Design and development of a cell marking system in transgenic mice.

Douglas J. C. Strathdee

Beatson Institute for Cancer Research, Glasgow.

Thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy.

January 1995.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To Mum and Dad.

Profound Quotation Page.

" There it goes ....."

Archibald Macpherson M. B. E.
5.20 p.m. Saturday May 16th, 1987

Table of Contents.

List of figures.

Previous page.

Acknowledgements.

Special thanks go to Stormin' Norman ; Bobby G, the King ; Hong Kong Phooey ; Basher ; Tindersticks ; PSG ; Rockingbirds ; Ello ; Boris Karloff ; Macca ; Ricky G ; Psycho ; Dibble ; Colo Litt ; Afghan Whigs ; Weezer ; Jayhawks ; Pat ; Bobby D ; Pigeonhed ; MC5 ; Vince ; the one and only Peripheral ; Gordo ; Wonder Stuff ; Rocks c/w Funky Jam ; Christina the Amazing ; James Brown ; Top Cat ; Bonehead ; Redd Kross ; AMC ; Steady Eddie ; Scooby Doo ; JD ; Mudhoney ; Bruce ; J Lizard ; NYC ; Pauls I 'n'II ; 808 All the way from Manchester on t' bus ; The Complete Moron With The Melon On His Head ; Gary ; Dinosaurs ; Free Scotland Now! ; King Duncan and the Sheep ; Andy Pandy ; The Meat Beat Manifesto ; Cats ; Cranberries ; Athletics The Fridge mutilated by Evil C ; Whatever, If Only ; Jackson Hole ; Consolidated ; SS Lazio ; Power Monger ; Nicky C ; Hope Sandoval ; SLG ; Throneberry ; Pavement ; Sarah Cairns-Smith ; The Freak Scene ; The Colour Wheel ; Tommy K ; Mum and Dad Strathdee ; Allan Mac ; Kojak ; Alice C ; Claude ; M Monroe and Hanoi ; Steve, the Blonde Dred and the Emotion ; The Reverend HH ; Dillon Fence ; Alien Jourgensen ; Movin' On Up ; Shonen Knife ; M*A*S*H ; RevCos ; Dodgy ; Terry Pratchett ; Lush ; Mets ; KMFDM ; Kris Bradley and Indian Head ; Ulrike ; The Drug Squad ; Mary Chain ; Pixies ; Debs ; 49'ers ; Jinky ; The Wonderful Thing About Tiggers ; Smashing P ; Milan AC ; Phillies ; Starsky and Hutch and Huggy ; Feel Burnt ; Breeders ; Gorgeous SBG ; Les Miserables ; Transglobal Underground ; Luscious J, the Unobtainable ; Tom and Jerry ; Garfield ; God Machine, my alter ego ; and always remember "I make music for mo'f-ers to get down to. If you gonna criticise it, it ain't made for you." - Tha Snoop.

No thanks go to Norbert Llachlittle Cromwell " Put 'em up, put 'em up!" and Christimoffits Kamper "I'm a loser baby, so why don't you kill me?"

Extra special thanks go to

Shiela Bryson for teaching me how to make transgenic mice.
Wendy Lambie for teaching me all about cloning and stuff.
Everybody else in R7, past and present, and the rest of the Beatson who has helped me along the way.
Andrew McGuigan and Stephen Bell for the time and tremendous effort they put into caring for my mice.
Allan Balmain for guiding me along the treacherous path to insanity and without whom none of the following would have been possible.
The Cancer Research Campaign who provided the funding throughout the course of my PhD.

All the work described in this thesis was completed by myself.

Abbreviations.

A	adenosine
AAF	acetoxyaminofluorene
AF	aminofluorene
APRT	adenosine phosphoribosyl transferase
bFGF	basic fibroblast growth factor
BK	bovine keratin
bp	base pairs
C	cytosine
cDNA	complimentary deoxyribonucleic acid
C. elegans	*Caenorhabditis elegans*
CFU-S	colony forming unit-spleen
cpm	counts per minute
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
D. melanogaster	*Drosophila melanogaster*
DMH	dimethyl hydrazine
DNA	deoxyribonucleic acid
dTTP	deoxythymidine triphosphate
E. coli	*Escherichia coli*
EMS	ethyl methansulphonate
en	*engrailed*
ENU	ethyl nitrosourea
ES cell	embryonic stem cell
FACS	fluoresence activated cell sorter
G	guanosine
G6PD	glucose-6-phosphate dehydogenase
GMCSF	granulocyte/macrophage colony stimulating factor
hCG	human chorionic gonadotrophin
hGH	human growth hormone
hh	*hedgehog*
ICM	inner cell mass
kb	kilobase pairs
kg	kilogrammes
kV	kilovolts
LRC	label retaining cell
LTR	long terminal repeat
mg	milligramme
ml	millilitre
µg	microgramme
µl	microlitre
Min	*multiple intestinal neoplasia*
MNNG	N-nitro-N'-methyl-N- nitrosoguanidine
MNU	methyl nitrosourea

mwh	*multiple wing hairs*
NAc-AAF	N-acetoxy-acetoxyaminofluorene
ng	nanogramme
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
Pgk-1	*phosphoglycerate kinase*
PMS	pregnant mare's serum
Rho	rhodamine
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
tRNA	transfer ribonucleic acid
Sca	stem cell antigen
SCF	stem cell factor
T	thymidine
TPA	12-*0*-tetradecanoylphorbyl-13-acetate
w/v	weight/volume
v/v	volume/volume
X-gal	5-bromo-4-chloro-3-indolyl β-D-galactoside

Abstract.

The aim of this project was to develop a marker system to analyse cell lineage in mouse skin. This should provide useful information on the usual size and organisation of clones in the epidermis and give some clues as to how cell replacement is normally controlled in the epidermis and to the possible location of the stem and progenitor cells. As there is no marker gene currently available in the murine epidermis a marker gene had to be designed and introduced. This was accomplished by generating transgenic mice with a novel marker gene inserted into their DNA. An *Escherichia coli lac Z* gene was used as a cell marker and activated in single cells following treating the mice with a chemical mutagen. The marker could be activated by a point mutation in the sequence of an oligonucleotide inserted at the 5' end of the *lac Z* gene, replacing the ATG initiation codon of the gene. Two different marking strategies were devised. One specific for the chemical mutagen N-nitro-N'-methyl-N-nitosoguanidine (MNNG) and the other for the chemical mutagen acetoxyaminofluorene (AAF). Oligonucleotides were designed, synthesised and cloned in front of the *lac Z* gene to generate oligo/*lac Z* fusion marker genes. Test constructs without initiation codons and control constructs with initiation codons were designed and cloned at the same time for both the MNNG and AAF strategies. The fusion genes were then tested *in vitro* by transfection in to murine epidermal cell lines. No expression was detected in the inactive marker gene designed to be activated by MNNG, whereas high expression was obtained from two control constructs which both contained and in-frame ATG initiation codon in the sequence of the oligonucleotide cloned into the *lac Z* gene. The MNNG-responsive marker gene was shown to be activated in stably transfected murine epidermal cell lines in response to treatment with MNNG at a frequency of approximately 4.3×10^{-3} per clonogenic cell. Subsequently the work with the AAF strategy was discontinued as there was insufficient time to pursue both strategies and positive results had been obtained with the MNNG construct. Twenty-two transgenic founders were generated with three marker gene constructs (Act-LacZ^{A}, K5-LacZ^{D} and K5-LacZ^{E}) by pronuclear microinjection of fertilised embryos. Lines of transgenic mice were bred from the founders and the expression of the marker genes were analysed. Expression from a control *lac Z* fusion gene driven by a *BKIII* (keratin 5) promoter was detected by staining skin tissue with the chromogenic β-galactosidase substrate 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal). The expression of the β-galactosidase protein was observed in the epidermal basal layer and the hair follicles of transgenic mice. The expression was highest in newborn mice and gradually

diminished as the mice reached adulthood. It was subsequently shown that targeting the expression of an activated H-*ras* oncogene to the epidermis with the same promoter results in a high frequency of malignant tumours. The implications of these results to the biology of the skin are discussed. Expression of the MNNG-activated test marker gene under the control of a β-actin promoter was also detected. Activation of the marker gene was detected in serial sections of murine brain tissue from transgenic mice treated topically with MNNG at birth. No activation was detected in other tissues probably due to the low level of expression of the marker gene. The explanation of these results and the implications for developing a marker system to work efficiently in the epidermis are discussed.

CHAPTER 1

Introduction.

1. Introduction.

1.1 Stem cells and hierarchies.

During adult life tissues are characterised by differing rates of cell division (Lajtha, 1979). The purpose of cell division is to maintain the number of cells, in a tissue, at a constant level by replacing those lost by death and injury.

Some tissues lose their capacity to divide almost totally. Most of the cells, in such organs, are formed during foetal development and are retained throughout adult life. This occurs in skeletal muscle and nerve tissues (Hall and Watt, 1989; Stockdale, 1992; Marvin and MacKay, 1992). Other body organs, *e.g.* the liver, retain the capacity to divide in response to the appropriate stimuli. The cell turnover and division rate in these tissues is usually low. The last category of tissues, such as blood, testis or stratified epithelia are constantly renewing. In these systems continual loss of mature functional cells is compensated by the division of more primitive cells.

There is considerable evidence that cell replacement, in body tissues, occurs via a proliferative hierarchy, which is populated by the division of stem cells. When these cells divide they produce progenitor cells which, although they retain the capacity to divide, are now committed to differentiate and form fully functional, mature cells. The progenitor population has the effect of multiplying each stem cell division and is therefore designated the transit amplifying population. These cells divide a number of times before becoming terminally differentiated, mature cells.

It is likely that the stem cells of renewing populations are multipotential. There is clear evidence in the haemopoietic system that a single undifferentiated cell is capable of giving rise to all the differentiated cell types of the blood including erythrocytes, lymphocytes, macrophages and polymorphonuclear leukocytes (Lemischka, 1992). The small intestine is another example of a renewing cell population. In this system there is evidence that columnar, mucous, enteroendocrine and Paneth cells are all descended from the same origin (Winton and Ponder, 1991).

Even in the epidermis there is tentative evidence that the stem cells are multipotential. It seems likely that cells responsible for populating the interfollicular epidermis are also capable of populating the hair follicles. When epidermis is transplanted from sites of the body which do not produce hair on to dermis from hair-producing sites the transplanted

epidermis will start to produce hair (Sengel, 1986; Reynolds and Jahoda, 1992). Also cells which seem to be responsible for re-population and maintenance of the hair follicles are capable of differentiating *in vitro* to produce interfollicular epidermal cells (Lenoir et al., 1988).

The functions of these tissues are very similar. They are all designed to produce large numbers of mature functional cells which have a short life-span. Hence it may not be surprising that these systems should all operate in the same fashion. The haemopoietic system and the small intestine certainly both operate via a similar proliferative hierarchy despite markedly differing tissue organisation. That both of these tissues have evolved to operate in the same fashion suggests that these mechanisms provide some advantage to the tissue.

There are probably several reasons why tissues should operate via a hierarchy descended from a multipotential stem cell. Initially all the separate cell types in each tissue evolved from a common precursor cell *i.e.* a fertilised egg. The hierarchy probably allows a finer degree of control of the amount and type of cells produced. Signalling a stem cell to divide should result in the production of more cells than the division of a transit amplifying cell. If one cell type was more urgently required than another, then the balance of differentiation could be altered in favour of the required cell.

A system that operates via a multipotential stem cell also has the advantage of reducing the size of the stem cell compartment. Stem cells are by definition long-lived. This could potentially make them very sensitive to carcinogenic transformation. The smaller the stem cell pool, therefore, the fewer the number of targets in which this transformation can occur. Also if each stem-cell division is amplified by subsequent division of the committed progenitor cell this reduces the number of divisions through which the stem-cell must pass. This in turn reduces the chance of DNA damage to the stem cell during division.

In both the haemopoietic system and the small intestine the stem cells are kept in an area separated from the main functional area of the tissue. Stem cells for the blood are located in the bone marrow while the small intestinal stem-cell is located in the crypts. Both these tissues are to different degrees exposed to the environment. Keeping the stem-cells in an area removed from this exposure to the environment protects them from damage that could be inflicted by environmental factors. Separation of the stem-cells from the tissue allows the environment of the stem-cells to be very carefully controlled. Creation of a tightly controlled niche for the stem-cells allows fine control of their division and differentiation to properly re-populate the tissue. This carefully controlled

environment is more easily generated for a small number of cells. This may be another factor which contributes to the generation of a hierarchy. As the stem-cells are vital to the renewal capacity of the tissue it is understandable that they should be well protected from environmental damage and kept under tight control.

Comparison of the skin to other systems would suggest there should be a proliferative hierarchy in the skin and that the stem cells for this hierarchy should be located in a well-protected niche, preventing environmental damage to the cells. The existence of a hierarchy can be confirmed by measuring the cell cycle times of keratinocytes in the basal layer. Analysis of the population by [^{3}H]-thymidine labelling demonstrates the existence of populations of quickly and slowly cycling cells (Potten, 1981). Slowly cycling cells have been found in both the basal layer of the epidermis and in the hair follicle. The basal layer of the epidermis must be subject to extreme stress from the environment. The hair follicle may, therefore, represent a better niche in which groups of stem cells can be carefully controlled. There is evidence too that cells from the hair follicle are at least bipotential. They are able to differentiate to form interfollicular epidermis as well as hair follicle cells. There are however areas of skin without hair follicles and it is unlikely that these areas are regenerated from follicle stem cells. Tissues in areas without hair, such as monkey palm epidermis, are often organised in a more complicated fashion than the simple stratified epithelium found on the back of a mouse. This still suggests that follicle stem cells may not be required for epidermal maintenance. How homeostasis is maintained in the epidermis is still not fully understood.

One of the main problems in studying stem and progenitor cell populations is that they are defined by their potential, *i.e.* how they will behave. Removing these cells from their environment may alter how they behave. One way to get round this problem is to study cells *in situ*, by using an *in vivo* marking system. When cells responsible for replacement in the epidermis are marked they should give rise to areas of marked cells which are maintained throughout the lifetime of the animal. If the cells marked are not ultimately responsible for epidermal re-population, the marked clones produced by these cells will eventually be lost from the skin as these cells are replaced by the division of progenitor cells. Whether or not the cells identified by this technique are "true stem cells" is open to question. It may be that in the event of injury or damage that the lineage may go further back than the cells identified by this technique. A marking system should be able to identify cells which are responsible for epidermal cell maintenance and those cells which are around for long enough to act as targets for carcinogens.

The aim of this project is to use an *E. coli lac Z* gene as a marker to determine epidermal cell lineage in murine epidermis *in vivo*. The marker gene will be activated at fixed times by application of a chemical mutagen. Lineage analysis will be performed *in vivo* on the skin of transgenic mice carrying the marker gene. By examining the skin tissue of the mouse at varying lengths of time after gene activation in discrete cells, it should be possible to determine the normal patterns of proliferation in the epidermis. This will allow the identification of the cell population normally responsible for epidermal replacement.

1.2 Embryos and embryonic stem cells.

Nowhere is the concept of stem cells more clearly demonstrated than the development of an entire organism from a single egg cell. The process of embryonic development requires co-ordinated diversification of a multitude of different cell types from a single fertilised egg (Wolpert, 1988; Smith A. G., 1992; Gardner and Beddington, 1988). The first divisions of this cell produce a number of identical cells known as blastomeres each of which has the capacity to produce mature cells of every lineage. There is no evidence that, in mammals, the regional specification of the embryo is controlled by maternally derived determinants deposited in the egg. The subsequent division of the blastomere cells and the generation of cellular diversity is controlled by interactions between the cells and between the embryo and the environment.

The initial stages of embryonic development involve the creation and expansion of the extra-embryonic tissues and of the pluripotential founder cell lineage. This is supported by evidence from cell marking studies in which the ability of different populations of cells to contribute to tissues was analysed in chimeric mice created by fusing genetically distinct embryos from different mouse strains. The first separation which occurs is between the trophectoderm and the inner cell mass (ICM) (Johnson and Ziomek, 1981; Gardner, 1983). The cells of the trophoectoderm only generate the trophoblast (part of the placenta). The cells of the ICM are able to generate all the embryonic and extra-embryonic tissues except the trophoblast (Chisholm *et al.*, 1985). The ICM is then divided into epiblast and primitive endoderm. The epiblast cells are found throughout the embryo except in the trophoblast and the extra-embryonic endoderm whereas the primitive endoderm cells are restricted to the colonisation of the parietal and visceral yolk-sac endoderm (Snell and Stevens, 1966; Gardner, 1985; Beddington, 1986).

Despite their complexity, mammalian embryos exhibit a greater degree of plasticity than simpler embryos. It is possible to aggregate several murine pre-implantation embryos and generate a normally sized viable foetus. It is also possible to kill a large proportion of the cells of a murine blastocyst or egg cylinder and the embryos will still form a viable foetus (Snow and Tam, 1979). Thus the embryo has a large capacity to repair damage or injury. The repair capacity is dependent on the presence of pluripotential stem cells whose fate can be reprogrammed and the existence of a flexible signalling system to modulate the behaviour of the cells.

Grafting experiments show that these cells are not irreversibly committed to the formation of particular tissues and that the cells may contribute to any cell lineage (Grobstein, 1952; Beddington, 1982). Teratocarcinomas may be readily generated by ectopic grafts of cells from epiblast tissues from early embryos taken from inbred mouse strains (Stevens, 1970; Solter *et al.*, 1970). These tumours are able to form multiple differentiated cell types and may be benign or malignant. The earlier the cells are removed from the embryo the more chance there is of malignancy (Damjanov *et al.*, 1970). These tumours are formed with such a high frequency it seems that the process is unlikely to involve a transformation event but instead is a result of the disturbance of the normal growth and differentiation of these cells (Stevens, 1983).

The teratocarcinomas contain a readily identifiable stem cell population known as embryonal carcinoma cells (EC cells). These cells may be readily propagated in culture and clonal lines can be established. In culture the cells often remain multipotent and can produces tumours containing a variety of cell types (Evans, 1972: Bernstine *et al.*, 1973; Martin & Evans 1975). The cells in culture remain morphologically very similar to the ICM and epiblast cells (Martin, 1975; Martin, 1980). Some of these cell lines can be introduced and incorporated into a new blastocyst (Brinster, 1974; Mintz and Illmensee, 1975; Papaioannou *et al.*, 1975). The EC cells can then participate in normal development, contribute extensively to a variety of different tissues and can produce normal healthy chimeric animals. Despite the obvious pluripotentiality of these cells they rarely contribute to the germ line. This is probably because most of the EC cell lines are aneuploid and often show a reduced developmental capacity *in vitro*. As these cells are not normal their use in developmental studies in the mouse is obviously limited.

It was subsequently discovered that cells could be cultured directly from blastocysts when plated out onto a feeder layer of mitotically inactivated murine embryonic fibroblasts (Evans and Kauffman, 1981; Martin, 1981). These cells can be established from whole blastocysts as well as isolated epiblast cells and are now known as

embryonic stem cells (ES cells). They are morphologically very similar to the EC cells possessing a prominent nucleus and very little cytoplasm. The cells are capable of differentiating to produce multiple cell types *in vitro* (Doetschman *et al.*, 1985) and *in vivo* in teratocarcinomas (Evans and Kaufman, 1981). The cells can colonise blastocysts and contribute extensively to chimeric mice without generating tumours and very often have a normal diploid karyotype. The cells are capable of colonising the mouse germline and transmitting the ES cell genotype to normal offspring (Bradley *et al.*, 1985). These cells seem to be a non-transformed pluripotential stem cell population.

Embryonic stem cells are able to integrate into a host blastocyst and participate in normal development (Beddington and Robertson, 1989; Lallemand and Brûlet, 1990). The cells are therefore able to respond to, and generate the full repertoire of regulatory signals which control murine embryogenesis. These cells may be exploited in trying to characterise regulators of stem cells and their mode of action. It is possible that these regulators may be important not only in the control of stem cells during embryogenesis but may also be important in the control of stem cell populations in the adult.

Another important application of ES cells is their ability to generate transgenic mice (Robertson, 1986; Capecci, 1989). This is made possible as the ES cells will colonise the germ-line of mice when introduced into blastocysts. One of the main advantages of using this system is that defined mutations in specific genes can be generated. The selection of rare homologous integration events is possible when DNA is transfected into a murine embryonic stem cell line.

The study of embryonic stem cells has, along with renewing cell populations such as the haemopoietic system has established a number of principles for the maintenance of tissues by stem cells. The most important implication of the work with the embryonic stem cells is that as all tissues are originally derived from pluripotential stem cells it is likely they will be maintained by a similar cells to the ones from which the tissues were originally derived *i.e.* stem cells. This principle certainly holds true for the haemopoietic system and it is becoming increasingly obvious that other tissues with slower cell turnover rates (*e.g.* the liver and part of the central nervous system) are also maintained by the division of stem cells. The maintenance of tissues by stem cells is fast becoming a general rule of developmental biology rather than an exception for tissues with a high cell turnover rate. As the epidermis consists of a renewing cell population with a high turnover rate it seems likely that it will be maintained by the division of stem cells and there already exists a great deal of evidence that this is the case.

1.3 Haemopoiesis and regulation of stem cells.

The haemopoietic system must be capable of regenerating cells throughout the lifetime of an animal. The turnover of cells in this system is very high. In the human this system needs produce as many as 3.7×10^{11} cells per day to replace cells lost by natural wastage (Dexter and Spooncer, 1987; Whetton and Dexter, 1993). The mature functional cells produced by this system have a high degree of specialisation. Erythrocytes are designed to carry oxygen and carbon dioxide. T and B lymphocytes are involved in the generation of specific immune responses. The function of platelets is to help blood clot. Macrophages and granulocytes are the general scavengers which also act as accessory cells for the immune system. Obviously, considering the high numbers of cells produced and the incredible diversity of function among these cells a tightly controlled system is required to regulate cell production.

This basis of this system is the production of all the cell types from a common set of multipotential stem cells. The stem cells originate in the yolk sac of the embryo and migrate to the foetal liver before moving on to the bone marrow, where most of the haemopoietic cell production occurs in the adult (Dexter and Spooncer, 1987).The number of stem cells is relatively small but they exist throughout life by generating a transit amplifying population which multiplies every stem cell division.

There is no easy assay for the presence of stem cells but evidence of their existence has been established. Following ablation of an animal's haemopoietic system with radiation or cytotoxic chemicals the entire tissue can be replaced if the animal is infused with normal bone marrow cells. Studies have shown that very few infused cells can replace the entire haemopoietic system of the animal. As little as 0.01% of the total number of bone marrow cells infused into a mouse can replace its entire complement of haemopoietic cells and the recipient animal will have a normal life span with a apparently normal haemopoietic system (Boggs *et al.*, 1982; Schofield *et al.*, 1986).

The mature cells in these radiation chimera experiments described above have been demonstrated to come from the transplanted donor cells (Abramson *et al.*, 1977). Radiation-induced karyotype markers were used to tag the donor cells in a transplant experiment. The results showed conclusively that cells of both the myeloid and lymphoid lineages were derived from the same engrafted bone marrow cells. This finding was extended more recently when retroviruses were used to mark the progeny of

single stem cells (Joyner *et al.*, 1983; Williams *et al.*, 1984; Lemischka *et al.*, 1986; Lemischka, 1992). These pieces of evidence demonstrate that pluripotent stem cells of the haemopoietic system actually exist.

Attempts to isolate haemopoietic stem cells date back to the development of the *in vivo* spleen colony forming assay (CFU-S), which can be used to quantify stem cells (Till and McCulloch, 1961; Wu *et al.*, 1968). This development of the CFU-S assay was followed a few years later by the description of an *in vitro* equivalent (Bradley and Metcalf, 1966). The limitations of these assays were recognised very shortly after they were described. It is possible by velocity or equilibrium centrifugation to separate the cells which give rise to splenic colonies from those which give rise to colonies *in vitro* (Worton *et al.*, 1969).

Specific antisera raised against cell surface molecules provided a more sensitive method of separating distinct populations of haemopoietic cells at different developmental stages. This in combination with the fluorescence-activated cell sorter (FACS) allowed the populations identified by specific antisera to be separated from other cells and used in functional assays (Herzenberg and Sweet, 1976). This technology along with clonogenic assays similar to those described above allowed the isolation of a population of cells proposed to contain stem cells. This cell population was originally isolated from murine cells on the basis of expression of stem cell antigen (Sca-1), low expression of Thy-1 and the absence of a range of molecules found on maturing and mature haemopoietic cells (collectively called lin^{neg}). These cells were termed the Thy-$1^{lo}Lin^{neg}Sca\text{-}1^{+}$ subset (Spangrude *et al.*, 1988; Spangrude, 1989).

It has subsequently been shown that these cells can be further divided on the basis of rhodamine 123 fluorescence (Bertoncello *et al.*, 1985; Spangrude and Johnston, 1990). The Thy-$1^{lo}Lin^{neg}Sca\text{-}1^{+}Rho^{dull}$ subset of cells does contain haemopoietic stem cells (Spangrude and Scollay, 1990; Smith *et al.*, 1991). Single Thy-$1^{lo}Lin^{neg}Sca\text{-}1^{+}Rho^{dull}$ cells have been injected into lethally irradiated mice and been shown to generate clones containing T and B lymphocytes, monocytes and granulocytes. Long-term repopulating ability and secondary transfer was observed in a subset of the cells demonstrating that some of these cells satisfy the strictest definition of stem cells.

It is still not clear how individual stem cells actually contribute to the process of haemopoiesis. It is clear however that the stem cells seem to evolve. Stem cells purified from foetal liver by the same Thy-$1^{lo}Lin^{neg}Sca\text{-}1^{+}$ criteria as from adult bone marrow seem to have a different developmental potential from stem cells purified from adult

bone marrow (Ikuta *et al.*, 1990). For instance, only foetal stem cells have the capacity to differentiate into $V_\gamma3^+$ T-cells in the foetal thymic microenvironment (Harvan and Allison, 1988; Harvan and Allison, 1990).

A number of different approaches have been used to attempt to discover how individual stem cells contribute to steady state haemopoiesis during the lifetime of an animal (*e.g.* Mintz *et al.*, 1984; Smith *et al.*, 1991). If a small number of Thy-1^{lo}Lin^{neg}Sca-1^+ cells are introduced to a lethally irradiated animal almost 10% of the cells contribute to the regenerative haemopoiesis. If 100 Thy-1^{lo}Lin^{neg}Sca-1^+ cells are injected usually about 8 of them will contribute to the regeneration. Most of the clones generated by these cells contain multiple lineages including B and T lymphocytes, monocytes and granulocytes. The clones accounted for between 1 to 65% of circulating white blood cells and could be maintained for long periods of up to 1 year or more. There was no correlation between the clone size and the length of time for which the clone persisted. The reasons for the heterogeneity of this response remains unclear and may be due to an inherent difference in the self-renewal capacity of the cells or their opportunity to interact with different microenvironments that promotes a capacity for self-renewal.

When unique retroviral insertion sites are used as clonal markers for haemopoietic stem cells similar results were obtained (Keller *et al.*, 1985; Dick *et al.*, 1985; Lemischka *et al.*, 1986). This work confirmed that stem cells could give rise to large multilineage clones which could account for the majority of peripheral blood cells. Analysis of the evolution of these clones has shown that the haemopoietic system can be repopulated oligoclonally after radiation. Initially, in the first few months after irradiation there seems to be extensive fluctuation of the repopulating clones. After a few months though an oligoclonal system is established which seems to persist indefinitely (Jordan and Lemischka, 1990: Capel *et al.*, 1989; Keller and Snodgrass, 1990). In contrast when this experiment was repeated using a competitive repopulation assay the results indicated that the number of stem cells contributing to haemopoiesis was proportional to the number of cells introduced (Harrison *et al.*, 1988 ; Ansell and Micklem, 1986).

In normal individuals, who have not been subjected to radiation the number of stem cells present is much larger than the dose given in any of the transplant experiments described above. The clonal composition of haemopoietic cells in females heterozygous for an X-linked marker gene was analysed. Analysis of humans (Fialkow, 1973) and mice (Micklem *et al.*, 1983) in this way suggests usually about 20 clones contribute to haemopoiesis at any one time. Repopulation of the haemopoietic system can be achieved with small numbers of stem cells however it is likely that oligoclonal

repopulation has been observed because limiting numbers of stem cells have been introduced.

In a normal mouse there are approximately 20 000 Thy-$1^{lo}Lin^{neg}Sca\text{-}1^{+}$ cells only 20 of which seem to contribute to steady state haemopoiesis. However when these Thy-$1^{lo}Lin^{neg}Sca\text{-}1^{+}$ cells are injected into lethally irradiated hosts about 1 in 10 cells can contribute to the regeneration of the haemopoietic system. The cells must be activated in response to signals from the host, probably growth factors.

The processes of differentiation and commitment of cells in the haemopoietic system have been extensively studied by *in vitro* clonogenic soft gel cultures (Metcalf, 1977). Using such systems it is possible to distinguish multipotent stem cells from more committed progenitor cells. The way that all these cells respond in tissue culture is governed by growth factors. The development of both multipotential stem cells and lineage restricted precursor cells occurs only in the presence of growth factors. Cells will die irrespective of the stage of development they have reached if growth factors are removed from the culture (Metcalf, 1977; Dexter and Spooncer, 1987).

It is obvious that the growth factors are serving several separate functions, helping the cells to survive, differentiate and proliferate. A large number of the growth factors involved in these processes have now been purified and identified (*e.g.* GM-CSF: Gough *et al.*, 1984). After identification the factors can be cloned and by making use of recombinant DNA technology large amounts can be produced and purified to homogeneity. This has allowed an assessment of the roles of different growth factors in haemopoiesis *in vitro*. The elucidation of the structure and function of growth factors and receptors will be important in understanding how stem cells make decisions about self-renewal and differentiation.

In analysing the behaviour of cells in a proliferative hierarchy such as the haemopoietic system it is important to realise that the environment or niche in which the cells are located plays a very important role in the development of the cells (Whetton and Dexter, 1993; Graham and Pragnell, 1992; Ogawa, 1993). A vital component of this niche is the interaction of the cell with the surrounding network of stromal cells in the bone marrow (Dexter, 1982). Strong support for this concept comes from work done *in vitro* with long term marrow cultures (Dexter *et al.*, 1984). Bone marrow derived stromal cells can support haemopoiesis in the absence of added growth factors, whereas stromal cells from other tissues such as liver are unable to support the system. Hence the bone marrow stromal cells must be providing the right extracellular matrix and growth factors

to support the growth and development of the haemopoietic cells. Some myeloproliferative disorders have been shown to be caused by defects in the stromal cells rather than the haemopoietic target cells e.g. the steel (Sl) and dominant white spotting (W) loci (McCulloch *et al.*, 1964; McCulloch *et al.*, 1965). These loci have subsequently been discovered to encode the haemopoietic stem cell growth factor, SCF, and its receptor *c-kit* (Geissler *et al.*, 1988; Zsebo *et al.*, 1990; Huang *et al.*, 1990).

It has now been established that stromal cells and resident bone marrow macrophages produce a variety of cytokines (Shirai *et al.*, 1993; Temeles, 1993). Some of these, such as stem cell factor (SCF), can be produced in both soluble and membrane bound forms and may have a dual role to play acting as both an adhesion molecule and a growth factor (Williams *et al.*, 1992). Many of the cytokines produced by the bone marrow stromal cells interact specifically with components of the extra-cellular matrix (Gordon,1991; Gordon *et al.*, 1987). The interactions of the cytokine with components of the extra-cellular matrix such as heparin sulphate proteoglycans may be important in preventing proteolysis and help promote specific interactions between the cytokine and its target cell (Roberts *et al.*, 1988).

Haemopoietic progenitor cells have been shown to bind directly to extracellular matrix components which are produced by stromal cells (Siczkowski *et al.*, 1992). The modification of the proteoglycans alters the affinity with which the progenitor cell bind to the extra-cellular matrix. This may allow the creation of specific microenvironments which can help to control the behaviour of the progenitor cells. The idea of distinctive microenvironment within the bone marrow stroma is supported electron microscopy studies of femoral tissue. A 110 kDa protein was localised to a small proportion of stromal cells which are associated with sites of B-cell lymphopoiesis (Jacobsen *et al.*, 1992).

The growth factors which help to control haemopoiesis and their receptors are now being identified and purified. The effect of the growth factors have on the behaviour of stem cells and other cells in the proliferative hierarchy is now being characterised. Some of these effects are probably quite complex and cytokines will act synergistically in combination with other growth factors. The story is far from complete but it seems likely that other hierarchical systems of cell replacement may be controlled by growth factors in a similar fashion. It may be possible to better understand the control of cell growth in the epidermis by understanding how the growth control is achieved in other systems such as haemopoiesis. Already there is evidence that a protein, MIP-1α, acts as an

inhibitor of both haemopoietic stem cells (Graham *et al.*, 1990) and epidermal stem cells *in vitro* (Parkinson *et al.*, 1993).

1.4 Cell hierarchies in the crypts of the small intestine.

The adult mouse intestinal epithelium contains four principal differentiated cell types. These cell types are columnar enterocytes, goblet cell, enteroendocrine cells and Paneth cells. These cells are produced and constantly renewed in a well ordered manner. Studies performed on mice indicate that the four principal cell types are produced from a common multipotential stem cell located near the base of the epithelial structures known as the crypts of Lieberkühn (Cheng and LeBlond, 1974b; Kirkland, 1988).

There are approximately 300 to 550 cells in each crypt depending on its location in the small intestine (Gordon *et al.*, 1992; Potten and Loeffler, 1990). Cell division in the intestine is confined to the crypt. About two thirds of the cells in the crypt divide every 12 hours and the crypt produces between 300 and 400 new cells per day. About 13 to 16 cells leave the crypt every hour and migrate up the surrounding villus as they differentiate.

Cell differentiation and migration are tightly linked in the mouse intestine. Enterocytes, goblet cells and endocrine cells mature as they migrate up the villus in an orderly fashion from the crypt to the villus apex. The cells are subsequently shed into the intestinal lumen. The whole process of migration takes around 2 to 3 days (Wright and Irwin, 1982). Paneth cells differentiate as they descend to the bottom of the crypt. The Paneth cells remain at the bottom of the crypt for about three weeks before they are removed by phagocytosis (Cheng and Leblond, 1974a). Again the number of crypts per villus is dependent on the location of the villus. In the proximal duodenum there are about 14 crypts per villus whereas in the ileum there are only six crypts per villus. In the colon there an no villi and the cells migrate out into a hexagonal surface epithelial cuff surrounding the orifices of the crypts (Schmidt *et al.*, 1985a).

Tritiated thymidine labelling and clonogenic radiation assays have provided insights into the number and location of stem cells in the mouse intestinal epithelium (Potten and Morris, 1988; Potten and Loeffler, 1990). It is obvious that the highly functional Paneth cells at the bottom of the small intestine cannot be the stem cells. The cells leaving the crypt no longer divide and cannot form the stem cell compartment. The stem cells must

be located within the 200 - 350 cells in the crypt which are actively dividing. Analysis of the morphology of the crypt leads to the conclusion that the stem cells are most likely to be located in the ring of epithelial cells which is positioned just above the Paneth cells at the bottom of the crypt.

A number of studies have looked at the cell cycle times in each of the different positions in the crypt (Potten, 1986). The length of the cell cycle time decreases as cells move up the crypt into the villus. After a continuous labelling with [^{3}H]-thymidine the cells which retain label for the longest time seem to be located towards the base of the crypt suggesting that the stem cells which should be more slowly cycling than the transit amplifying population are more likely to be located in the lower part of the crypt.

There are a number of potential problems associated with this type of kinetic data. As cells are continually moving up the crypt it is hard to determine the original position of the cell when it was initially labelled. It is also hard to estimate the fraction of non-proliferative cells at a given position in the crypt. The small number of mitoses at the bottom of the crypt make it hard to obtain statistically significant data. It is also possible that the result could be affected by Circadian fluctuations and that the radioactive label may disturb the kinetics of the tissue. It has been shown that some of the cells near the base of the crypt may be especially sensitive to irradiation.

