, A., & Stice, S. 2003, "Human embryonic stem cell lines derived from discarded embryos", *Stem Cells*, vol. 21, no. 5, pp. 521-526.

Mitalipova, M. M., Rao, R. R., Hoyer, D. M., Johnson, J. A., Meisner, L. F., Jones, K. L., Dalton, S., & Stice, S. L. 2005, "Preserving the genetic integrity of human embryonic stem cells", *Nat.Biotechnol.*, vol. 23, no. 1, pp. 19-20.

Mizushima, S. & Nagata, S. 1990, "pEF-BOS, a powerful mammalian expression vector", *Nucleic Acids Res.*, vol. 18, no. 17, p. 5322.

Moore, R. C., Redhead, N. J., Selfridge, J., Hope, J., Manson, J. C., & Melton, D. W. 1995, "Double replacement gene targeting for the production of a series of mouse strains with different prion protein gene alterations", *Biotechnology (N.Y.)*, vol. 13, no. 9, pp. 999-1004.

Moreau, J. F., Donaldson, D. D., Bennett, F., Witek-Giannotti, J., Clark, S. C., & Wong, G. G. 1988, "Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells", *Nature*, vol. 336, no. 6200, pp. 690-692.

Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M., Chambers, I., & Smith, A. 1994, "Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 10, pp. 4303-4307.

Muller-Ehmsen, J., Whittaker, P., Kloner, R. A., Dow, J. S., Sakoda, T., Long, T. I., Laird, P. W., & Kedes, L. 2002, "Survival and development of neonatal rat cardiomyocytes transplanted into adult myocardium", *J.Mol.Cell Cardiol.*, vol. 34, no. 2, pp. 107-116.

Muller HJ. 1930, "Types of visible variations induced by X-rays in *Drosophila*.", Journal of Genetics, 2: 299-334.

Muller, M., Fleischmann, B. K., Selbert, S., Ji, G. J., Endl, E., Middeler, G., Muller, O. J., Schlenke, P., Frese, S., Wobus, A. M., Hescheler, J., Katus, H. A., & Franz, W. M. 2000, "Selection of ventricular-like cardiomyocytes from ES cells in vitro", *FASEB J.*, vol. 14, no. 15, pp. 2540-2548.

Munroe, R. J., Bergstrom, R. A., Zheng, Q. Y., Libby, B., Smith, R., John, S. W., Schimenti, K. J., Browning, V. L., & Schimenti, J. C. 2000, "Mouse mutants from chemically mutagenized embryonic stem cells", *Nat.Genet.*, vol. 24, no. 3, pp. 318-321.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., & Smith, A. 1998, "Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4", *Cell*, vol. 95, no. 3, pp. 379-391.

Nickoloff, J. A. & Reynolds, R. J. 1990, "Transcription stimulates homologous recombination in mammalian cells", *Mol.Cell Biol.*, vol. 10, no. 9, pp. 4837-4845.

Nickoloff, J. A. 1992, "Transcription enhances intrachromosomal homologous recombination in mammalian cells", *Mol.Cell Biol.*, vol. 12, no. 12, pp. 5311-5318.

Nielsen, L. B., Kahn, D., Duell, T., Weier, H. U., Taylor, S., & Young, S. G. 1998, "Apolipoprotein B gene expression in a series of human apolipoprotein B transgenic mice generated with recA-assisted restriction endonuclease cleavage-modified bacterial artificial chromosomes. An intestine-specific enhancer element is located between 54 and 62 kilobases 5' to the structural gene", *J.Biol.Chem.*, vol. 273, no. 34, pp. 21800-21807.

Niemann, H., Kues, W., & Carnwath, J. W. 2005, "Transgenic farm animals: present and future", *Rev.Sci.Tech.*, vol. 24, no. 1, pp. 285-298.

Niwa, H., Yamamura, K., & Miyazaki, J. 1991, "Efficient selection for high-expression transfectants with a novel eukaryotic vector", *Gene*, vol. 108, no. 2, pp. 193-199.

Nonaka, I. 1998, "Animal models of muscular dystrophies", *Lab Anim Sci.*, vol. 48, no. 1, pp. 8-17.

Oakley, B. R. 1992, "Gamma-tubulin: the microtubule organizer?", *Trends Cell Biol.*, vol. 2, no. 1, pp. 1-5.

Oka, M., Tagoku, K., Russell, T. L., Nakano, Y., Hamazaki, T., Meyer, E. M., Yokota, T., & Terada, N. 2002, "CD9 is associated with leukemia inhibitory factor-mediated maintenance of embryonic stem cells", *Mol.Biol.Cell*, vol. 13, no. 4, pp. 1274-1281.

Orr-Weaver, T. L., Szostak, J. W., & Rothstein, R. J. 1981, "Yeast transformation: a model system for the study of recombination", *Proc.Natl.Acad.Sci.U.S.A*, vol. 78, no. 10, pp. 6354-6358.

Orr-Weaver, T. L. & Szostak, J. W. 1983, "Yeast recombination: the association between double-strand gap repair and crossing-over", *Proc.Natl.Acad.Sci.U.S.A*, vol. 80, no. 14, pp. 4417-4421.

Paabo, S. & Wilson, A. C. 1988, "Polymerase chain reaction reveals cloning artefacts", *Nature*, vol. 334, no. 6181, pp. 387-388.

Pells, S., Di Domenico, A. I., Callagher, E. J., & McWhir, J. 2002, "Multipotentiality of neuronal cells after spontaneous fusion with embryonic stem cells and nuclear reprogramming in vitro", *Cloning Stem Cells*, vol. 4, no. 4, pp. 331-338.

Pesce, M., Anastassiadis, K., & Scholer, H. R. 1999, "Oct-4: lessons of totipotency from embryonic stem cells", *Cells Tissues.Organs*, vol. 165, no. 3-4, pp. 144-152.

Pfeifer, A., Ikawa, M., Dayn, Y., & Verma, I. M. 2002, "Transgenesis by lentiviral vectors: lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos", *Proc.Natl.Acad.Sci.U.S.A*, vol. 99, no. 4, pp. 2140-2145.

Pickering, S. J., Braude, P. R., Patel, M., Burns, C. J., Trussler, J., Bolton, V., & Minger, S. 2003, "Preimplantation genetic diagnosis as a novel source of embryos for stem cell research", *Reprod.Biomed.Online.*, vol. 7, no. 3, pp. 353-364.

Pikaart, M. J., Recillas-Targa, F., & Felsenfeld, G. 1998, "Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators", *Genes Dev.*, vol. 12, no. 18, pp. 2852-2862.

Pittenger, M. F., Mackay, A. M., Beck, S. C., Jaiswal, R. K., Douglas, R., Mosca, J. D., Moorman, M. A., Simonetti, D. W., Craig, S., & Marshak, D. R. 1999, "Multilineage potential of adult human mesenchymal stem cells", *Science*, vol. 284, no. 5411, pp. 143-147.

Porter, A. 1998, "Controlling your losses: conditional gene silencing in mammals", *Trends Genet.*, vol. 14, no. 2, pp. 73-79.

Porteus, M. H. & Baltimore, D. 2003, "Chimeric nucleases stimulate gene targeting in human cells", *Science*, vol. 300, no. 5620, p. 763.

Potts, W., Tucker, D., Wood, H., & Martin, C. 2000, "Chicken beta-globin 5'HS4 insulators function to reduce variability in transgenic founder mice", *Biochem.Biophys.Res.Commun.*, vol. 273, no. 3, pp. 1015-1018.

Pouton, C. W. & Haynes, J. M. 2005, "Pharmaceutical applications of embryonic stem cells", *Adv.Drug Deliv.Rev.*, vol. 57, no. 13, pp. 1918-1934.