Regenerative clonogenic assays have been used to attempt to determine the number of stem cells in each crypt (*e.g.* Withers and Elkin, 1969; Withers and Elkin, 1970; Potten and Hendry, 1985; Hornsey, 1985; Potten *et al.*, 1983). The results from the use of this technique really estimates the number of potential stem cells in the tissue and this may not actually represent the number of cells which would normally act as stem cells in the unmanipulated tissue. The estimate obtained from this method is that each crypt contains about 30 - 40 potential stem cells (Potten *et al.*, 1987). The number of potential stem cells identified by this technique changes depending on the dose of radiation administered. As the radiation dose is increased so the estimate of the stem cell number increases. This suggests that the more severe the tissue injury, the greater the number of cells which can respond to the damage by proliferation. This could be explained by more differentiated transit cells replacing the damaged stem cells to repair the injury. Crypts can be sterilised by exposure of the bottom part of the structure to weak β particle radiation (Hendry *et al.*, 1989). This confirms that the majority of cells in the upper portion of the crypt are unable to act as stem cells in the tissue. It is very likely that the number of actual stem cells in the tissue is less than the number of potential stem cells in the tissue. The best estimate from the kinetic data is that each

crypt contains about 4 - 16 stem cells located directly above the Paneth cells (Loeffler and Grossman, 1990; Potten and Loeffler, 1990).

Cell marking techniques have also been used to analyse the organisation of stem cells in the mouse intestinal epithelium and produce a different estimate of stem cell number in the intestine. One method of cell lineage marking that has been adopted is X-inactivation in mice. X-inactivation is a random process which occurs in female mice by embryonic day 6. Stem cells and their descendants will express only the allele of a gene which is on the chromosome which is not inactivated (MacMahon *et al.*, 1983). Female mice were used which were heterozygous for a two phosphoglycerate kinase alleles (*Pgk-1*^{a} / *Pgk-1*^{b}) which code for different isozymes that can be separated by electophoresis. When whole crypts are microdissected from adult heterozygous female mice each crypt only expressed one allele or the other (Ponder *et al.*, 1985). This result was confirmed using another enzyme marker located on the X-chromosome glucose-6-phosphate dehydrogenase (G6PD). This alleles of this enzyme G6PD and G6PDX are separated *in situ* by a histochemical stain. Small intestinal crypts from heterozygous females were examined (Griffiths *et al.*, 1988). Analysis of 400 000 crypts demonstrated that the structures were monoclonal, derived from a single stem cell.

This hypothesis has been supported by other cell lineage markers located on other chromosomes. Aggregation chimeras were prepared by fusion of blastocysts of two strains of mice which varied at the major histocompatibility locus H-2. The embryos fused contained either the *H-2*^{k} or the *H-2*^{b} allele (Ponder *et al.*, 1983). Specific antibodies were used on the tissues of the chimeric mice to differentiate between the cells descended from each progenitor population (Ponder *et al.*, 1986). Again during the analysis of several hundred thousand crypts in adult mice no mixed crypts were observed suggesting the crypts were monoclonal in origin.

The results of the chimeric mouse experiments were extended in small intestine by application of a different marker system. The Dlb-1 locus on chromosome 11 codes for the expression of a binding site for the carbohydrate lectin, *Dolichos biflorus* (Ponder *et al.*, 1985b; Uiterdijk *et al.*, 1986). The gene has two codominant alleles, *Dlb-1*^{a} determines expression of the carbohydrate locus on the vascular epithelium but not on intestinal epithelium. The other allele *Dlb-1*^{b} determines expression of the carbohydrate on the intestinal epithelium but not in the vascular (Ponder *et al.*, 1983; Schmidt *et al.*, 1985b). Aggregation chimeras were constructed by fusion of inbred mouse strain embryos that were homozygous for the *Dlb-1*^{a} allele with embryos homozygous for the *Dlb-1*^{b} allele. The cells in the small intestine expressing the *Dlb-1*^{b} allele are easily

detected by staining the tissue with a peroxidase conjugated lectin. Bands of staining cells were seen to stretch out of the DBA-positive crypts to the top of the adjacent villus. Villi in tissue preparations from adult mice are composed of cells from several crypts and were seen to contain bands of cells from wholly positive or wholly negative crypts. Crypts containing DBA-positive and DBA-negative cells were never observed. It could be shown that the three cell types enterocytes, goblet cells and Paneth cells in a crypt had the same DBA phenotype (Schmidt *et al.*, 1985a; Ponder *et al.*, 1985a). This work agrees with earlier studies that suggest that intestinal crypts are monoclonal in origin and also shows that the villi are polyclonal, composed of cells derived from several different crypts.

Although the use mouse embryo aggregation chimeras is an extremely powerful tool for looking at the organisation of stem cell hierarchies the technique has a few drawbacks. It is possible that cells of a similar genotype may tend to group together and produce an artificial degree of monoclonality in the crypts (MacLaren, 1976). Also it is not possible to mark cells at various stages of development.

These problems were overcome by the development of a mutation assay. Breeding of the mouse strains which are homozygous for the different alleles results in the production of mice which are heterozygous and have one copy of each allele of the gene. These heterozygote mice can then be treated with a mutagen or under go spontaneous mutation which will oblate expression of the gene on the intestinal epithelium by mutating the *Dlb-1*^{b} locus in the occasional cell (Winton *et al.*, 1988; Winton *et al.*, 1989; Schmidt *et al.*, 1990; Winton *et al.*, 1992).

If the cell mutated is a multipotential stem cell then after a sufficient time period all the cells in each crypt would either stain with the lectin or not stain with the lectin. This would imply that each crypt is derived from a single cell. If crypts were derived from multiple stem cells a mutation in the *Dlb-1*^{b} locus in one stem cell should only prevent expression in some of the cells in each crypt. If a shorter time period is left after mutagen treatment of the intestine clones may be present which only cover part of each crypt. This would probably be the result of mutation of cells in the transit amplifying population which are eventually replaced by division of the stem cell at which time the crypt cell population becomes uniform. It would also be possible to detect stem cells present in the embryo which may generate large patches of wholly positive or wholly negative cells.

These expectations were confirmed by analysis of the spontaneous mutation rates of the allele of heterozygous mice in whole mounts of the intestine stained with the peroxidase-

conjugated lectin (Winton and Ponder, 1990; Winton *et al.*, 1991; O'Sullivan *et al.*, 1991). In young adult animals, single or groups of two or more negative crypts were seen at a frequency of approximately 4 crypts per intestine examined. As the animals aged the frequency of single negative crypts but not of groups of negative crypts increased. The increase is in the region of 1-10 x 10^5 crypts per year in each intestine. Wholly negative crypts are seen to accumulate at a similar rate to mixed crypts in serial section of intestinal epithelia. The presence of mixed crypts indicates that there is probably a transit amplifying population present.

Chemical mutagens can be used to increase the number of mutations observed and analyse the maintenance of the stem cell hierarchies in the adult mouse. In the small intestinal crypt, mutations were induced in heterozygote *Dlb-1*^{a} / *Dlb-1*^{b} mice with the point mutagen ethlynitrosourea (ENU). ENU is an ideal mutagen for this purpose as it has a high mutation rate, relatively little cytotoxicity and it is metabolised quickly (Schmidt *et al.*, 1990). The results generated led to the conclusion that a single multipotent stem cell maintains each small intestinal crypt in the adult mouse. This type of analysis was also performed in the colonic crypt epithelium with the X-inactivated G6PD gene (Griffiths *et al.*, 1988). Patches of low enzyme activity were observed in normal C3H mice treated with dimethyl hydrazine (DMH). These patches were similar to that observed in the colon of heterozygous G6PD-deficient mice. The patches of low enzyme activity were restricted to whole crypts. This suggests that the colonic crypt in the adult is also maintained by a single stem cell. No mixed crypts were observed when using this technique which suggests that there were no mutations in the transit amplifying population or that the transit cells were cleared very quickly.

Earlier studies with chimeric animals suggested that as the crypts develop in the postnatal animal they are gradually organised into monoclonal structures. Crypts of mixed origin are found in the epithelium of 2 day old chimeric animals (Schmidt *et al.*, 1988). The villus ribbons associated with these crypts are not coherent but consist of patches of similar cells. As the animal matures these crypts are gradually organised into the monophenotypic pattern described in the adult. By postanatal day 10 the epithelium resembles that found in the adult, all the crypts are derived from single progenitor cell types and the crypts are associated with ribbons of cells stretching up to the tip of the villus. The mechanism by which this occurs has yet to be determined. It is possible that a single stem cell takes over the crypt or that the stem cells are distributed when new crypts are formed by invagination of the intervillus epithelium or by crypt fission.

Transgenic mice have also been used to analyse the stem cell hierarchy present in the intestine of the mouse. The liver fatty acid binding protein (L-FABP) is expressed in both hepatocyes and enterocytes (Sweetser *et al.*, 1988; Roth *et al.*, 1990). It is activated in enterocytes as the cells begin to differentiate (Roth *et al.*, 1991a; Cohn *et al.*, 1991). It is expressed at relatively high levels in the proximal small intestine and expression declines progressively towards the distal ileum (Sweetser *et al*, 1988). The promoter used was part of the rat *Fabpl* gene (Sweetser *et al.*, 1986). The expression of the transgene in the intestinal epithelium was more widespread than that of the wild-type gene (Roth *et al.*, 1990; Sweetser, *et al.*, 1988). Inappropriate expression was observed in a variety of cell types and in the large intestine of adult transgenic animals. As the mice aged expression was gradually switched off. This started at the distal colon and gradually progressed to the proximal segments of the intestine.

All cells of an individual crypt were seen to express the transgene at the same level (Roth *et al.*, 1991b). Mixed crypts containing cells with varying expression levels were never observed. Furthermore the crypts were often grouped into small patches of tissue which had similar expression levels. These observations are consistent with the hypothesis that a single multipotent stem cell is responsible for populating the intestinal crypts. The expression of the transgene in co-ordinated patches may imply that the crypts are derived from a single progenitor cell. This cell may even maintain clusters of crypts in the adult mouse. It is also possible that patches of cells may have been derived from one crypt by the process of crypt fission.

Another property of the stem cells present in the murine intestine is that these cells will maintain their identity if they are removed from their normal environment. The cells from different sections of the intestine undergo slightly distinct programs of differentiation. If tissue is transplanted from various sections of intestine from mouse embryos on day 15 to 16 of development and grafted into subcutaneous tissue of nude mice the tissue is vascularised and continues development. The cells in the graft express markers which are appropriate for the tissue of origin of the graft (Rubin *et al.*, 1991; Rubin *et al.*, 1992). The same conclusion is reached when analysing the tumours of multiple intestinal neoplasia (*Min*) mice (Moser *et al*, 1990). Traces of differentiating cells are found in the spontaneous tumours which arise throughout the intestine of these mice. The differentiation programme of the cells in these tumours is always appropriate to the area of the intestine in which the tumour arose (Moser *et al.*, 1992). These studies suggest that the stem cells are able to express positional information irrespective of its surrounding environment.

1.5 The existence of stem cells in adult liver.

The existence of stem cell hierarchies have long been postulated in rapidly proliferating tissues such as the adult haemopoietic system described above. Recently though similar hierarchies have been proposed to exist in more quiescent tissues such as adult liver (Sigal *et al.*, 1992; Thorgeirsson, 1993). Liver is characterised by a long cellular life span and although the tissue retains its capacity to divide in response to the appropriate stimuli, the turnover rate in this tissue is very low.

The structure of the liver is very well defined. The tissue is divided up into a large number of functional units of hepatic parenchyma. The functional unit, called the acinus, is composed of a cylinder of hepatic parenchymal cells surrounding a central vein, or terminal hepatic venule (Jones and Spring-Mills, 1983). The acinus is divided into three zones which are based on the microcirculatory pattern (Campara and Reynolds, 1989). *Zone 1* is the region nearest to the portal triad and is termed the periportal region, *zone 2* is the central region and *zone 3* is the region surrounding the terminal hepatic venule. There are marked morphological, biochemical and functional differences between the hepatocytes located in each of the three zones (Jungermann, 1986; Traber *et al.*, 1988). Cells in *zone 1* are smaller than the cells in *zone 2* which are of intermediate size. The cells in *zone 3* are larger than the cells in the other zones (Loud, 1968). Other differences are observed in the morphology of mitochondria, endoplasmic reticulum and glycogen granules. Cells in *zone 1* are usually diploid but in *zone 2* and *3* the cells are increasingly polyploid (Wilson and Leduc, 1948). Cells from *zone 1* have a greater proliferative capacity than those in *zone 3* (Gebhardt, 1988). There are differences in gene expression between cells in the different zones. Enzymes for glycolysis and glycogen synthesis are restricted to *zone* 3 (Andersen *et al.*, 1983; Jungermann, 1986) whereas expression of α-fetoprotein is restricted to the periportal cells. Other changes in gene expression between the cells in different zones have also been observed.

The explanation of the differences between these cells has often been assumed to be that the liver cells are adapting to the environment which surrounds them (Andersen *et al.*, 1983; Jungermann, 1986; Traber *et al.*, 1988). As blood flows unidirectionally through the sinusoid from the portal triad to the terminal venule the cells in *zone 3* must be exposed to different concentrations of metabolites after cells in *zones 1* and *2* have removed them from the blood. The cells in the various zones then adapt to the various concentrations of oxygen, substrates and hormones (Bennet *et al.*, 1987; Shiojiri *et al.*, 1991; Kanamura and Asada-Kubota, 1980).

This idea is supported by a study which shows that the direction of blood flow in the liver does alter the expression of many genes, including glycolysis, gluconeogenesis and oxygen consumption. This study was performed by retrograde perfusion of the liver which reverses the direction of the blood flow (Thurman and Kauffman, 1985). However expression of other genes such as α-fetoprotein or glutamine synthase are not effected (Bennet *et al*, 1987; Gaasbeek-Janzen *et al.*, 1987). If *zone 3* hepatocytes are transplanted to the spleen they retain their capacity to produce cytochromes *P-450b* and *e* in response to sodium phenobarbitol (Maganto *et al.*, 1990). *Zone 1* cells do not usually produce these proteins and upon transplantation to the spleen the cells remain unable to respond to phenobarbitol. This response is not dependent the direction of blood flow or the sinusoidal microenvironment.

A possible explanation of the remaining differences between the cells in different zones is that the cells are at different stages of maturation. Rare mitoses seem to be present equally spread throughout the zone (Messier and LeBlond, 1960). It is usually thought that all hepatocytes have an equal capacity for proliferation, despite the heterogeneity in gene expression. Arber *et al.* (1988) have proposed that in adult liver stem cells give rise to a cell hierarchy similar to that found characterised in other tissues. This is supported by the evidence suggesting cells migrate from the periportal region to the central region. After injection of rats with tritiated thymidine labelled cells are first observed next to the portal triad within 300 μm of the rim of the acinus. The cells then move at a speed of 2.5 μm per day until they reach the terminal hepatic vein in the centre of the acinus about 148 days later.

This not only supports the idea of a stem cell compartment but seems to confirm that the heterogeneity of gene expression observed in the cells occurs because these cells are at different stages of maturation. The liver stem cells do not seem to expand quickly in response to liver damage (Farber, 1984), however this could be due to a lag phase before stem cell activation. Stem cell activation is only likely to occur when there is severe damage to the mature cells.

Studies of carcinogenesis provide further evidence of the existence of stem cells (e. g. Evarts *et al.*, 1987; Germain *et al.*, 1985; Germain *et al.*, 1988a; Snell, 1983). Transformation occurs usually in a population of small cell in the periportal region. These cells have been termed oval cells and consist of scant cytoplasm and a relatively large ovoid nucleus. These cells have been proposed as the stem cells in the adult liver, although the normal cellular couterparts to the transformed oval cells have yet to be identified. The oval cells are morphologically similar to bile duct cells but possess

features of hepatocytes. The oval cells are able to divide to produce hepatocytes *in vitro* (Evarts *et al.*, 1987) and can form hepatocyte colonies *in vivo* when transplanted into the liver (Evarts *et al.*, 1989). These cells have been shown to differentiate to bile duct cells *in vitro* (Hayner *et al.*, 1984). This suggests that the liver stem cells may be multipotential and suggests that the cells may have a similar pattern of differentiation to hepatoblasts isolated from embryonic liver (Germain *et al.*, 1988b).

This idea is further enhanced by work which suggests that oval cells similar to those found in the liver are also found in the pancreas (Reddy *et al.*, 1984; Rao *et al.*, 1990). The oval cells from the pancreas show very similar patterns of gene expression to the oval cells of the liver. The oval cells in the pancreas proliferate when copper deficiency is induced in experimental rats. The proliferation of oval cells is accompanied by the atrophy of acinar cells in the pancreas. When the rats are returned to a normal diet the pancreas is seen to contain a number of hepatocytes, suggesting that the oval cells of the pancreas are also able to differentiate to the hepatocyte lineage.

1.6 Lineage studies and stem cells of the nervous system.

During embryonic development changes in the morphology and gene expression reflect the division of the neuroepithelium into regions which correspond to the major subdivisions of the adult central nervous system. Neurons and glia are formed during development on a precise schedule (McKay, 1989; Marvin and McKay, 1992). The neurons are generated first followed by the different types of glia. Neuronal precursor cells are found only in the developing brain in mammals however quiescent glial precursors are retained in the adult brain. In this system the stem cells seem to have a limited life span *in vivo* as the stem cells for the majority of the tissue are lost soon after birth.

Thymidine labelling studies carried out on the cerebral cortex show that it develops from the inside out (Angevine and Sidman, 1961; Rakic, 1988). Cell division takes place in the ventricular zone and the progeny of the dividing stem cells migrate outward once they have become post-mitotic. When new neuronal cells are created they must migrate past the older cells to their assigned positions. This migration is disrupted in the *reeler* mutant mouse which leads to the formation of a cerebral cortex with inverted layers which are highly disorganised (Caviness and Sidman, 1973). The cells are still able to make the connections appropriate for their birthdates (Caviness, 1976; Drager, 1981;

Lemmon and Pearlman, 1981). The mechanism by which neurons acquire their identity is therefore independent of the process of migration in these cells (McConnell, 1991). A series of elegant transplantation experiments demonstrated that the neuronal cells become committed to a particular fate within 4 hours of their final S-phase and migrate accordingly (McConnell, 1988; McConnell and Kaznowski, 1991). These results in combination with the results obtained in the *reeler* mutant mice indicate that multipotent neuronal stem cells become committed to a specific fate, in response to extracellular signals, when the cells are located in the proliferative ventricular zone of the neuroepithelium.

Cell lineage analysis, along with cell culture and transplantation, has proved to be a powerful technique for analysis of the potential of neuronal stem cells (Sanes *et al.*, 1986; Price *et al.*, 1987). One technique that has been used for this is the injection of a tracer dye into single cells. This has been applied to the study of vertebrate systems including zebra fish (Kimmel and Warga, 1987) and *Xenopus* (Hartenstein, 1989; Jacobsen 1985; Holt *et al.*, 1988). It is restricted to large, accessible cells and has not been feasible in the study of the mammalian nervous system. The dye also become diluted as the cell divides making it hard to detect after a number of cell divisions. Another approach used has been the generation of chimeras by the fusion of primitive cells or embryos (Balaban *et al.*, 1988; Goldowitz, 1989; Herrup, 1986). This technique does not really provide information about the lineage of single cells. The technique most widely used in the analysis of the mammalian nervous system has been the infection of the cells with retroviruses. Retroviral marking exploits the fact that retroviruses integrate a DNA copy of their genome into the DNA of the host. By using replication incompetent retroviruses any gene of interest can be transferred to a dividing cell (Cepko *et al.*, 1984; Cepko, 1988; Sanes, 1989; Walsh and Cepko, 1990). A virally transmitted β-galactosidase gene can be used as a simple, histochemical marker which is not diluted and passed on to all the progeny of the infected cell (Price *et al.*, 1987; Price, 1987; Sanes *et al.*, 1986) .

In many parts of the nervous system individual ventricular cells can give rise to a wide range of neuronal cell types (Turner and Cepko, 1988; Turner *et al.*, 1990). Single progenitor cells can give rise to all types of neuron in the retina (Wetts and Fraser, 1988) and this is also the case in the optic tectum of the chick (Gray *et al.*, 1988; Galileo *et al.*, 1990). In the spinal chord of both *Xenopus* (Hartenstein, 1989) and chick lineage (Leber *et al.*, 1990) mapping experiments show that progenitor cells can produce both motor neurons and interneurons. The fate of these cells can be altered by transplantation showing that the precursor cells are probably uncommitted (Placzek *et al.*, 1990;

Yamada *et al.*, 1991). In the cerebral cortex clusters of neurons with a wide variety of morphologies located in different layers have been described although some lines of evidence suggest that early progenitors may give rise preferentially to cells in deep or superficial layers (Price and Thurlow, 1988; Walsh and Cepko, 1992; Austin and Cepko, 1990; Walsh and Cepko, 1990).

Progenitor cells are not only capable of differentiating to produce many neuronal cell types but even glial cells as well (Luskin *et al.*, 1988; Williams *et al.*, 1989). Retinal neurons and Müller glial cells can differentiate from a common precursor cell even during the final cell division (Turner and Cepko, 1987; Wetts and Fraser, 1988). Common precursors have been detected in the chick optic tectum and the spinal cord. In the optic tectum clones have been observed which include neurons, astrocytes and radial glia (Galileo *et al.*, 1990). In cerebral cortex the case is somewhat different. If a precursor cell exists for both neurons and glia it must represent a very small percentage of the precursor population. Most neurons are derived from cells which give rise to neurons alone (Luskin *et al.*, 1988).

Primary culture of cells from the rat optic nerve has led to the identification of glial progenitor cells which may act as stem cells involved in cell replacement in the adult central nervous system (Raff and Lillien, 1988; Noble *et al.*, 1992). In this system there is evidence that the stem cells in the adult tissue are different to the progenitors present during tissue development and that the stem cells of this system may be capable of asymmetric division.

In culture the O-2A cell can give rise to both oligodendrocytes and type-2 astrocytes (Raff *et al.*, 1983a,b; Ffrench-Constant and Raff, 1986). O-2A cells taken from the perinatal rat optic nerve can be induced to divide or differentiate by co-culture with type-1 astrocytes or by addition of platelet-derived growth factor (PDGF)and basic fibroblastic growth factor (bFGF) to the culture medium (Bögler *et al.*, 1990; Wolswijk *et al.*, 1989). Similar O-2A progenitor cells can be identified in tissue from both perinatal and adult rat optic nerve (Wolswijk and Noble,1989; Wolswijk *et al.*, 1990).

Comparison of the two cell populations show that they possess different properties. When cultured with type-1 astrocytes the O-2A^{adult} cells have a unipolar morphology whereas the O-2A^{perinatal} cells are usually bipolar (Wolswijk and Noble, 1989; Small *et al.*, 1987). The adult progenitors had a longer cell cycle time and were less motile than their counterparts derived from perinatal tissue (Wolswijk *et al.*, 1989; Noble *et al.*, 1988). Evidence suggest as the animals mature the O-2A^{perinatal} cells are replaced by

cells of the O-2A^{adult} type (Wolswijk *et al.*, 1990) and work done *in vitro* indicates that some O-2A^{perinatal} cells can be induced to generate O-2A^{adult} progenitor like cells when co-cultured with purified cortical astrocytes (Wren *et al.*, 1992). Cells of the O-2A^{perinatal} divide symmetrically to produce larger populations of progenitor cells whereas their adult counterparts divide asymmetrically to generate one stem cell and on differentiated cell (Temple and Raff, 1986; Wren *et al.*, 1992).

These results suggest that the O-2A^{adult} cells represent a stem cell population present in the adult nervous system (Noble *et al.*, 1992). The cells are continuously present throughout the life of the animal. They may be multipotential, giving rise to both oligodendrocytes and type-2 astrocytes although the existence of the latter cell type *in vivo* is disputed. There is also evidence that the O-2A^{adult} cells can divide asymmetrically which is another property which stem cells have been postulated to possess (Lajtha, 1979). The cells can also be distinguished from their counterparts which are present during embryonic development.

1.7 Stem cell concepts.

The previous sections of this manuscript have discussed the properties of stem cells in two of the tissues, intestinal epithelium and the haemopoietic system in which they have been most widely studied. Also the prospects are examined for the application of the knowledge acquired about the behaviour of stem cells to other tissues in which the stem cells are less well characterised such as the liver and the central nervous system. The work on ES cells has shown that all tissues are originally derived from a common set of multipotential stem cells during embryonic development.

As well as the tissues described above a number of other tissues have been proposed to contain stem cells in the adult mouse. These include renewing tissues such as the testis (De Rooij, 1988) and in addition to skin other epidermal tissues such as the tongue (Potten, 1981), the limbal epithelium of the eye (Cotsarelis *et al.*, 1989; Thoft *et al.*, 1991; Tseng 1991) and the epithelia of the breast (Rudland and Barraclough, 1988; Barraclough and Rudland, 1990) and the uterus (Padykula, 1991). Also other tissues such as cartilage (Solursh, 1989) have been proposed to contain stem cells.

It seems to be becoming more clear that cell replacement in most if not all tissues occurs via a proliferative hierarchy. This is true not only of the tissues with a fast turnover rate

such as the haemopoietic system but also in tissues which have slower rates of cell division such as the liver. A great deal of work has been done on studying the proliferative hierarchy of cells and the control of this hierarchy in the haemopoietic system by growth factors (*e.g.* Sachs, 1987; Dexter and Spooncer, 1987; Spangrude *et al.*, 1991; Whetton and Dexter, 1993; Ogawa, 1993). It is likely that there will be parallels between the haemopoietic system and other tissues proposed to contain stem cells in both the organisation and control of the proliferative hierarchy. The lessons learned from the study of haemopoiesis and other tissues such as intestinal epithelium should be take into account when analysing the behaviour of stem cells in other tissues.

As the epidermis of the mouse is a renewing tissue with a fast cell turnover rate this make it a good candidate tissue for the study of stem cells. Stem cells have already been proposed to exist in the epidermis. It is now almost 20 years since the epidermal proliferative unit was proposed by Potten (1974) and evidence for the organisation of epidermal keratinocytes into a proliferative hierarchy was discovered. The location of the stem cells and the organisation of the proliferative hierarchy in this tissue still remains a mystery. Contradictory lines of evidence suggest that the stem cells may be either located in the epidermal basal layer or in the hair follicles. The study of the epidermis has an advantage over that of the haemopoietic system as all the cells in the tissue remain in the one location and are not constantly moving round the body. This allows an *in situ* analysis of cell lineage to be used to analyse the proliferative hierarchy in the skin. The disadvantage is that the cells are not so easily dissaggregated and sorted.

The project described in this thesis has been directed toward the generation of an *in situ* marking system which can be used to analyse cell lineage in the skin. The remainder of the introduction discusses the evidence for the existence and the location of epidermal stem cells.

1.8 Epidermal structure.

The epidermis is arranged into a series of cell layers (Matoltsy, 1986). The outer layers are composed of horny keratin-filled squames, which protect the organism from the environment. Cells in the inner layers are undergoing the process of differentiation to form the corneocytes. The most primitive cells in the innermost layers are capable of division and divide to replace cells lost from the top layers.

Epithelial cells are the major components of the epidermis, constituting about 85% of the cells in the basal layer, but other cells are present in smaller numbers (Potten, 1974; Potten, 1976; Potten, 1981). The other cells include Langerhans' cells and melanocytes. The epithelial cells are arranged into four distinguishable layers (Matoltsy, 1986). These layers have been called the basal, spinous, granular and squamous layers on the basis of the morphology of the cells in each layer. The basal layer is located on the basement membrane, directly on top of the dermis. Cells move up through the layers as they differentiate to form an anucleated keratin-filled squame.

The epithelial cells in the basal layer are small cells with a high nuclear : cytoplasmic size ratio. These cells are involved in proliferating to provide cells for the differentiating layers. Cells leave the basal layer and move up into the spinous layers. As they do there is a large increase in the cytoplasmic volume of the cell and the keratinocytes start to synthesize the specific proteins and lipids which help to perform the protective functions of the squamous layers. As the cells continue to move up they pass through the granular layer where keratohyalin granules become visible in the cytoplasm of the cell. As the cells move up into the stratum corneum they release enzymes which attack and degrade all the organelles. The soluble components of the cell flow out into the extracellular space. The cellular keratin filaments fold up as the cell compacts to form a tough squame and interdigitates with others as it does so.

In simple stratified epithelia, such as the trunk skin of the mouse, the differentiating cells are arranged in columns. Each column is made up of a simple stack of 10 - 15 single differentiating cells (MacKenzie, 1969). Each of these columns rests on top of a group of 10 - 12 smaller basal cells (MacKenzie, 1970). This is obvious when looking in transverse section through stained epithelial sheets. The hexagonal outline a single differentiating cell can be seen overlying a circular group of basal cells (Potten, 1974). In this system the only site of cell division is the basal layer and hence it is the obvious potential location for stem cells.

In more complex epithelia, such as monkey palm epidermis, there are more cell layers. These layers undulate into alternate shallow and deep ridges (Lavker and Sun, 1983). Cell division in this tissue occurs not only in the basal layer but, more frequently, in the suprabasal layers. On the basis of morphology the cells at the bottom of deep ridges are considered to be more primitive than the cells elsewhere in the epithelium. They have a higher nuclear : cytoplasmic size ratio, fewer keratin filaments and a larger number of melanosomes. The melanosomes are believed to provide protection for the cells against damage by ultraviolet radiation. On this basis these cells have been proposed to represent the stem cells for this tissue.

This hypothesis can be strengthened by measuring the cell cycle times. The first cells to incorporate [^{3}H]-thymidine are cells in the suprabasal layers. This indicates that these cells have a relatively fast cycle time. If the labelling is continued for a long period of time until most of the cells are labelled and left for a long chase period, the cells which remain labelled for the longest period of time are the cells at the bottom of the deep ridges. This suggests that these cells are the most slowly-cycling population. An obvious conclusion from this data is that the more slowly cycling cells on the basal layer are the progenitors of a transit amplifying population found in the suprabasal layer.

1.9 Kinetic analysis of epidermal cell renewal.

1.9.1 Epidermal proliferative unit.

The basal layer of the epidermis is the only site of cell division in simple stratified epithelia as found on the trunk of a mouse (Potten, 1974). This population of cells was classically considered to be homogeneous, *i.e.* all the basal cells had an equal capacity to proliferate. Examination of the cells found here show that 15% of the cells are non-keratinocytes and will not contribute to the epidermal cell compartment. The remaining population of keratinocytes has been shown to consist of cells which are at various stages of the cell cycle, as well as some cells which may be post-mitotic (Wright, 1983).

It can be clearly demonstrated that cells in the mouse epidermis are organised in columns. If the epidermis is removed carefully so the columns are intact, it is possible to see a group of approximately 10 - 12 basal cell underlying a single column of differentiating cells (Potten, 1974; MacKenzie, 1969; MacKenzie, 1970). This figure can be confirmed by counting the number of differentiating cell columns per mm^2 and comparing this with the number of basal cell nuclei per mm^2. If this is done the number of basal cell nuclei is approximately 10 times the number of differentiating cell columns in the same area. This supports the idea that a single differentiating column overlies about 10 basal cells. As cells in the skin are relatively immobile, Potten proposed that the differentiating cell column was the product of the division of the basal cells beneath the column. According to this theory each differentiating column was populated only by the underlying group

of basal cells. The structure was termed an epidermal proliferative unit and it was predicted that each separate unit should contain its own stem cell.

1.9.2 Proliferative indices.

There are many ways to examine the population kinetics of a tissue such as the epidermis. The simplest of these methods are the mitotic index or the labelling index. These measurements indicate the proportion of cells which are at a particular stage of the cell cycle. The mitotic index is easily measured by counting the number of mitoses in an epidermal section and is usually somewhere in the vicinity of 1 - 2% for cells in the basal layer of murine epidermis (Iverson *et al.*, 1968; Wright, 1983).

The labelling index is calculated by injecting [^{3}H]-thymidine into the mouse either systemically or intradermally. The [^{3}H]-thymidine is taken up by the dividing cells and identifies the number of cells in S phase. The thymidine is completely broken down after 40 minutes and this technique represents a flash exposure to the label. The number of cells in S phase identified by this method is usually in the range of 2 - 10%, although this if complicated by a pronounced circadian rhythm of division in the epidermis (Weinstein *et al.*, 1984). The problem with these methods is that they require assumptions about other phases of the cell cycle *i.e.* an increase in the number of cells in S phase could be due to an increase in the length of this phase as well as an increased proliferative rate. These methods are not very informative for analysis of the behaviour of individual cells.

1.9.3 Cell cycle analyses.

A more informative method of looking at the proliferation is to look at the rate of entry of cells into a particular phase of the cell cycle (Duffill *et al.*, 1977; Wright & Appleton, 1980). The mitotic rate or birth rate can be measured by local or systemic exposure to a drug, such as vincristine, which has the capacity to arrest cells in metaphase. It does this by interfering with the function of the mitotic spindle. By looking at the accumulation of cells which are arrested in metaphase it is possible to calculate the birth rate of cells in the tissue. When this technique is used to look at mouse skin the birth rate is around 0.0093 cells cell $^{-1}$ hour $^{-1}$. It is also possible to look at the entry of cells into S phase by labelling twice with tritiated thymidine.

Another method used to examine cell proliferation in the epidermis is to look at transit time of [^{3}H]-thymidine labelled cells to the cornified layers of the epidermis. For mouse epidermis this is about 5 - 9 days, although this method has a tendency towards minimum values as the first cells to reach the upper layers of the epidermis will be counted (Epstein and Maibach, 1965). There is a discrepancy when comparing the time for flash-labelled cells to reach the upper layers of the epidermis and the known transit rate of cells through the cell layers. This is consistent with the cells spending 2 - 3 days on the basal layer, maturing before migrating to the suprabasal compartment. This is further evidence of the heterogeneity of cells on the basal layer of the epidermis.

Perhaps the most powerful technique for estimating the length of the cell cycle in the skin is to determine the fraction of labelled mitoses. This method is based on counting the number of dividing cells which are labelled after a pulse of tritiated thymidine. Samples are taken at regular intervals after administration of a pulse of label and the percentage of mitoses which are labelled, is calculated. There is an initial rise as the cells labelled in S phase begin to go through mitosis. The lag time of this curve represents the length of S phase for the labelled cells. This is followed by a reduction in the percentage of labelled dividing cells which increases again when then initially labelled cells, which have already divided once, divide again. The length of time between the occurrence of the two peaks provides an estimate of the cell cycle time in the tissue.

The two published estimates of the cell cycle time in murine trunk epidermis predict a cell cycle time of 80-110 hours (Hegazy and Fowler, 1973; Wright, 1983). In the first of these as much as 80% of the cells labelled in the first peak appeared to contribute to the second peak although this is considerably less in the second experiment which estimates about 20-30% of the cells in the first peak contribute to the second. Some cells contributing to the first peak in this experiment do not ever go back into S phase. Obviously there is insufficient evidence to make any firm conclusions from this type of experiment. The results do suggest some variability in the length of cell cycle times in the epidermis. This can be explained by the existence of a possible hierarchy in which the progenitor cells have a slower cycle time.

There are problems in interpreting the data from such an experiment. This approach assumes that the cell population in the epidermis is asynchronous *i.e.* all the cells are at varying stages of the cell cycle. In the epidermis this is not the case as there is a pronounced circadian rhythm in this tissue. Another problem occurs when labelled cells migrate to the cornified layers. Here the cells undergo nuclear degradation and release their breakdown products into the extracellular spaces. The breakdown products can be

reutilised by other cells which may become labelled if they use tritiated thymidine from other cells.

1.9.4 Label retaining cells.

Another method to analyse cell production rates *in vivo* is to use a continuous labelling technique. This involves continuous exposure to [^{3}H]-thymidine until all the cells in the epidermal basal layer become labelled. The labelled cells are analysed at a series of time points after administration of [^{3}H]-thymidine has been stopped. The faster cycling cells effectively dilute the label left in their DNA as they divide while more slowly cycling cells appear to retain label for longer periods. Hence this method can be used to identify slowly cycling cells in the epidermis. After about 100 hours of continuous exposure approximately 95% of cells in the basal layer are labelled (Hamilton and Potten, 1972). This is surprising as 15% of the cells are non-keratinocyte cells which are assumed to be post-mitotic. This technique works optimally when used on newborn mice, although it is still impossible to label all the cells and examination of labelled skin reveals that some of the unlabelled cells are almost certainly keratinocytes (Potten, 1981).

If skin is transplanted and examined *in vitro* it can be shown that some cells acquire label at a significantly lower rate than others, suggesting that these cells have a slower cycle rate (Barrandon and Green, 1987; Fusenig, 1986; Dover and Potten, 1983; Jensen *et al.*, 1985; Watt, 1988). These cells appear to be located in the centre of the postulated epidermal proliferative unit. This result can be confirmed by labelling epidermis *in vivo*. Approximately one month after epidermis has been continually labelled for a week a low number of cells can be seen to retain label (Morris *et al.*, 1985). These cells are presumably slow-cycling and are located towards the centre of the epidermal proliferative unit. Electron microscopy performed on skin treated in this way demonstrates that most, though not all, the label-retaining cells (LRCs) are keratinocytes (Morris *et al.*, 1986).

It is impossible to exclude the possibility that these cells are prevented from cycling by damage inflicted by the [^{3}H]-thymidine uptake. It is unlikely that this would delay cells for more than a few hours. 24 hours after treating skin with 12-*0*-tetradecanoylphorbol-13-acetate (TPA) it is possible to distinguish labelled mitoses. This suggests that it is still possible for these LRCs to divide. It can be shown by double-labelling experiments with [^{3}H]-thymidine and [^{14}C]-benzo(a)pyrene that the slowly cycling cells also retain carcinogen for long periods of time. One unexplained feature of the label retaining cells

is that they seem to be more strongly labelled than expected. There is greater than a two-fold difference between the labelling in the LRCs and other cells in the skin suggesting that the cells must have a very slow cell cycle or that mechanisms operate in these cells to conserve the label (Potten, 1981).

If skin containing LRCs is disaggregated and the cells plated out in culture then the colonies formed usually contain several lightly labelled cells in the middle (Potten and Morris, 1988). This suggests that the LRCs have divided in the formation of the cell cluster. If the same procedure is carried out on skin which has been pulse-labelled then the resultant colonies, if they contain labelled cells at all, contain only single labelled cells and very often, do not include any. Obviously this indicates the cells labelled by pulse-labelling do not divide in the production of keratinocyte cultures *in vitro*. This suggests that the LRCs have a greater proliferative capacity than other cells in the epidermal basal layer.

Kinetic data reveals that there is considerable heterogeneity among keratinocytes in the epidermal basal layer of simple stratified epithelia. There is considerable evidence to support the existence of a cell hierarchy presumably populated by the occasional division of slowly cycling stem cells. In continuous labelling experiments it is never possible to label all the cells in the epidermal basal layer. It is impossible to tell if these are post-mitotic mature cells or very slowly cycling cells. Most of the experiments are compatible with the idea of the epidermal proliferative unit as proposed by Potten. However kinetic labelling experiments do not provide definitive evidence of the existence of a proliferative unit. The result of kinetic experiments in the crypts of the small intestine estimated that the stem cell compartment consists of between 4 - 16 actual stem cells per crypt (Potten and Loeffler, 1990). However there is quite clear evidence from lineage analysis in this tissue that a single stem cell is sufficient to populate an entire crypt and there was no suggestion that crypts are usually populated by more than a single cell. Attempts using aggregation chimeras to examine patch size in the skin failed to provide any visual evidence of the existence of discrete proliferative units in the skin. In summary, although kinetic data strongly suggest the existence of proliferative units in the skin they do not provide definitive evidence and the possibility of stem cells in other locations in the skin further complicates this debate.

1.10 Hair follicle stem cells.

1.10.1 Outer root sheath cells.

Slowly cycling cells have also been shown to exist in the hair follicles. These cells retain label in the same fashion as the cells located in the epidermal basal layer. Administration of [^{3}H]-thymidine twice daily for one week from day 3 after birth is sufficient to label nearly all the cells in both the epidermal basal layer and in the hair follicles. 4 weeks after labelling LRCs were found in the bulge area of the outer root sheath (Cotsarelis *et al.*, 1990). This is in the sheet of cells surrounding the hair , just underneath the sebaceous gland. No LRCs were found in the matrix area of the hair near the dermal papilla.

This suggests that the slowly cycling cells in the hair follicle are located in the bulge area of the outer root sheath and that the cells surrounding the hair bulb near the dermal papilla are the transit amplifying population. It had previously been assumed that stem cells for the hair follicle would be located near the bulb of the hair beside the dermal papilla. The existence of LRCs in the bulge area of the outer root sheath in newborn mice can be confirmed in adult mice using a continuous supply of [^{3}H]-thymidine administered by an osmotic pump. If skin was treated with TPA during the initial stages of the labelling process more cells in the bulge area were seen to be labelled. This suggests that bulge cells are stimulated to divide by application of TPA.

The bulge cells have also been distinguished as carcinogen-retaining cells and may be important target cells during multi-stage chemical carcinogenesis. The formation of tumours during this process has been strongly correlated with the stage of the hair cycle the mice were undergoing at the start of treatment. The yield of tumours increases when the treatment is started during the resting phase of the cell cycle when the cells of the outer root sheath would probably be more accessible to carcinogenic chemicals. This also implies that a large number of the tumours derived in this system come from the hair follicle.

Cells in the bulge area have also been characterised ultrastructurally by electron microscopy. They seem to be primitive cells expressing very few differentiation markers. These cells have a high nuclear : cytoplasmic ratio and very sparse expression of keratin proteins. There would be advantages of locating stem cells in the bulge area of the outer root sheath. The stem cells would be located in the permanent portion of

the hair follicle whereas the cells near the dermal papilla undergo cyclic degradation during the catagenic phase of the hair cycle. This area may be resistant to plucking as cells from the outer root sheath are often left behind and the bulge area is further away from the hair shaft that cells surrounding the papilla. The bulge area is a well vascularised area of the follicle which should help maintain the stem cells in a stable, well-nourished, environment.

In a previous report it had been shown that outer root sheath cells were capable of shifting their differentiation to form cells which have an interfollicular epidermal phenotype when cultured under suitable conditions (Coulomb *et al.*, 1986; Lenoir *et al*, 1988). When plucked human scalp hair follicles are implanted on dermal equivalents the regeneration of epidermis seems to be better than epidermis formed by disaggregated basal keratinocytes treated in the same fashion (Asselineau *et al.*, 1986). It seems unlikely that this could be due to contamination of the preparation with interfollicular keratinocytes, as both the upper and lower portions of the outer root sheath were equally capable of generating epidermis *in vitro*. If the result was due to contaminating keratinocytes a greater number should be present in the upper portion of the outer root sheath which is closer to the rest of the epidermis. It is to impossible to detect any interfollicular keratin production in any of the cells in the outer root sheath.

The bulge area of human scalp is much less pronounced than that found in hairs from either the human trunk or the skin of mice. It is quite possible that cells from the bulge area are more easily removed by plucking and it could be these cells which are responsible for generating the interfollicular type epidermis.

This finding was confirmed and extended in another study which demonstrated that a shift to interfollicular phenotype was also observed when disaggregated outer root sheath cells were cultured on dermal rafts at the air-liquid interface (Limat *et al.*, 1991). In this case the epidermis produced was not quite as well differentiated as that produced by the keratinocytes treated in the same manner. However the keratinocyte cultures differentiated more quickly resulting in the loss of viable cells. Cultures set up from the outer root sheath cells appeared to be more primitive , but maintained a viable epidermis for longer than the keratinocyte cultures. The cells from the outer root sheath started to express markers of keratinocyte differentiation which they do not normally express.

The findings of the reports cited above are quite compatible and suggest a location for epidermal stem cells other than the basal layer of the epidermis. It is obvious that there are slow-cycling potential stem cells located in the permanent portion of the hair follicle. It is equally obvious that these cells are quite capable of generating interfollicular

keratinocytes. The bulge area of the outer root sheath represents a good environment for stem cells in a well nourished and probably easily controllable area. Cells located in the epidermal basal layer must be exposed to considerable environmental stress in a highly changeable surrounding. Cells in the basal layer may also be susceptible to mechanical damage more readily than cells in the hair follicle. Hence the bulge area of the outer root sheath seems an equally probable location for epidermal stem cells as the basal layer. It seems possible that the basal layer stem cells are temporary stem cells which are replaced by the division of cells in the follicles.

1.10.2 Germinative cells.

Although there is considerable evidence which identifies the outer root sheath as the location of stem cells in the hair follicle, other work suggests another possible location (Reynolds and Jahoda, 1991a, b). This work suggests that the stem cells may be located in the cell layer which is directly opposed to the dermal papilla. This layer has been termed the germinative layer. Cells from this layer are capable of proliferating and forming organotypic structures when cultured with dermal papilla cells. Other hair cell types such as outer root sheath cells or basal epidermal cells do not show any response when cultured with dermal papilla cells, but attached to dishes and proliferated when cultured with skin fibroblasts.

There may be several problems in designating these cells as stem cells for the hair follicle. The cells do not seem to be slow cycling. The only slow cycling cells found in the hair follicle which retain radioactive label for long periods were located in the outer root sheath. The germinative cells are not located in the permanent portion of the hair follicle. As the hair goes through its normal cycle part of the follicle is lost and then regenerated.

The germinative cells seem to be located in the part of the follicle which is lost during the cycle. Secondly it seems likely that outer root sheath cells are capable of regenerating the population of germinative cells once the lower end of the follicle bulb is removed. A piece of evidence which strongly supports this idea is the induction of hair follicles by the dermal papilla in glabrous epidermis (Jahoda and Reynolds, 1993; Jahoda, 1992). This suggests that epidermal stem cells are capable of forming both the interfollicular epidermis and hair follicles.

There is no evidence to suggest that germinative cells are capable of differentiating towards the interfollicular epidermal phenotype although work from the same group indicates that hair follicles can be induced in glabrous epidermis (Reynolds and Jahoda, 1992). When cultured along with skin fibroblasts these cells remained inactive. It therefore seems unlikely that these germinative cells could be the stem cells for both the follicle and the interfollicular epidermis. It is possible however that the environment generated in culture was not suitable to support the transdifferentiation of these cells from the follicular to the epidermal phenotype. Another potential problem with this study is that while it was necessary to prepare germinative cells by careful dissection the outer root sheath cells were prepared from plucked hairs. Obviously this may have damaged or left behind any potential stem cells.

One possible explanation of these observations is that the germinative cells could be a transit amplifying population derived from the outer root sheath cells. The transit amplifying population then become able to respond more easily to the signals from the dermal papilla. It is perhaps not very surprising that dermal papilla cells are capable of inducing the transdifferentiation of glabrous adult epidermis. After all it seems unlikely that the epidermis in different areas of the body would be composed of different cells types. The differences between epidermis in different body sites are probably generated in response to the different environments. Epidermis in all areas of the body is likely to contain the same type of multipotential stem cell. It would obviously be of interest to attempt to discover which cells in glabrous epidermis are affected by the dermal papilla. It would be also be interesting to discover whether dermal papilla cells are capable of inducing new follicle production in normal interfollicular epidermis.

1.11 Identification of clonogenic cells in the epidermis.

Another method of analysing the division capacity of epidermal cells is to look at the response of cells to damage and how the damage is repaired by the tissue. If all cells in the basal layer were equally capable of acting as stem cells then it would be expected that all cells would be equally capable of responding to damage by proliferating (Potten and Hendry, 1973; Potten, 1981). This idea can be tested by treating skin with radiation *in vivo*.

This technique was originally developed *in vitro*. Cells plated out at low density divide without very much movement and hence form exponentially growing colonies. By

applying a size and time limit on the number of divisions required to produce a colony it is possible to estimate the number of clonogenic cells which have been killed by the treatment. This technique can also be applied to *in vivo* situations.

It is hard to determine the proper shape of the curve *in vivo*, because only the very high dose regions can be determined. This usually produces a straight line. This straight line can be worked back to a zero dose to give a figure for the total numbers of clonogenic cells in the epidermis. This does not take into account however that there would a shoulder on the curve as cells would be capable of surviving and repairing a certain amount of radiation damage. The shoulder on the curve can be estimated by analysing split level data *i.e.* giving two separate doses of radiation which are equivalent to a single dose given at the same time and analysing the difference between the two. Sufficient time must be must be allowed to elapse for cells to repair the damage inflicted by the first dose. This should give an estimate of the repair capacity of the cells.

The number of cells in the basal layer of the epidermis usually works out to be approximately 1.5×10^4 cells per mm^2 in the interfollicular epidermis, and approximately the same number in the hair follicles. This gives a figure of about 3.0×10^4 total basal cells per mm^2. Extrapolation of the straight line obtained for cell response to radiation results in a figure of about 1.0×10^6 clonogenic cells per mm^2. When the figure of clonogenic cells per mm^2 is estimated by using split level data it works out between 309 to 1107. This represents about 2 to 7% of the total basal cells and roughly correlates with the number of proliferative units as proposed by Potten (1974). This suggests that there may be one clonogenic cell and hence one stem cell per proliferative unit.

In the same studies it was noticed that there are some especially efficient clonogenic cells located in the hair follicles. These cells were often seen to produce bigger colonies than their interfollicular counterparts and colonies derived from follicles were often able to produce hair. These data again indicate the possibility that hair follicle cells may be more primitive, less well differentiated cells than interfollicular keratinocytes and may be better able to act as stem cells for the interfollicular epidermis.

1.12 Cell lineage analysis and marking .

The ability to trace cell lineages is critical to the study of many developmental phenomena. Analysis of cell lineages allows us to determine the number of cells that give rise to a given structure, the rate at which cells grow and the origin of the various

cell types that form a tissue in the mature animal (Crick and Lawrence, 1975; Ponder, 1987). Cell marking is a technique which can be used to identify single cells or groups of cells and their descendant clones in an animal. It can be used to identify cell lineages and to analyse the contribution a cell makes to a tissue during normal or abnormal development.

This technique has been used to great effect in tissues such as the intestinal crypts, haemopoietic system and the central nervous system as described earlier (Hall and Watt, 1989; Winton and Ponder, 1991; Marvin and MacKay, 1992). The application of this technique to the analysis of cell division and replacement in the skin should help to reveal how cell turnover is controlled. Analysis of clone size should contradict or support the current theories on the location of epidermal stem cells either in the basal layer of the epidermis or in the outer root sheath of the hair follicle.

There are several properties a marker should possess in order to make it useful (Ponder 1987). Firstly it should be cell-autonomous *i.e.* its expression should be limited to cells of the appropriate genotype and not diffusible into other cells. Ideally it should be stably expressed under all conditions. It is important that the marker should not affect the behaviour of the marked cells in any way. It must also be detectable *in situ* if the marker is to be used to analyse the organisation of single cells or groups of cells in a tissue.

An ideal marker would be expressed in all tissues at all stages of development, however it is unlikely that this would ever be achieved in practice. This is not required if the study is limited to a particular tissue. It is also helpful if both populations are marked otherwise it is necessary to assume that the unmarked cells belong to the other population which is not necessarily the case. If attempting to look for small groups of cells in a tissue it is much better that these cells be the marked cells for positive identification.

1.12.1 Cell lineage analysis in *Caenorhabditis* and *Drosophila*.

Cell lineage experiments in invertebrates have been designed to elaborate how the cells of the embryo are capable of generating the adult. Intuitively there are two possible ways that cells may be induced to form the patterns required for the development of adult organs. Cells may divide to generate large clumps of cells used to make adult organs with no regard to their ancestry. An alternative method of pattern formation would be that the entire lineage of a cell would be invariant. Cells would divide in a

fixed an pre-programmed pattern to generate organs arising from specific sets of related cells.

The latter method is almost exactly what happens during development in the small free-living nematode *Caenorhabditis elegans* (Gilbert, 1994; Schierenberg 1989). The work done with this organism demonstrates the enormous potential of lineage studies. These studies along with the extensive work on the genetics of the organism have ensured that *C. elegans* has become a well-known model organism for developmental biologists. This organism has a number of favourable properties for the study of embryogenesis. The embryo can develop outside the mother in a drop of water. The embryo is small and the eggshell is transparent allowing development to be observed conveniently under the microscope. The developmental process is rapid and take less than 12 hours from fertilisation to hatching.

These advantages in combination with the simplicity and reproducibility of the development of this organism has allowed the entire cell lineage of the nematode to be determined both throughout embryonic development and post-embryonic stages (Sulston and Horvitz, 1977; Sulston *et al.*, 1983; Kimble and Hirsch, 1979). It is possible to write this out in the form of a branching tree. Cell lineages were visualised in the developing embryo under a microscope with Nomarski optics. The embryonic cell lineage is essentially invariant giving rise to a total of 671 cells during development of the nematode, 113 of which undergo programmed cell death. In the nematode the patterns of division, programmed death and terminal differentiation are constant from one individual to another. During post-embryonic development approximately 10% of the cells continue to divide eventually generating about 1000 somatic cells in the adult and about 2000 germ cells a few days after hatching.

Analysis of normal embryogenesis in *C. elegans* has produced some insights into the developmental principle which underlie the formation of the embryo (Strome and Wood, 1983; Sulston *et al.*, 1983). In the first hour after the onset of cleavage in the embryo five somatic founder cells (termed AB, MS, E, C and D) and the primordial germ cell (P_4) are created by a series of unequal divisions. Most of the founder cells contribute to a number of different tissues. Most tissues are derived from more than one founder cell. The gut and the germ line however arise solely from single founder cells, E and P_4 respectively. Bilateral symmetry seems to originate very early within the establishment of individual lineages. This happens as early as the first division in the MS, C and D lineages. Even during the last division of a cell it is possible for it to give rise to two different cell types.

Programmed cell death occurs during the formation of the nematode from the embryo. More than 100 cells die at specific positions during embryogenesis. Mutations which prevent this type of programmed cell death do not cause abnormal phenotypes on the adult worms (Ellis and Horvitz, 1986). Cell death is not therefore a prerequisite for normal development. The patterns of division and establishment of sublineages are repeated throughout the maturation of the embryo. This makes it likely that the different lineages were established by a process of duplication and divergent evolution.

Two fundamentally different cleavage patterns are observed during the process of development in the nematode embryo. During generation of the germline an unequal division of cells is observed similar to that proposed for stem cells. Conversely throughout development of the soma, cell division produces cells of similar size and behaviour. The mechanisms which control development can be further revealed by disruption of the causal chain of events controlling embryogenesis. Embryos have been manipulated in various ways and the effects on cellular development can be interpreted in the light of the lineage information already gathered. Numerous mutants have been isolated which show effects during embryogenesis (Kenyon, 1994; Ambros & Moss, 1994; Gilbert, 1994). Studies on these mutants have concluded how important maternal gene expression to supply the egg with essential gene products is to embryogenesis in the nematode. This is also similar in many other organisms such as *Drosophila*. Only a small region of the cytoplasm in the egg is vital for correct development. The embryo can compensate for lost cytoplasm and experimentally induced dwarf eggs can generate the full complement of founder cells and develop normally.

Another technique that can be used to disrupt embryogenesis in order to study the resultant effects on embryogenesis is cell ablation (Shierenberg, 1989). A number of cell ablation experiments have been performed in the embryo of *C. elegans* by repeatedly pulsing the nucleus of selected blastomeres with a laser beam. The E_p cell is the posterior daughter of the founder cell E and is the precursor of the posterior half of the nematode intestine. When this cell is ablated the worm hatches with a normal anterior half-gut which ends blindly and no posterior half (Laufer et al., 1990). The ablated cell remains intact and is located in the posterior half of the animal. These worms are still able to move and feed normally but die after a few days without reaching adulthood. This demonstrates a general principle of development in the nematode. The lineage of cell is invariant and if a precursor cell is removed then all the progeny of the ablated cell are missing when the worm hatches. The control of development in this organism is largely cell autonomous.

The invariant development of the nematode is probably exceptional and development in other organisms including flies, mice and humans is much more indeterminate (Gurdon, 1992; Kimmel and Warga, 1987; Garcia-Bellido et al., 1979). The origin and destiny of cells can not so easily be visualised in other organism as they are in the nematode worm *C. elegans* and hence the cell lineages can not be analysed. To analyse lineage in an organism such as *Drosophila melanogaster* an objective method that will put an indelible mark on an embryonic cell which will be displayed in all of its descendants in the adult (Crick and Lawrence, 1975; Akam, 1987; Ingham, 1988). In *Drosophila*, lineage analysis has been accomplished primarily though the generation of mosaic animals be genetic techniques.

Mosaic flies have been commonly generated through mitotic recombination (Lawrence *et al.*, 1986; Harrison and Perrimon, 1993). X-rays are used to induce chromosome breaks which result in the exchange of parts between homologous chromatids. This exchange occurs very rarely and a picture of cell lineage is built up bit by bit, by looking at clones found in a large number of individuals. One such marker gene is *multiple wing hairs* (*mwh*). A mutation in this locus (mwh^-) results in the production of a number of hairs by wing epidermal cells which usually only produce one hair. The mutant flies remain viable and fertile. Heterozygote flies are produced by crossing mwh^- and mwh^+ flies to produce F1 heterozygote flies with an mwh^- / mwh^+ genotype. As the F1 flies mature they are irradiated *en masse* with 1000 rads and the surviving adults are searched for clones with mwh^- cells which are seen as patches of cells producing extra hairs.

A great deal of information can be elucidated by the study of clones. The later the irradiation, the greater the number of target cells and hence the higher the frequency and the smaller the size of the clones produced. These data allow the average frequency of cell division and the mean growth rate of the tissue to be calculated. When sister cells become labelled it is possible to compare the number of divisions followed by each cell. It is also possible to assess to which parts of the tissue the descendants of a marked cell contribute. By using this technique it was possible to show that late in development the clones mark only the bristles or the cuticle of the wing epidermis and that these cell types must be derived from separate sets of cells (Bryant, 1970). When clones are made earlier they can contain both cell types.

One of the most interesting aspects of cell marking and lineage studies in *Drosophila* is that the cell marker mutant genes have been combined with genes which effect cell behaviour. These mutations can either affect the developmental fate of the clone of cells such as *engrailed* (*en*) or the speed of growth of the cells such as *Minute* (Crick and

Lawrence, 1975; Morata and Lawrence, 1975). *Minute*$^+$ homozygote cell clones grow faster than *Minute*$^+$/*Minute*$^-$ heterozygote cells. The use of the latter mutation led to the discovery of cell compartments and competition. If a marked cell grows faster than its neighbours it will make larger patches in the adult fly. It is obviously of interest to determine whether or not there are limits to the territory that can be occupied by the descendants of a single cell.

It was discovered by the use of the *Minute* growth mutation that marked clones are confined to precisely determined domains which have been termed compartments (Lawrence, 1981, 1988; Ingham and Martinez Arias, 1992). A *Minute*$^+$ clone generated in a heterozygote animal grows more quickly and competes out the cells around it until it comes up against a compartment border. When embryos are irradiated at the blastoderm stage a marked daughter cell is formed by a cell which has undergone homologous chromatid exchange, at the next mitosis in developmental stage 9. The cell even if it grows relatively rapidly only colonises one precisely defined part of each segment at either the anterior or the posterior of the developmental segment. The cells will never colonise both anterior and posterior compartments (Garcia-Bellido et al., 1973). Hence even as early as the blastomere stage the cells of the embryo must be allocated to make either the anterior or posterior part of each segment.

If a similar experiment is performed before the embryo reaches the blastoderm stage a fast-growing clone can colonise several compartments. Hence it is possible to determine the precise timing of cell allocation in the developing embryo. These experiments indicate that the allocation of cells to the anterior or posterior segment compartments occurs some time after the beginning of the blastoderm and before the epidermal cells enter a second post-blastoderm mitosis. Cells are still able to colonise different appendages such as leg and wing but are confined to either the anterior or posterior compartment in both (Morata and Lawrence, 1977). Colonies generated later are confined to either the leg or wing. At even later stages the colonies in the wing are restricted to either the dorsal or ventral side. These results indicate how compartments develop in a progressive fashion and show that in the young embryo the wing and leg in the same segment have a common origin.

Lineage studies lead to the conclusion that following the blastoderm stage the embryo is divide into stripes of determined cells. These cells will generate either the anterior or posterior compartments of the adult segments. The lineage studies also seem to indicate that all the cells are allocated and that there are no intermediate cell capable of contributing to more than one compartments.

The work on compartments has been confirmed and extended by the investigation into the function of another mutant fly gene, *engrailed* (DiNardo *et al.*, 1994; Heemskerk et al., 1991; Vincent and O'Farrell, 1992). A great deal of work on this gene has depended largely on the same sorts of cell lineage experiments as described above. The function of this gene can be investigated in flies in combination with a clonal marker to establish the requirement for the gene in cells (Morata and Lawrence, 1975, 1979; Lawrence and Morata, 1976; Morata *et al.*, 1983). The results of these experiments determine that *engrailed* was a selector gene necessary for all posterior cells, not only for determining the pattern that the cells form but also in maintenance of a separate identity from the anterior cells. When the gene was cloned and antibodies raised against the protein the gene was determined to be expressed in a pattern of stripes which first appear in the young embryo and persist indefinitely in the cell of the posterior compartments.

The *engrailed* mutation was combined with the *Minute* mutation as well as a cell marker and the effect of the removal of the wild type *engrailed* allele from clones of cell was tested. Another advantage of using this system is that strong *engrailed* alleles that are lethal to the whole animal can be tested in clones. Removal of *engrailed* expression from the anterior cells cell has no effect. When the expression is removed from posterior cells changes only those cells which no longer express the gene. The posterior cells acquire some characteristics of the anterior cells but are not organised in exactly similar patterns. When stronger *engrailed* alleles were used the pattern of marked posterior clones was considerably more disorganised (Kornberg, 1981).

It seems that the removal of *engrailed* expression from posterior cell affects their 'affinity' for mixing with anterior cells, although this is not a direct effect of engrailed which is a nuclear protein. The degree of intermingling at the compartment border is reduced an this separates the anterior cells from the posterior and prevents clones of cell spanning the compartment border. When *engrailed* expression is removed from posterior cells the cells acquire anterior-like adhesive properties. The cell can then cross the compartment border displacing anterior cells as they go. All these experiments serve to illustrate that the anterior cells are different to the posterior cells and that these differences are dependent on the *engrailed* gene.

Recently a novel method of marking cells in *Drosophila* has been described (Struhl and Basler, 1993; Harrison and Perrimon, 1993). This method circumvents the problems with using X-irradiation for clonal induction. The typical dose of X-rays used in such experiments gives a low rate of clone formation and a high rate of cell lethality. Two

different strategies have been used which exploit the site-specific yeast recombinase, FLP. The first method uses mitotic exchange in the same manner as X-ray induced recombination. Two homologous chromosomes carry an FLP recombination target site (FRT) at the same chromosomal position, with a mutation on one homologue. The FLP recombinase gene is expressed under the control of a heat shock promoter and catalyses exchange of the chromosome arms specifically at the FRTs. This can generate patches of mutant tissue in an otherwise normal animal much the same way as the x-ray induced recombination.

The other variation of this method is the 'flp-out' technique. A bacterial *lac Z* gene is expressed under the control of a constitutive promoter. The *lac Z* gene contains a 5' transcriptional stop sequence flanked by two FRTs, inserted into the coding sequence. Expression of the FLP recombinase again under the control of a heat shock promoter results in specific excision of the transcriptional stop sequence by recombination occurring at the FRTs. This restores the reading frame of the *lac Z* gene which is now expressed under the control of its own promoter. The expression of the gene is passed on to the progeny of the marked cell and creates clones useful for analysis of cell lineage.

This second technique has already been used to great effect in the analysis of the function of the function of the *wingless* (*wg*, DWnt-1)and *hedgehog* (*hh*) genes in the creation of ventral patterns in the *Drosophila* leg (Struhl and Basler, 1993; Basler and Struhl, 1994). Clones of cell were randomly induced to express the wingless protein in the leg imaginal discs. When these clones are located in the prospective ventral portions of the leg disc the cell develop normally. When the induced cells are located in the prosepective dorsal parts of the disc the clones give rise to ventrolateral patterns and seem to exert a ventralising influence on neighbouring wild-type cells. Some work has been performed in the application of this technique to mice although it has not been widely used to generate lineage information in this organism yet (Wahl *et al.*, 1991). Despite not having been extensively used in cell marking, this type of technique has been now used to generate a tissue specific knockout in mice (Gu et al., 1994).

1.12.2 Cell marking in mammals.

One method of indelibly marking cell in mouse tissues is to generate aggregation chimeras with mouse embryos. This can be done in several separate ways, either by

fusing two embryos (Schmidt *et al.*, 1987, 1988), injecting cells into a blastocyst (Lallemand *et al.*, 1990) or repopulation of a tissue after damage *i.e.* a bone marrow transplant (Joyner *et al.*, 1983). Embryo aggregation chimeras have been used to analyse the clone size of cells in the skin. Embryos from mouse strains which have different H-2 haplotypes were fused. The two parental cell types were distinguished in skin whole mounts by using monoclonal antibodies directed against the H-2 molecules. This study found no evidence for the existence of an epidermal proliferative unit populated by a single stem cell as, at areas of overlap between the patches of parental cells, columns of differentiating cells were found which contained both parental cell types. If each column was populated by a single stem cell each differentiating column should be of one parental cell type or the other. This was also found to be true of hair follicle which were seen to contain a mixture of parental cell types. It is not really possible to analyse the behaviour of single cells in this system or to tell whether a group of cell is descended from a single progenitor.

Another method for cell marking which is able to analyse the behaviour of single clones of cells is to use retroviruses (Soriano and Jaenisch, 1986). Cells can be infected with a replication-incompetent retrovirus which contains a gene which is not normally present in the tissue and can be detected *in situ* with a histochemical stain. One gene which can be used in this fashion is the *E. coli lac Z* gene (Price, 1987). This type of marking has been used to analyse the lineage of cells in the nervous system (Price *et al.*, 1987; Walsh and Cepko, 1988; Austin and Cepko, 1990) and the retina (Turner and Cepko, 1987) as well as in post-implantation mouse embryos (Sanes *et al.*, 1987) and in renal development (Herzlinger *et al.*, 1992). The study on post-implantation mouse embryos helped to define some of the lineage paths for the skin (Sanes *et al.*, 1987). By infecting embryos at a series of timepoints after fertilisation of the embryos it was possible to demonstrate when the precursor cells which form both the epidermis and the periderm (the outermost layer of epidermal cells which is shed before birth) become restricted. The embryos were infected with the virus on day 7, 9 and 11 after fertilisation. When the clones were analysed on day 13 it was obvious that clones from embryos injected on day 9 included both epidermal and peridermal cells. Clones from embryos injected on day 11 were restricted to either the epidermis or the periderm.

There would be problems in applying this technique to analysis of clone size in adult murine epidermis. It may require cell division before the virus genome can integrate an the cell becomes marked (Ponder, 1987). If the virus is only present for a short time and the stem cell cycle is very long then it may prove very hard to mark a stem cell. Also there is the problem of access of the virus to the cells, if the virus does not have access

to the cells it cannot infect them. At the very least it would be necessary to inflict serious tissue damage to allow the virus access to all the epidermal cells. This might affect the result obtained from the experiment.

1.13 Aim of the project.

The aim of this project is to develop and use a marker system to analyse cell lineage in the skin. This should provide useful information on the usual size and organisation of clones in the epidermis and give some clues as to how cell replacement is normally controlled in the epidermis and to the possible location of the stem and progenitor cells.

As there is no marker gene currently available in the epidermis a marker gene can be introduced. This can be accomplished by generating transgenic mice with a novel marker gene inserted into their DNA. Potentially every cell of the mouse may express the marker gene. As the *E. coli lac Z* gene is not expressed by mammals this makes this gene very suitable for the purpose (Price, 1987). The gene can be used as a cell marker by altering the expression of the gene by treating the mice with a chemical mutagen. Altering expression of the transgene in this fashion would provide a means of marking cells in a similar way to the experiment performed on the small intestinal epithelium with lectin markers (e.g. Winton *et al.*, 1988). The *lac Z* gene codes for a protein, β-galactosidase which is easily detectable *in situ* by a simple histochemical stain. The stain is suitable for use on cultured cells as well as tissue sections and whole mount preparations.

It is important to consider that expression of a gene in transgenic mice is unlikely to be ubiquitous. Ideally the transgenic marker should be expressed in all tissues at all stages of development. It is unlikely that this would be achieved in practice. Transgenes are often subject to alteration of their expression by methylation of the genes and by imprinting. This will probably lead to variable patchy expression of the transgene *in vivo*. The project involves looking at a small minority of cells in the population. This combined with the probable variation in expression levels of the marker means that it would be advantageous if the cells which were marked could be positively identified *i.e.* the cells that express the *lac Z* gene.

Cells can be positively identified if the marker gene is switched on in marked cells rather than switched off. If the transgene coding for the β-galactosidase protein, includes a point mutation in the coding sequence then expression of the gene can be switched on, in

the cells to be marked, by reversing the mutation. It is necessary that the point mutation must prevent expression of the gene. When then mutation is reversed the coding sequence of the gene is restored and expression of the protein is allowed. Detection of the intact protein by the histochemical stain for β-galactosidase gene allows identification of the cell as a marked cell.

Using this reverse mutation strategy also gets round another potential problem in generating transgenic mice for this purpose. Usually when transgenic mice are generated, by injecting fertilised embryos, multiple copies of the injected DNA integrates into the mouse genome (Palmiter and Brinster, 1987; Westphal and Gruss, 1989). Attempting to mutate every copy in one cell in order to mark it, would be next to impossible. Using the reverse mutation strategy, mutation in only one of the integrated transgenes would be required to cause expression of the protein and mark the cell. In this case high transgene copy number may be an advantage as this may increase the frequency with which marked cells are generated by increasing the number of targets for the mutagen.

Point mutation is a rare event and a specific reverse mutation to activate a gene would be even less frequent. There may be problems generating sufficient marked cells to detect *in vivo*. In order to attempt to increase the frequency of this reactivation it was decided to design a mutation in the marker gene which correlates with the specificity of particular chemical mutagen. It has since been shown that using this type of approach a reverse mutation rate of 1×10^{-3} can be achieved when treating cells in tissue culture with mutagens (Schaff *et al.*, 1990).

In order to increase the mutation frequency still further it is possible to engineer a construct which has a number of target sites, a mutation at any one of which will result in gene reactivation. This can be done by replacing the ATG translation initiation codon with an oligonucleotide which encodes a number of target sites for a chemical mutagen. Mutation at any one of these target sites results in the production of an in-frame start codon which will begin translation of the protein in the oligonucleotide and read on into the *lac Z* gene. This should produce a fully functional β-galactosidase fusion protein which has the first few amino acids of the normal protein replaced by amino acids encoded by the oligonucleotide. A mutation in this site is also likely to prevent expression of the protein whereas amino acid substitutions in the coding sequence of the gene may not completely inhibit protein function. Predicting which amino acid substitutions would prevent expression could prove to be very tricky. Stop codons in the coding sequence may also prove to be leaky and allow some expression of the protein.

1.14 Mutagenesis.

In order to increase the chances of developing a useful marker system it was decided to design two different oligonucleotides. Each oligonucleotide is designed so it can be activated by a separate chemical mutagen. The mutagens were selected because they have defined specificity of mutation and induce different types of mutational events.