Prado, F., J. I. Piruat, and A. Aguilera, 1997, "Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation", *EMBO J.*, v. 16, no. 10, p. 2826-2835.

Proudfoot, N. J. 1986, "Transcriptional interference and termination between duplicated alpha-globin gene constructs suggests a novel mechanism for gene regulation", *Nature*, vol. 322, no. 6079, pp. 562-565.

Przyborski, S. A. 2005, "Differentiation of human embryonic stem cells after transplantation in immune-deficient mice", *Stem Cells*, vol. 23, no. 9, pp. 1242-1250.

Ramezani, A., Hawley, T. S., & Hawley, R. G. 2003, "Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator", *Blood*, vol. 101, no. 12, pp. 4717-4724.

Ramirez-Solis, R., Liu, P., & Bradley, A. 1995, "Chromosome engineering in mice", *Nature*, vol. 378, no. 6558, pp. 720-724.

Rasmussen, L. K., Due, H. A., & Petersen, T. E. 1995, "Human alpha s1-casein: purification and characterization", *Comp Biochem.Physiol B Biochem.Mol.Biol*, vol. 111, no. 1, pp. 75-81.

Rayssiguier, C., Thaler, D. S., & Radman, M. 1989, "The barrier to recombination between Escherichia coli and Salmonella typhimurium is disrupted in mismatch-repair mutants", *Nature*, vol. 342, no. 6248, pp. 396-401.

Recillas-Targa, F. & Razin, S. V. 2001, "Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes", *Crit Rev.Eukaryot.Gene Expr.*, vol. 11, no. 1-3, pp. 227-242.

Recillas-Targa, F., Pikaart, M. J., Burgess-Beusse, B., Bell, A. C., Litt, M. D., West, A. G., Gaszner, M., & Felsenfeld, G. 2002, "Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities", *Proc.Natl.Acad.Sci.U.S.A*, vol. 99, no. 10, pp. 6883-6888.

Recillas-Targa, F., Valadez-Graham, V., & Farrell, C. M. 2004, "Prospects and implications of using chromatin insulators in gene therapy and transgenesis", *Bioessays*, vol. 26, no. 7, pp. 796-807.

Reid, L. H., Gregg, R. G., Smithies, O., & Koller, B. H. 1990, "Regulatory elements in the introns of the human HPRT gene are necessary for its expression in embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 87, no. 11, pp. 4299-4303.

Reid, L. H., Shesely, E. G., Kim, H. S., & Smithies, O. 1991, "Cotransformation and gene targeting in mouse embryonic stem cells", *Mol.Cell Biol.*, vol. 11, no. 5, pp. 2769-2777.

Restifo, N. P., Kawakami, Y., Marincola, F., Shamamian, P., Taggarse, A., Esquivel, F., & Rosenberg, S. A. 1993, "Molecular mechanisms used by tumors to escape immune recognition: immunogenetherapy and the cell biology of major histocompatibility complex class I", *J.Immunother.*, vol. 14, no. 3, pp. 182-190.

Reubinoff, B. E., Pera, M. F., Fong, C. Y., Trounson, A., & Bongso, A. 2000, "Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro", *Nat.Biotechnol.*, vol. 18, no. 4, pp. 399-404.

Reubinoff, B. E., Itsykson, P., Turetsky, T., Pera, M. F., Reinhartz, E., Itzik, A., & Ben Hur, T. 2001, "Neural progenitors from human embryonic stem cells", *Nat.Biotechnol.*, vol. 19, no. 12, pp. 1134-1140.

Reyes, M. & Verfaillie, C. M. 2001, "Characterization of multipotent adult progenitor cells, a subpopulation of mesenchymal stem cells", *Ann.N.Y.Acad.Sci.*, vol. 938, pp. 231-233.

Richards, M., Fong, C. Y., Chan, W. K., Wong, P. C., & Bongso, A. 2002, "Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells", *Nat.Biotechnol.*, vol. 20, no. 9, pp. 933-936.

Rideout, W. M., III, Hochedlinger, K., Kyba, M., Daley, G. Q., & Jaenisch, R. 2002, "Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy", *Cell*, vol. 109, no. 1, pp. 17-27.

Rio, M. D., Larcher, F., Meana, A., Segovia, J., Alvarez, A., & Jorcano, J. 1999, "Nonviral transfer of genes to pig primary keratinocytes. Induction of angiogenesis by composite grafts of modified keratinocytes overexpressing VEGF driven by a keratin promoter", *Gene Ther.*, vol. 6, no. 10, pp. 1734-1741.

Ripps, M. E., Huntley, G. W., Hof, P. R., Morrison, J. H., & Gordon, J. W. 1995, "Transgenic mice expressing an altered murine superoxide dismutase gene provide an animal model of amyotrophic lateral sclerosis", *Proc.Natl.Acad.Sci.U.S.A*, vol. 92, no. 3, pp. 689-693.

Rivella, S., Callegari, J. A., May, C., Tan, C. W., & Sadelain, M. 2000, "The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites", *J.Virol.*, vol. 74, no. 10, pp. 4679-4687.

Rosen, D. R. 1993, "Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis", *Nature*, vol. 364, no. 6435, p. 362.

Rosler, E. S., Fisk, G. J., Ares, X., Irving, J., Miura, T., Rao, M. S., & Carpenter, M. K. 2004, "Long-term culture of human embryonic stem cells in feeder-free conditions", *Dev.Dyn.*, vol. 229, no. 2, pp. 259-274.

Rucker, E. B. & Piedrahita, J. A. 1997, "Cre-mediated recombination at the murine whey acidic protein (mWAP) locus", *Mol.Reprod.Dev.*, vol. 48, no. 3, pp. 324-331.

Rudolph, U., Brabet, P., Hasty, P., Bradley, A., & Birnbaumer, L. 1993, "Disruption of the G(i2) alpha locus in embryonic stem cells and mice: a modified hit and run strategy with detection by a PCR dependent on gap repair", *Transgenic Res.*, vol. 2, no. 6, pp. 345-355.

Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G., Sherry, S., Mullikin, J. C., Mortimore, B. J., Willey, D. L., Hunt, S. E., Cole, C. G., Coggill, P. C., Rice, C. M., Ning, Z., Rogers, J., Bentley, D. R., Kwok, P. Y., Mardis, E. R., Yeh, R. T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R. H., McPherson, J. D., Gilman, B., Schaffner, S., Van Etten, W. J., Reich, D., Higgins, J., Daly, M. J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M. C., Linton, L., Lander, E. S., & Altshuler, D. 2001, "A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms", *Nature*, vol. 409, no. 6822, pp. 928-933.

Sadelain, M., Wang, C. H., Antoniou, M., Grosveld, F., & Mulligan, R. C. 1995, "Generation of a high-titer retroviral vector capable of expressing high levels of the human beta-globin gene", *Proc.Natl.Acad.Sci.U.S.A*, vol. 92, no. 15, pp. 6728-6732.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., & Erlich, H. A. 1988, "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase", *Science*, vol. 239, no. 4839, pp. 487-491.

Saito, Y. & Takashima, S. 2000, "Neurotransmitter changes in the pathophysiology of Lesch-Nyhan syndrome", *Brain Dev.*, vol. 22 Suppl 1, p. S122-S131.

Sambrook j, Fritsch EF, Maniatis T. 1989, "Molecular Cloning. A laboratory manual (second edition)" Cold Spring Harbor Laboratory Press 1989.

Sauer, B. & Henderson, N. 1988, "Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1", *Proc.Natl.Acad.Sci.U.S.A*, vol. 85, no. 14, pp. 5166-5170.

Sauer, B. 1998, "Inducible gene targeting in mice using the Cre/lox system", *Methods*, vol. 14, no. 4, pp. 381-392.

Schedl, A., Montoliu, L., Kelsey, G., & Schutz, G. 1993, "A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice", *Nature*, vol. 362, no. 6417, pp. 258-261.

Scheerer, J. B. & Adair, G. M. 1994, "Homology dependence of targeted recombination at the Chinese hamster APRT locus", *Mol.Cell Biol.*, vol. 14, no. 10, pp. 6663-6673.

Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D. A., & Benvenisty, N. 2000, "Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 97, no. 21, pp. 11307-11312.

Schuldiner, M., Eiges, R., Eden, A., Yanuka, O., Itskovitz-Eldor, J., Goldstein, R. S., & Benvenisty, N. 2001, "Induced neuronal differentiation of human embryonic stem cells", *Brain Res.*, vol. 913, no. 2, pp. 201-205.

Schuldiner, M., Itskovitz-Eldor, J., & Benvenisty, N. 2003, "Selective ablation of human embryonic stem cells expressing a "suicide" gene", *Stem Cells*, vol. 21, no. 3, pp. 257-265.

Schwartz, R. E., Reyes, M., Koodie, L., Jiang, Y., Blackstad, M., Lund, T., Lenvik, T., Johnson, S., Hu, W. S., & Verfaillie, C. M. 2002, "Multipotent adult progenitor cells from bone marrow differentiate into functional hepatocyte-like cells", *J.Clin.Invest*, vol. 109, no. 10, pp. 1291-1302.

Schwartzberg, P. L., Robertson, E. J., & Goff, S. P. 1990, "Targeted gene disruption of the endogenous c-abl locus by homologous recombination with DNA encoding a selectable fusion protein", *Proc.Natl.Acad.Sci.U.S.A*, vol. 87, no. 8, pp. 3210-3214.

Schwenk, F., Baron, U., & Rajewsky, K. 1995, "A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells", *Nucleic Acids Res.*, vol. 23, no. 24, pp. 5080-5081.

Sedivy, J. M. & Sharp, P. A. 1989, "Positive genetic selection for gene disruption in mammalian cells by homologous recombination", *Proc.Natl.Acad.Sci.U.S.A*, vol. 86, no. 1, pp. 227-231.

Sedivy, J.M., Joyner, A.L. 1993, "Gene targeting", Oxford University Press USA, Publ., New York.

Sedivy, J.M., Vogelstein, B., Liber, H.L., Hendrickson, E.A., Rosmarin, A. 1999, "Gene targeting in human cells without isogenic DNA", Science 1999 1 january vol 283:9a.

Seibler, J., Schubeler, D., Fiering, S., Groudine, M., & Bode, J. 1998, "DNA cassette exchange in ES cells mediated by Flp recombinase: an efficient strategy for repeated modification of tagged loci by marker-free constructs", *Biochemistry*, vol. 37, no. 18, pp. 6229-6234.

Selfridge, J., Pow, A. M., McWhir, J., Magin, T. M., & Melton, D. W. 1992, "Gene targeting using a mouse HPRT minigene/HPRT-deficient embryonic stem cell system: inactivation of the mouse ERCC-1 gene", *Somat.Cell Mol.Genet.*, vol. 18, no. 4, pp. 325-336.

Shamblott, M. J., Axelman, J., Wang, S., Bugg, E. M., Littlefield, J. W., Donovan, P. J., Blumenthal, P. D., Huggins, G. R., & Gearhart, J. D. 1998, "Derivation of pluripotent stem cells from cultured human primordial germ cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 95, no. 23, pp. 13726-13731.

Sharp, J. D., Capecchi, N. E., & Capecchi, M. R. 1973, "Altered enzymes in drug-resistant variants of mammalian tissue culture cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 70, no. 11, pp. 3145-3149.

Shastry, B. S. 1998, "Gene disruption in mice: models of development and disease", *Mol.Cell Biochem.*, vol. 181, no. 1-2, pp. 163-179.

Shaw-White, J. R., Denko, N., Albers, L., Doetschman, T. C., & Stringer, J. R. 1993, "Expression of the lacZ gene targeted to the HPRT locus in embryonic stem cells and their derivatives", *Transgenic Res.*, vol. 2, no. 1, pp. 1-13.

Shaw, C. E., Enayat, Z. E., Chioza, B. A., Al-Chalabi, A., Radunovic, A., Powell, J. F., & Leigh, P. N. 1998, "Mutations in all five exons of SOD-1 may cause ALS", *Ann.Neurol.*, vol. 43, no. 3, pp. 390-394.

Shin, S., Dalton, S., & Stice, S. L. 2005, "Human motor neuron differentiation from human embryonic stem cells", *Stem Cells Dev.*, vol. 14, no. 3, pp. 266-269.

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., & Muller, M. 2005, "Strong and ubiquitous expression of transgenes targeted into the beta-actin locus by Cre/lox cassette replacement", *Genesis.*, vol. 42, no. 4, pp. 229-235.

Shulman, M. J., Nissen, L., & Collins, C. 1990, "Homologous recombination in hybridoma cells: dependence on time and fragment length", *Mol.Cell Biol.*, vol. 10, no. 9, pp. 4466-4472.

Singh, R. N., Nakano, T., Xuing, L., Kang, J., Nedergaard, M., & Goldman, S. A. 2005, "Enhancer-specified GFP-based FACS purification of human spinal motor neurons from embryonic stem cells", *Exp.Neurol.*, vol. 196, no. 2, pp. 224-234.

Smith-Arica, J. R., Thomson, A. J., Ansell, R., Chiorini, J., Davidson, B., & McWhir, J. 2003, "Infection efficiency of human and mouse embryonic stem cells using adenoviral and adeno-associated viral vectors", *Cloning Stem Cells*, vol. 5, no. 1, pp. 51-62.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., & Rogers, D. 1988, "Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides", *Nature*, vol. 336, no. 6200, pp. 688-690.

Smith, A. J., De Sousa, M. A., Kwabi-Addo, B., Heppell-Parton, A., Impey, H., & Rabbitts, P. 1995, "A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination", *Nat.Genet.*, vol. 9, no. 4, pp. 376-385.

Smith, K. 2001, "Theoretical mechanisms in targeted and random integration of transgene DNA", *Reprod.Nutr.Dev.*, vol. 41, no. 6, pp. 465-485.

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., & Kucherlapati, R. S. 1985, "Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination", *Nature*, vol. 317, no. 6034, pp. 230-234.

Solter, D. & Knowles, B. B. 1975, "Immunosurgery of mouse blastocyst", *Proc.Natl.Acad.Sci.U.S.A*, vol. 72, no. 12, pp. 5099-5102.

Soriano, P., Cone, R. D., Mulligan, R. C., & Jaenisch, R. 1986, "Tissue-specific and ectopic expression of genes introduced into transgenic mice by retroviruses", *Science*, vol. 234, no. 4782, pp. 1409-1413.

Sottile, V., Thomson, A., & McWhir, J. 2003, "In vitro osteogenic differentiation of human ES cells", *Cloning Stem Cells*, vol. 5, no. 2, pp. 149-155.

Soukharev, S., Miller, J. L., & Sauer, B. 1999, "Segmental genomic replacement in embryonic stem cells by double lox targeting", *Nucleic Acids Res.*, vol. 27, no. 18, p. e21.

Southern, E. M. 1975, "Detection of specific sequences among DNA fragments separated by gel electrophoresis", *J.Mol.Biol.*, vol. 98, no. 3, pp. 503-517.

Southern, P. J. & Berg, P. 1982, "Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter", *J.Mol.Appl.Genet.*, vol. 1, no. 4, pp. 327-341.

Southgate, T. D., Bain, D., Fairbanks, L. D., Morelli, A. E., Larregina, A. T., Simmonds, H. A., Castro, M. G., & Lowenstein, P. R. 1999, "Adenoviruses encoding HPRT correct biochemical abnormalities of HPRT-deficient cells and allow their survival in negative selection medium", *Metab Brain Dis.* , vol. 14, no. 4, pp. 205-221.