1.14.1 N-nitro-N'-methyl-N-nitrosoguanidine.

N-nitro-N'-methyl-N-nitrosoguanidine (MNNG) is a mutagenic alkylating agent. It requires metabolic activation before producing a methyl diazonium ion which is a reactive species capable of methylating DNA (Burns *et al.*, 1987). MNNG is a powerful mutagen and is more powerful than other common DNA alkylating agents methylnitrosourea (MNU) and methyl- (MMS) by 100 and 500 times respectively, in causing cell transformation (Doniger *et al.*, 1985; Cupples and Miller, 1989). The potential of alkylating agents to transform cells and hence presumably to cause mutations correlates with the amount of O^6-methylguanine adduct they produce. MNNG can cause other DNA lesions but the O^6-methylguanine adduct seems to be the major mutagenic lesion. This adduct probably causes mutations in the DNA by mispairing with thymidine during DNA synthesis to produce G:C to A:T transitions.

MNNG is active in epidermal cells and is capable of transforming epithelial cells *in vitro* and is capable of inducing tumours *in vivo* (*e.g.* Rhim *et al.*, 1986; Pai *et al.*, 1983; Weissman *et al.*, 1989). The compound also induces unscheduled DNA synthesis in primary cultures of murine epidermal cells which suggests the compound is actively damaging DNA in these cells (Sawyer *et al.*, 1988).

In a forward mutation assay in the *lac I* gene of *E. coli* it can be shown that > 97% of the mutations caused by MNNG are G:C to A:T transitions (Burns *et al.*, 1987). If the influence of the surrounding sequence is taken into account it can be shown that mutations are far more likely to occur at a guanine residue flanked on the 5' side by another purine. This was despite the fact that inactivating mutations could have occurred at G residues flanked by either purine or pyrimidine residues.

Mutation was even more likely to occur if the 5' flanking base is a G residue. The mutation rate at a guanine residue flanked by a 5' G was 9 and 2 times higher respectively than at guanine residues preceded by a 5' pyrimidine or 5' adenine. This

specificity of mutation has been shown to apply to other alkylating agents as well (Glickman *et al.*, 1987), and is also maintained in mammalian cells. Cells were transfected with a shuttle vector encoding the *lac I* gene. Stable transfectants were treated with the alkylating agent nitrosomethylurea (NMU). When mutations in the shuttle vector were analysed > 95% of the mutations were G:C to A:T transitions and the same bias towards mutations at residues flanked by a 5' guanine was noted.

1.14.2 N2-acetylaminofluorene.

As the first oligonucleotide was designed as the target for a point mutagen it was decided to design the second for a mutagen which caused an alternative type of mutational event. N2-acetylaminofluorene (AAF) is a potent liver carcinogen (*e.g.* Poirier *et al.*, 1991). When it is metabolically activated in male rat liver it is capable of producing two major DNA adducts which both occur on C-8 position of guanine (Beland *et al.*, 1987). The most common adduct produced is an aminoflourene (-AF) adduct which accounts for 80% of the total DNA adducts. The next most common adduct is the acetylated form acetylaminofluorene (-AAF) which accounts for 15 - 20% of the remaining adducts.

When DNA is reacted *in vitro* with N-acetoxy-acetylaminofluorene (NAc-AAF) a reactive metabolite of AAF, -AAF adducts are formed on the C-8 position of guanine (Burnouf and Fuchs, 1985; Shibutani *et al.*, 1991). In a forward mutation assay when the DNA containing the adducts is transformed into *E. coli* > 90% of the mutations recovered were frameshift mutations (Burnouf and Fuchs, 1985; Burnouf *et al.*, 1989; Seeberg and Fuchs, 1990). Analysis of the sequence at which these mutations occur shows that they primarily occur at repetitive sequences and at specific sequences. A specific sequence at which the -AAF adduct causes a frameshift is the *Nar I* restriction site $^{5'}$GGCGCC$^{3'}$ in the tetracycline resistance gene of pBR 322. The -AAF adduct causes a frameshift mutation by removal of a 2 base pair sequence GC to go from $^{5'}$GGCGCC$^{3'}$ to $^{5'}$GGCC$^{3'}$. This specificity can be further analysed by adding an extra GC base pair to the sequence to form a *Bss HII* restriction site $^{5'}$GCGCGC$^{3'}$. This reverts to the wild type sequence when exposed to NAc-AAF and transformed into *E. coli* in a similar fashion to the forward mutation assay. A reversion frequency of about 1 x 10^{-3} is observed.

When the same analysis was performed with the -AF adduct the change observed was primarily base substitutions (Biachiara and Fuchs, 1985). The -AAF adduct is larger

than the -AF adduct and may cause local DNA distortion. This distortion could cause slippage of the DNA polymerase during DNA synthesis resulting in a deletion of a GC dinucleotide from the sequence.

CHAPTER 2.

Materials and methods.

2. Materials and methods.

2.1 Materials.

Restriction enzymes were obtained from Boehringer Mannheim, Lewes, East Sussex; New England Biolabs (CP Labs, Herts.); Northumbria Biologicals Limited (NBL), Northumberland; or from GibcoBRL, Paisley. Proteinase K, RNase A and Klenow fragment of *Escherichia coli (E. coli)* were obtained from NBL. Polynucleotide kinase was obtained from Pharmacia Ltd., Milton Keynes, Buckinghamshire, and *Taq* polymerase from Promega, Southampton.

The Sequenase™ kit used in double-strand sequencing reactions was from Cambridge Biosciences, Cambridge, and the SP6/T7 transcription kit used in the RNase protection assay was obtained from Boehringer Mannheim. The Geneclean™ kit for DNA fragment isolation was obtained from Stratech Scientific Ltd., Luton, Beds. Deoxynucleotides and mixed hexanucleotides were from Pharmacia. All radioisotopes came from Amersham International P.L.C., Amersham, Buckinghamshire.

DNA size markers (Lambda/Hind III, ΦX174/Hae III) for gel electrophoresis, and the RNA ladder used in Northern blotting were supplied by Gibco/Bethesda Research Laboratories, Paisley, as were the Ultrapure™ agarose and urea. Nitro-cellulose and nylon membranes were supplied by Sartorius Instruments Ltd., Belmont, Surrey; Gelman Sciences Ltd., Broadmills Northampton and Amersham International.

Bacto-tryptone, bacto-agar and yeast extract were from DIFCO Laboratories, Detroit, Michigan, U.S.A. Phenol was delivered as a water-saturated liquid from Rathburn Chemicals Ltd., Walkerburn, Peebleshire. Ethanol was supplied by James Burroughs (F.A.D) Ltd., Witham, Essex. All other chemicals were obtained from BDH Laboratory Supplies, Poole, or Sigma Chemical Co. Ltd., Poole, Dorset.

Serum, media and supplements for cell culture were obtained from GibcoBRL, Paisley. Plastic-ware for cell culture was supplied by Nunc Intermed, Roskilde, Denmark; Sterlin Ltd., Feltham, Middlesex; or by Falcon™, Becton Dickinson, Dublin. Mice were from Harlan Olac Ltd., Bichester, Oxon.

Plasmid pMC1871 was obtained from Pharmacia and the plasmid pIrv-Neo-Act-LacZ was obtained from R. Beddington, Developmental Biology Unit, Department of Zoology, University of Oxford.

2.2 Oligonucleotide synthesis.

Oligonucleotide were synthesised on an Applied Biosystems 381A Synthesiser, using the manufacturer's protocols and reagents. After deprotection in ammonia at 55 °C overnight, the oligonucleotides were precipitated and used without further purification.

2.3 Analysis of oligonucleotides by polyacrylamide gel electrophoresis.

Oligonucleotide integrity was checked by electrophoresis on a polyacrylamide gel. 5 μg of oligonucleotide was dissolved in 10 μl dH_2O and 3 μl of sequencing buffer was added. The solution was heated to 90 °C for 5 minutes. The sample was then run on a 16% denaturing polyacrylamide gel. Oligonucleotides were visualised by shadowing on a TLC plate.

2.4 Restriction enzyme digestion of DNA.

Plasmid DNA was digested with 5 - 10 units enzyme/μg DNA for 1 - 5 hours depending on suppliers instructions. Genomic DNA was digested under overnight conditions with 10 units/μg DNA. Digestions were analysed for completeness, by gel electrophoresis.

2.5 Agarose gel electrophoresis.

Agarose gel electrophoresis of DNA was performed using the appropriate horizontal gel apparatus, with agarose concentrations between 0.7 - 2.0% w/v, depending on the

desired fragment resolution. The agarose was dissolved and cast routinely in TAE buffer (40 mM Tris·Cl pH 7.8; 20 mM Sodium acetate; 1 mM EDTA) which was also used as the running buffer. Ethidium bromide (1 μgml^{-1}) was either added direct to the molten agarose, or in the case of Southern blotting, the gel was soaked in an ethidium bromide solution, prior to UV illumination and photography.

2.6 General cloning techniques.

2.6.1 Purification of Fragments from Agarose Gels.

Using UV illumination the fragment of interest was excised with the minimum amount of agarose. Purification was carried out by two methods depending on fragment size. For fragments 500 bp - 3 kb the Geneclean™ technique (Stratech Scientific Ltd.) was used efficiently (see manufacturers instructions). For slightly larger DNA fragments 2 - 10 kb Spin-X tubes (supplied by Costar) were found to be most effective.

2.6.2 Removal of 5' Phosphate Groups from DNA.

The dephosphorylation of DNA for the purposes of minimising recircularisation of vector DNA was accomplished by treating the restricted DNA with calf intestinal phosphatase (Boehringer Mannheim). 10 - 20 μg of plasmid DNA was digested with a two - threefold excess of restriction enzyme until digestion was complete. Completion of digestion was checked by running a small aliquot of DNA on an agarose gel. The sample was then extracted with an equal volume of phenol:chloroform and the remainder of the DNA was then precipitated by addition of 2.5 volumes of ethanol. After 20 minutes on ice the DNA was recovered by centrifugation for 30 minutes at 12 000 rpm.

The DNA was then resuspended in 90 μl TE buffer (pH 8.0). 10 μl of 10 x CIP dephosphorylation buffer (10 mM $ZnCl_2$; 10 mM $MgCl_2$; 100 mM Tris·Cl (pH 8.3)) was then added to the sample along with 2 units of CIP enzyme. This was incubated at 37 °C for 30 minutes when another 2 units of CIP was added and the incubation continued for another 30 minutes. After incubation EDTA (pH 8.0) was added to a final concentration of 5 mM and the reaction heated to 85 °C for 15 minutes to inactivate the

CIP. The solution was then extracted with phenol and phenol:chloroform. DNA was then precipitated by addition of 0.1 volumes of 3M sodium acetate (pH 7.0) and 2.5 volumes of ethanol. The DNA was recovered by centrifugation at 12 000 rpm. The pellet was washed in 70% ethanol, dried and the DNA resuspended in TE buffer (pH 8.0).

2.6.3 Ligation of Fragments.

The ligation of DNA fragments was performed using bacteriophage T4 DNA ligase. Fragments with sticky complementary ends were ligated overnight at 14 °C, whereas, blunt-end ligations were carried out at 4 °C for at least 24hrs. The total quantity of DNA per reaction varied, but the molar ratio of insert to vector was always greater than 3:1. The quantitation of vector to insert DNA was accomplished roughly, by gel electrophoresis. The vector DNA was usually phosphatased, according to Section 2.6.2, to minimise vector recircularisation. Fresh commercial 5 × ligation buffer was routinely used. The volume of the reaction was kept to a minimum, usually 5-10 μl.

2.7 Oligonucleotide cloning technique.

2.7.1 Annealing of single stranded oligonucleotides.

Oligonucleotides were synthesised and purified as described above and were then checked by polyacrylamide gel electrophoresis. Complementary oligonucleotides to be annealed were mixed in an 1.5 ml Eppendorf tube at total volume of 100 μl, at a concentration of 20 $\mu g ml^{-1}$. The solution was the covered with an equal volume of liquid paraffin to prevent evaporation. The oligos were heated for 5 minutes at 90 °C in a beaker of water and left to cool slowly in the beaker which was placed in a polystyrene cover to retain heat. When the beaker had cooled to room temperature the paraffin was removed from the oligonucleotide solution by extraction with chloroform : isoamylalchohol. This step was repeated twice. The annealing was checked by polyacrylamide gel electrophoresis as described above for single stranded oligos.

2.7.2 Direct cloning of double stranded oligonucleotides into pMC1871.

The double-stranded oligonucleotides were cloned blunt-ended into the *Sma I* site of the plasmid *pMC1871*. The ligation was performed as described in section 2.6.3 but with a greater than ten-fold molar excess of oligonucleotide. After ligation the T4 DNA ligase was inactivated by heating to 65 °C for 15 minutes. KCl was added to a final concentration of 50 mM and the ligation was redigested with *Sma I* enzyme to remove recircularised plasmid. After digestion, the restriction enzyme was inactivated by heat (65 °C for 30 minutes).

2.7.3 Sequence of oligonucleotides used for direct cloning.

Oligonucleotides were synthesised single stranded as described above Complementary pairs were annealed before cloning directly into plasmid.

A (1) +ve strand	$^{5'}$ AAA GTG GTG GTG GTG GTG GTG GTG GTG GTG $^{3'}$
A (2) -ve strand	$^{5'}$ CAC CAC CAC CAC CAC CAC CAC CAC CAC TTT $^{3'}$
B (1) +ve strand	$^{5'}$ AAA ATG GTG GTG GTG GTG GTG GTG GTG GTG $^{3'}$
B (2) -ve stand	$^{5'}$ CAC CAC CAC CAC CAC CAC CAC CAC CAT TTT $^{3'}$
C (1) +ve strand	$^{5'}$ AAA GTG GTG GTG GTG GTG GTG GTG GTG ATG $^{3'}$
C (2) -ve strand	$^{5'}$ CAT CAC CAC CAC CAC CAC CAC CAC CAC TTT $^{3'}$
D (1) +ve strand	$^{5'}$ AAA ATG GCG CGC GCG CGC GCG CGC GC $^{3'}$
D (2) -ve strand	$^{5'}$ GCG CGC GCG CGC GCG CGC GCC ATT TT $^{3'}$
E (1) +ve strand	$^{5'}$ AAA ATG GCG CGC GCG CGC GCG CGC $^{3'}$
E (2) -ve strand	$^{5'}$ GCG CGC GCG CGC GCG CGC CAT TTT $^{3'}$

All oligos used in cloning were purified by ethanol precipitation. 0.1 volumes of 3 M sodium acetate was added to the ammonia solution containing the oligonucleotide and the DNA was precipitated by addition of 3 volumes of ethanol. The sample was left on dry ice for 60 minutes and the centrifuged at 12 000 rpm, 4 °C for 15 minutes. The oligonucleotide was then resuspended in a smaller volume and reprecipitated. After the second precipitation the pellet was washed with cold 80% ethanol and recentrifuged. The pellet was then dried under vacuum and resuspended. The concentration of the solution was assessed by spectrophotometry.

2.8 Bacterial transformation.

Commercial competent cells, DH5α^{TM}, supplied by GibcoBRL, were used for most routine cloning, otherwise competent cells were made according to Mandel and Higa (1970). All ligations were carried out in Falcon™ 2063 polypropylene tubes. DNA ligations were diluted 5-fold in 10 mM Tris·Cl (pH 7.5) and 1 mM Na_2EDTA, and 1 - 5 μl (approximately 10 ng) of the dilution was added to 50 or 100 μl of thawed competent cells kept on ice. The cells were left on ice for 30 minutes, then heat shocked at 42 °C for 45 seconds. The cells were left to recover for 2 minutes on ice, then incubated with 450 or 900 μl of L broth (supplemented with 20 mM glucose) at 37 °C for 1 hr. The tubes were centrifuged at 1 000 rpm for 5 minutes, the supernatant poured off, and the pellet gently resuspended in the residual supernatant. This was spread onto LB plates (1.5% w/v bacto-agar in L-broth) supplemented with normally 100 μgml^{-1} ampicillin or 25 μgml^{-1} tetracycline. The plates were incubated overnight at 37 °C.

SURE competent cells (Stratagene, Cambridge) were used to clone repetitive sequence oligonucleotides into the vector pBRlacZ. All transformations into these cells were carried out in these cells were performed in Falcon™ 2059 polypropylene tubes, according to the manufacturer's protocol.

2.9 Purification of plasmid DNA.

2.8.1 Rapid small scale plasmid preparation for restriction enzyme analysis.

1-5 ml of L broth (with 100 μgml^{-1} ampicillin) was inoculated using a loop. The cultures were incubated overnight at 37 °C. Cells were centrifuged in 2 ml Eppendorf tubes. The pellets were resuspended in 100 μl of lysis solution (25 mM Tris·Cl(pH 8.0); 10mM EDTA; 50 mM glucose; 0.5 mg/ml lysozyme). 200 μl of alkaline/SDS solution (0.2N NaOH, 1% SDS), was added to the tubes and mixed. 150 μl of NaOAc (pH 4.8) was then added and the tubes vortexed. After incubation on ice for 5 minutes, the tubes were centrifuged, and the supernatant added to a fresh tube containing 400 μl of isopropanol. The tubes were centrifuged at 10 000 rpm for 15 minutes. The pellet was

left to air dry and resuspended in 200 μl of dH_2O. A further precipitation step with 100% ethanol was optional. RNase A (100 μgml^{-1} f/c) was included at the time of restriction enzyme digestion.

2.8.2 Medium scale preparation of high quality DNA for sequencing.

A 500 μl sample was taken from a miniprep culture and was then added to 50 ml of L-broth containing antibiotic (100 μgml^{-1} ampicillin) and incubated for at least 16 hrs on a rotary shaker at 37 °C. The culture was then centrifuged for 5 minutes at 8 000rpm for 5 minutes in a refrigerated centrifuge. After the supernatant was discarded, the pellet was resuspended in 5 ml STET buffer (8% sucrose, 5% Triton X-100, 50 mM EDTA, 50 mM Tris·Cl pH 7.5). Boiled RNase and fresh lysozyme were added to concentrations of 10 μgml^{-1} and 0.5 $mgml^{-1}$ respectively. The mixture was then decanted into a 30 ml polycarbonate tube (Sorvall) and after leaving for 10 minutes room temperature the sample was boiled for one minute and left on ice for a further 10 minutes. The tube was then centrifuged at 12 000 rpm for 10 minutes at 4 °C. The crude plasmid DNA was the precipitated by addition of 4 ml of ice cold isopropanol and wad pelleted by centrifugation as above. The pellet was resupended in 600 μl TE buffer (10 mM Tris·Cl pH 7.5, 50 mM EDTA). Protein contaminants are removed by extraction once with phenol:chloroform, then once with chloroform alone. After extraction the DNA was transferred into a 10 ml polycarbonate tube and precipitated by a the addition of an equal volume of 7.5 M ammonium acetate and 2.5 volumes of -20 °C ethanol. After 15 minutes on ice the sample was centrifuged at 12 000 rpm for 10 minutes at 4 °C. The pellet was the rinsed twice with 70% ethanol, dried under vacuum and resupended in 300 μl TE buffer.

2.8.3 Large scale preparation of high quality DNA.

A 10 ml starter culture was prepared overnight. This was then added to 500 ml of L-broth containing antibiotic (100 μgml^{-1} ampicillin) and incubated for at least 16 hrs. The culture was centrifuged in 500 ml Sorvall bottles at 5 000 rpm for 5 minutes at 4 °C. The pellet was resuspended in 25 mls of ice cold lysis solution (as Section 2.3.1). After 15 minutes on ice 50 mls of alkaline /SDS solution (as Section 2.3.1) was added, and the solution mixed vigorously by swirling. Finally 37.5 mls of potassium acetate solution (60 mls 5M potassium acetate, 11.5 mls glacial acetic acid and 28.5 mls dH_2O) was

added and the bottles incubated on ice for approximately 15 minutes. The bottles were centrifuged at 6 000 rpm at 4 °C for 15 minutes. The supernatant was decanted through gauze and isopropanol precipitated with 0.6 volumes of isopropanol. The pellet was resuspended in TE (10 mM Tris·Cl pH 8.0; 1 mM EDTA pH 8.0). The plasmid DNA was the further purified by equilibrium centrifugation in a continuous CsCl-EtBr gradient. 1 gml^{-1} solid CsCl was dissolved in the DNA solution. Ethidium bromide was added to a final concentration of 740 $\mu g ml^{-1}$. The refractive index of the solution was checked to ensure the density of the solution was correct (1.55 gml^{-1} ; refractive index = 1.3860). The DNA sample was then loaded into polycarbonate tubes and spun at 40 000 rpm for 60 hours at 20 °C or at 80 000 for 16 hours at 20 °C. After centrifugation the lower red band containing the plasmid was removed from the gradient with a syringe. The ethidium bromide was removed from the DNA by extraction with 1-butanol. This was repeated at least 6 times and was continued until the red colour of the ethidium bromide had disappeared. The CsCl was then removed from the DNA by dialysis for 48 hours against several changes of TE buffer (pH 8.0).

2.10 DNA sequencing of cloned oligonucleotides.

Sequencing of double-stranded DNA was carried out using the Sequenase™ T7 DNA polymerase kit and [α-^{32}P] dATP. 10 μg of plasmid DNA was linearised with the appropriate restriction enzyme, ethanol precipitated and resuspended in 20 μl dH_2O. The primer used sequence oligonucleotides inserted in front of the *lac Z* gene was the 17 bp M13 universal sequencing primer. The sequencing reaction was performed with end-labelled primer. The end-labelling reaction was carried out in a total volume of 10 μl (50 mM Tris·Cl pH 7.5, 10 mM $MgCl_2$, 1 mM EDTA, 5 ng primer, 5 units T4 DNA kinase and approximately 50 μCi [α-^{32}P] dATP). This reaction mixture was then incubated at 37 °C for 30 minutes and stopped by addition of 1 μl of 0.1M EDTA. The kinase was then inactivated by heating to 70 °C for 15 minutes. 7 μl of the precipitated DNA template was combined with 3 μl of Sequenase™ sequencing buffer and 5 μl of the end-labelled primer. The remainder of the sequencing reaction was carried out with the reagents and the protocol supplied in the USB (United States Biochemical) Sequenase™ kit.

2.11 General culture of mammalian cell lines.

2.11.1 Medium and supplements.

Epidermal cell lines (C5O) were grown routinely in Special Liquid Medium (SLM, GibcoBRL) supplemented with 10% foetal bovine serum (FBS) and 4 mM glutamine. Cells were grown in 25, 75 or 175 cm^2 flasks or 9 cm dishes gassed with 5% CO_2, and placed in 5% CO_2 ; 37 °C incubators. Cells were washed with phosphate buffered saline or PBS (0.14 M NaCl; 27 mM KCl; 10 mM Na_2HPO_4; and 15 mM K_2HPO_4). Trypsin for cell disaggregation was used at 0.025% w/v in PE (1 mM EDTA pH 7.8 in PBS). Trypsin was inactivated with 10-fold volume of serum containing medium. Cell numbers were estimated using a Coulter counter.

Fibroblast cell lines (NIH 3T3, Ψ2) were grown in Dulbecco modified Eagles medium (DMEM), (for 100 mls :10 mls 10 x DMEM; 1 ml 200 mM glutamine; 1 ml 100 mM pyruvate; 4.2 ml 7.5% sodium bicarbonate; 1 ml non-essential amino acids, in dH_2O).

2.11.2 Freezing and storage of cell lines.

Cells for freezing were grown to near confluence, typsinised and centrifuged at 1 000 rpm for 5 minutes. The cells were resuspended in equal proportions of growth medium and freezing medium (25% FBS; 20% DMSO; 55% growth medium). Cells were stored in 1-2 ml Nunc™ vials in liquid nitrogen.

2.12 Transfection of mammalian cell lines.

2.11.2 Calcium Phosphate Technique.

5-10 μg of supercoiled plasmid DNA was mixed with 35 μg of carrier DNA (salmon sperm genomic DNA) in a solution of TE (1.0 mM Tris pH 8.0; 0.1 mM EDTA pH 8.0) to a final volume of 500 μl. To this was added 400 μl of TE (as above) and 100 μl of

2.5 M $CaCl_2$ (refractive index 1.401). This mixture was added slowly to 1 ml of 2 x HBS (280 mM NaCl; 50 mM HEPES; 1.5 mM $Na_2HPO_4.2H_2O$ pH 7.12) and left for 30 minutes at room temperature to form a precipitate. This was then added to 4 ml of medium, mixed gently and placed over a sub-confluent culture of cells. After 24 hours the transfection medium was replaced with fresh growth medium and incubated at 37 °C for a further 1-2 days. The cells were then split 1:10 and selection applied as appropriate. Transfected cells, containing the neomycin resistance gene (*neo*), were selected in medium containing G418 (Gibco BRL, Paisley) at a concentration of 400 μgml^{-1}.

2.12.2 Lipofection Technique.

The technique of cationic liposome mediated transfection was used to introduce DNA into a number of mouse epithelial cell lines. The protocol was as for the DOTAP™ reagent supplied by Boehringer Mannheim.

2.12.3 Electroporation Technique.

A 175 cm^2 flask of confluent cells were trypsinised and the trypsin inactivated with serum containing medium. The cells were washed with phosphate buffered saline, PBS and resuspended in about 0.4 ml of PBS. In an Eppendorf tube 75 -150 μg of linearised plasmid DNA was resuspended in 0.4 ml of PBS and mixed with the cell suspension. The contents were transferred to a sterile Bio-Rad electroporation cuvette and electroporated in a Bio-Rad Gene Pulser™. The capacitance was set at 3 μF (not on extender), whilst the voltage was set at 0.8 kV. The cuvette was removed and the contents added to 50 - 100 ml of prewarmed growth medium and plated into 75 cm^2 flasks. Selection with G418 was started 24 - 48 hours after electroporation.

2.13 Generation of replication deficient retroviruses.

Retroviruses containing the *lac Z* fusion genes were generated by $CaPO_4$ transfection of the murine viral producer cell line ψ2. After transfection cells were selected in DMEM medium containing G418 at a concentration of 1 mgml^{-1}. The medium was changed

every 3 days for 2 weeks. Surviving cells were pooled and used as viral producer cell lines as described below.

2.14 Preparation of retroviruses and infection of cultured cells.

Virus producer cell lines were grown until 80% confluent. The medium was replaced with the minimum volume of growth medium to ensure coverage. This was left for 24 hours removed and filtered through a series of syringe filters 1.2 - 0.2 microns. 24 hours prior to harvesting the viral supernatant, $5x10^5$ cells (for infection) were seeded per 25 cm^2 flask. To each flask was added a mixture of 50 μl of viral supernatant, 150 μl of polybrene (0.8 $mgml^{-1}$), made up to 5 ml with growth medium. After 48 hours the cells were split into 75 cm^2 flasks and selection with G418 commenced if appropriate.

2.15 Staining of cultured cell lines for β-galactosidase activity.

Cells were stained in tissue culture flasks. The medium was removed from flasks and the cells were rinsed twice in PBS. Cells were then fixed in 2% formaldehyde (McQuilken, Glasgow), O.2% gluteraldehyde (McQuilken, Glasgow) in PBS for 1 hour on ice. After fixation the cells were again rinsed twice, in PBS. The β-galactosidase stain consists of: 1 $mgml^{-1}$ 5-bromo-4-chloro-3-indolyl β-D-galactoside (from a stock solution of 40 $mgml^{-1}$ in dimethylsulphoxide); 5 mM $K_3Fe(CN)_6$; 5 mM $K_4Fe(CN)_6.6H_2O$ and 2 mM $MgCl_2$ in PBS. This stain was then added to the flask and the cells were cultured overnight at 37 °C in an atmosphere of 5% CO_2. After staining the cells were then rinsed three times in PBS before visualisation by microscopy.

2.16 Treatment of cells with the chemical mutagen MNNG.

Cells to be treated were cultured in 175 cm^2 until approaching confluence. The growth medium from the flask was removed and the cells washed in PBS (pre-warmed to 37 °C) three times and then once in serum-free SLM. 20 ml serum-free SLM with glutamine

was then added to the flask and the medium. Serial dilutions on the stock solution of MNNG (3 $mgml^{-1}$) were prepared in acetone. From these dilutions MNNG was added to the flask at concentrations ranging from 0.1 to 10 μgml^{-1}. Control flasks were set up containing acetone alone. The flasks were sealed and left at 37 °C for 3 hours. The cells were then washed and growth medium containing serum was replaced. The culture then left for 48 hours at 37 °C to recover from the mutagen treatment. The treated cells were then trypsinised, to remove them from the flask and counted. Cells were then plated out into 10 cm dishes at a density of 5 x 10^5 cells per dish. A small number of cells were also plated out into 25 cm^2 flasks to test the number of the cells. After 2-3 days the number the cells in 10 cm dishes were stained with X-gal as described in section 2.14. The numbers of lac Z positive cells on each dish were then counted. The number of clonogenic cells plated out onto each dish was determined by counting the number of colonies formed by the 200 cells plated into the 25 cm^2 flasks.

2.17 Preparation of genomic DNA from cultured cells.

Genomic DNA was prepared by treating adherent cultured cells grown in 75 or 175 cm^2 flasks with 2 or 5 mls of lysis buffer (50 mM Tris pH 8.0; 50 mM EDTA; 100 mM NaCl; 5 mM dithiothreitol; 1% SDS and 0.5 mM spermidine) respectively. The cell lysate was scraped from the flask using a sterile scraper, and placed in a 15 ml Falcon tube, to which 200 μl (10 $mgml^{-1}$ solution) of proteinase K was added. The tubes were incubated at 37 °C in a rotator, overnight. The contents were then phenol/chloroform extracted, and subsequently isopropanol precipitated. The viscous pellet was spooled onto a sealed glass capillary and washed in 70% then 100% ethanol, then allowed to air dry. The genomic DNA was finally resuspended in 200-400 μl of TE (10 mM Tris pH 7.4; and 1 mM EDTA).

2.18 Preparation of RNA from cultured cells.

RNA was prepared using the RNAzol™ protocol (Biogenesis Ltd.). To 10^6 adherent cells, 0.2 ml of RNAzol was added directly to the flask, and the lysate scraped off using a sterile scraper. This was placed in a 15 ml Falcon™ 2059 tube and one tenth volume of chloroform was added. After shaking, the tube was centrifuged at about 5 000 rpm at 4 °C for 15 minutes. The colourless upper phase was transferred to a fresh tube, and the

RNA precipitated using an equal volume of isopropanol. The RNA was then pelleted and washed with 75% ethanol and allowed to air dry. The pellet was then resuspended in 50% ethanol in diethylpyrocarbonate (DEPC) treated water. Care was taken to ensure that all solutions used were RNase free. RNA was stored at -20 °C.

2.19 Southern blot transfer of DNA.

DNA fragments were separated by agarose gel electrophoresis and transferred to nylon membranes (Hybond-N™ or Biotrace-RP™) by the method of Rigaud *et al.*, 1987. The gel was soaked in ethidium bromide solution and photographed. It was then soaked in 1.5M NaCl, 0.5M NaOH for 2 × 20 minutes, and then in 1M NH_4OAc, 0.02M NaOH for 2 × 30 minutes. The gel was then transferred onto a nylon membrane (Hybond, Amersham, or Biotrace RP), by capillary transfer overnight in NH_4OAc/NaOH transfer solution. Alternatively vacuum transfer of the gel was used according to manufacturers' guidelines. The membrane was then fixed by baking at 80 °C for 2 hours and hybridised as described in Section 2.22.

2.20 Northern blot transfer of RNA.

10-20 μg of total RNA was freeze dried and resuspended in freshly made sample buffer (50% formamide; 1 × MOPS; 16.6% formaldehyde up to 100% with dH_2O). This was incubated at 60 °C for 15 minutes, then on ice for 5 minutes. 5 μl of northern loading buffer (50% glycerol, 10 μM Sodium phosphate pH 7.0, and a few bromophenol blue crystals) was added to the samples prior to loading the gel (1% agarose; 1 × MOPS; and 18% formaldehyde). The gel was run in 1 × MOPS (for 1 litre of 10 × MOPS: 42g MOPS; 6.8g NaOAc; 3.7g EDTA pH 7.0). The RNA ladder was supplied by GibcoBRL. The RNA was blotted onto nitro-cellulose membrane (Sartorius™) and hybridised as below, in Section 2.22.

2.21 Preparation of radioactively labelled probes.

Labelled probes from double-stranded DNA were made by the random priming method (Feinberg and Vogelstein, 1983,1984). Approximately 50 ng of purified insert DNA was boiled for 10 minutes to ensure denaturation, and then labelled in a total volume of 50 μl, containing 10 μl of oligonucleotide labelling buffer (OLB) and 2 μl of BSA (10 μg/ml stock). Finally, 1.85 $\times$ 10^6 MBequerel of [α-^{32}P] dCTP and 5 units of Klenow enzyme were added, and the reaction incubated at room temperature for at least 6 hours. The unincorporated nucleotides were removed by running the probe through a NICK-column™ (Pharmacia), and the specific activity estimated using a scintillation counter.

The recipe for oligonucleotide labelling buffer (OLB) was : 62 μl 1 M Tris pH 7.4; 6 μl 1 M $MgCl_2$; 1 μl 2-Mercaptoethanol; 2 μl 15 mM dATP; 2 μl 15 mM dGTP; 2 μl 15 mM dTTP; 125 μl 2 M HEPES pH 6.6 and 50 μl 150 OD/ml calf thymus hexanucleotides (Pharmacia). It was stored in aliquots at -20 °C.

2.22 Hybridisation of radioactive probes.

Filters were prehybridised at 42 °C for minimum of 4 hours in sealed plastic containers with 20-50 ml of heated pre-hybridisation solution (50% formamide; 4 $\times$ SSPE (20 $\times$ SSPE was 3.6M NaCl; 200 mM NaH_2PO_4 pH 7.4); 0.5% w/v dried milk (Marvel); 1% SDS; 10% dextran sulphate; 1 $mgml^{-1}$ denatured salmon sperm DNA). After the allotted prehybridisation time, radio-labelled probe (see section 2.20) was denatured by boiling for 10 minutes and added direct to the solution. Blots were hybridised for a 16 - 20 hours at 42 °C. Filters were washed on a shaker with 1 $\times$ SSC / 0.1% SDS, then 0.5 $\times$ SSC / 0.1% SDS, at room temperature each wash for 15 minutes. The final wash was with 0.1 $\times$ SSC / 0.1% SDS at 50-65 °C for 10-30 minutes with constant monitoring of the background with a Geiger counter. The filters were exposed to Kodak X-OMAT AR or X-OMAT S film at -70 °C in cassettes containing intensifying screens.

2.23 Amplification of DNA and RNA by the polymerase chain reaction (PCR).

2.23.1 Amplification from genomic DNA

DNA was amplified according to Saiki et al. (1988). A 50 μl reaction containing up to 1 μg of genomic DNA, or in the case of plasmid DNA, concentrations were in the picomolar range. Commercial 10x buffer supplied by Promega (500 mM KCl; 100 mM Tris·Cl pH 9.0 at 25 °C; 15 mM $MgCl_2$; 0.1% gelatin w/v; 1% Triton X-100), each amplimer at 140 μgml^{-1} and each dNTP (2.5 mM stock mix, 4 μl in 50 μl reaction) were combined in a 0.5 ml PCR tube. 2.5 units of thermostable DNA polymerase from *Thermus aquaticus* (Taq polymerase), (also supplied by Promega) was finally added to the mixture. Thermal cycling was controlled by a programmable heating block (Perkin Elmer Cetus, DNA Thermal Cycler 480 or GeneAmp PCR System 9600). Three-step temperature cycling conditions varied, depending on GC content of amplimers, and length of amplification.

Optimum amplimers were designed by a computer program (Oligo 4.0-2008 Primer Analysis Software), and made as in Section 2.13. In order, to avoid cross contamination of samples, aerosol resistant tips (ART™ supplied by BCL) were employed. In addition, positive controls were set up last.