St Clair, M. H., Lambe, C. U., & Furman, P. A. 1987, "Inhibition by ganciclovir of cell growth and DNA synthesis of cells biochemically transformed with herpesvirus genetic information", *Antimicrob.Agents Chemother.*, vol. 31, no. 6, pp. 844-849.

Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A., & Melton, D. W. 1994, "Use of double-replacement gene targeting to replace the murine alpha-lactalbumin gene with its human counterpart in embryonic stem cells and mice", *Mol.Cell Biol*, vol. 14, no. 2, pp. 1009-1016.

Stanford, W. L., Cohn, J. B., & Cordes, S. P. 2001, "Gene-trap mutagenesis: past, present and beyond", *Nat.Rev.Genet.*, vol. 2, no. 10, pp. 756-768.

Sternberg, N. & Hamilton, D. 1981, "Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites", *J.Mol.Biol.*, vol. 150, no. 4, pp. 467-486.

Stief, A., Winter, D. M., Stratling, W. H., & Sippel, A. E. 1989, "A nuclear DNA attachment element mediates elevated and position-independent gene activity", *Nature*, vol. 341, no. 6240, pp. 343-345.

Storkus, W. J., Howell, D. N., Salter, R. D., Dawson, J. R., & Cresswell, P. 1987, "NK susceptibility varies inversely with target cell class I HLA antigen expression", *J.Immunol.*, vol. 138, no. 6, pp. 1657-1659.

Stout, J. T. & Caskey, C. T. 1988, "The Lesch-Nyhan syndrome: clinical, molecular and genetic aspects", *Trends Genet.*, vol. 4, no. 6, pp. 175-178.

Straub, T. 2003, "Heterochromatin dynamics", *PLoS.Biol.*, vol. 1, no. 1, p. E14.

Strauss, W. M., Dausman, J., Beard, C., Johnson, C., Lawrence, J. B., & Jaenisch, R. 1993, "Germ line transmission of a yeast artificial chromosome spanning the murine alpha 1(I) collagen locus", *Science*, vol. 259, no. 5103, pp. 1904-1907.

Stutts, P. & Brockman, R. W. 1963, "A biochemical basis for resistance of L1210 mouse leukemia to 6-thioguanine", *Biochem.Pharmacol.*, vol. 12, pp. 97-104.

Sutherland, H. G., Kearns, M., Morgan, H. D., Headley, A. P., Morris, C., Martin, D. I., & Whitelaw, E. 2000, "Reactivation of heritably silenced gene expression in mice", *Mamm.Genome*, vol. 11, no. 5, pp. 347-355.

Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., & Stahl, F. W. 1983, "The double-strand-break repair model for recombination", *Cell* , vol. 33, no. 1, pp. 25-35.

Szybalski, W. & Szybalska, E. H. 1962, "Drug sensitivity as a genetic marker for human cell lines", *Med.Bull.(Ann.Arbor)*, vol. 28, pp. 277-293.

Taboit-Dameron, F., Malassagne, B., Viglietta, C., Puissant, C., Leroux-Coyau, M., Chereau, C., Attal, J., Weill, B., & Houdebine, L. M. 1999, "Association of the 5'HS4 sequence of the chicken beta-globin locus control region with human EF1 alpha gene promoter induces ubiquitous and high expression of human CD55 and CD59 cDNAs in transgenic rabbits", *Transgenic Res.*, vol. 8, no. 3, pp. 223-235.

Tajbakhsh, S. & Buckingham, M. E. 1994, "Mouse limb muscle is determined in the absence of the earliest myogenic factor myf-5", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 2, pp. 747-751.

Taniguchi, M., Sanbo, M., Watanabe, S., Naruse, I., Mishina, M., & Yagi, T. 1998, "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.*, vol. 26, no. 2, pp. 679-680.

Tartof, K. D. 1994, "Position effect variegation in yeast", *Bioessays*, vol. 16, no. 10, pp. 713-714.

te Riele. H., Maandag, E. R., Clarke, A., Hooper, M., & Berns, A. 1990, "Consecutive inactivation of both alleles of the pim-1 proto-oncogene by homologous recombination in embryonic stem cells", *Nature*, vol. 348, no. 6302, pp. 649-651.

te Riele. H., Maandag, E. R., & Berns, A. 1992, "Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs", *Proc.Natl.Acad.Sci.U.S.A*, vol. 89, no. 11, pp. 5128-5132.

Terada, N., Hamazaki, T., Oka, M., Hoki, M., Mastalerz, D. M., Nakano, Y., Meyer, E. M., Morel, L., Petersen, B. E., & Scott, E. W. 2002, "Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion", *Nature*, vol. 416, no. 6880, pp. 542-545.

Thomas, B. J. & Rothstein, R. 1989, "Elevated recombination rates in transcriptionally active DNA", *Cell*, vol. 56, no. 4, pp. 619-630.

Thomas, K. R. & Capecchi, M. R. 1987, "Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells", *Cell*, vol. 51, no. 3, pp. 503-512.

Thompson, S., Stern, P. L., Webb, M., Walsh, F. S., Engstrom, W., Evans, E. P., Shi, W. K., Hopkins, B., & Graham, C. F. 1984, "Cloned human teratoma cells differentiate into neuron-like cells and other cell types in retinoic acid", *J.Cell Sci.*, vol. 72, pp. 37-64.

Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L., & Melton, D. W. 1989, "Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells", *Cell*, vol. 56, no. 2, pp. 313-321.

Thomson, J. A., Kalishman, J., Golos, T. G., Durning, M., Harris, C. P., Becker, R. A., & Hearn, J. P. 1995, "Isolation of a primate embryonic stem cell line", *Proc.Natl.Acad.Sci.U.S.A*, vol. 92, no. 17, pp. 7844-7848.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. 1998, "Embryonic stem cell lines derived from human blastocysts", *Science*, vol. 282, no. 5391, pp. 1145-1147.

Thomson, J. A. & Marshall, V. S. 1998, "Primate embryonic stem cells", *Curr.Top.Dev.Biol.*, vol. 38, pp. 133-165.

Thyagarajan, B., Johnson, B. L., & Campbell, C. 1995, "The effect of target site transcription on gene targeting in human cells in vitro", *Nucleic Acids Res.*, vol. 23, no. 14, pp. 2784-2790.

Trinh, K. R. & Morrison, S. L. 2000, "Site-specific and directional gene replacement mediated by Cre recombinase", *J.Immunol.Methods*, vol. 244, no. 1-2, pp. 185-193.

Tsai, T. F., Jiang, Y. H., Bressler, J., Armstrong, D., & Beaudet, A. L. 1999, "Paternal deletion from Snrpn to Ube3a in the mouse causes hypotonia, growth retardation and partial lethality and provides evidence for a gene contributing to Prader-Willi syndrome", *Hum.Mol.Genet.*, vol. 8, no. 8, pp. 1357-1364.

Tsuda, H., Maynard-Currie, C. E., Reid, L. H., Yoshida, T., Edamura, K., Maeda, N., Smithies, O., & Jakobovits, A. 1997, "Inactivation of the mouse HPRT locus by a 203-bp retroposon insertion and a 55-kb gene-targeted deletion: establishment of new HPRT-deficient mouse embryonic stem cell lines", *Genomics*, vol. 42, no. 3, pp. 413-421.

Urbach, A., Schuldiner, M., & Benvenisty, N. 2004, "Modeling for Lesch-Nyhan disease by gene targeting in human embryonic stem cells", *Stem Cells*, vol. 22, no. 4, pp. 635-641.

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., & Holmes, M. C. 2005, "Highly efficient endogenous human gene correction using designed zinc-finger nucleases", *Nature*, vol. 435, no. 7042, pp. 646-651.

Valancius, V. & Smithies, O. 1991, "Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells", *Mol.Cell Biol.*, vol. 11, no. 3, pp. 1402-1408.

Vallier, L., Rugg-Gunn, P. J., Bouhon, I. A., Andersson, F. K., Sadler, A. J., & Pedersen, R. A. 2004, "Enhancing and diminishing gene function in human embryonic stem cells", *Stem Cells*, vol. 22, no. 1, pp. 2-11.

van Deursen. J., Fornerod, M., Van, R. B., & Grosveld, G. 1995, "Cre-mediated site-specific translocation between nonhomologous mouse chromosomes", *Proc.Natl.Acad.Sci.U.S.A*, vol. 92, no. 16, pp. 7376-7380.

van Deursen, J. & Wieringa, B. 1992, "Targeting of the creatine kinase M gene in embryonic stem cells using isogenic and nonisogenic vectors", *Nucleic Acids Res.*, vol. 20, no. 15, pp. 3815-3820.

Van Duyne, G. D. 2001, "A structural view of cre-loxp site-specific recombination", *Annu.Rev.Biophys.Biomol.Struct.*, vol. 30, pp. 87-104.

Verhoeyen, E., Hauser, H., & Wirth, D. 2001, "Evaluation of retroviral vector design in defined chromosomal loci by Flp-mediated cassette replacement", *Hum.Gene Ther.*, vol. 12, no. 8, pp. 933-944.

Villemure, J. F., Savard, N., & Belmaaza, A. 2001, "Promoter suppression in cultured mammalian cells can be blocked by the chicken beta-globin chromatin insulator 5'HS4 and matrix/scaffold attachment regions", *J.Mol.Biol.*, vol. 312, no. 5, pp. 963-974.

Visser, J. E., Bar, P. R., & Jinnah, H. A. 2000, "Lesch-Nyhan disease and the basal ganglia", *Brain Res.Brain Res.Rev.*, vol. 32, no. 2-3, pp. 449-475.

Vivian, J. L., Klein, W. H., & Hasty, P. 1999, "Temporal, spatial and tissue-specific expression of a myogenin-lacZ transgene targeted to the Hprt locus in mice", *Biotechniques*, vol. 27, no. 1, pp. 154-162.

Wade-Martins, R., White, R. E., Kimura, H., Cook, P. R., & James, M. R. 2000, "Stable correction of a genetic deficiency in human cells by an episome carrying a 115 kb genomic transgene", *Nat.Biotechnol.*, vol. 18, no. 12, pp. 1311-1314.

Wahl, G. M., Hughes, S. H., & Capecchi, M. R. 1975, "Immunological characterization of hypoxanthine-guanine phosphoribosyl transferase mutants of mouse L cells: evidence for mutations at different loci in the HGPRT gene", *J.Cell Physiol*, vol. 85, no. 2 Pt 1, pp. 307-320.

Waldman, A. S. & Liskay, R. M. 1987, "Differential effects of base-pair mismatch on intrachromosomal versus extrachromosomal recombination in mouse cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 84, no. 15, pp. 5340-5344.

Waldman, A. S. & Liskay, R. M. 1988, "Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology", *Mol.Cell Biol.*, vol. 8, no. 12, pp. 5350-5357.

Wallace, H., Ansell, R., Clark, J., & McWhir, J. 2000, "Pre-selection of integration sites imparts repeatable transgene expression", *Nucleic Acids Res.*, vol. 28, no. 6, pp. 1455-1464.

Wallrath, L. L. & Elgin, S. C. 1995, "Position effect variegation in Drosophila is associated with an altered chromatin structure", *Genes Dev.*, vol. 9, no. 10, pp. 1263-1277.

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M. S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T. J., Lipshutz, R., Chee, M., & Lander, E. S. 1998, "Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome", *Science*, vol. 280, no. 5366, pp. 1077-1082.

Wang, Y., DeMayo, F. J., Tsai, S. Y., & O'Malley, B. W. 1997, "Ligand-inducible and liver-specific target gene expression in transgenic mice", *Nat.Biotechnol.*, vol. 15, no. 3, pp. 239-243.

Wang, Y., Spatz, M. K., Kannan, K., Hayk, H., Avivi, A., Gorivodsky, M., Pines, M., Yayon, A., Lonai, P., & Givol, D. 1999, "A mouse model for achondroplasia produced by targeting fibroblast growth factor receptor 3", *Proc.Natl.Acad.Sci.U.S.A*, vol. 96, no. 8, pp. 4455-4460.

Watts, R. W., Harkness, R. A., Spellacy, E., & Taylor, N. F. 1987, "Lesch-Nyhan syndrome: growth delay, testicular atrophy and a partial failure of the 11 beta-hydroxylation of steroids", *J.Inherit.Metab Dis.*, vol. 10, no. 3, pp. 210-223.

Weiler-Guettler, H., Aird, W. C., Husain, M., Rayburn, H., & Rosenberg, R. D. 1996, "Targeting of transgene expression to the vascular endothelium of mice by homologous recombination at the thrombomodulin locus", *Circ.Res.*, vol. 78, no. 2, pp. 180-187.

Wells, W. A. 2002, "Is transdifferentiation in trouble?", *J.Cell Biol.*, vol. 157, no. 1, pp. 15-18.

West, A. G., Gaszner, M., & Felsenfeld, G. 2002, "Insulators: many functions, many mechanisms", *Genes Dev.*, vol. 16, no. 3, pp. 271-288.

Westervelt, P., Lane, A. A., Pollock, J. L., Oldfather, K., Holt, M. S., Zimonjic, D. B., Popescu, N. C., DiPersio, J. F., & Ley, T. J. 2003, "High-penetrance mouse model of acute promyelocytic leukemia with very low levels of PML-RARalpha expression", *Blood*, vol. 102, no. 5, pp. 1857-1865.

Wilson, C., Bellen, H. J., & Gehring, W. J. 1990, "Position effects on eukaryotic gene expression", *Annu.Rev.Cell Biol.*, vol. 6, pp. 679-714.

Wu, H., Liu, X., & Jaenisch, R. 1994, "Double replacement: strategy for efficient introduction of subtle mutations into the murine Col1a-1 gene by homologous recombination in embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 7, pp. 2819-2823.

Wurmser, A. E. & Gage, F. H. 2002, "Stem cells: cell fusion causes confusion", *Nature*, vol. 416, no. 6880, pp. 485-487.

Wurst, W., Rossant, J., Prideaux, V., Kownacka, M., Joyner, A., Hill, D. P., Guillemot, F., Gasca, S., Cado, D., Auerbach, A., & . 1995, "A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice", *Genetics*, vol. 139, no. 2, pp. 889-899.

Xu, C., Inokuma, M. S., Denham, J., Golds, K., Kundu, P., Gold, J. D., & Carpenter, M. K. 2001, "Feeder-free growth of undifferentiated human embryonic stem cells", *Nat.Biotechnol.*, vol. 19, no. 10, pp. 971-974.

Xu, C., Police, S., Rao, N., & Carpenter, M. K. 2002, "Characterization and enrichment of cardiomyocytes derived from human embryonic stem cells", *Circ.Res.*, vol. 91, no. 6, pp. 501-508.

Yamaguchi, H., Yamamoto, C., & Tanaka, N. 1965, "Inhibition of protein synthesis by blasticidin S. I. Studies with cell-free systems from bacterial and mammalian cells", *J.Biochem.(Tokyo)*, vol. 57, no. 5, pp. 667-677.

Yanez, R. J. & Porter, A. C. 1999, "Influence of DNA delivery method on gene targeting frequencies in human cells", *Somat.Cell Mol.Genet.*, vol. 25, no. 1, pp. 27-31.