2.23.2 Amplification from RNA.

Amplification of RNA was carried out according to the method of Bartek *et al.* (1991). 10 μg of RNA was dried in a Speedivac vacuum drier. The dried RNA pellet was resuspended in 10 μl of 2 x annealing buffer (800 mM KCl, 20 mM Pipes pH 6.6), along with 1 μl RNasin (RNase inhibitor, Pharmacia), 1 μl 3' PCR amplimer (140 μgml^{-1}) and 8 μl of DEPC-treated dH_2O. The sample was heated to 85 °C for 5 minutes the 65 °C for a further 15 minutes. 40 μl of 5 x reverse transcriptase reaction buffer was then added to the mixture along with 10 μl 2.5 mM dNTP mix, 20 μl 0.1 M dithiothreitol (DTT), 9μl RNasin and 71 μl ddH_2O. After mixing the sample was split into 2 x 100μl aliquots. 2 μl of reverse transcriptase enzyme was added to one aliquot and the other

was used as a control for DNA contamination. Both aliquots are heated to 42 °C for 60 minutes and then heated to 85 °C for 5 minutes to inactivate enzyme. After the tubes have been cooled on ice 5 μl of each reaction was used in a standard PCR reaction as described above.

2.23.3 Sequence of commonly used amplimers.

Sequence of amplimers used for screening genomic DNA and RNA for the presence of the *lac Z* gene.

LacZA	5' AAT CGC CTT GCA GCA CAT CCC 3'
LacZB	5' CAT CGT AAC CGT GCA TCT GCC 3'
LzU1	5' TGT CGT CGT CCC CTC AAA CTG 3'
LzL1	5' TCA CCC TGC CAT AAA GAA ACT 3'
LzU2	5' GCA TAA ACC GAC TAC ACA AAT 3'
LzL2	5' GCT TCA TCC ACC ACA TAC AGG 3'
LzU3	5' ATC TGG TCG CTG GGG AAT GAA 3'
LzL3	5' AAA CGG AAC TGG AAA AAC TGC 3'
LzU4	5' CAG TCA GGC TTT CTT TCA CAG 3'
LzL4	5' GCG GTC GGG ATA GTT TTC TTG 3'
LzU5	5' GCA GGT AGC AGA GCG GGT AAA3'
LzL5	5' AGT AAG GCG GTC GGG ATA GTT 3'

All oligonucleotides for PCR were purified by precipitation by butanol as described by Kemp *et al.* (1993). The oligonucleotide was resuspended and the concentration of the solution was assessed by spectrophotometry. All stock solutions of oligonucleotides used in PCR were kept at a concentration of 140 μgml^{-1}.

2.24 Ribonuclease protection assay.

The probe for the RNase protection assay was generated using the SP6/T7 transcription kit (Boehringer Mannheim) following the manufacturers instructions. 10-20 μg of total RNA was dried down in a Speedivac, along with 1 × 10^6 cpm of probe. This was resuspended in 20 μl of 80% FAB (80% formamide; 400 mM NaCl; 40 mM Pipes pH 6.4;

1 mM EDTA). This was incubated for 16 hours at 50 °C. The samples were then placed on ice and 300 μl of digestion buffer (10mM Tris pH 7.6; 5 mM EDTA; 0.3M NaAc pH 7.0; and 40 μgml^{-1} RNaseA, 150 U/ml RNase T1 f/c). The samples were digested for 1 hour at 30 °C and then 2.5 μl of 10 mgml^{-1} Proteinase K and 3.2 μl of 20% SDS were added and the mixture incubated at 37 °C for 15 minutes. 5 μl of yeast tRNA (1 mgml^{-1}) carrier was added to each sample. The samples were phenol/chloroform extracted and precipitated with 100% ethanol followed by 70% and 100% washes. The RNA digested samples were freeze dried and resuspended in 5 μl of loading buffer (8 ml formamide, 1 ml 10 × TBE, 20 μl 0.5M EDTA, 10 mg xylenecyanol, 10 mg bromophenol blue, and 1 ml of dH_2O). The samples may be stored at -70 °C, or run on an 8% polyacrylamide gel (8M urea; 1 × TBE; and dilution of 40% acrylogel solution, with ammonium persulphate and TEMED added prior to pouring). The samples were heated to 90 °C for 5 minutes prior to loading. Gel electrophoresis was at 40-50W for 2 hours. The gel was then dried down under heated vacuum and exposed to film (Kodak X-OMAT AR or S).

2.25 Preparation of DNA for microinjection.

DNA used for microinjection is linearised with as much vector sequence as possible removed. The presence of vector sequences can reduce expression of the transgene by up to 1000 times. Contamination of the preparation with organic solvents such as phenol or chloroform can have severely deleterious effects on the developing embryos so use of these chemicals was avoided.

The DNA construct was digested by restriction enzymes and loaded onto a standard agarose gel until good separation of the fragments is achieved. The DNA band in the low melting point agarose was then excised and heated to 70 °C until the agarose completely melts. 5 volumes of 0.5 M NaCl in TE (10 mM Tris·Cl pH 7.5; 1 mM EDTA) is added and the solution completely mixed. Incubation at 70 °C is continued for a further 10 minutes.

The DNA\agarose mixture was then run through a disposable ion exchange column, NACS-52 Prepac (BRL) according to the manufacturer's instructions. This procedure was carried out in a hot room at 37 °C to prevent the agarose solution solidifying. In addition all solutions were maintained at 42 °C in a waterbath. The eluted DNA is precipitated by addition of 10 μg of carrier tRNA and 0.6 ml of cold 95% ethanol. After incubation on ice for 15 minutes the sample is centrifuged at 12 000 rpm, 4 °C for 30

minutes. The DNA pellet is then washed several times with 70% ice cold ethanol. The pellet is then dried and resuspended in a small volume of TE buffer. At this stage the DNA can be checked by ultra-violet light spectrophotometry and agarose gel electrophoresis. The DNA was diluted to the desired concentration (~ 1-2 ngμl^{-1}) with injection buffer (10 mM Tris·Cl pH 7.6; 1 mM EDTA). The sample was then extensively dialysed against several changes of a large volume of injection buffer at 4 °C over 48 hours. The DNA preparation is stored in 20 µl aliquots at -20 °C.

2.26 Generation of transgenic mice.

Transgenic mice were produced using similar materials and methods to the protocols described by in *Manipulation of the mouse embryo : A laboratory manual* (Hogan,B., Constantini, F., and Lacy, E. eds.). The method employed to transfer genes into mice during the course of this project was microinjection of DNA directly into the pronuclei of fertilised mouse eggs.

2.26.1 Pipettes, injection needles and injection chambers.

All glass capillaries for making pipettes and injection needles were obtained from Clark Electromedical Instruments, Reading. Handling pipettes were pulled over a Bunsen burner flame. Holding and transfer pipettes were manufactured from GC100 and GC120 glass capillaries respectively on a microforge (Model MF-9, Narishige Co. Ltd., obtained from Microinstruments Ltd., Oxford). The injection needles were made from GC100TF capillaries, pulled on a Flaming/Brown Micropipette Puller (Model P87, Sutter Instrument Co.)

2.26.2 Superovulation of female mice.

Three week old (C57BL/6J x CBA/J)F1 female mice were superovulated by injection of hormones. This procedure is performed to increase the number of fertilised eggs obtained from mice. It involves sequential injection of two hormones pregnant mare's serum (PMS) and human chorionic gonadotrophin (hCG). Pregnant mare serum was dissolved in 0.9% (w/v) NaCl at a concentration of 50 IUml^{-1} and stored in aliquots at -

20 °C. Each mouse was injected intraperitoneally with 100 µl of PMS. The second hormone hCG was dissolved in dH_2O at a concentration of 500 $IUml^{-1}$ and was divided in 100 µl aliquots, lyophilised and stored, protected from light, at -20 °C. Before use the hormone was resuspended in 1 ml of sterile 0.9% NaCl to give a final concentration of 50 $IUml^{-1}$. 100 μl (5 IU)of the hCG was injected into the mice 46 - 48 hours after injection of the PMS. The female mice were then mated to MF 1 stud male mice overnight. Females were checked the morning after for coital plugs. Plugged female mice were used as embryo donors for the gene transfer.

2.26.3 Collection and culture of fertilised embryos.

Microdrop cultures were set up first of all. Four or five 30 μl drops of equilibrated M16 medium were placed in a 35 mm petri dish. These drops were overlaid by 5 ml of liquid paraffin, which flooded the whole dish. The whole culture is then placed in a humidified, gassed (5% C02, 95% air) incubator and left to equilibrate.

Plugged (C57BL/6J x CBA/J)F_1 female mice were humanely killed by cervical dislocation. The two oviducts, each with a short segment of uterus were then dissected out of the animals. The oviduct was then placed in a 50 µl drop of M2 culture medium (Biological Industries, Haemek, Israel) supplemented with 4 $mgml^{-1}$ bovine serum albumin (Sigma), in a 35 mm petri dish. The oviducts are visualised under a stereomicroscope at a magnification of approximately 20 times. 10 µl of 10 $mgml^{-1}$ solution of hyaluronidase (Sigma) was then added to a the drop of medium containing the embryos. The embryos were left for 1 - 2 minutes until the cumulus cells had become detached. The embryos were then removed from the medium and the cell debris and washed through six drops of M2 medium. The eggs were then washed once in M16 medium and transferred to the drops of medium in the microdrop culture set up earlier and cultured for 2 hours at 37 °C prior to injection.

2.26.4 Pronuclear microinjection of fertilised eggs.

Mouse egg pronuclei were injected using a Nikon Diaphot TMD inverted microscope with Nomarski optics. Separate Leitz micromanipulators were used to direct the holding pipette and the injection needle. The holding pipette was controlled by an inverted micrometer syringe (Beaudolin, supplied by Microinstruments Ltd., Oxford). The entire

system including the holding pipette itself was filled with Fluorinert FC77 (Sigma, Dorset). Injection needles were filled by capillary action and were then backfilled with Fluorinert FC77. Injection was regulated by air pressure which was controlled by a 50 ml ground glass syringe.

An injection chamber was prepared for use as described in section 2.25.1 and 10 - 20 fertilised mouse eggs were placed in the chamber. DNA solution is injected into the pronucleus by depressing the plunger on the 50 ml glass syringe. If the pronucleus is injected properly it should swell visibly. The needle is the removed quickly to prevent damage to the pronucleus and the injected egg is placed to one side of the injection chamber, away from the uninjected embryos.

Once all the eggs in the chamber have been injected the viable embryos are removed from the injection chamber, washed in M16 medium and replaced in another microdrop culture. This procedure is repeated until approximately 100 fertilised eggs are injected. The embryos are the culture overnight in a humidified, gassed incubator at 37 °C. Overnight the embryos should divide to reach the two-cell stage.

2.26.5 Transfer of embryos into pseudopregnant females.

MF 1 female mice (0.5 days p.c.) approximately 6-12 weeks old were used as recipients for injected embryos. Recipient mice were mated the night before with vasectomised MF1 stud males (obtained already vasectomised from Harlan Olac). Mice were visually inspected and those judged to be in oestrus were selected and mated with the stud males. The next day recipient mice were checked for coital plugs. Plugged mice were used as recipients for eggs injected with DNA on the previous day.

Recipient mice were anaesthetised by injection of 100 μl of anaesthetic. The anaesthetic was made by mixing 1 volume of Hypnorm™ (Janssen Pharmeceutica, Tilburg) with 1 volume of Hypnovel™ (Roche) and diluting with 2 volumes of sterile dH_2O. The lower back of the mouse was then shaved. A transfer pipette (made previously as described in section 2.25.1) was then loaded with embryos. The eggs to be transferred were removed from the microdrop culture and washed twice in M2 medium. About 10 - 20 embryos were transferred into each of the two oviducts in a recipient mouse. An incision in the skin and body wall is made just above the ovary. The body wall is then

stretched to stop any bleeding. The fat pad attached to the ovary is then pulled through the incision and clipped with a serafine to hold it outside the body.

The mouse is the placed under a stereomicroscope. The swollen opening to oviduct, the infundibulum, was identified and a small tear was made in the bursa above it. The transfer pipette was inserted and the embryos were blown into the oviduct. The ovary was replaced and the incisions in the skin and body wall were then closed with stitches. The whole procedure was then repeated on the other oviduct. At the end of the operation the recipient mouse was left in a warm area to recover. Mice were monitored to see if they became pregnant and litters were assessed for the presence of transgenes.

2.27 Preparation of genomic DNA from mouse biopsies.

Genomic DNA was isolated from mice of various ages by biopsies of tail. Approximately 1 cm of tail was cut and placed in a microfuge tube containing 700 μl of lysis buffer (50 mM Tris·Cl pH 8.0, 100 mM EDTA, 100 mM NaCl and 1% (w/v) SDS). 30 μl of a 12.5 $mgml^{-1}$ solution of Proteinase K was added to the tube and it was incubated for 10 - 16 hours at 56 °C. After cooling and mixing 25 μl RNase (12.5 $mgml^{-1}$) was added and the sample and the tube was incubated at 37 °C for 3 hours. 750 μl of phenol was then added to the tube and it was placed on a vertical rotator for 30 minutes. The sample was then spun in a microfuge for 15 minutes and the separated aqueous phase was transferred to a fresh tube. This extraction procedure was repeated with phenol : chloroform and with chloroform : isoamylalchohol. The DNA was then precipitated with 1.5 volumes of 100 % ethanol. The DNA was then spooled onto the sealed end of a glass micropipette and was washed in 70% and then 100% ethanol. The end of the pipette was broken off and placed in a microfuge tube containing 250 µl TE buffer (10 mM Tris·Cl pH 8.0; 1 mM EDTA). The tube was then left overnight at 37 °C to ensure the DNA was properly dissolved.

2.28 Preparation of RNA from transgenic mouse biopsies.

A discontinuous CsCl gradient was used to prepare RNA from various mouse tissues. All solutions used in the protocol were DEPC-treated or made up with DEPC-treated dH_2O. The tissue sample was first of all ground in liquid N_2 using a mortar and pestle.

The powder was then scraped into a universal containing 8 ml of 5 M guanidinium solution (5 M guanidinium thiocyanate, 50 mM Tris·Cl pH 7.0, 50 mM EDTA, 5% (v/v) β-mercaptoethanol, 2% (w/v) sodium laurylsarcosine). A 14 ml polycarbonate tube was washed out with DEPC-treated dH_2O and then filled with 2 mls of CsCl solution I (~ 5.7 M CsCl, 50 mM EDTA pH 7.0 ; refractive index = 1.3925). This CsCl solution was then underlaid by 2 ml of a CsCl solution II (~ 5.7 M CsCl, 50 mM EDTA pH 7.0 ; refractive index = 1.4025). The DNA solution was then carefully overlaid on top of the CsCl gradient. The tube was spun for 60 - 72 hours in an ultracentrifuge at 25 000 rpm and 20 °C. After centrifugation the supernatant was removed from the tube leaving the RNA pellet at the bottom of the tube. The bottom of the tube was removed to allow access to the RNA pellet and the pellet was resuspended in 600 µl of dH_2O. The RNA was then precipitated with 60 µl of 3 M sodium acetate (pH 7.0) and 1.7 ml of ethanol. This was cooled to -20 °C for 3 hours and then centrifuged for 15 minutes at 12 000 rpm, 4 °C. The RNA pellet was washed in 70 % ethanol and then resuspended in DEPC dH_2O. Ethanol was added to the RNA solution to a final concentration of 50% and the sample was stored at -20 °C.

2.29 Staining of tissue for β-galactosidase activity.

2.29.1 Staining whole murine tissues.

Staining of tissues was done according to the methods of Whiting *et al.* (1991). Tissue samples to be stained were first of all fixed in 1% formaldehyde, 0.2% glutaraldehyde, 2 mM $MgCl_2$, 5 mM EGTA, and 0.02% NP-40 in PBS at 4 °C for 30 to 90 minutes, depending on size. Very large tissue pieces were partially dissected to permit full penetration of reagents. The samples were then washed in three changes of PBS plus 0.02% NP-40 at room temperature for 30 minutes each. The samples were then stained in 1 $mgml^{-1}$ 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal), 5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6.6H_2O$ and 2 mM $MgCl_2$, 0.01% sodium deoxycholate, and 0.02% NP-40 in PBS at room temperature overnight in the dark. The samples were then washed again as described above and fixed overnight in 4% paraformaldehyde at 4 °C.

The fixed and stained tissue samples were then sent to the Pathology Department at Glasgow University Veterinary School. They were dehydrated, and embedded in

paraffin wax. Sections were cut from the samples and the sections were dewaxed and counterstained with eosin. The sections were mounted and visualised by light microscopy.

2.29.2 Staining frozen sections.

Frozen sections were cut by Iain McMillan in the Pathology Department at Glasgow University Veterinary School. The frozen sections were washed with 0.02% NP-40 in PBS twice for 5 minutes. The samples were then fixed in 0.2% glutaraldehyde, 2 mM $MgCl_2$, 5 mM EGTA, and 0.02% NP-40 in PBS for 5 minutes. The samples were then washed twice with 0.02% NP-40 in PBS. The sections were then stained with 1 $mgml^{-1}$ X-gal, 5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6.6H_2O$ and 2 mM $MgCl_2$, 0.01% sodium deoxycholate, and 0.02% NP-40 in PBS. The staining was carried out for 36 hours in a humid 37 °C incubator in the dark. After staining the sections were again washed twice with 0.02% NP-40 in PBS for 5 minutes. The samples were then returned to the Pathology Department at Glasgow University Veterinary School. The sections were stained with eosin, dehydrated and mounted. The tissue sections were visualised by light microscopy.

2.30 Treatment of mice with chemical mutagen MNNG.

In order to reactivate the transgenic marker gene, mice were treated with the chemical carcinogen MNNG either by topical application or intra-peritoneal injection. For topical application a stock solution of 3 $mgml^{-1}$ MNNG in acetone was made. A dose of 200 μl (a total dose of 600 μg) of this solution was applied to the back of the mice in a fume hood. This dose was repeated as often as twice weekly. For intra-peritoneal injection a stock solution of 2.5 $mgml^{-1}$ MNNG in tricaprylan (Fluka) was made. This was injected in to pregnant mice at day 8 - day 15 of pregnancy. Mice to be injected were initially weighed and a volume of the stock solution was injected such that the mice received a total dose of 5-25 $mgkg^{-1}$.(Inouye and Murakami, 1978; Faustman *et al.*, 1989)

CHAPTER 3

Results.

3. Results.

3.1 Design of strategy for cell marking.

Design is of paramount importance to the success of cell marking experiments. Care was taken to design a flexible system with the potential to be applied to all tissues although the main focus of the research was murine epidermis. There are a number of methods previously used for cell marking and lineage analysis in mice as discussed above in Section 1.12.2.

The application of a classical cell marking technique to the clonal analysis of murine epidermis has yet to be successfully attempted. Comparison of the epidermis to other tissues such as the intestinal epithelia and the haemopoietic system indicates that cell marking would contribute a great deal of useful information to the understanding of the development and maintenance of this tissue. No cell marking technique used has addressed the relationship between stem cell in the basal layer of the epidermis and the stem cells of the hair follicle. Marking of stem cells in the hair follicle could establish whether these cells usually contribute or are capable of contributing to the regeneration of interfollicular epidermis.

An attempt at classical cell marking had already been made in this laboratory. Recombinant replication-deficient retroviruses containing the *Escherichia coli* (*E. coli*) *lac Z* gene as a marker were applied to skin which had been disrupted by abrasion (Boeger-Brown, 1993). The same technique has been used to generate tumours in the epidermis when the applied retroviruses contained mutant H-*ras* oncogenes. The tumours generated contained the retroviruses applied. Very few marked cells were observed when skin treated with *lac Z*-containing retroviruses was stained for the activity of β-galactosidase. This may be due to a low rate of infection or perhaps a low expression of the *lac Z* gene in skin cells *in situ* although the retrovirus construct expressed *lac Z* at high levels in tissue culture.

One of the main problems in designing a useful marker for murine epidermal cells is the lack of naturally occurring marker genes which have been used to great effect in cell lineage studies in Drosophila and the intestinal epithelium. For the purposes of this project it was decided to introduce a marker gene into transgenic mice. Expression of the gene in the epidermis would allow the epidermal cell to be marked and their lineage

and contribution to the maintenance of the tissue to be determined. The use of transgenic mice would allow the development of a marker system in the epidermis which could also be applied to the analysis of cell lineage in different mouse tissues. This could be done by targeting the expression of the marker gene to the tissue of interest by altering the promoter which controls the gene. A ubiquitous promoter would allow expression of the gene in a wide variety of tissues and be useful in determining lineage throughout the mouse. A tissue-specific promoter could be used to engender high level expression of the marker gene in a tissue of interest although the marking in this case would be restricted to only one tissue.

The major problem encountered when using transgenic mice as a vehicle for cell marking is the expression of the introduced gene. It is vital to consider that the expression of an introduced transgenic marker gene could be patchy. The transgene may be expressed at a high level in one cell but in the neighbouring cell be expressed at a different level or even not at all. When the expression of the marker gene is patchy it is necessary to ensure that the gene must be activated to mark the cell and not inactivated. If the maker gene was to be inactivated to mark cells it would be impossible to distinguish the marked cell clones from cells which had randomly shut down expression of the gene.

When the marker gene is activated by a genetic event the marked cell clones are distinguished as the only cells that express the gene in the tissue. There is no interference from background expression in other cells in the tissue as they do not express the gene. This makes identification of cell clones much more simple. The effect of patchy expression may be to disguise some marked cells by preventing expression of the marker gene. It will not cause cells to be wrongly identified as marked cells. There is a possibility that all cells in a clone may have similar levels of expression. This would mean that clones of cells would be composed entirely of cells expressing the gene or entirely of cells not expressing the gene. This would prevent any problems with identification of only parts of clones rather than the whole clones. The effect of patchy expression of a transgene in mice would then be to reduce the frequency of clones generated in the mouse but not to affect the outcome of the clonal identification.

When transgenic mice are generated by pronuclear microinjection of DNA usually a number of copies of the transgene integrate in a head to tail array. This would potentially create problems if trying to inactivate a gene to mark a cell. Obviously if a cell contains a number of copies of a marker gene it would be necessary to inactivate all copies of the gene to mark the cell and distinguish it from the rest of the population. This is not a problem if the marker gene requires to be activated and not inactivated. If

a cell contains a number of copies of a marker gene which must be activated in order to distinguish it from the rest of the population the increased copy number may become an advantage. As the copy number is increased so the number of potential targets for activation is increased. This could lead to an increased frequency of marked cells generated when the mice are exposed to the activating agent.

3.1.1 Design of cell marking gene for use in transgenic mice.

The design of any marker gene must take into account the concepts discussed above. Initially a marker gene is required which can be expressed in all or most tissues of a transgenic mouse. A mechanism is required to activate this marker in single cells at a low enough frequency to distinguish individual cell clones. The frequency of detection must still be high enough to allow easy identification of clones in mouse tissues. Thirdly to perform *in situ* analysis of clones of cells populating mouse tissues the marker gene must be easily distinguishable in tissue sections or whole mount preparations.

The *Escherichia coli* (*E. coli*) *lac Z* gene was selected as a suitable candidate for a marker gene in the tissues of transgenic mice. The *lac Z* gene is not expressed in mammalian tissue and so marked cells expressing the gene would be easily distinguishable from the rest of the tissue where the gene could not be expressed. It had already been widely used in analysis of cell clones and gene expression in a wide variety of murine tissues. These analyses have demonstrated that the protein product of the *lac Z* gene is non-toxic and does not affect cell behaviour. This is a vital property of any marker gene used to analyse the population of normal mouse tissue by cell marking. If cell behaviour was affected by the expression of a marker gene the normal patterns of cell division may be disturbed, perhaps not only in the cells expressing the gene but also in the surrounding cells. This would render any conclusions reached by the use of the cell marking technique invalid, as the patterns of cell clones observed in murine tissue containing cells expressing the marker may be different from the patterns in normal tissue. The *lac Z* gene does not seem to influence cell behaviour and therefore circumvents this potential problem.

The *E. coli lac Z* gene codes for a protein product, β-galactosidase which is an enzyme. This enzyme is easily detectable *in situ* with a simple histochemical stain containing 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal). This makes the *lac Z* gene potentially very useful as a cell marker for analysis of the patterns of cell division in tissues

in vivo. Neither the protein nor the products of the histochemical stain are diffusible. This is also of tremendous importance. Diffusion of either the enzyme or the products of the stain could lead to the identification of cells which do not express the gene as marked cells because they acquire the marker or the stain from neighbouring marked cells.

Once a suitable marker gene for transgenic mice was determined it was then necessary to design a system for activation of the gene. Work going on in the laboratory at the time suggested that point mutation may provide a good system for activating marker genes. Point mutation was responsible for the activation of H-*ras* oncogenes in mouse skin tumours. An average of 12 and up to 50 skin tumours can be produced on the back skin of one mouse by treatment of the skin with chemical carcinogens as initiators and promoters (Kemp et al., 1993; C. Kemp, personal communication). Furthermore the activation of H-*ras* oncogene was dependent on specific changes in the gene sequence which could only be produced by treatment of the mice by specific carcinogenic chemicals.

After initiation of the tumours by activation of the *ras* oncogene the initiated cells must then undergo further changes before they can produce visible tumours on the back of the mouse. These changes are probably happen at a low frequency in cells and it is unlikely that every initiated cell produces a tumour. This means that if 50 tumours are produced the number of initiated cells which have suffered a point mutation at a specific base in the *ras* proto-oncogene by a single treatment with an initiating chemical carcinogen must be considerably higher than 50. It seems likely therefore that specific point mutation could be used as a method to activate marker genes at a reasonably high rate in murine epidermis.

The gene must be inactivated by a point mutation in its sequence before it can be used as a marker gene in this system. When this point mutation is reversed by the action of the chemical mutagen the marker gene is activated and distinguishes the cell from the background cells in the tissue. In the case of an enzyme (such as the *lac Z* gene selected as a marker) this point mutation must interfere with the activity of the enzyme. This could be for instance be achieved by a point mutation which alters an amino acid at the active site of the enzyme.

The frequency of activation of the marker gene could be increased if the point mutation required to activate the gene correlated with the specificity of a particular chemical sequence. Chemical mutagens often induce certain mutations at a higher frequency than

others. This occurs because the local DNA environment affects the interaction of the mutagenic chemical with the DNA and the shape distortion caused by the binding of the chemical to the DNA affects the resultant base substitution. Mutagenic chemicals will often mutate specific bases more frequently than others. N-nitro-N'-methyl-N-nitrosoguanidine (MNNG), for instance, methylates guanine residues more readily than to other nucleotide bases (the specificity of this chemical is more extensively discussed in Section 1.14.1).

It was decided to target the translation initiation codon ($^{5'}$ATG$^{3'}$) of the *lac Z* marker gene to prevent translation of the protein from the mRNA. This method could provide two separate advantages over mutation of a single amino acid residue. A mutation at this residue would be very likely to prevent wild-type function of the gene. If the gene is not translated from RNA to protein it can not function at all as an enzyme. Obviously there is still a possibility that such a mutation would be leaky as translation may be initiated at an alternate site. The gene would still have to be tested in tissue culture before use in transgenic mice as a marker gene.

Another reason for selecting the *lac Z* $^{5'}$ATG$^{3'}$ initiation codon as the target for mutagenesis was that it allowed the insertion of more than one target site for the chemical mutagen. This could be accomplished by cloning an oligonucleotide containing a number of target sites for a chemical mutagen in front of the gene substituting for the initiation codon (Figure 1). Mutation at any one of the sites in the oligonucleotide would result in the production of an in-frame initiation codon which would result in the translation of the gene and activation of the gene. Mutation in the oligonucleotide may attach a few amino acids on to the amino terminus of the gene. It was thought that these few amino acids would be unlikely to affect the function of the gene as the β-galactosidase protein has been used frequently as a fusion protein. This does not seem to affect the enzymatic function of the gene.

3.1.2 Design of the oligonucleotides for the fusion marker genes.

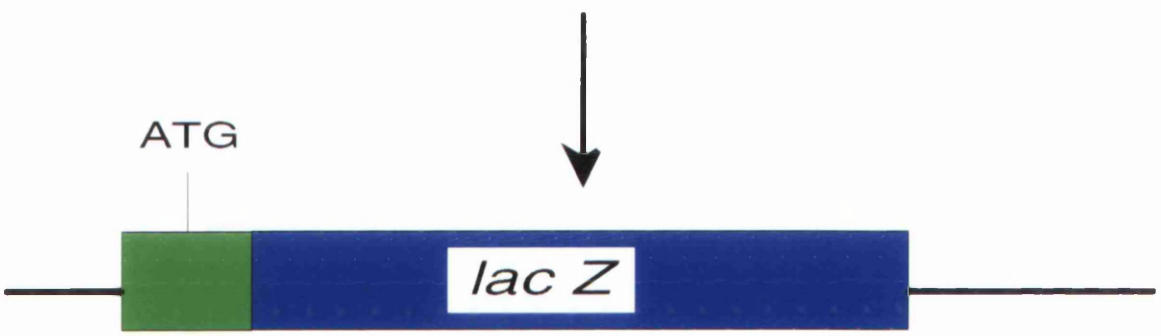
To increase the chances of producing a useful marker gene by the construction of the oligonucleotide-*lac Z* fusions it was decided to design two different oligonucleotides as potential mutagen target sites. Each of these two oligonucleotides was constructed to contain target sites for a different type of chemical mutagen. One oligonucleotide was designed to be activated by a point mutagen, the other was constructed to by activated

Figure 1. Diagram of the *lac Z* marker gene strategy.

An oligonucleotide inserted in front of a *lac Z* gene with no initiation codon is used to activate the gene.

Lac Z gene with no
ATG initiation codon

Oligonucleotide with target site for chemical
mutagen inserted in front of gene



Exposure to the chemical mutagen causes
a mutation in the sequence of the
oligonucleotide which results in the
production of an in-frame ATG initation codon

by a frameshift mutagen. The point mutagen selected was N-nitro-N'-methyl-N-nitrosoguanidine (MNNG - see Sections 1.17.1 and 3.1.3 for a discussion of the activity and oligonucleotide construction). This was selected because it is very mutagenic has a highly specific site of action. The mutations it causes have been characterised and it has been used successfully to induce H-ras mutations in murine epidermal cells *in vivo*. The frameshift mutagen selected was N2-acetylaminofluorene (AAF - see Sections 1.14.1 and 3.1.4 for a discussion of the activity and oligonucleotide construction). This was selected because it also has a well characterised mutagenic activity and a highly defined target sequence. It is a potent liver carcinogen and although it is debatable whether or not it will have any mutagenic activity in epidermal cells. Treatment of mice with activated derivatives such as N-acetoxy-acetylaminofluorene should produce the appropriate DNA adducts. This will hopefully extend the range of tissues to which this technique can be applied.

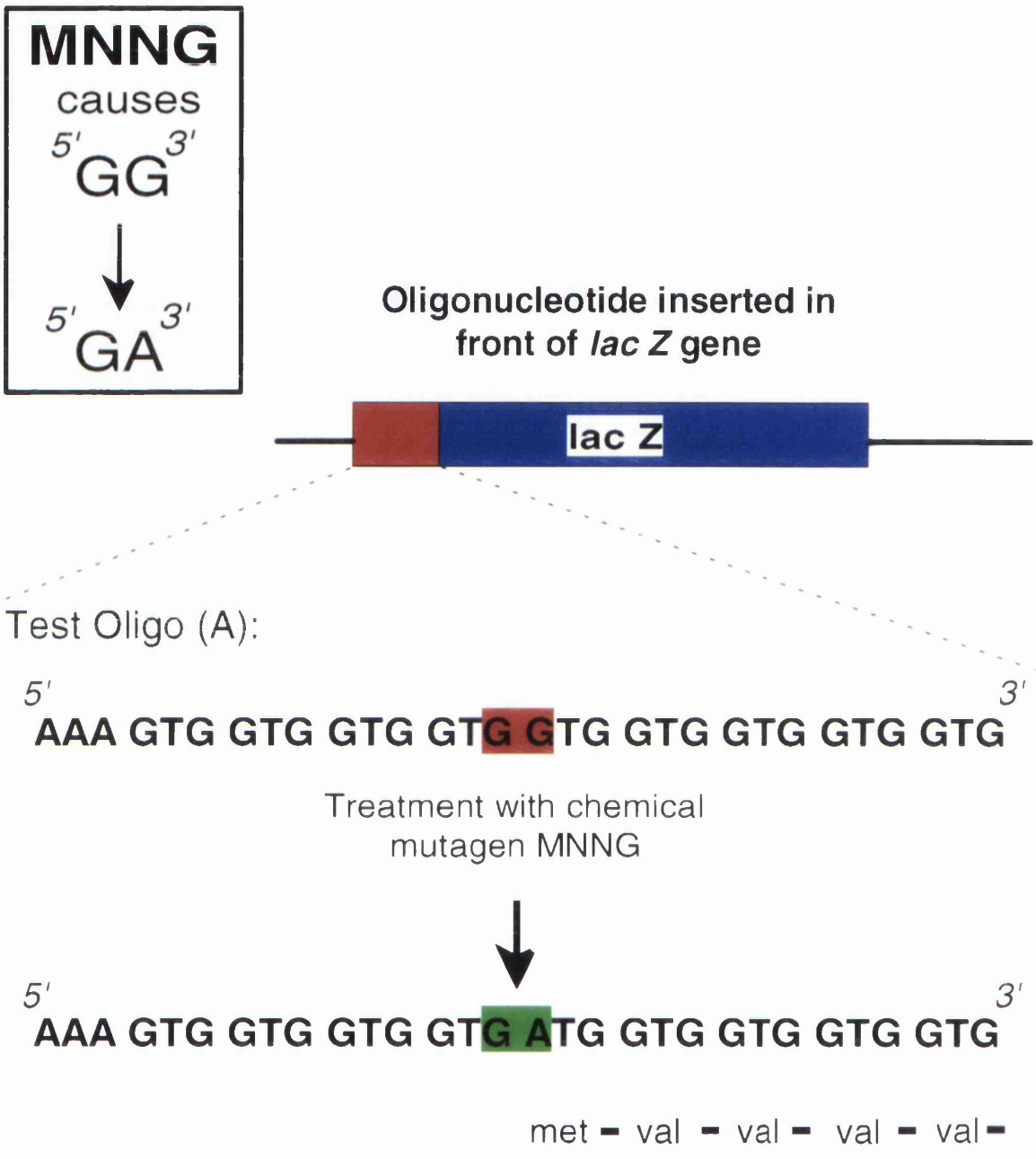
Because of the lack of available restriction sites and the availability of *lac Z* fusion vectors it was decided to clone these oligonucleotides blunt-ended into the fusion vector pMC 1871 (obtained from Pharmacia). The design and structure of the oligonucleotides is discussed below. At the same time as designing oligonucleotides with target sites for chemical mutagens it was necessary to engineer oligonucleotides which contained the predicted changes that would activate the gene. These would act as controls and allow testing to confirm that once the predicted mutation occurred it actually resulted in the activation of the *lac Z* marker gene.

3.1.3 Structure of oligonucleotide activated by MNNG.

For the purposes of this project it was desirable to construct an oligonucleotide which contains as many target sequences as possible for MNNG, mutation at any one of which would result in the production of an in-frame ATG initiation codon. This can be done by constructing an oligonucleotide with a series of $^{5'}GTG^{3'}$ codons and putting these sequences in frame with the gene (Figure 2). Direct repetition of these sequences ensures that a series of $^{5'}GG^{3'}$ mutagen target sites are produced. Mutation of the 3' G residue to an A will result in the production of an in-frame ATG start codon. The GTG codon specifies the amino acid valine so the reading frame will be maintained throughout the length of oligonucleotide. If mutation occurs anywhere along the length of the oligonucleotide at a predicted target sequence the protein should be translated with a small number of valine residues at the amino terminus. The oligonucleotide starts with a

Figure 2. Structure of the oligonucleotide activated by MNNG.

The most frequent mutation induced by the chemical MNNG is a G to A transition at the 3' sequence G residue of a $^{5'}GG^{3'}$ sequence. This known specificity was used to design an oligonucleotide which would activate a *lac Z* gene with no initiation codon when mutated by MNNG to generate an in-frame initiation codon. The sequence of the test (A) and control oligonucleotides (B and C) are shown.