Ying, Q. L., Nichols, J., Evans, E. P., & Smith, A. G. 2002, "Changing potency by spontaneous fusion", *Nature*, vol. 416, no. 6880, pp. 545-548.

Yoo, S. J., Yoon, B. S., Kim, J. M., Song, J. M., Roh, S., You, S., & Yoon, H. S. 2005, "Efficient culture system for human embryonic stem cells using autologous human embryonic stem cell-derived feeder cells", *Exp.Mol.Med.*, vol. 37, no. 5, pp. 399-407.

Young, H. E., Steele, T. A., Bray, R. A., Hudson, J., Floyd, J. A., Hawkins, K., Thomas, K., Austin, T., Edwards, C., Cuzzourt, J., Duenzl, M., Lucas, P. A., & Black, A. C., Jr. 2001, "Human reserve pluripotent mesenchymal stem cells are present in the connective tissues of skeletal muscle and dermis derived from fetal, adult, and geriatric donors", *Anat.Rec.*, vol. 264, no. 1, pp. 51-62.

Young, R. G., Butler, D. L., Weber, W., Caplan, A. I., Gordon, S. L., & Fink, D. J. 1998, "Use of mesenchymal stem cells in a collagen matrix for Achilles tendon repair", *J.Orthop.Res.*, vol. 16, no. 4, pp. 406-413.

Zeidler, R., Eissner, G., Meissner, P., Uebel, S., Tampe, R., Lazis, S., & Hammerschmidt, W. 1997, "Downregulation of TAP1 in B lymphocytes by cellular and Epstein-Barr virus-encoded interleukin-10", *Blood*, vol. 90, no. 6, pp. 2390-2397.

Zeng, X., Miura, T., Luo, Y., Bhattacharya, B., Condie, B., Chen, J., Ginis, I., Lyons, I., Mejido, J., Puri, R. K., Rao, M. S., & Freed, W. J. 2004, "Properties of pluripotent human embryonic stem cells BG01 and BG02", *Stem Cells*, vol. 22, no. 3, pp. 292-312.

Zhang, S. C., Wernig, M., Duncan, I. D., Brustle, O., & Thomson, J. A. 2001, "In vitro differentiation of transplantable neural precursors from human embryonic stem cells", *Nat.Biotechnol.*, vol. 19, no. 12, pp. 1129-1133.

Zhao, Z., Fu, Y. X., Hewett-Emmett, D., & Boerwinkle, E. 2003, "Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution", *Gene*, vol. 312, pp. 207-213.

Zheng, H., Hasty, P., Brenneman, M. A., Grompe, M., Gibbs, R. A., Wilson, J. H., & Bradley, A. 1991, "Fidelity of targeted recombination in human fibroblasts and murine embryonic stem cells", *Proc.Natl.Acad.Sci.U.S.A* , vol. 88, no. 18, pp. 8067-8071.

Zhou, Q. Y. & Palmiter, R. D. 1995, "Dopamine-deficient mice are severely hypoactive, adipsic, and aphagic", *Cell*, vol. 83, no. 7, pp. 1197-1209.

Zijlstra, M., Li, E., Sajjadi, F., Subramani, S., & Jaenisch, R. 1989, "Germ-line transmission of a disrupted beta 2-microglobulin gene produced by homologous recombination in embryonic stem cells", *Nature*, vol. 342, no. 6248, pp. 435-438.

Zuk, P. A., Zhu, M., Mizuno, H., Huang, J., Futrell, J. W., Katz, A. J., Benhaim, P., Lorenz, H. P., & Hedrick, M. H. 2001, "Multilineage cells from human adipose tissue: implications for cell-based therapies", *Tissue Eng*, vol. 7, no. 2, pp. 211-228.

Zwaka, T. P. & Thomson, J. A. 2003, "Homologous recombination in human embryonic stem cells", *Nat.Biotechnol.*, vol. 21, no. 3, pp. 319-321.

List of Tables & Figures

Abbreviations

AFP	α-feto protein
A.T.F	Absolute targeting frequency
ATG	Translational initiation start
bFGF	Basic fibroblast growth factor
Bp	Base pair
BSA	Bovine serum albumin
BSD	*Blasticidin*
CM	Conditioned medium
CDNA	Complementary DNA
CHO	Chinese hamster ovary
dH_20	Distilled water
DEAE	Diethylaminoethylamine
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
Dnase	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
DSBs	Double-strand breaks
EB	Embryoid body
EC cells	Embryonic carcinoma cells
EDTA	Ethylenediaminetetraacetic acid
EF1-α	Elongation factor 1α
EG cells	Embryonic germ cells
EGTA	Ethyleneglycol-bis (β-aminoethyl)-N,N,N',N'-tetraacetic acid
ES cells	Embryonic stem cells
EtBr	Ethidium bromide
E.T.F	Effective targeting frequency
F	Farad
FAC	Fluorescence-activated cell sorting
F.C	Final concentration
FCS	Foetal calf serum
gDNA	Genomic DNA

GFP	Green fluorescent protein
GMEM	Glasgow modified Eagle's medium
HAT	Hypoxanthine, Aminopterin, Thymidine
HEK	Human embryonic kidney cells
HES	Human embryonic stem cells
Hprt	*Hypoxanthine-guanine phosphoribosyl transferase*
HR	Homologous recombination
HSV-tk	Herpes simplex virus-thymidine kinase
ICM	Inner cell mass
Kb	Kilobase pair
KO-DMEM	Knock-out DMEM
LB	Luria Bertani broth
LIF	Leukemia inhibitory factor
LNS	Lesch-Nyhan syndrome
mM	Millimolar
M	Molar
MEF	Mouse embryonic fibroblasts
MES	Murine embryonic stem cells
MHC	Major histocompatibility complex
mRNA	Messenger RNA
NaAc	Sodium acetate
NBCS	New born calf serum
Neo	*Neomycine phosphotransferase*
NGS	Normal goat serum
NHEJ	nonhomologous end joining
NT	Nuclear transfer
OD	Optical density
OS	Osteogenic supplements
P	Passage
PARP	Poly(ADP-ribose) polymerase
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction

PEV	Position effect variegation
PFA	Paraformadehyde
PolyA	Poly adenylation
PRPP	5' phosphoribosyl-1-pyrophosphate
Rag2	Recombination-activating gene 2
RMCE	Recombinase-mediated cassette exchange
RNA	Ribonucleic acid
Rnase	Ribonuclease
RT	Room temperature
SCID	Severe combined immuno deficiency
SCNT	Somatic cell nuclear transfer
SDS	Sodium dodecyl sulphate
SE	Subcloning efficiency
SNPs	Single-nucleotide polymorphisms
Ss	Single strand
SV40	Simian virus 40
SR	Serum replacement
U	Unit
UV	Ultraviolet
TAE	Tris base glacial acetic acid EDTA
TE	Tris-HCl EDTA
TEG	Trypsin/EGTA
6-TG	6-Thioguanine
TI	Transcriptional interference
Tk	Thymidine kinase
Tris	Tris (hydroxymethyl) aminoethane
UV	Ultraviolet
V	Volts
V/V	Volume by volume
WT	Wild-type
W/V	Weight by volume