Control Oligos:

(B) **AAA ATG GTG GTG GTG GTG GTG GTG GTG GTG**

(C) **AAA GTG GTG GTG GTG GTG GTG GTG GTG ATG**

series of three A residues in order to ensure that it has a proper ribosome binding site so the protein will be efficiently translated in the event of mutation (Kozak, 1983).

Control oligonucleotides were also designed which included ATG sequences at either the first or the last GTG codon. These were made to demonstrate that the gene would be expressed in the event of a mutation in the oligo. If mutations at the first and last GTG codons are capable of activating the gene this implies that mutation anywhere along the sequence of the oligo will activate the gene. The MNNG-mutation test oligonucleotide was named oligonucleotide A, the 5' ATG control was named oligonucleotide B and the 3' control was named oligonucleotide C.

3.1.4 Structure of oligonucleotide activated by AAF.

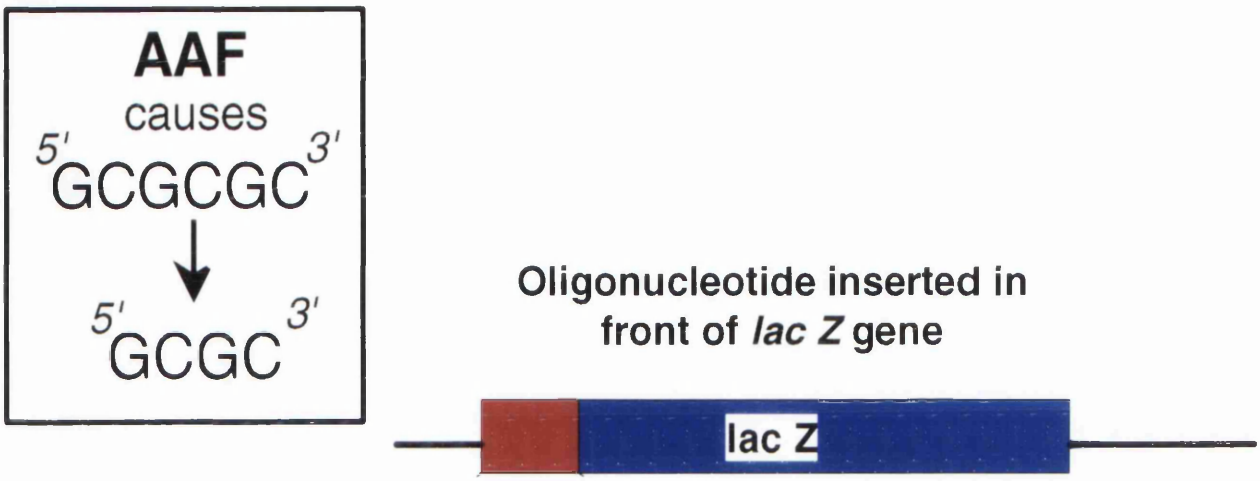
An oligonucleotide was designed to take advantage of the mutational specificity of the -AAF adduct (Figure 3). The oligo consists of an ATG initiation codon followed by a run of GC dinucleotides such that the ATG codon is two bases out of frame with the coding sequence of the gene. An -AAF adduct at any one of these should result in the deletion of a dinucleotide which will bring the initiation codon into frame with the rest of the gene. The coding sequence is maintained throughout the sequence of the oligo which should code for six alternating alanine and arginine residues before reaching the coding sequence of the gene. To ensure there is a proper ribosome binding site there is a sequence of three A nucleotides before the ATG (Kozak, 1983). A control oligonucleotide was also designed to ensure the gene would be expressed if the deletion took place and that the short amino acid fusion does not interfere with protein function.

Treatment of rats with AAF results in the production of the acetylated adduct on guanine residues in liver cell DNA (Beland *et al.*, 1982; Poirier *et al.*, 1983, 1992). However epidermal cells do not contain the enzymatic pathway which is responsible for generating the reactive metabolites in the liver (Kreik, 1974). In the liver the reactive species is formed by the action of a sulphotransferase enzyme which forms a sulphate ester from the oxygenated N-hydroxy-acetylaminofluorene (NH-AAF). Formation of this sulphate ester bond allows the generation of a nitrenium ion which is capable of reacting with DNA. In other tissues the acetylated adduct can be produced via the formation of the formation of another ester N-acetoxy-AAF.

As AAF cannot be metabolised in the skin it is not a potent carcinogen in this tissue. There is evidence that AAF does not damage DNA in epidermal cells as it does not

Figure 3. Structure of the oligonucleotide activated by AAF.

The carcinogenic chemical AAF will cause the deletion of two base pairs in a $^{5'}$GCGCGC$^{3'}$ sequence. When the test oligonucleotide is exposed to the metabolically activated form of the AAF chemical a resultant two base pair deletion will bring the oligo into frame with the *lac Z* gene. The sequence of the control oligonucleotide (E) is also shown.



Test Oligo (D):

5' **AAA ATG GCG CGC GCG CGC GCG CGC GC** 3'

Treatment with chemical
mutagen AAF

5' **AAA ATG GCG CGC GCG CGC GCG CGC** 3'

met - ala - arg - ala - arg - ala - arg -

Control Oligo:

(E) **AAA ATG GCG CGC GCG CGC GCG CGC**

cause unscheduled DNA synthesis in these cells (Sayer *et al.*, 1988). Treatment of epidermal cells with the reactive metabolite of AAF, N-acetoxy-acetylaminofluorene, does result in the production of -AAF adducts on the DNA in murine epidermal cell lines. These adducts are only present as 3 - 5% of the total DNA adducts formed (Poirier et al., 1979, 1980). The presence of these adducts in epidermal cells though indicates that this compound should be able to cause the reverse mutation in the oligonucleotide which will activate the *lac Z* gene. This should be true not only in the epidermis but also in other tissues. The 26 bp AAF test oligo was named oligonucleotide D and the 24 bp control oligo with the in-frame ATG sequence was named oligonucleotide E.

3.2 Cloning of fusion genes and expression vectors.

3.2.1 Synthesis and analysis of oligonucleotides.

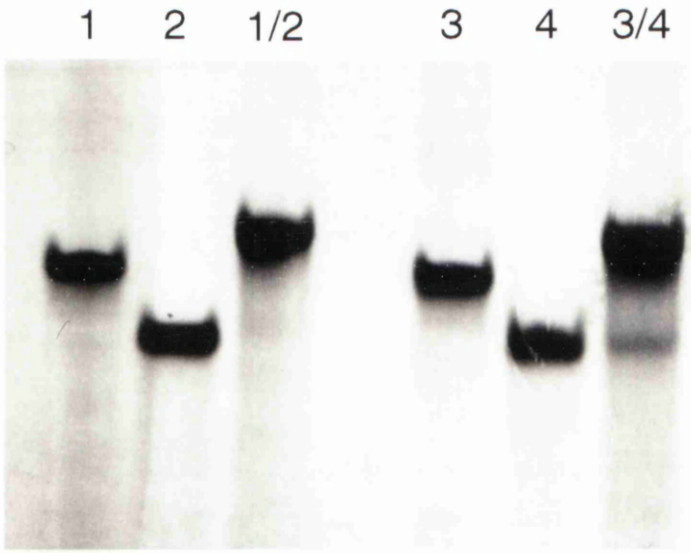
The first step in construction of the mutant *lac Z* fusion genes was the synthesis of the specific oligonucleotides which were to act as target sequences for the chemical mutagens to activate expression of the *lac Z* gene. The oligonucleotides were synthesised single-stranded on an Applied Biosystems 381A Synthesiser. A total of 10 oligonucleotides were synthesised (the sequences of these oligonucleotides are described in Section 2.7.3). These represented the positive and negative strands of the two test and three control oligos to be cloned into the *lac Z* fusion vector. The oligonucleotides were visualised in polyacrylamide gels by UV shadowing. Using this technique all the 10 synthesised oligonucleotides were visualised and shown to be complete.

Once good preparations of all 10 oligonucleotides had been obtained the corresponding positive and negative strands were annealed (described in Section 2.7.1). Figure 4 shows a selection of these oligos. Each of the oligonucleotides is represented by a single band which shows that each one of the oligos is intact. When the oligos have been mixed and annealed the band shifts to a new location. Each of the annealed oligo mixtures seems to consist mainly of annealed double-stranded oligos and very little single stranded product. Inspection of the gel perhaps suggest that the annealing of oligos 9 and 10 may have worked less well than the others. This may well be due to the

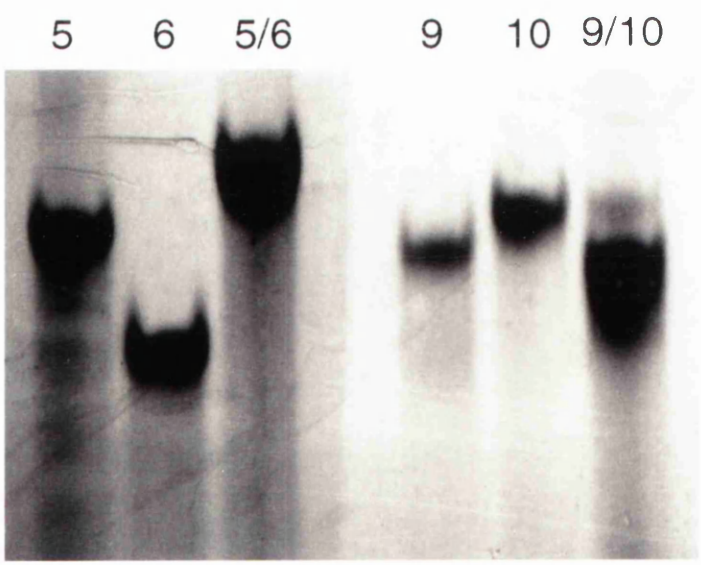
Figure 4. U. V. shadowing of the oligonucleotides.

The oligonucleotides were run on polyacrylamide gels and visualised by U. V. shadowing to check the integrity and the annealing. **(a)** Gel showing oligos 1, 2, 3 and 4 before and after annealing to their corresponding partners. **(b)** Oligos 5, 6, 9 and 10 before and after annealing to their corresponding partners.

(a)



(b)



large number of GC repeats found in the oligos and the fact that these oligos would be able to self-anneal.

3.2.2 Cloning of oligonucleotides in pMC1871.

Once double-stranded oligos representing the two test and three control oligonucleotides had been obtained the next stage was to clone these in front of the *lac Z* gene. This was done by cloning the oligos blunt-ended into the *lac Z* fusion vector pMC1871. The structure of the pMC1871 plasmid and the cloning strategy are outlined in Figure 5. This vector contains an entire *lac Z* gene which is missing the first eight codons of the *lac Z* gene.

As the vector is missing the first eight codons it has no $^{5'}$ATG$^{3'}$ initiation codon which makes it perfect for use in this system. The oligos were cloned into the unique *Sma I* site of the fusion vector which is at the start of the *lac Z* gene. Great care was taken to ensure that the oligos were cloned so that the ATGs or potential ATGs were in-frame with the rest of the *lac Z* gene and that they would be capable of initiating translation of the *lac Z* gene once in place. The pMC1871 vector was restricted with *Sma I* restriction enzyme. The oligos were then cloned directly into the plasmid vector without modification.

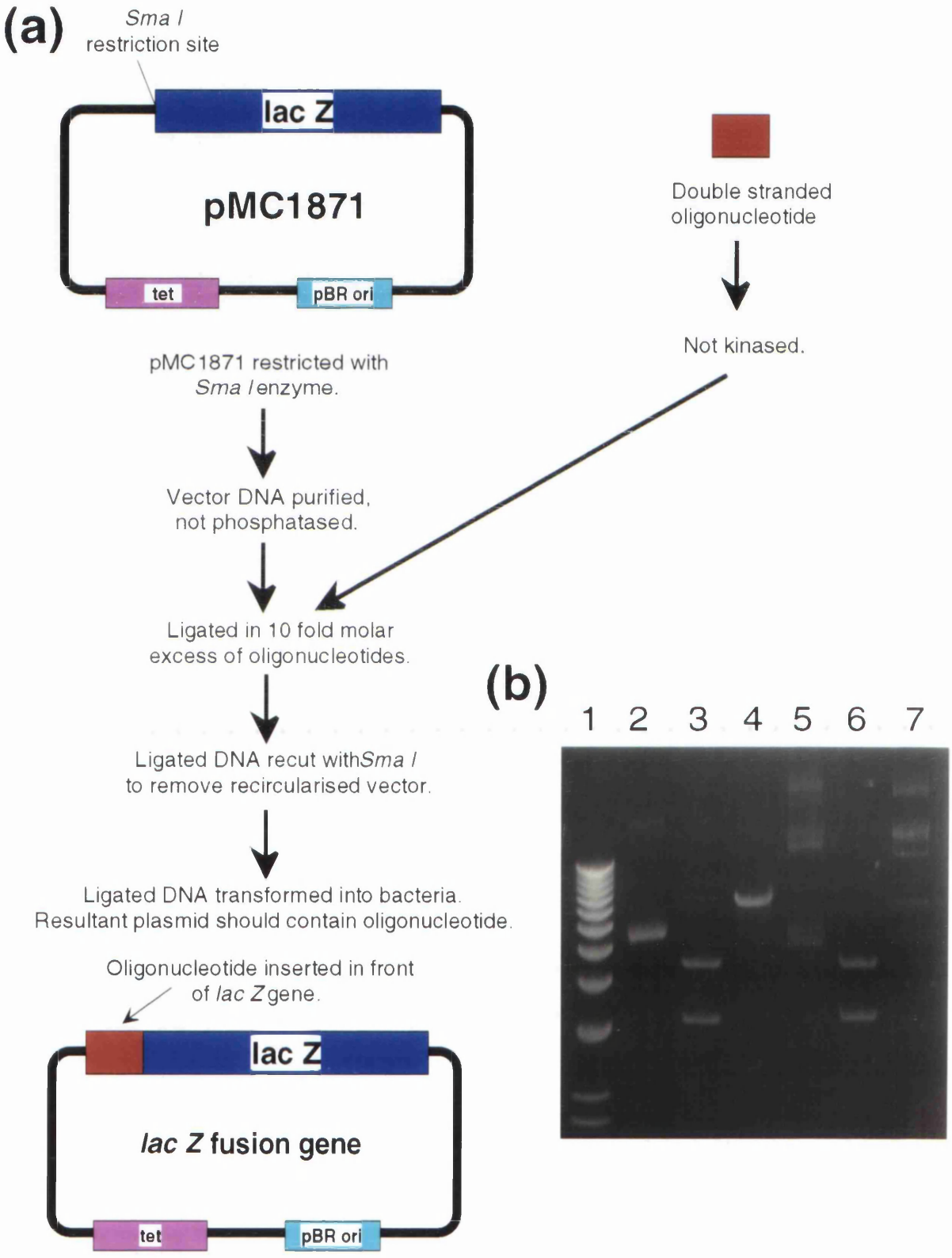
After ligation the mixture was redigested *Sma I* enzyme to remove unwanted recircularised bacteria and then transformed into bacteria as described in Section 2.8. The cloning strategy performed very well. A large number of colonies were obtained and when minpreps were performed on these clones 100% of the plasmids obtained were resistant to restriction with the *Sma I* endonuclease.

3.2.3 Mutations in the sequence of the oligos cloned into vector.

All of the miniprep DNA samples checked were resistant to cutting with the *Sma I* restriction enzyme. This indicated that they may contain oligonucleotides. A few of these minipreps were selected for sequencing. These were then sequenced to reveal the orientation of the oligonucleotides and if the oligonucleotides were intact. DNA was prepared from the selected bacterial clones by the medium scale method described in

Figure 5. Cloning of double stranded oligonucleotides into pMC1871.

(a) Outline of the cloning strategy. **(b)** Restriction endonuclease digestion of the plasmid. **1)** 1 kb ladder **2)** Undigested pMC1871 **3)** pMC1871 digested with *Sal I.* Cuts out a 3.1 kb *lac Z* gene. **4)** pMC1871 digested with *Sma I.* linearises 7.5 kb plasmid. **5)** Undigested pMC1871 containing ligated oligonucleotide. **6)** pMC1871 containing oligo digested with *Sal I.* **7)** pMC1871 containing oligonucleotide digested with *Sma I.* It does not cut.


(a)
Sma I
restriction site
lac Z
pMC1871
tet
pBR ori
Double stranded
oligonucleotide
Not kinased.
pMC1871 restricted with
Sma I enzyme.
Vector DNA purified,
not phosphatased.
Ligated in 10 fold molar
excess of oligonucleotides.
Ligated DNA recut with Sma I
to remove recircularised vector.
Ligated DNA transformed into bacteria.
Resultant plasmid should contain oligonucleotide.
Oligonucleotide inserted in front
of lac Z gene.
lac Z
lac Z fusion gene
tet
pBR ori
(b)
1 2 3 4 5 6 7

Section 2.9.2 or occasionally by the caesium chloride method described in section 2.9.3. Both methods were provided DNA of sufficient purity to give good sequencing results.

Unfortunately when the sequences of the oligos cloned into the plasmid pMC1871 were determined a very high mutation rate of the oligos was discovered. By this method a total of five oligonuctleotide A inserts and twelve oligonucleotide E inserts were sequenced. The oligonucleotide A sequence was designed to be activated by MNNG and consists of a run of GTG trinucleotides. Of the five oligonucleotide A inserts sequenced four were in the wrong orientation and one was in the correct orientation. One of the five oligo A inserts displayed the correct sequence but was in the wrong orientation. The other four oligonucleotides sequenced displayed mutations.

The most significant mutagenesis was observed in the cloning of the oligonucleotide D inserts. This oligo was designed to be activated by AAF and consists of a run of GC dinucleotides. Of the 12 oligo D inserts sequenced five were in the correct orientation but all 12 displayed a mutant sequence. Three mutant oligonucleotide sequences are shown in Figure 6. Some sequence compression is visible in both these sequences. All the mutations in the oligo D inserts sequenced were invariably addition or deletions in the run of GC dinucleotides. Ten of the mutations were deletions and the remaining two were addition mutations. Although it may have been possible to clone the oligonucleotide sequences based on a run of GTG trinucleotides by this system it seemed unlikely that the oligonucleotides based on the GC dinucleotide repeat would be cloned using this system.

3.2.4 Use of SURE™ bacteria to clone repetitive oligo sequences.

As it was proving impossible to clone the oligonucleotides by the method previously described a strain of bacteria called SURE™ (Stop Unwanted Rearrangement Events) was purchased from Stratagene to enable the cloning of these repetitive sequences. These bacteria allow the cloning of pieces of DNA which are 'unclonable' in other strains of *E. coli*. Host bacteria can often delete foreign pieces of DNA which take up unusual conformations. The SURE™ bacteria have been engineered to carry mutations inactivating the pathways which catalyse the rearrangement of these DNA pieces. These include mutations in the *rec B*, *rec C*, *uvr C* and *umu C* genes. These mutations help to allow the cloning of highly unstable DNA structures such as cruciforms formed by inverted repeat sequences and Z-DNA. The Z-DNA conformation is formed by

Figure 6. Sequencing of mutant oligos in pMC1871.

Sequencing of oligonucleotides in pMC1871. The oligonucleotides shown have all undergone rearrangements during the cloning procedure. **(i)** Mutant oligonucleotide B in reverse orientation. Nine base pairs have been deleted from the oligo. **(ii)** Mutant oligonucleotide D. Four base pairs have been added to the sequence. **(iii)** Mutant oligonucleotide E. Twelve base pairs have been deleted from the oligonucleotide sequence.

(i)

CAC CAC CAC CAC CAC CAT TTT

(iii)

GCG CGC GCG CGC

(ii)

AAA ATG GCG CGC GGC GGC GCG CGC GCG CGC

stretches of alternating purine and pyrimidine residues. The oligonucleotide D and E sequences is based largely of dinucleotide GC repeats and therefore consist mainly of alternating purine and pyrimidine residues which might be expected to adopt the Z-DNA conformation. This may be the explanation for the deletions of the bases in the D and E oligonucleotides in the DH5α bacteria initially used.

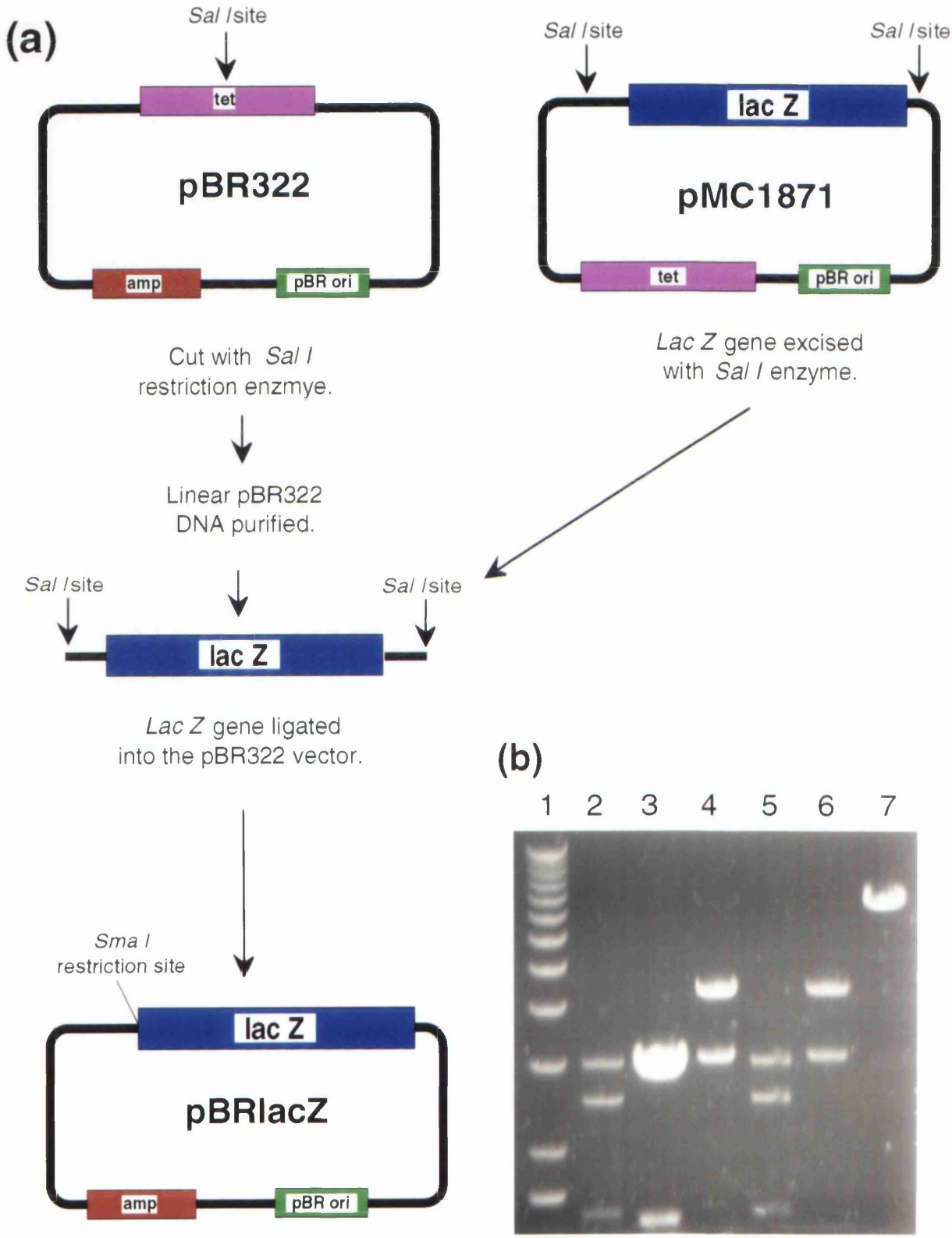
Unfortunately the SURE™ E. coli carry an F' plasmid which carries the tetracycline resistance gene. As the selection for the pMC1871 plasmid is based on the tetracycline resistance gene it was necessary to reclone the *lac Z* sequences into another vector which carried another resistance gene. The vector chosen for this purpose was pBR322. The cloning strategy for this purpose is described in Figure 7. The *lac Z* sequences from the pMC1871 vector were excised with the restriction enzyme *Sal I* and ligated into the *Sal I* site in the tetracycline resistance gene of the pBR322 vector (described in Section 2.7). After ligation the mixture was transformed into DH5α™ cells. The bacterial colonies were picked and replica-plated onto duplicate agar plates, one of which contained ampicillin as the selective antibiotic and the other contained tetracycline. A few colonies which grew on the ampicillin plate but not the tetracycline plate were selected. One of these plasmids contained the correct *lac Z* insert. This plasmid could now be used to clone the oligonucleotides in front of the *lac Z* gene in the SURE™ bacteria.

The new vector, named pBRlacZ, contains no *Sma I* restriction enzyme site other than the one directly juxtaposed to the front of the *lac Z* gene and therefore still allowed the cloning of the oligonucleotides into the unique *Sma I* site in front of the *lac Z* gene. The same cloning strategy was used to clone the oligos into pBRlacZ as described above for pMC1871 (see Section 3.2.2 and Figure 5).

The new cloning strategy still did not work perfectly. Only 50% of the minipreps from bacterial colonies produced by transformations of the SURE bacteria were indigestible by the *Sma I* enzyme and therefore only 50% could contain inserts. It is possible that these bacteria may be more easily transformed by linear DNA although this was not tested. Alternatively the conditions for digestion of the plasmid may not have been perfect. Of these 50% approximately 75-90% proved to be impossible to sequence. This was despite the fact that the DNA preparations were intact and digested with restriction enzymes other than *Sma I*. Eventually detailed restriction analysis of the clones which could not be sequenced demonstrated that 100% of these clones had suffered a deletion of >200 bp of the plasmid sequence at the beginning of the *lac Z* gene, around the oligonucleotide. This deletion had also included the binding site for the

Figure 7. Cloning of the pBRLacZ vector.

(a) Diagram of the cloning strategy. **(b)** Restriction enzyme digestion of the plasmids. **1)** 1 kb ladder. **2)** pMC1871 digested with *Hpa I*. **3)** pMC1871 digested with *Pst I*. **4)** pMC1871 digested with *Sal I*. **5)** pBRLacZ digested with *Hpa I*. **6)** pBRLacZ digested with *Sal I*. **7)** pBRLacZ digested with *Pst I*.


(a)
Sal I site
tet
pBR322
amp
pBR ori
Cut with Sal I
restriction enzmye.
Linear pBR322
DNA purified.
Sal I site
Sal I site
lac Z
Lac Z gene ligated
into the pBR322 vector.
Sma I
restriction site
lac Z
pBRlacZ
amp
pBR ori
Sal I site
Sal I site
lac Z
pMC1871
tet
pBR ori
Lac Z gene excised
with Sal I enzyme.
(b)
1 2 3 4 5 6 7

universal primer and therefore as the primer site was deleted it was impossible to sequence these clones. The reason for these extremely frequent deletions remains a mystery. It is possible that they were caused by recombination with DNA sequences present in these SURE™ bacteria.

The remainder of the clones could however be sequenced. Of these approximately 75% contained oligonucleotide inserts. Remarkably, >90% of the oligonucleotide sequences were intact and had suffered no mutation. All of the 5 oligo inserts were eventually cloned by this method. Figure 8 shows the sequencing of a pBRlacZ clones containing oligos A, B, D and E. All of the oligos were present in the correct orientation and the entire sequence of the oligo can be read from the sequencing gel flanked by the appropriate plasmid sequences. Similar sequencing results have been obtained for the other oligonucleotide C demonstrating that it was cloned in the correct orientation in front of the *lac Z* gene and is intact.

Each of the *lac Z* fusion gene was named after the oligonucleotide that it contained. The fusion genes, named $LacZ^{A}$, $LacZ^{B}$, $LacZ^{C}$, $LacZ^{D}$ and $LacZ^{E}$ contained the oligonucleotides A, B, C, D and E respectively (the oligonucleotides are described in Figures 2, 3 and 5, and in Section 2.7.3 and the names of the constructs are outlined in Figure 9).

3.2.5 Cloning of *lac Z* fusion genes into pIrv-Neo-Act.

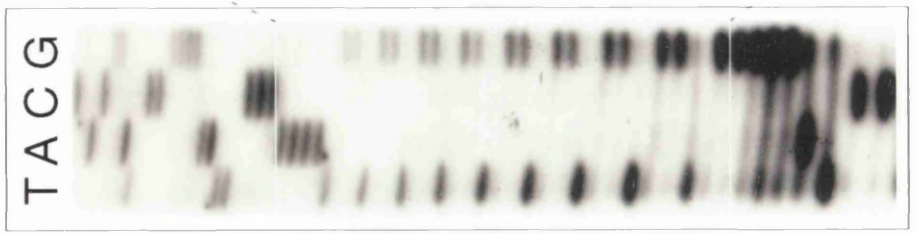
After constructing the oligonucleotide-*lac Z* fusion gene the next step was to clone all of the fusion genes into expression vectors to test the marker genes in tissue culture. The plasmid vector pIrv-Neo-Act-LacZ was selected for use as an expression vector. The vector is a replication-incompetent Muloney murine leukaemia virus (MoMuLV). The use of this vector had a number of advantages for the purposes of this project.

Firstly the expression of the *lac Z* gene was driven by a 400 bp rat β-actin promoter. This would allow expression of the marker genes in a wide variety of cell types. The vector also contained a *neo* gene which confers resistance to the antibiotic G418. This would allow the selection of stably transfected cell lines to test the effectiveness of the marker gene approach. In addition the vector contains both a cap site and a poly A tail for the mRNA which should help the expression of the cloned genes. The whole construct was part of a replication deficient retrovirus which would allow the production

Figure 8. Sequencing of oligonucleotides in pBRLacZ.

Intact oligonucleotide sequences cloned into pBRLacZ. All sequences are in the correct orientation. **(i)** Intact oligonucleotide A. **(ii)** Intact oligonucleotide B **(iii)** Intact oligonucleotide B. **(iv)** Intact oligonucleotide E.

AAA ATG GTG GTG GTG GTG GTG GTG GTG GTG

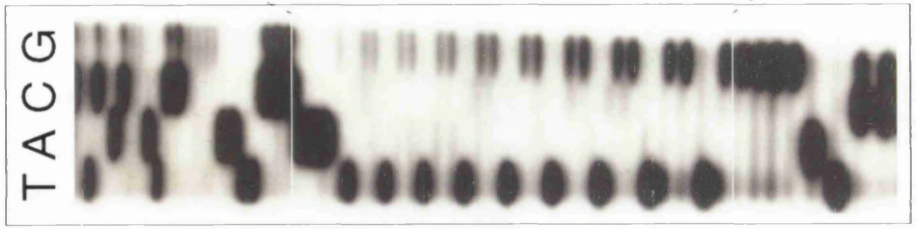


(ii)

AAA ATG GCG CGC GCG CGC GCG CGC

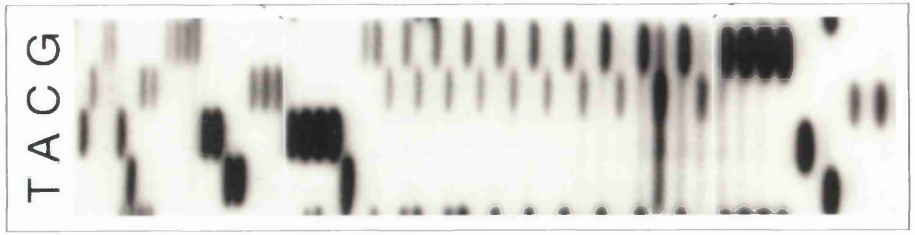
(iv)

AAA GTG GTG GTG GTG GTG GTG GTG GTG GTG



(i)

AAA ATG GCG CGC GCG CGC GCG CGC GC



(iii)

Figure 9. Table of names of the constructs generated.

The *lac Z* fusion genes were named after the oligonucleotides that they contain. The expression vectors are referred to by the name of the promoter that controlled the expression of the *lac Z* fusion gene.

Oligonucleotide	Mutagen	Activity	Start codon	Lac Z	pIrv-Neo-Act	pIrv-Neo-K5
A	MNNG	Test	No	$LacZ^{A}$	$Act\text{-}LacZ^{A}$	☒
B	MNNG	Control	Yes	$LacZ^{B}$	$Act\text{-}LacZ^{B}$	☒
C	MNNG	Control	Yes	$LacZ^{C}$	$Act\text{-}LacZ^{C}$	☒
D	AAF	Test	No	$LacZ^{D}$	$Act\text{-}LacZ^{D}$	$K5\text{-}LacZ^{D}$
E	AAF	Control	Yes	$LacZ^{E}$	$Act\text{-}LacZ^{E}$	$K5\text{-}LacZ^{E}$

of retroviruses containing the marker genes for infection of cells with single copies of the marker genes. This vector had previously been used to generate transgenic mice with the *lac Z* gene and wide expression of the transgene in embryonic mice (Beddington et al., 1990). Expression of the transgene was almost ubiquitous in the early embryo and the construct would certainly be useful for analysis of lineage in the tissue of prenatal mice. Hopefully this would be useful for testing the marker gene *in vivo* in as well as in tissue culture.

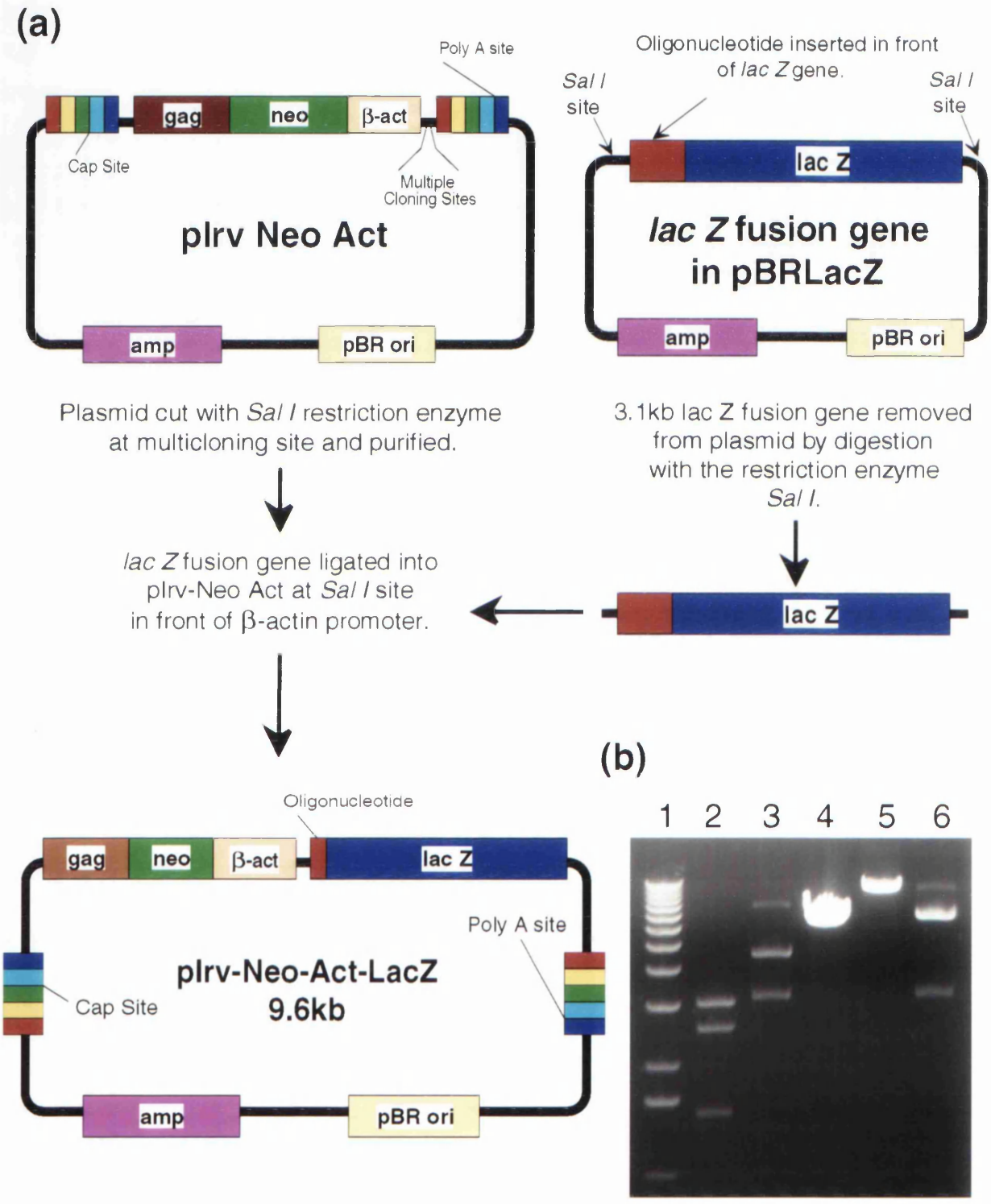
The pIrv-Neo-Act-LacZ vector was obtained from Rosa Beddington in Oxford. The vector already contained a wild-type *lac Z* gene and could be used as a positive control in transfection experiments. It is subsequently referred to as Act-LacZ$^{wild\text{-}type}$. In order to clone the marker genes into the vector the wild-type *lac Z* gene was removed by digestion with the restriction endonuclease *Bam HI*. The vector was then religated the resultant plasmid pIrv-Neo-Act vector was used for the cloning of the marker gene constructs in pBRLacZ. These constructs were cloned into the *Sal I* site in the polylinker in front of the β-actin promoter. Figure 10 shows a diagram of the pIrv construct and the cloning strategy used to insert the marker genes in front of the promoter.