APPENDIX 1

1. Sequencing of *Lox* sites

1.1. pCASEfN

TCTAATACTTGGGGACAGGGCCAGAGTATCACTGTCCTCCAGTCATTGAAGGGCGAATTCGATATCAA
GCTTATCGATACCGTCGAGGCCGC**ATAACTTCGTATAAAGTATCCTATACGAAGTTAT**TCTAGAACTA
GTGGATCCCCCGGGCTGCAGGAATTCAACAGGCATCTACTGAGTGGACCCAACGCATGAGAGGACAGT
GCCAAGCAAGCAACTCAAATGTCCCACCGGTTGGGCATGGCCAGGTAGCCTATGCTGTGTCTGGACGT
CCTCCTGCTGGTATAGTTATTTTAAAATCAGAAGGACAGGGAAGGGAGCAGTGGTTCACGCCTGTAAT
CCCAGCAATTTGGGAGGCCAAGGTGGGTAGATCACCTGAGATTANGAGTTGGAGACCAGCCTGCCAAT
ATGGTGAAACCCCGTCTCTACCAAAAAAACAAAAATTAA
CTTGAGCCTGGTCATGCATGCCTGGAATCCAACAACTCCGGAGGCTGAGGCANGAGAATCCTTTGACC
CCCGAGGCGGGGATTGCAATGAGCCCAAAATTGTGNCCCTGCCTCCAACTTGGTTTCCAAANACCCCC
CNGCCCTCCANGGGGGTTTNCANATTGCCCTTGTTTATACCCCCCCTCACCCCTTAATNTTTNTAAGG
CNCTTTTCNACANAANGTTGCCCTTANTTTTTTTGGCAANGGTGNGGGNCTGGGGGGGACTTCCCGCC
CNAAAGCCNTNGGGGGGNCCAGGTTCCCCCCCCNCGG

TTCNTATNGGGAGACCCAAGCTGGCTAGGTAAGCTATCTGCGGCCTAGCTAGTCGAGGTCGAGTACCC
GGACCG**ATAACTTCGTATAGCATACATTATACGAAGTTAT**CTCGAGGTCGACGGATCCCCACTTAAGT
AAGTAAAATATGGGAAGATCCGGTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGA
GAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT
CAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGAC
GAGGCAGCGCGGCTATCGTGGCTGGCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACT
GAAGCCGGGAAGGGACTGGCTTGCTATTGGGCGAANTGCCCGGGCAGGATCTCCTGTCATCTCACCTT
GCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACNCTTGATCCCGCTAC
CTGCCATTCGACCACCAAAGCGAAACATCGCATTCGAGCCAACACCTACTCGGATTGGAACCCGGCNT
TGTCGATCAAGGATATCTTGGACNAAAAACATNAAGGGCTCGCGCCANCCCAACTGTTCGCCAGCTTC
AAGGGCGCATGNCCCACGGGGAGGATCTCTNTGACCCATGGGATGCCTNTTTCCCAAANTATGGGGNA
AAGGCCCTTTTTGGNTTTTNACTTGGGCCN

- ***Lox2272***
- ***LoxP***

1.2. pHPRTEfN

AAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCCTTCTGATTTTAAAATAACTATA
CCAGCAGGAGGACGTCCAGACACAGCATAGGCTACCTGGCCATGCCCAACCGGTGGGACATTTGAGTT
GCTTGCTTGGCACTGTCCTCTCATGCGTTGGGTCCACTCAGTAGATGCCTGTTGAATTCCTGCAGCCC
GGGGGATCCACTAGTTCTAGA**A**ATAACTTCGTATAGGATACTTTATACGAAGTTATGCGGCCTCGACGG
TATCGCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGAC
TGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGA
GCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCA
TCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGGGGATCGGCAATAAAAAGACAGAATAAAACGC
ACGGGTGTTGGGTCGTTTGTTCGGATCCGTCGAGCAAGACGTTCAGTCCTACAGAAATAAAATCAGGA
ATTTAATAGAAAGTTTCATACATTAAACTTTATAACAAACACCTCTTATCATTAAACTTCCCACCAAC
CTGGGCNATATAGTGAGACCCNNGCCTGCCNAAAAAAAAAAAAAAACAGCCTGGNGG

NNGCTGTTCCTATAGGGAGACCCAAGCTGGCTAGGTAAGCTATCTGCGGCCTAGCTAGTCGAGGTCGA
GTACCCGGACCG**ATAACTTCGTATAGCATACATTATACGAAGTTAT**CTCGAGGTCGACGGATCCCCAC
TTAAGTAAGTAAAATATGGGAAGATCCGGTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTG
GGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCC
GGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTG
CAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGT
TGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTC
ACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCG
GCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGG
TCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCACCGAACTGTTCGCCAGCTC
AAGGCGCGCATGCCCCACGGCAAGGATCTNNTCNTGACCCATGGCGATGCCTGCTTGCNAANANA

- ***Lox2272***
- ***LoxP***

1.3. pfBSDOctGFP

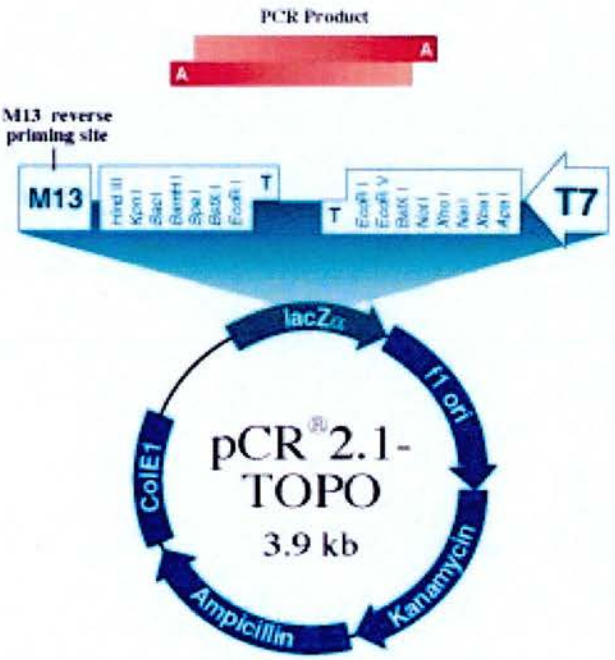
GNNNGTTTNCgAAtTGGagCTCCaCCGCGGTGGCGGCCGC**ATAACTTCGTATAAAGTATCCTATACGA**
AGTTATTCTAGAACTAGTGGATCAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAG
GTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCTAGAGTC
GCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCA
GGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTG
TCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCT
GCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGA
TATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAG
TCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTT
GTAGTTGCCCGCGTCCTTGAANAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGGNGGACTTG
AANANTCAGCCTGCTTCANGTGNCGGGNANCGGCTGAACANTGNCCCCGTAGTNAGGCGTNANNAGNN
NGGNCNAGAANGNANCTNNCNNGGGAGANNAAATAAGCCCNTNCNNNGGGATNCCNCCCCNCCNACCC
NATTNNGNNNTCC

GNNNGTTTNNNAAAAACTGGTACCCGGACCG**ATAACTTCGTATAGCATACATTATACGAAGTTAT**CTC
GAGGTCGACGGTATCGATAAGCTTTCTCATTGTCAATATGCCTTTGTCTCAAGAAGAATCCACCCTCA
TTGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCT
CTCTCTAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTACTGGGGGACCTTGTGCAGA
ACTCGTGGTGCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAA
ATGAGAACAGGGGCATCTTGAGCCCCTGCGGACGGTGCCGACAGGTGCTTCTCGATCTGCATCCTGGG
ATCAAAGCCATAGTGAAGGACAGTGATGGACAGCCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTC
TGGTTATGTGTGGGAGGGCTAAGCACCAGCCATTGATCTTGTACATTCAAGCTACTTTCTGTTACTAG
ATGAACCGAGCATGAGTTGATGGTGACTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCCGGATC
TTTGTGAAGGAACCTTACTTCTGNGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTC
TTAAGNAAATATAAAATTTTAAGTGTTTAAGGGNTAAACAACCGATTCTAATGGNTGGGNTTTANAAT
CCAACCANGAATNNAANTGGACANNGNGNAGCNTTATNNAAANCNNTTTTAAAAAAAANACTNNGNAN
NANNCCNNNCCTNNNN