The cloning of the marker genes into the pIrv-Neo-Act vector was performed according to the methods outlined in Section 2.7. The vector was restricted with *Sal I* and phosphatased to prevent self-religation. The marker genes were removed from the pBRlacZ vector by restriction with *Sal I* and the fragment separated on a low-melting point agarose gel. The band containing the 3.1kb *lac Z* insert was isolated using Geneclean™. The pIrv-Neo-Act vector was mixed with inserts at various concentrations and ligated with T4 DNA ligase in appropriate conditions overnight at 14°C. All ligations were transformed into SURE™ competent bacteria in order to maintain the integrity of the oligonucleotide sequences.

Despite these precautions the cloning of these vectors proved to be very inefficient. A very low number of marker gene inserts were found in the plasmids obtained from minipreps of colonies obtained from transformation of the ligated vector insert mixtures into the SURE™ competent cells. The marker gene inserts seemed to act as inhibitors and prevented religation of vector alone. A lower number of colonies was frequently obtained with transformations vector/insert mixtures than vector alone. This did not seem to be due to a contaminant of the insert preparations as insert were ligated into the vector with very little problem. The low frequency of ligation was perhaps caused by the same mechanism which caused the deletions in the pBRlacZ/oligo ligations described

Figure 10. Cloning of the *lac Z* fusion genes into pIrv-Neo-Act.

(a) Outline of the cloning strategy. **(b)** Restriction enzyme digestion of the plasmids. **1)** 1 kb ladder **2)** pBRLacZ^A digested with *Hpa I*. **3)** pBRLacZ^A digested with *Sal I*. 3.1 kb band is the *lac Z* gene. **4)** pIrv-Neo-Act linearised with *Sal I*. **5)** Act-LacZ^A construct linearised by *Not I*. **6)** Act-LacZ^A construct cut with *Sal I*. Removes *lac Z* from the construct.


(a)
Poly A site
gag
neo
β-act
Cap Site
Multiple
Cloning Sites
pIrv Neo Act
amp
pBR ori
Oligonucleotide inserted in front
of lac Z gene.
Sal I
site
Sal I
site
lac Z
lac Z fusion gene
in pBRLacZ
amp
pBR ori
Plasmid cut with Sal I restriction enzyme
at multicloning site and purified.
3.1kb lac Z fusion gene removed
from plasmid by digestion
with the restriction enzyme
Sal I.
lac Z
lac Z fusion gene ligated into
pIrv-Neo Act at Sal I site
in front of β-actin promoter.
(b)
1 2 3 4 5 6
Oligonucleotide
gag
neo
β-act
lac Z
Poly A site
pIrv-Neo-Act-LacZ
9.6kb
Cap Site
amp
pBR ori

in Section 3.2.4. Digestions of miniprep samples with restriction enzymes often revealed unusual band sizes which suggested that rearrangement of the vector was occurring. When the number of colonies obtained from the transformation of the vector/marker gene ligation were less than that of the vector alone no plasmids containing the correct insert were obtained. Even when the number of colonies obtained by transformation of the vector/marker gene ligation was greater than that of vector alone as little as 1% of the miniprep samples analysed contained the correct insert.

Eventually all the marker gene constructs and the control construct were cloned into the pIrv-Neo-Act vector. Once the marker gene constructs were cloned into the pIrv-Neo-Act vector the oligonucleotides were sequenced to check that their integrity had been maintained (Figure 11). All the constructs contained oligonucleotides which were intact and maintained their sequence. No mutations were observed in the oligonucleotide sequences during this cloning step.

Each vector was named after the *lac Z* fusion gene it contained. The vectors were called Act-LacZ^{A}, Act-LacZ^{B}, Act-LacZ^{C}, Act-LacZ^{D} and Act-LacZ^{E} and contained the fusion genes LacZ^{A}, LacZ^{B}, LacZ^{C}, LacZ^{D} and LacZ^{E} respectively (Figure 9 describes the names of the constructs generated). The original vector which contains an intact *lac Z* gene under the control of a β-actin promoter was used as a positive control when testing the oligo/*lac Z* fusion gene constructs and is referred to as Act-LacZ$^{wild\text{-}type}$.

At the same time as the construction of the marker gene constructs another vector was generated as a negative control for the transfection experiments discussed in Section 3.3. This vector included the *lac Z* gene with no ATG sequence and no oligonucleotide inserted in front of it. As the gene has no ATG start site inserted in front of the coding sequence of the *lac Z* gene the mRNA transcribed from this gene should not be translated and the cells should not express the β-galactosidase protein. This *lac Z* gene was cloned directly from the vector pMC1871 into the *Sal I* site in front of the β-actin promoter in the pIrv-Neo-Act plasmid. This was achieved using exactly the same strategy used for cloning the oligo/*lac Z* fusion genes into the same site as described above. This negative control construct was designated Act-LacZ^{pMC1871}.

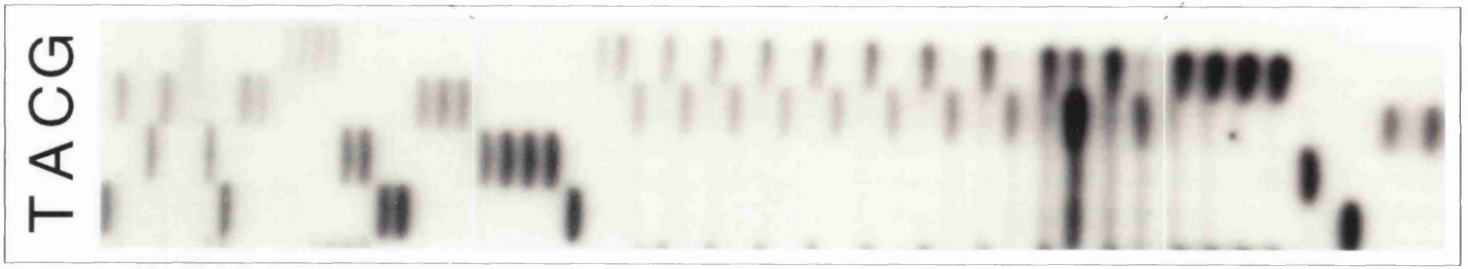
Figure 11. Sequencing of oligonucleotides in pIrv-Neo-Act and pIrv-Neo-K5 constructs.

The oligonucleotides retained their integrity after subcloning into of the *lac Z* fusion genes into expression vectors. **(i)** Sequencing of oligonucleotide D after subcloning of the $LacZ^{D}$ fusion gene into pIrv-Neo Act. **(ii)** Sequencing of oligonucleotide E after subcloning into the pIrv-Neo-K5 construct.

AAA ATG GCG CGC GCG CGC GCG CGC

(ii)

AAA ATG GCG CGC GCG CGC GCG CGC GC



(i)

3.2.6 Generation of pIrv-Neo-K5 vector.

In order to obtain expression of the marker gene constructs specifically in epidermal cells principally in transgenic mice but also in epidermal cells *in vitro*, a vector containing an epidermal specific keratin promoter was constructed. This vector was constructed from the pIrv-Neo-Act expression plasmid described in Section 3.2.5. This was used as a basis for the construction of a new expression vector because it contained a number of useful elements. This vector was based on a replication incompetent retrovirus and would allow the production of retroviruses and the vector contains the neo gene under the control of the viral LTR region which confers resistance to the antibiotic G418.

The promoter chosen for the epidermal specific vector was a 1.4 kb piece of the 5' sequence of the bovine keratin III *(BK III)* gene. This gene is equivalent to the human keratin 5 gene and is expressed in the basal layer of the epidermis in mouse (Byrne *et al.*, 1994). This promoter had already been shown to express at high levels in murine epidermal cell lines during the course of other projects going on in the laboratory. The new epidermal specific vector, named pIrv-Neo-K5, was constructed by swapping the 1.4 kb fragment of the *BK III* gene for the 400 bp β-actin promoter present in the pIrv-Neo-Act construct. The cloning strategy used to construct the pIrv-Neo-K5 expression vector is detailed in Figure 12.

The *BKIII* was ligated directly into the *Bgl II/Bam HI* sites in the pIrv-Neo vector using the standard techniques described in Section 2.6. The pIrv-Neo vector was phosphatased and mixed with different concentrations of the BKIII insert and was then ligated with T4 DNA ligase. Part of each ligation reaction was then transformed into SURE™ competent bacteria. In contrast to the results obtained with the cloning of the oligo/*lac Z* fusion genes this cloning step worked very efficiently. Greater than 94% of the clones analysed contained the appropriate insert.

3.2.7 Cloning of *lac Z* fusion genes D and E into pIrv-Neo-K5.

The *lac Z* fusion genes were cloned into the pIrv-Neo-K5 construct in a similar fashion to the pIrv-Neo-Act vector (Figure 13). The *Sal I* enzyme site could not be used as this enzyme cut in the K5 promoter. Instead, an *Xho I* site located in the polylinker at the 3' end of the K5 promoter and cloned into the pIrv-Neo vector along with the K5

Figure 12. Generation of the pIrv-Neo-K5 expression vector.

(a) Diagram of the cloning strategy. **(b)** Restriction enzyme digestion of plasmids. **1)** 1 kb ladder. **2)** pIC-K5 construct linearised by digestion with *Bam HI*. **3)** 1.4 K5 promoter excised by digestion with *Bam HI/Bgl II* enzymes. **4)** pIrv-Neo-Act linearised with *Not I*. **5)** 400 bp actin promoter removed from pIrv-Neo-Act by digestion with *Bgl II/ Bam HI*. **6)** pIrv-Neo-K5 linearised by *Not I*. **7)** 1.4 kb K5 promoter removed from pIrv-Neo-K5 by digestion with *Bgl II*.

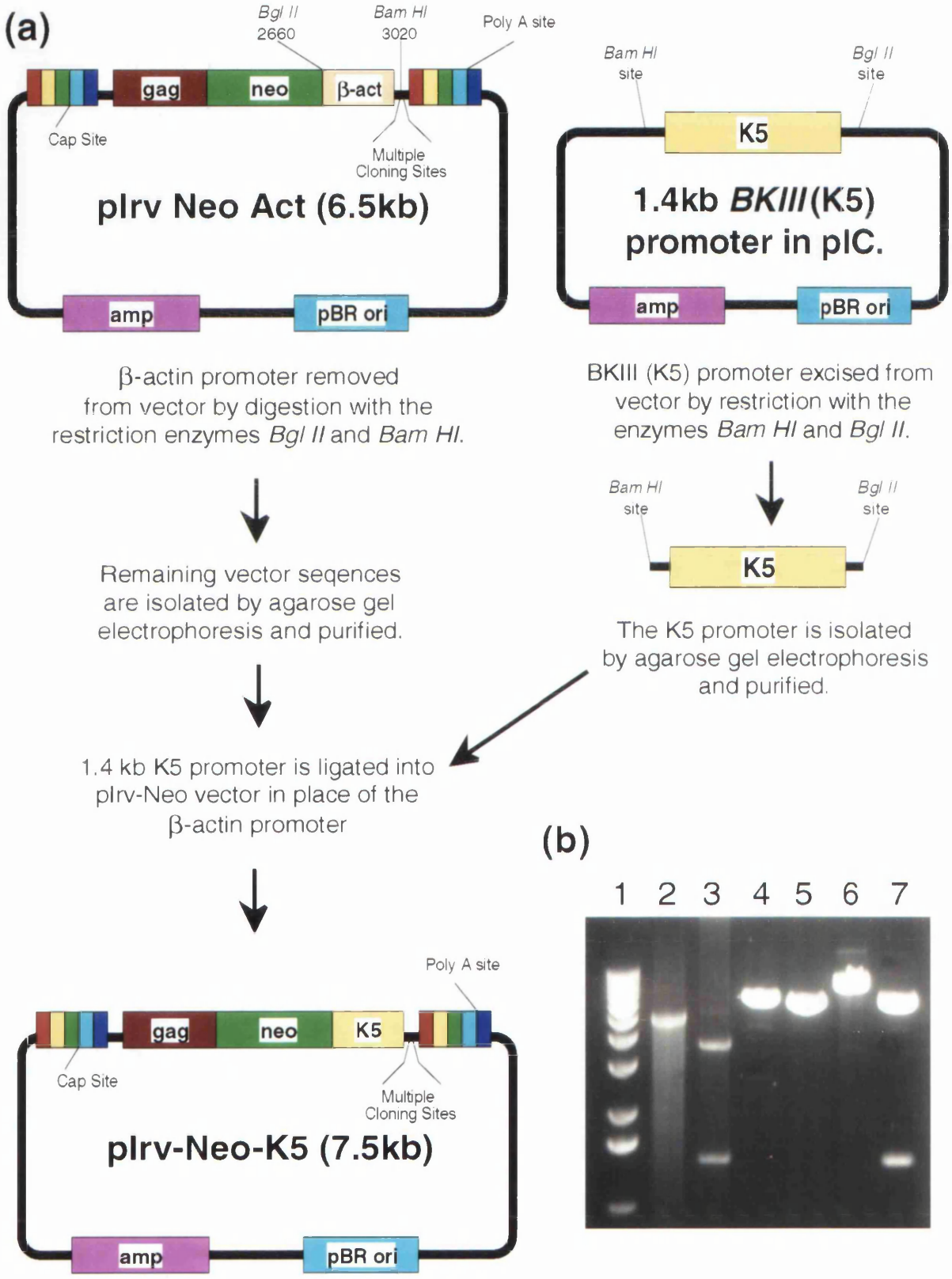
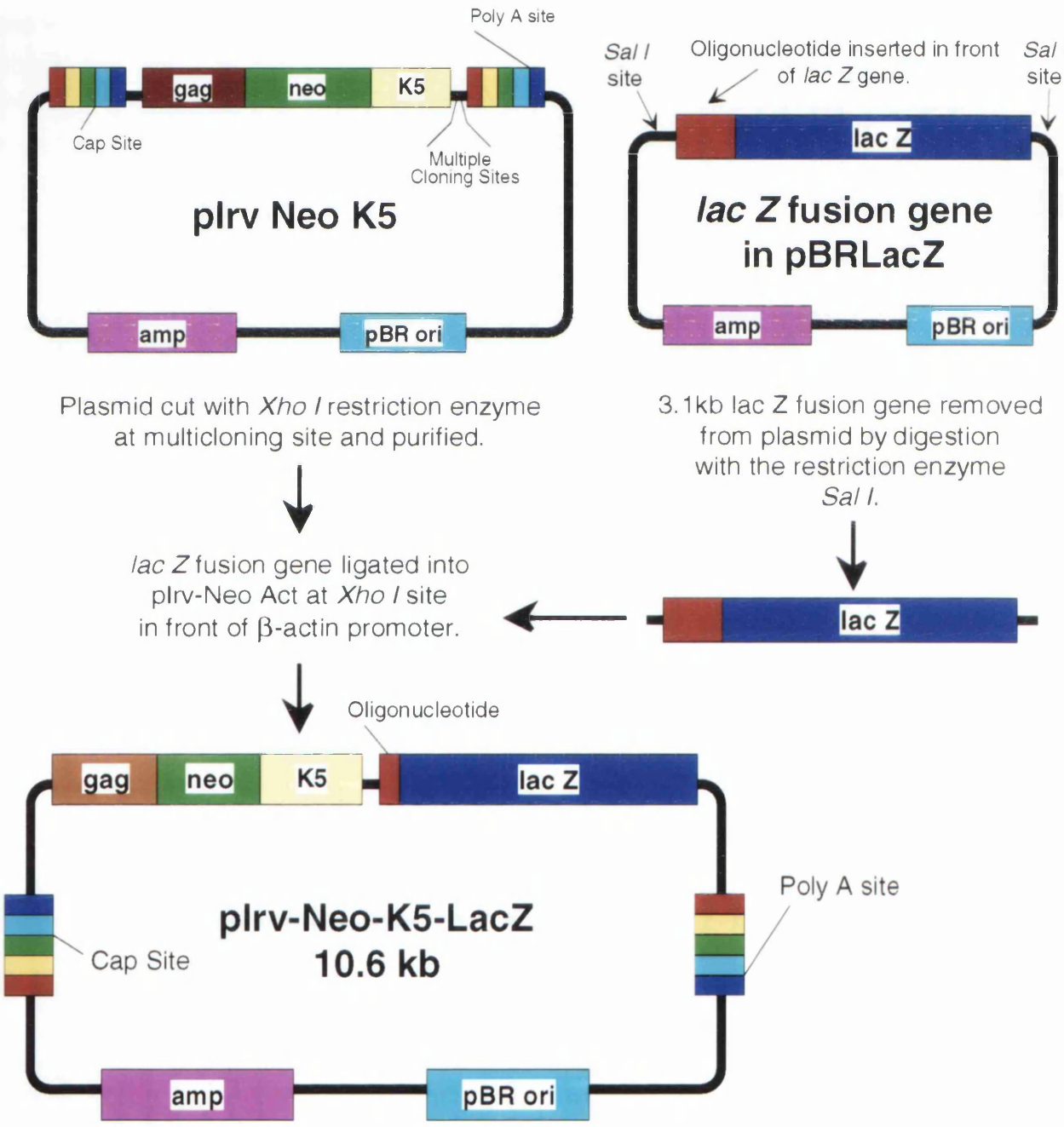

(a)
Bgl II
2660
Bam HI
3020
Poly A site
gag
neo
β-act
Cap Site
Multiple
Cloning Sites
pIrv Neo Act (6.5kb)
amp
pBR ori
β-actin promoter removed
from vector by digestion with the
restriction enzymes Bgl II and Bam HI.
Remaining vector seqences
are isolated by agarose gel
electrophoresis and purified.
1.4 kb K5 promoter is ligated into
pIrv-Neo vector in place of the
β-actin promoter
Poly A site
gag
neo
K5
Cap Site
Multiple
Cloning Sites
pIrv-Neo-K5 (7.5kb)
amp
pBR ori
Bam HI
site
Bgl II
site
K5
1.4kb BKIII(K5)
promoter in pIC.
amp
pBR ori
BKIII (K5) promoter excised from
vector by restriction with the
enzymes Bam HI and Bgl II.
Bam HI
site
Bgl II
site
K5
The K5 promoter is isolated
by agarose gel electrophoresis
and purified.
(b)
1 2 3 4 5 6 7

Figure 13. Cloning of *lac Z* fusion genes into pIrv-Neo-K5.

Outline of the cloning strategy used to clone the fusion genes.


Poly A site
gag
neo
K5
Cap Site
Multiple
Cloning Sites
pIrv Neo K5
amp
pBR ori
Sal I
site
Oligonucleotide inserted in front
of lac Z gene.
Sal I
site
lac Z
lac Z fusion gene
in pBRLacZ
amp
pBR ori
Plasmid cut with Xho I restriction enzyme
at multicloning site and purified.
3.1kb lac Z fusion gene removed
from plasmid by digestion
with the restriction enzyme
Sal I.
lac Z
lac Z fusion gene ligated into
pIrv-Neo Act at Xho I site
in front of β-actin promoter.
Oligonucleotide
gag
neo
K5
lac Z
Poly A site
pIrv-Neo-K5-LacZ
10.6 kb
Cap Site
amp
pBR ori

promoter was used instead. *Xho I* and *Sal I* have compatible restriction overhang sequences which allowed the *lac Z* fusion genes removed from the pBRLacZ vector by restriction with Sal I to be cloned directly into the *Xho I* site at the 3' end of the K5 promoter.

The cloning of the marker genes D and E into the pIrv-Neo-K5 vector was performed according to the methods outlined in Section 2.7. The vector was restricted with *Xho* and phosphatased to prevent self-religation. The marker gene-oligonucleotides were removed from the pBRlacZ vector by restriction with *Sal I* and the fragment separated on a low-melting point agarose gel. The pIrv-Neo-K5 vector was mixed marker gene inserts at various concentrations and ligated with T4 DNA ligase in appropriate conditions overnight at 14°C. All ligations were transformed into SURE™ competent bacteria in order to maintain the integrity of the oligonucleotide sequences. This cloning step proved to be slightly more efficient than the cloning of the *lac Z* fusion genes into the pIrv-Neo-Act vector. Approximately 5% of the minipreps prepared from colonies of transformed SURE™ *E. coli.* contained the appropriate insert.

Once the marker gene constructs were cloned into the pIrv-Neo-K5 vector the oligonucleotides were sequenced to check that their integrity had been maintained (Figure 11). Both the constructs contained oligonucleotides which were intact and maintained their sequence. The vectors were named after the *lac Z* fusion gene they contained. The vectors were called K5-LacZ^{D} and K5-LacZ^{E} and contained the fusion genes LacZ^{D} and LacZ^{E} respectively (Figure 9 describes the names of the constructs generated).

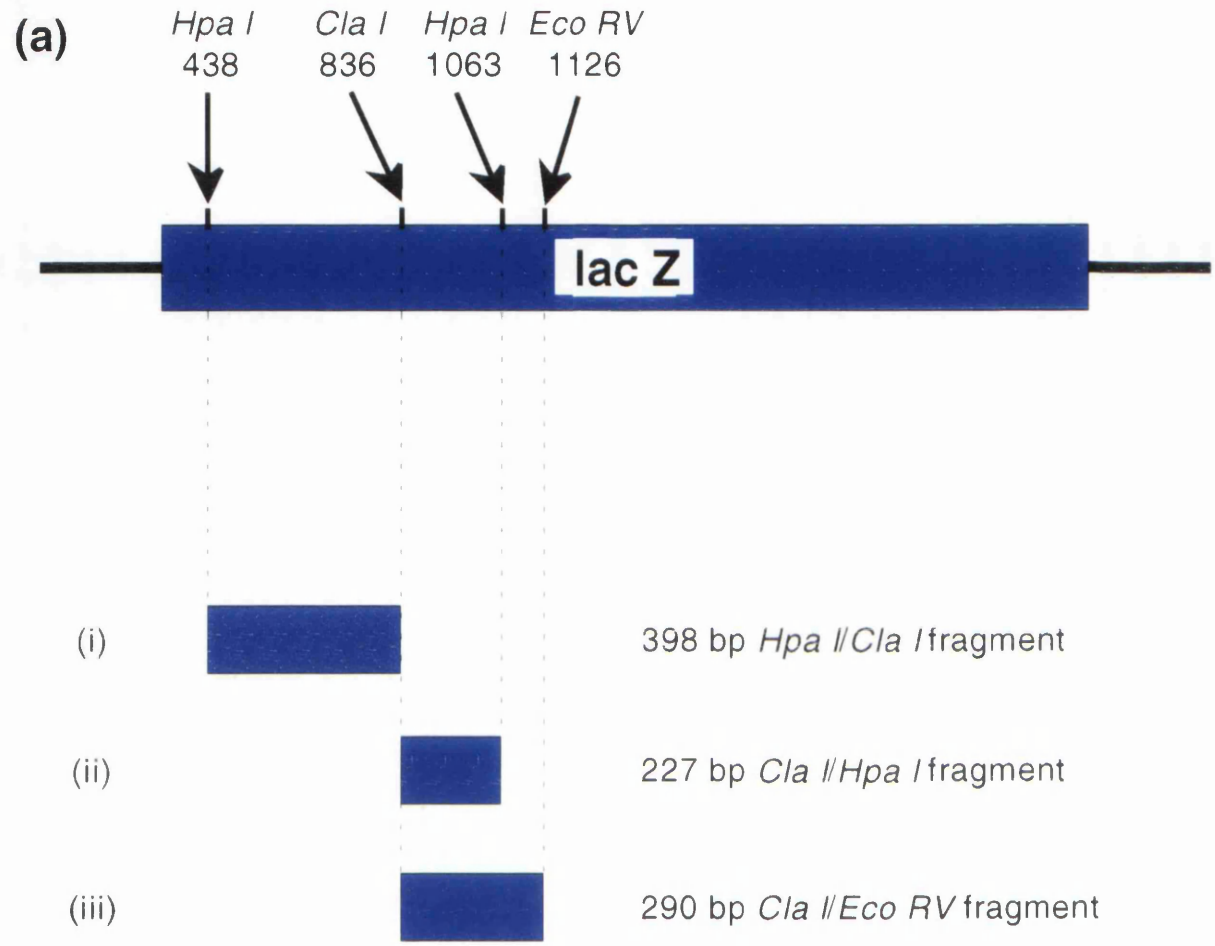
3.2.8 Cloning of fragments of *lac Z* gene into pBluescript for generation of and DNA and RNA probes.

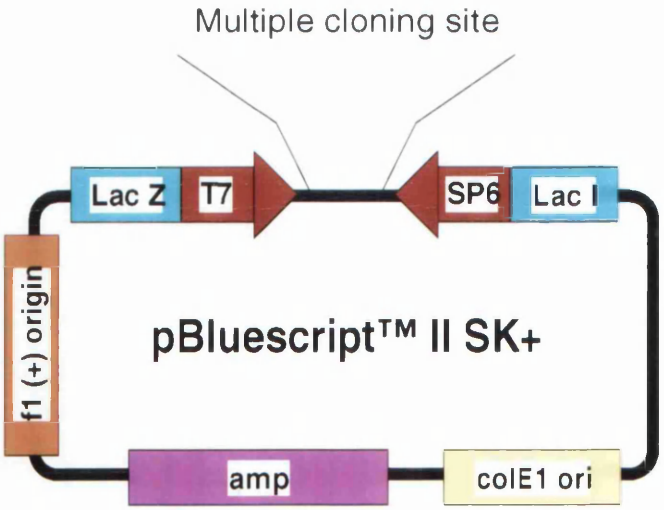
Three fragments of the *lac Z* gene were cloned into the pBluescript™ plasmid for use as RNA and DNA probes for the analysis of transgenic mouse tissues. The pBluescript™ vector was used because it contains sites for the T3 and T7 RNA polymerases. Figure 14 outlines the fragments of the *lac Z* gene cloned for use as probes.

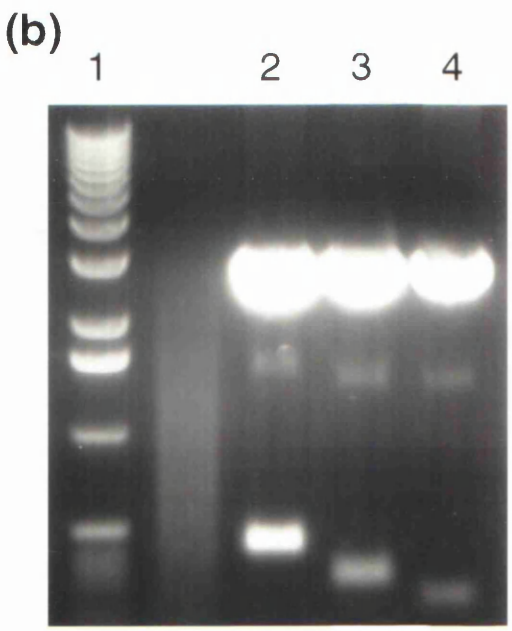
All the *lac Z* inserts prepared were excised from the *lac Z* gene by restriction with *Cla I* and a restriction enzyme that cleaved the DNA such that it is blunt ended (either *Hpa I* or *Eco RV*). The inserts were subsequently cloned into a the pBluescript™ SK+ vector

Figure 14. Cloning of *lac Z* fragments for DNA and RNA probes.

(a) Diagram of the *lac Z* fragments selected for use as probes and the pBluescript™ vector. **(b)** Restriction enzyme analysis of the resultant constructs. **1)** 1 kb ladder. **2)** 398 bp fragment excised from pBluescript by digestion with *Bgl II*. **3)** 290 bp fragment. **4)** 227 bp fragment.


(a)
Hpa I
438
Cla I
836
Hpa I
1063
Eco RV
1126
lac Z
(i)
398 bp Hpa I/Cla I fragment
(ii)
227 bp Cla I/Hpa I fragment
(iii)
290 bp Cla I/Eco RV fragment


Multiple cloning site
Lac Z
T7
SP6
Lac I
f1 (+) origin
pBluescript™ II SK+
amp
colE1 ori


(b)
1
2
3
4

that had been restricted with *Cla I* and the blunt-end cutter *Eco RV*. The cloning was performed according to the methods detailed in Section 2.7. This cloning step worked very efficiently and approximately 75% of the minipreps screened contained inserts. Three different sized inserts were prepared to allow different sizes of probes to be prepared.

3.3 Testing of marker system *in vitro*.

3.3.1 $CaPO_4$ transfection of Act-Lac Z fusion genes into a murine fibroblast cell line.

The pIrv-Neo-Act-LacZ fusion gene constructs were transfected into cells *in vitro* to test that the constructs allowed the appropriate expression of the β-galactosidase protein. The initial transfections were performed on the murine embryonic fibroblast cell line NIH 3T3. The NIH 3T3 cells were cultured according to the methods detailed in Section 2.11.1 and the transfections of the fusion gene expression constructs were carried out according to the $CaPO_4$ method detailed in Section 2.12.1. After transfection the cells were cultured for a further three days in growth medium. The cells then fixed and stained for β-galactosidase activity with a stain containing the chromogenic β-galactosidase substrate 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal) as detailed in Section 2.15.

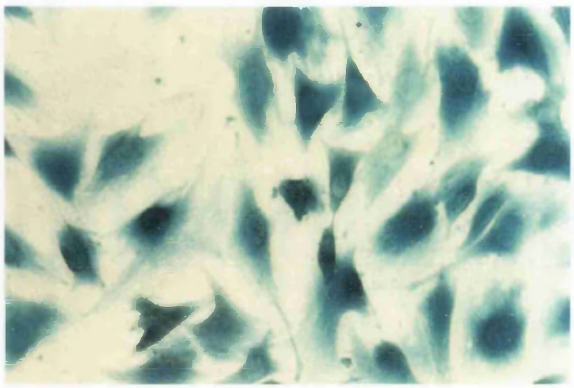
The results of these transfection experiments are detailed in Figure 15. Two negative control transfections were set up. The first of these transfections was with the pIrv-Neo-Act vector alone which has no *lac Z* gene sequence at all. None of the transfected cells with this vector stained with the X-gal stain. This demonstrates that the cells have no intrinsic background enzymatic activity to convert the X-gal from colourless to blue. For the other negative control NIH 3T3 cells were transfected with the vector Act-Lac$Z^{pMC1871}$ (described in section 3.2.4). This vector includes a *lac Z* gene with no endogenous ATG initiation codon and no oligonucleotide inserted in front of it. There was no staining in the cells transfected with this construct either. This indicates that β-galactosidase enzyme is not translated from a *lac Z* gene which has no ATG initiation codon.

Figure 15. Transfection of Act-LacZ constructs into NIH3T3 cells.

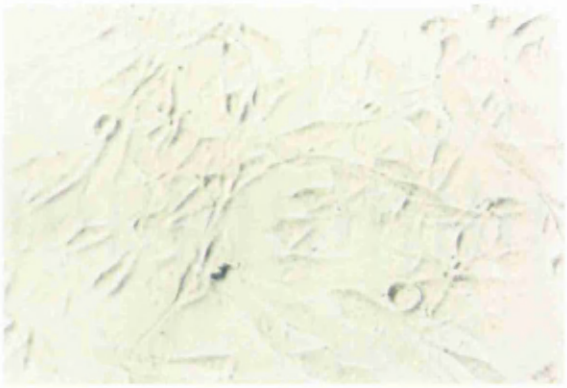
(a) Transfected NIH3T3 cell populations stained with X-gal for β-galactosidase activity. **(i)** Act-LacZ$^{wild\text{-}type}$. **(ii)** Act-LacZ^{A}. **(iii)**Act-LacZ^{B}. **(iv)** Act-LacZ^{D}. **(b)** Table of β-galactosidase activity in transient and stable transfections of NIH3T3 cells with the Act-LacZ constructs. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. Transiently transfected cells were stained two days after transfection. Pools of stably transfected cells were stained after four weeks selection in G418 when sufficient numbers of selected cells had been obtained. Flasks were screened under a light microscope. Positive indicates that a high proportion (>1%) of cells stained strongly with X-gal. Negative indicates no staining cells were seen.

(a)

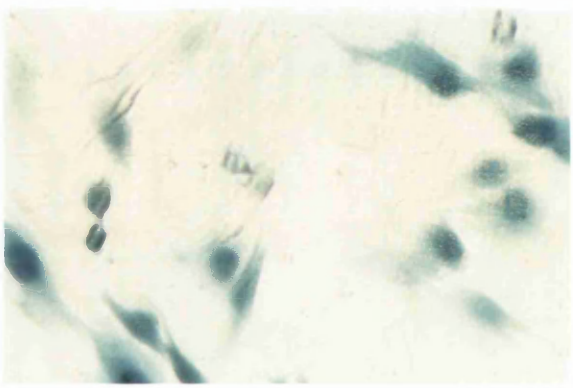
(i) Act-LacZ *wild-type*



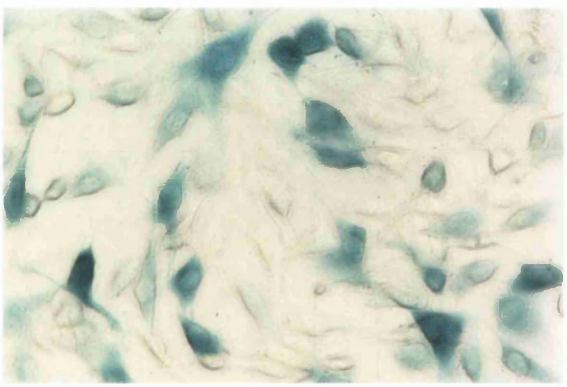
(ii) Act-LacZ *A*



(iii) Act-LacZ *B*



(iv) Act-LacZ *D*



(b)

Construct transfected.	Transient NIH3T3 transfection.	Stable NIH3T3 transfection.
pIrv-Neo-Act	Negative	Negative
pIrv-Neo-Act-LacZ *pMC1871*	Negative	Negative
pIrv-Neo-Act-LacZ *wild-type*	Positive	Positive
pIrv-Neo-Act-LacZ *A*	Negative	Negative
pIrv-Neo-Act-LacZ *B*	Positive	Positive
pIrv-Neo-Act-LacZ *C*	Positive	Positive
pIrv-Neo-Act-LacZ *D*	Positive	Positive
pIrv-Neo-Act-LacZ *E*	Positive	Positive

A high level of β-galactosidase expression was observed in the cell transfected with the Act-LacZ$^{wild\text{-}type}$ construct which contains the original intact *lac Z* gene. This demonstrates that the expression vector is functional and supports the expression of the β-galactosidase protein in fibroblast cells. These experiments also demonstrate that high levels of β-galactosidase expression can be achieved by transient transfection of the pIrv-Neo-Act constructs into fibroblast cells and that this system would be useful for testing the *lac Z* expression.