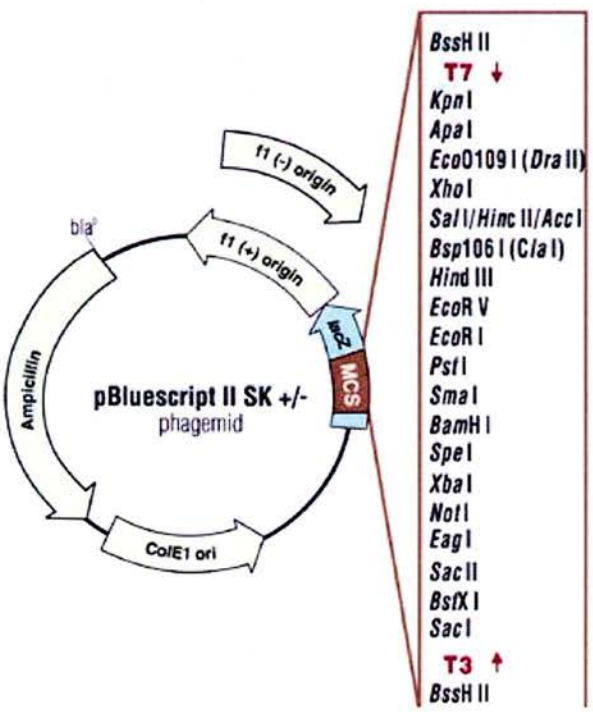
- ***Lox2272***
- ***LoxP***

2. Plasmid Maps

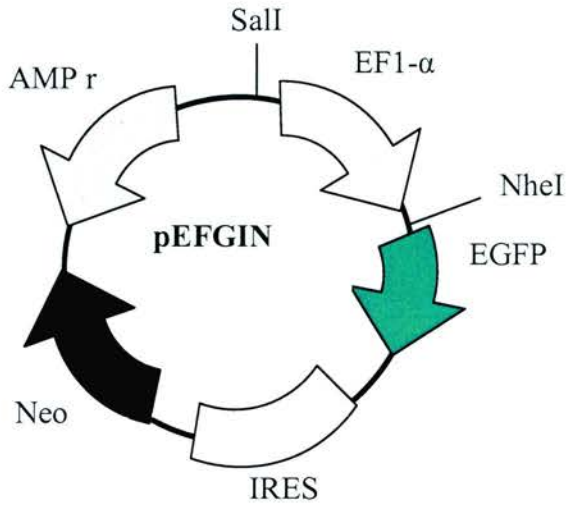
2.1. pCR2.1-TOPO TA (3.9 kb)



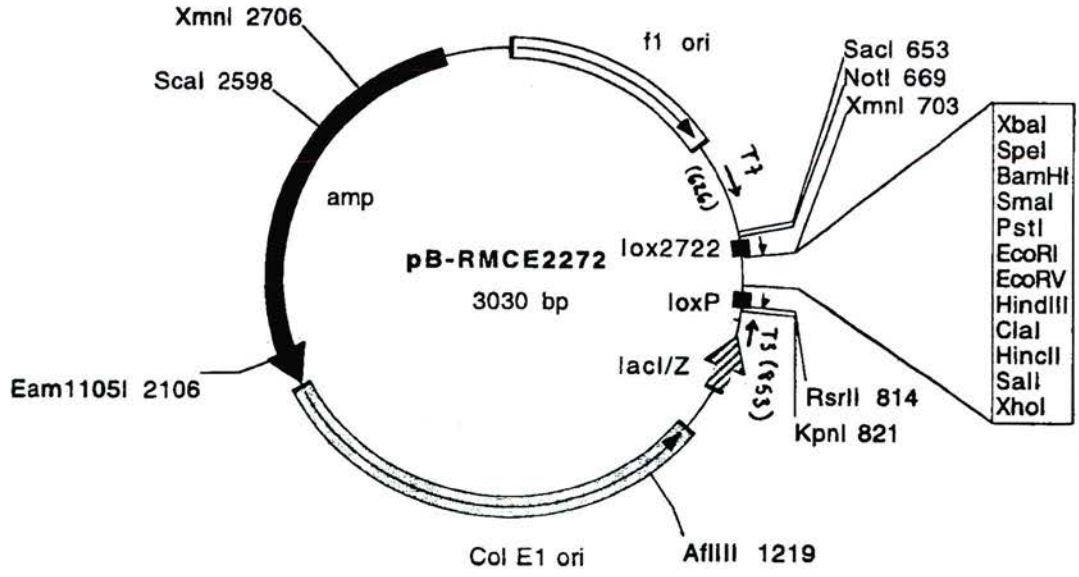
2.2. pBluescript II SK +/- (3 kb)



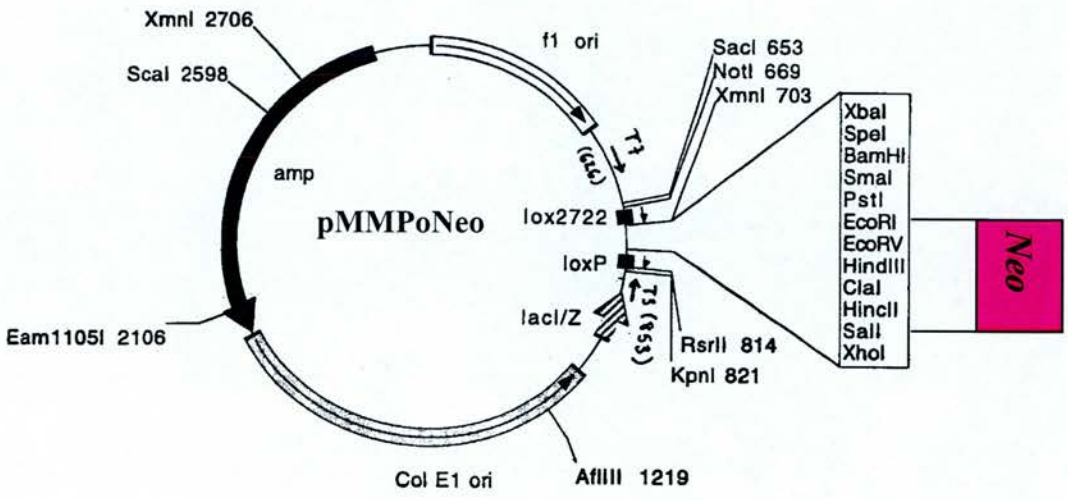
2.3. pEFGIN (6.7 kb)



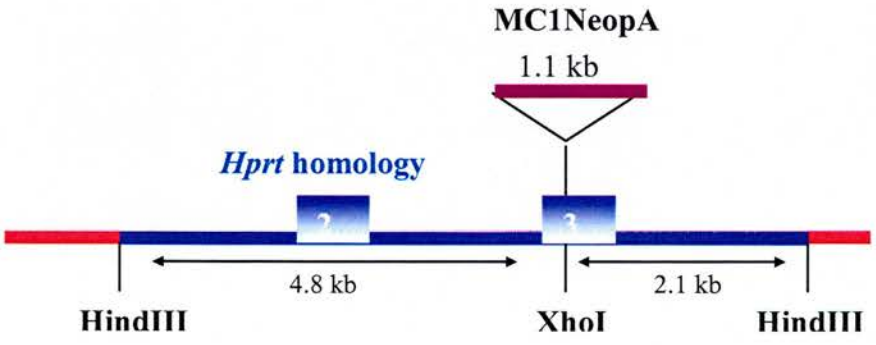
2.4. pB-RMCE2272 (3 kb)



2.5. pMMPoNeo (4.6 kb)



2.6. pE3ΔNeo (11 kb)



2.7. ploxBSD1B (5.2 kb)