When the cells were transfected with the pIrv-Neo-Act-LacZ^{A} construct there was no staining observed in any of the cells. This indicates that the addition of the oligonucleotide A sequences to the *lac Z* gene did not initiate translation of the protein. This oligonucleotide consists of a sequence of GTG trinucleotides in-frame with the *lac Z* coding region with no ATG initiation site. The possibility existed one of the GTG sequences in this oligo would act as an alternate start site in the absence of an ATG initiation codon and that translation would be started randomly within the sequence of this oligo. As there is no expression of the gene from the test oligo A sequences this suggests that it might be a useful marker gene.

Transfection of the Act-LacZ^{B} and Act-LacZ^{C} constructs into the fibroblast lines resulted in high expression of the β-galactosidase protein. These constructs contain the oligo/*lac Z* fusion genes constructed with the oligonucleotides B and C respectively. Oligonucleotides B and C are the two control oligonucleotides for the oligo to be activated by MNNG (detailed in Figure 2). Both these oligonucleotides contain a sequence of GTG trinucleotides but with one ATG trinucleotide in-frame with the sequence of the *lac Z* gene. Both of these support the expression of the β-galactosidase protein in cell *in vitro*. This demonstrates that the principal of the marking system works. A mutation in the sequence of the GTG oligonucleotide to give an ATG sequence in-frame with the coding region of the *lac Z* gene will cause translation of the mRNA and expression of the β-galactosidase protein. These results also indicate that the short amino acid sequences which to oligo adds to the coding sequence of the *lac Z* gene do not interfere significantly with the enzymatic activity of the β-galactosidase protein.

The oligonucleotide B has an ATG sequence at the 5' end of the GTG trinucleotide repeats and the oligonucleotide C has an ATG sequence at the 3' end of the GTG trinucleotide repeats. Both of these oligonucleotides support the expression of the β-galactosidase protein in transfected cells *in vitro*. Hence a start signal at either end of the inserted oligonucleotide is able to initiate translation of the protein. These results

indicate that an ATG sequence anywhere along the sequence of the oligonucleotide would be able to initiate translation of the gene. Therefore a mutation which generates an ATG at any point along the sequence of the oligonucleotide should be able to initiate translation of the gene. There are a number of potential target sites in the sequence for activation of the marker gene by the chemical mutagen MNNG. Thus the first set of transfections demonstrated that the LacZA fusion marker gene would be useful as a marker gene for lineage analysis.

The results obtained by transfection of the constructs, Act-LacZ^{D} and Act-LacZ^{E}, designed to be activated by the chemical mutagen AAF were not so encouraging. Transfections of the control construct Act-LacZ^{E} into the fibroblast cell line NIH 3T3 resulted in the expression of the β-galactosidase protein in these cells. This demonstrates that the expression of the gene can be supported by this oligonucleotide and that the short amino acid sequence added to the protein by the coding sequence of the oligonucleotide does not interfere with the activity of the β-galactosidase protein. This also demonstrated that a mutation in the sequence of the test oligo D (described in Figure 3 and Section 3.1.3) which deleted a GC dinucleotide and brought the ATG sequence into frame with the coding region of the *lac Z* gene would initiate translation of the gene.

Unfortunately transfection of the test construct Act-LacZ^{D} also resulted in high expression of the β-galactosidase in the transfected fibroblast cells. There were several possible explanations for the aberrant expression of the protein from this oligo. It is entirely possible that the oligonucleotide D did not prevent translation of the gene even though the ATG initiation codon is two bases out of frame with the rest of the *lac Z* gene. This would obviously make this oligo useless for the construction of a marker gene. Perhaps the oligonucleotide adopted a conformation in the cell that it allowed translation of the gene by missing out a few bases of the oligonucleotide bringing the ATG start codon into frame with the rest of the coding sequence of the *lac Z* gene.

The observed anomalous expression in the fibroblasts could also have been due to mutation in a single construct. Transfecting cells by the $CaPO_4$ method may introduce a large number of copies of a transfected piece of DNA into each cell. If one of the copies of the Act-LacZ^{D} construct was rearranged or mutated during transfection then expression of the *lac Z* gene may be detected in a number of transfected cells even though the majority of the constructs remain silent. Mutations often occur during transfection of constructs into cells. If for instance, the mutation rate was 5% and each transfected cell had on average 10 copies of the construct, the number of cells that

would be expected to express the gene is 50% of the transfected cell population. This rearrangement could even occur in the bacteria when the constructs were grown up in order to make the DNA. When cloning in bacteria these oligonucleotides they proved to be very unstable (see Section 3.2.4). The integration site of the DNA constructs may also affect the expression of the transfected genes. In order to test these hypotheses it was decided to introduce low copy numbers of the gene into cells (detailed in Sections 3.3.3 and 3.3.4). If the observed expression is due to a rearrangement in a small proportion of transfected pIrv-Neo-Act-LacZ^{D} vector then the majority of cells which contain a single copy of the gene should not express the β-galactosidase protein.

3.3.2 $CaPO_4$ transfection of Act-Lac Z fusion genes into a murine keratinocyte cell line.

As the marker genes were designed for use principally in epidermal cells, $CaPO_4$ transfections were carried out in murine keratinocyte cell lines in order to confirm that the conclusions reached from the results obtained from the transfections into fibroblast cells were also valid in epidermal cells. The murine epidermal cell line C5N was chosen as the recipient for the constructs as this cell line was not transformed and most closely resembled normal epidermal cells. These cells were cultured and maintained according to the methods detailed in Section 2.11.

The cells were transfected in duplicate both by $CaPO_4$ and lipofection for comparison (these techniques are described in Sections 2.12.1 and 2.12.2 respectively). 5 μg of DNA was transfected into cells in duplicate F75 flasks by each of these techniques in order to determine the efficiency of transfection in epidermal cells. Both techniques were almost equally effective for the transfection of epidermal cells, although the transfection efficiency obtained with the $CaPO_4$ techniques was probably slightly higher. The subsequent transfections described in this Section were performed by the $CaPO_4$ technique.

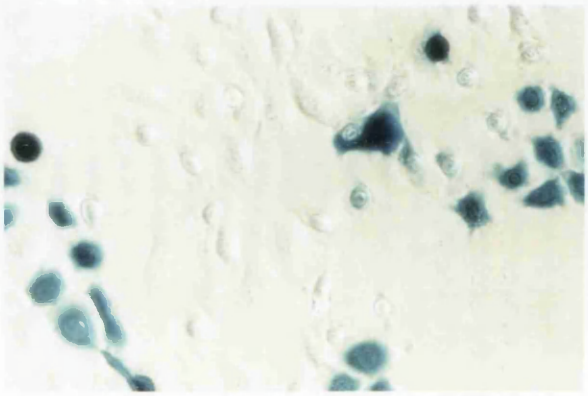
The $CaPO_4$ transient transfections into the C5N cells largely confirmed the results obtained by transfection of the NIH 3T3 cells (detailed in Figure 16). Transfection of the negative control constructs pIrv-Neo-Act and Act-LacZ^{pMC1871} produced no expression of the *lac Z* gene in the transfected cells whereas transfection of the positive control construct Act-LacZ$^{wild\text{-}type}$ again resulted in a high level of β-galactosidase

Figure 16. Transfection of Act-LacZ constructs into C5N cells.

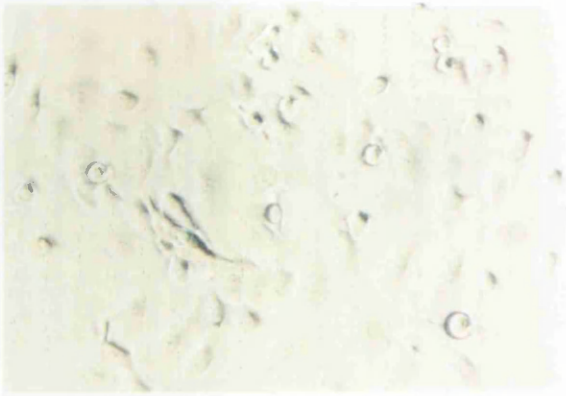
(a) Transfected C5N cell populations stained with X-gal for β-galactosidase activity. **(i)** Act-LacZ$^{wild\text{-}type}$. **(ii)** Act-LacZ^{A}. **(iii)**Act-LacZ^{C}. **(iv)** Act-LacZ^{E}. **(b)** Table of β-galactosidase activity in transient and stable transfections of C5N cells with the Act-LacZ constructs. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. Transiently transfected cells were stained two days after transfection. Pools of stably transfected cells were stained after four weeks selection in G418 when sufficient number of selected cells had been obtained. Flasks were screened under a light microscope. Positive indicates that a high proportion (>1%) of cells stained strongly with X-gal. Negative indicates no staining cells were seen.

(a)

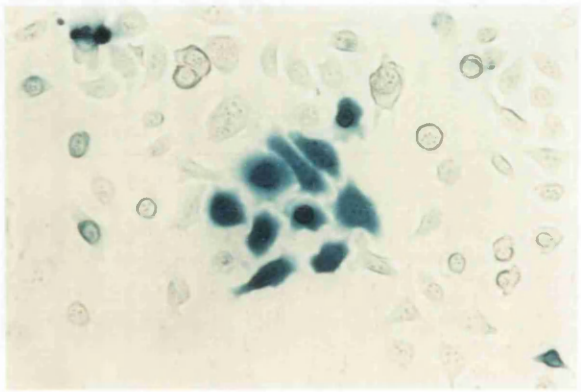
(i) Act-Lac Z$^{wild\text{-}type}$



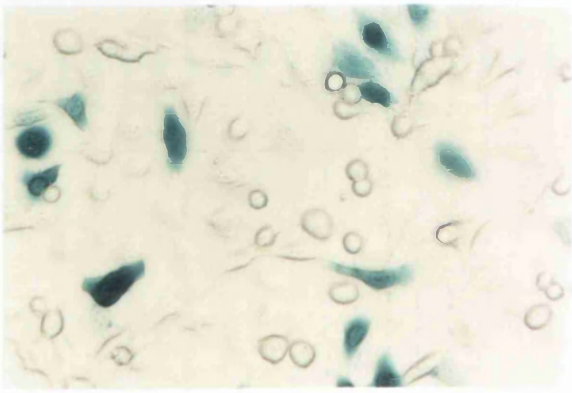
(ii) Act-Lac Z^{A}



(iii) Act-Lac Z^{C}



(iv) Act-Lac Z^{E}



(b)

Construct transfected.	Transient C5N.	Stable C5N.
pIrv-Neo-Act	Negative	Negative
Act-LacZ ^{pMC1871}	Negative	Negative
Act-LacZ $^{wild\text{-}type}$	Positive	Positive
Act-LacZ ^{A}	Negative	Negative
Act-LacZ ^{B}	Positive	Positive
Act-LacZ ^{C}	Positive	Positive
Act-LacZ ^{D}	Positive	Positive
Act-LacZ ^{E}	Positive	Positive

activity in the epidermal cells. The level of β-galactosidase activity was rarely as high in epidermal cells as it was in the fibroblasts.

When the marker construct Act-LacZ^{A} was transfected into epidermal cell lines no expression of the *lac Z* gene was observed, confirming the result obtained in the fibroblasts. Transfection of the control constructs Act-LacZ^{B} and Act-LacZ^{C} generated high levels of expression in the epidermal cells. In accordance with the transfection of the positive control Act-LacZ$^{wild\text{-}type}$ construct the expression observed in the epidermal cells was markedly lower than the expression observed in the fibroblast cell transfectants.

This confirmed the conclusions drawn from the transfections into NIH 3T3 cells. The test construct did not express while the control construct produced β-galactosidase expression. The next step is to test this hypothesis and determine whether the mutagen MNNG is able to activate the test construct Act-LacZ^{A} in epidermal cells *in vitro*. In order to achieve this a population of stable transfectants is required.

Selection of clones of stable transfectants should ensure that a high proportion of the cells express the RNA for the *lac Z* gene and would be available as a target population for the activation of the marker gene by MNNG. The levels of *lac Z* RNA expression in the stable transfectants could be determined and clones with high expression of *lac Z* RNA could be used as a target for reactivation by chemical mutagens. It would be helpful if the stably transfected cell population contained only one construct, this would allow a reasonable estimate of the mutation frequency to be calculated. If the cells contained more than one construct it would be impossible to tell how many of the transfected vectors were expressed and therefore acting as a target for reactivation by the mutagen.

The results of the transfections with the Act-LacZ^{D} and Act-LacZ^{E} constructs were also in accordance with the results obtained by transfection of these constructs into fibroblast cells. High levels of expression were obtained with both the test and control constructs. The expression from the oligonucleotide in the test Act-LacZ^{D} construct which does not have a in-frame initiation codon was unexpected. In order to test if the observed expression of the β-galactosidase protein is due to mutation or rearrangement in a small proportion of constructs it was decided to use a transfection technique which is known to introduce only one or a few constructs into each cell. If each cell contains only one construct and then only a small number of cell should express the protein.

The transfected C5N cell populations were cultured in G418 to select cells which are stably transfected. This was done to confirm that the expression patterns observed in the transient transfections were maintained, and not artefacts caused by the transfection efficiency of the various constructs. After the culture period of three weeks a control population of cells transfected with carrier DNA alone had been completely killed by the G418. Approximately the same number of cell colonies were obtained in the flasks from each of the transfected cell population. This suggested that the transfection efficiencies had been similar in all of the transfections. This also demonstrated that the product of the *neo* gene was expressed in the transfected cells and could be utilised for selecting transfected cell populations.

After three weeks of selection the resultant transfected cell clones were pooled. The cells were trypsinised and replated at high density into F75 flasks. After two days the cells were fixed and stained with X-gal to look at the activity of the β-galactosidase protein. The expression observed in the stably transfected epidermal cell populations mirrored that observed in the transient transfections of both the fibroblast and epidermal cell lines. No expression was detected in the C5N epidermal cells stably transfected with the negative control constructs pIrv-Neo-Act and Act-LacZ^{pMC1871}, as well as the MNNG-test oligo fusion construct Act-LacZ^{A}. High levels of expression were observed in the cells transfected with the positive control construct Act-LacZ^{pMC1871} and the MNNG-strategy control constructs Act-LacZ^{B} and Act-LacZ^{C}. In addition both the test and control constructs of the AAF-strategy fusion genes Act-LacZ^{D} and Act-LacZ^{E} were also observed to express β-galactosidase protein at high levels in these cells. This confirmed the results obtained from the transient transfections of both the fibroblast and keratinocyte cell lines.

3.3.3 Production of retroviruses by $CaPO_4$ transfection of Ψ2 cells with Act-LacZ fusion genes.

Retroviruses were produced from all of the Act-LacZ fusion genes by transfection of the murine embryonic fibroblast viral producer cell line, Ψ2. This was done principally to test the AAF-test oligonucleotide construct Act-LacZ^{D}. Expression of β-galactosidase protein is obtained unexpectedly in cells transfected with this construct. The oligonucleotide in this gene codes for an ATG initiation codon which is 2 bp out of frame with the coding sequence of the *lac Z* gene and should therefore not initiate translation of the protein.

The *lac Z* marker gene in the Act-LacZ^{D} construct was generated from an oligonucleotide sequence which proved to be highly unstable when cloned in bacteria (see Section 3.2.4). A mutation or rearrangement in the sequence of this oligonucleotide in a small proportion of the constructs transfected into cells by the $CaPO_4$ technique may cause expression in a significantly higher number of transfected cells. If the erroneous expression is caused by rearrangement or mutation in a small proportion of the constructs then, if each cell only contains one or a few transfected constructs, only a small proportion of the cells should express the β-galactosidase protein.

The Ψ2 cell line is a murine embryonic fibroblast line which contains a retrovirus with a DNA packaging deficiency. It contains all the elements necessary for the viral life cycle except that it lacks the correct sequences necessary for the packaging of the viral RNA genome into the viral particles. When a DNA construct is introduced into the Ψ2 cells which includes the appropriate packaging sequences, the introduced construct is packaged into the viral particles in place of the endogenous viral genome. The Act-LacZ fusion genes are based on a retroviral vector which has the appropriate packaging sequences and when introduced into Ψ2 cell the constructs will be incorporated into infectious viral particles.

Retroviruses were produced from each of the Act-LacZ fusion genes by transfection of the Ψ2 cells by the $CaPO_4$ techniques (described in Section 2.12.2). The Ψ2 cells are cultured in under the same conditions and using the same medium as the NIH 3T3 cells (described in section 2.11). A proportion of the transfected cells were examined for β-galactosidase activity by staining with X-gal (Figure 17). After G418 selection the surviving cells in the flasks transfected with the Act-LacZ fusion gene constructs were growing in discrete colonies. The transfected colonies were pooled and used as a source of retroviruses to infect murine epidermal cells as described below.

3.3.4 Infection of a murine keratinocyte cell line with retroviruses.

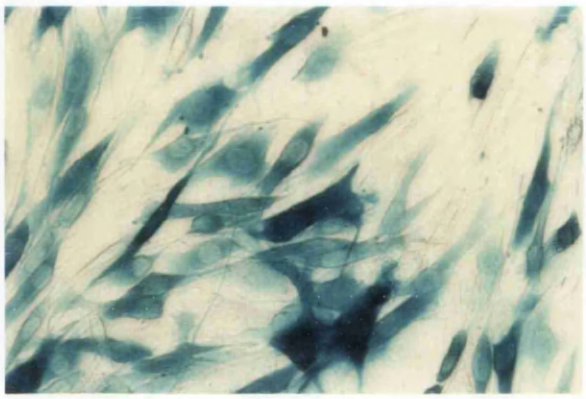
Retroviruses containing the Act-LacZ fusion genes were used to infect the murine epidermal cell line C50. The retroviruses were produced from pools of $CaPO_4$ transfected Ψ2 cells (described above in Section 3.3.3 and Figure 17) and used to infect the keratinocyte cell line as described in Sections 2.13 and 2.14. Pool of Ψ2 transfectants were used to generate retroviruses for each of the Act-LacZ fusion genes.

Figure 17. Transfection of Act-LacZ constructs into Ψ2 cells.

(a) Transfected Ψ2 cell populations stained with X-gal for β-galactosidase activity. **(i)** Act-LacZ$^{wild\text{-}type}$. **(ii)** Act-LacZ^{A}. **(iii)**Act-LacZ^{C}. **(iv)** Act-LacZ^{D}. **(b)** Table of β-galactosidase activity in transient and stable transfections of Ψ2 cells with the Act-LacZ constructs. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. Transiently transfected cells were stained two days after transfection. Pools of stably transfected cells were stained after four weeks selection in G418 when sufficient number of selected cells had been obtained. Flasks were screened under a light microscope. Positive indicates that a high proportion (>1%) of cells stained strongly with X-gal. Negative indicates no staining cells were seen.

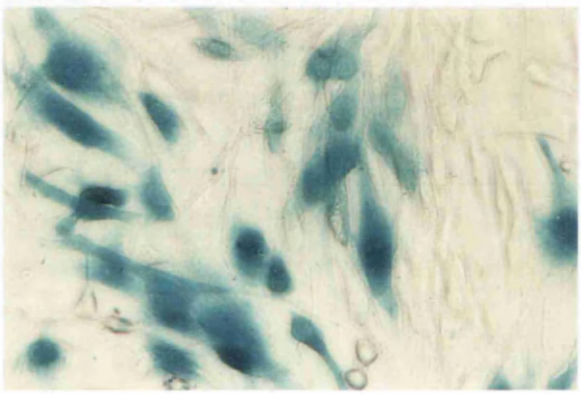
(a)

(i) Act-Lac Z$^{wild\text{-}type}$

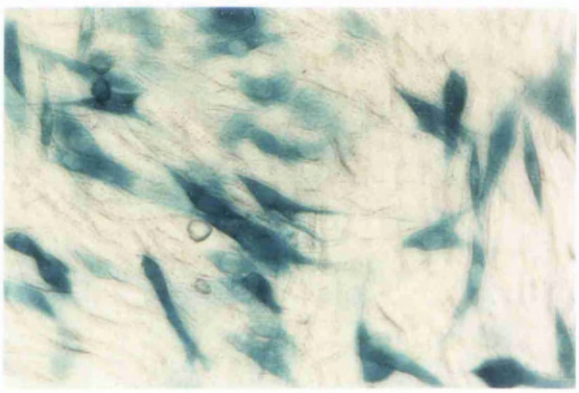


(ii) Act-Lac Z^{A}

(iii) Act-Lac Z^{C}



(iv) Act-Lac Z^{D}



(b)

Construct transfected.	Transient Ψ2.	Stable Ψ2.
pIrv-Neo-Act	Negative	Negative
Act-LacZ^{pMC1871}	Negative	Negative
Act-LacZ$^{wild\text{-}type}$	Positive	Positive
Act-LacZ^{A}	Negative	Negative
Act-LacZ^{B}	Positive	Positive
Act-LacZ^{C}	Positive	Positive
Act-LacZ^{D}	Positive	Positive
Act-LacZ^{E}	Positive	Positive

Two days after infection the cells were fixed and stained for β-galactosidase activity with X-gal. The results from these infection experiments are shown in Figure 18.

Again the negative control construct Act-LacZ^{pMC1871} showed no expression of β-galactosidase expression, while in the positive control cell population infected with the pIrv-Neo-Act-LacZ$^{wild\text{-}type}$ construct a high level of β-galactosidase expression was observed. Again the expression of the β-galactosidase protein produced from the MNNG-strategy test and control constructs was as expected. Cells infected with the test oligo construct Act-LacZ^{A} showed no sign of expression of the *lac Z* gene whereas cells infected with MNNG oligo control constructs Act-LacZ^{B} and Act-LacZ^{C} demonstrated consistently high levels of expression comparable to that observed in the positive control infections performed with the retrovirus containing the wild-type *lac Z* gene.

Disappointingly the expression of the AAF-strategy constructs in cells infected with retroviruses containing the LacZ^{D} and LacZ^{E} fusion genes also mirrored the results seen in previous transfection experiments. High levels of expression were seen in cell infected with retroviruses derived from both the Act-LacZ^{D} and the Act-LacZ^{E} constructs. The expression levels of the β-galactosidase protein observed in the cells infected with the construct was very similar when compared to each other and comparable to the positive control. More than 50% of the cells express the gene. These cells have not been selected and the remaining 50% of the cells which do not express the gene may not be infected. As the cells contain relatively few copies of the construct and a high number of the cells express β-galactosidase protein most of the Act-LacZ^{D} constructs must be expressing a functional enzyme. As the majority of the Act-LacZ^{D} constructs express the β-galactosidase protein the explanation of the aberrant expression of β-galactosidase from the oligo D/*lac Z* fusion gene is very unlikely to be mutation or rearrangement of the oligonucleotide.

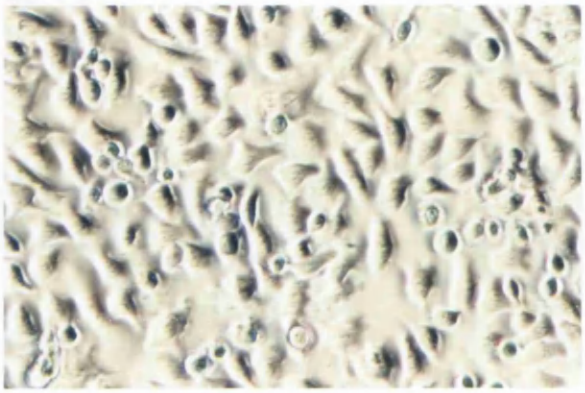
After this result the oligonucleotide in the Act-LacZ^{D} construct was sequenced again to check its integrity and ensure that it had not suffered a mutation resulting in the observed aberrant β-galactosidase expression. The sequencing revealed that the oligonucleotide D sequence is still intact but revealed that an ATG sequence is located in the 3' end of the β-actin promoter about 61 base pairs upstream of the oligonucleotide. This ATG sequence is in frame with the coding sequence of the *lac Z* gene and there is an open reading frame stretching from this ATG sequence through the oligonucleotide to the coding region of the *lac Z* gene. This ATG can probably initiate translation and adds about 27 amino acid residues onto the amino terminus of the β-galactosidase. This

Figure 18. Infection of C50 with Act-LacZ retroviruses.

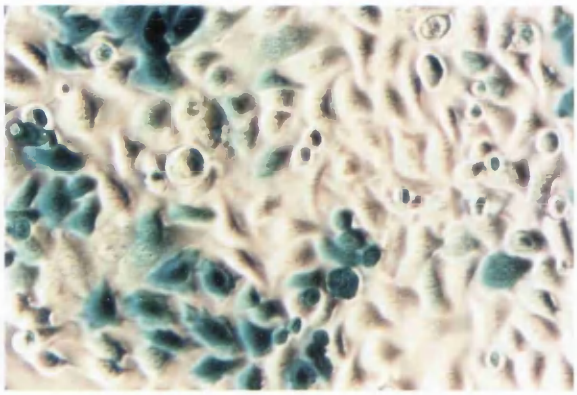
(a) Infected C5O cell populations stained with X-gal for β-galactosidase activity. **(i)** Act-LacZ^{A}. **(ii)** Act-LacZ^{B}. **(iii)**Act-LacZ^{D}. **(iv)** Act-LacZ^{E}. **(b)** Table of β-galactosidase activity in virally infected C5O cells with the Act-LacZ constructs. After infection cells were selected in G418. Pools of stably transfected cells were stained after four weeks selection in G418 when sufficient number of selected cells had been obtained. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. Flasks were screened under a light microscope. Positive indicates that a high proportion (>1%) of cells stained strongly with X-gal. Negative indicates no staining cells were seen.

(a)

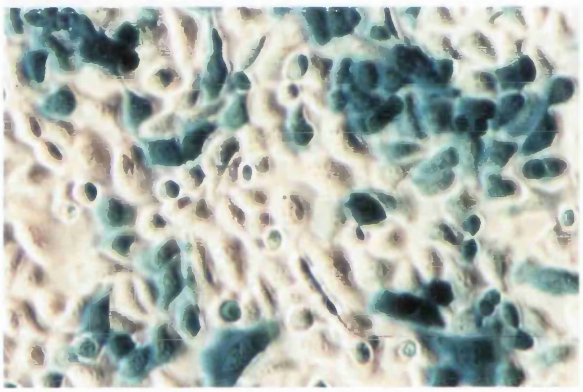
(i) Act-Lac Z^{A}



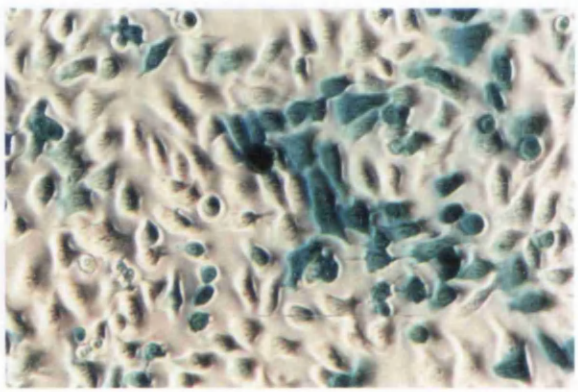
(ii) Act-Lac Z^{B}



(iii) Act-Lac Z^{D}



(iv) Act-Lac Z^{E}



(b)

Virus preparation.	**Expression in infected C5O cells.**
pIrv-Neo-Act	Negative
Act-LacZ^{pMC1871}	Negative
Act-LacZ$^{wild\text{-}type}$	Positive
Act-LacZ^{A}	Negative
Act-LacZ^{B}	Positive
Act-LacZ^{C}	Positive
Act-LacZ^{D}	Positive
Act-LacZ^{E}	Positive

probably explains the aberrant expression of the β-galactosidase protein from the Act-LacZ^{D} construct.

This upstream ATG only interfered with the expression of the Act-LacZ^{D} construct because it is only in-frame with the *lac Z* coding region if a 26 bp oligonucleotide is inserted in front of the *lac Z* gene. The oligonucleotides for the MNNG-test and control construct are 30 base pairs long and the AAF-control oligonucleotide in the LacZ^{E} fusion gene is 24 base pairs long and hence the upstream ATG sequence is not in-frame with the *lac Z* gene in these cases.

This problem was circumvented simply by recloning the LacZ^{D} gene in front of another promoter. Both the test and control fusion genes LacZ^{D} and the LacZ^{E} were recloned into the pIrv-Neo-K5 expression vector (described in Sections 3.2.6 and 3.2.7). The new constructs were then transiently transfected into the murine epidermal cell line C5O by the $CaPO_4$ technique. Cells transfected with the K5-LacZ^{E} construct demonstrated a high level of expression of the β-galactosidase protein. This construct contains the AAF-strategy control oligo/*lac Z* fusion gene. This demonstrated that the 1.4 kb fragment of the *BK III* promoter was able to support expression of the *lac Z* gene in the murine epidermal cell at a high levels. Transfections with the negative control construct Act-LacZ^{pMC1871} and the AAF-strategy test fusion gene construct K5-LacZ^{D} both resulted in no observed expression of the *lac Z* gene when the cells were stained for β-galactosidase activity.

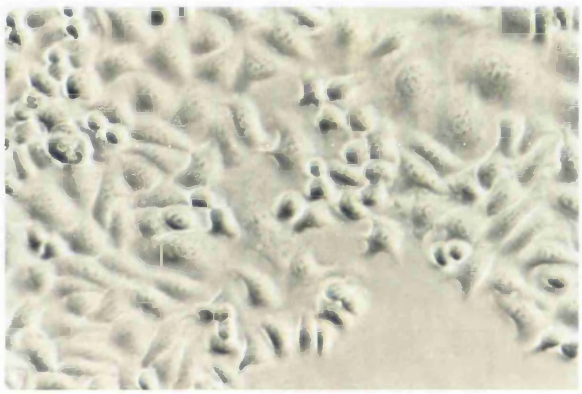
The transfections with the K5-LacZ fusion gene constructs indicated that the AAF-strategy test fusion gene may in fact be useful for cell marking as it does not itself initiate the translation of the β-galactosidase protein. This result also suggests that the expression of the lac Z gene observed in the cells transfected with the Act-LacZ^{D} construct was probably caused by the upstream in-frame ATG sequence which was discovered in the β-actin promoter sequences. The next step in the in testing the AAF-fusion strategy constructs is to generate stably transfected cell lines which can be shown to express the mRNA from the *lac Z* gene but not the β-galactosidase protein. Such cell lines should allow reactivation of the fusion genes to be tested *in vitro*.

Figure 19. Expression of β-galactosidase protein in Act-LacZ^A electroporation ringclones.

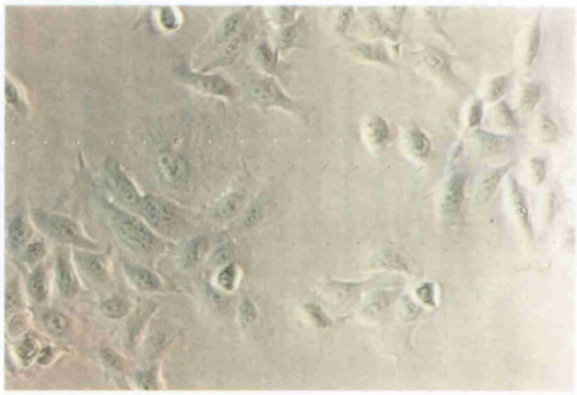
(a) Electroporated C5O cell ringclones stained with X-gal for β-galactosidase activity. **(i)** Act-LacZ^A clone 5. **(ii)** Act-LacZ^A clone 6. **(b)** Table of β-galactosidase activity in the Act-LacZ^A C5O electroporation ringclones. Transfected cells were selected in G418 for three weeks until discrete colonies were obtained. The colonies were then ringcloned and grown for a further week until sufficient numbers had been obtained to use in experimental proceedures. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. The percentage figure given above represents an estimate of the proportion of staining cells based on screening 5 x 10^5 cells in a 25cm^2 flask under an inverted light microscope.

(a)

(i) Act-Lac Z^A clone 5



(ii) Act-Lac Z^A clone 6



(b)

Clone	Expression of β-galactosidase.
A1	0.00%
A3	0.00%
A4	0.00%
A5	< 0.01%
A6	0.00%
A7	0.00%
A8	0.00%
A11	0.00%
A12	0.00%
A21	0.00%
A22	0.00%
A24	0.00%

3.3.5 Electroporation of Act-LacZ and K5-LacZ fusion genes into a murine keratinocyte cell line.

Electroporation of constructs into epidermal cell lines was performed in order to generate stably transfected cell lines for testing the reactivation of the marker genes. Electroporation was chosen as the transfection method to introduce constructs into the murine epidermal cell line C5O as it is known to introduce a small number of constructs into each transfected cell. This method was chosen in preference to the generation of retroviruses as it eliminates the need for the intermediate transfection of Ψ2 cells.

C5O is a murine epidermal cell line very similar to the C5N cell line used for the transient transfections described above in section 3.3.2. It is an untransformed cell line which is not tumorigenic when injected into nude mice *in vivo*. As such it represents a good model for the normal epidermal cell *in vitro*, the lineage of which the marker genes were designed to analyse. The C5O cell line was chosen in preference to the C5N cell line as previous work performed in the laboratory had shown that these cells were easily transfectable by the electroporation technique.

The constructs Act-LacZ^{A}, K5-LacZ^{D} and K5-LacZ^{E} were all electroporated into C5O cells and used to generate stably-transfected cell lines. This was done according to the methods described in Section 2.12.3. The electroporated cells containing the positive and negative control constructs Act-LacZ$^{wild\text{-}type}$ and Act-LacZ^{pMC1871} were fixed and stained after two days. The cells electroporated with the positive control constructs contained cell expressing the β-galactosidase protein at high levels but those which expressed the negative control construct Act-LacZ^{pMC1871} did not show β-galactosidase expression in the transfected cells.

After growth for 3 weeks in the selective medium the transfected cell were growing is discrete visible colonies on the bottom of the tissue culture flask. Individual colonies were separated by ringcloning. The cell lines were maintained in selective medium containing G418 throughout this process. A number of the separate cell lines which contained the constructs Act-LacZ^{A}, K5-LacZ^{D} and K5-LacZ^{E} were selected for analysis of marker gene expression.

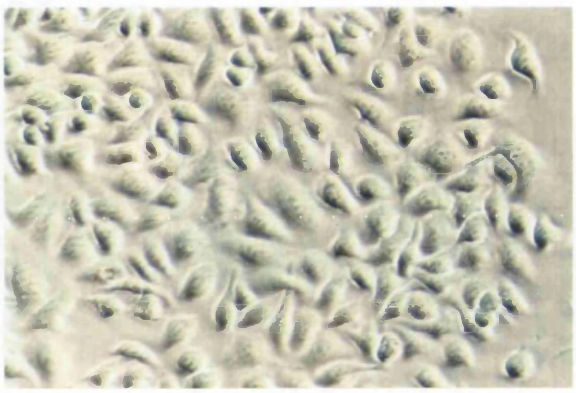
All of the ringcloned cell lines were grown up and analysed for the expression of the β-galactosidase protein by staining with X-gal. The cell lines ringcloned from the population transfected with the Act-LacZ^{A} construct showed very little expression of the

Figure 20. Expression of β-galactosidase protein in K5-LacZ^D electroporation ringclones.

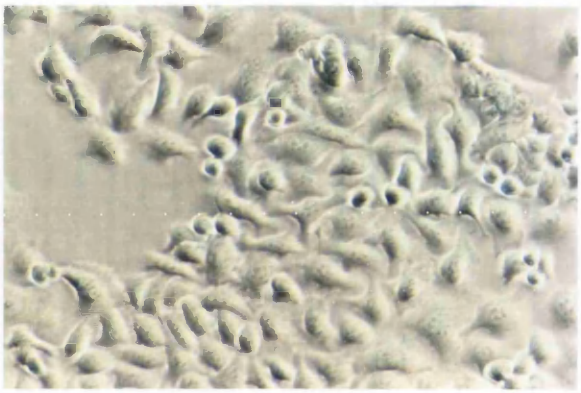
(a) Electroporated C5O cell ringclones stained with X-gal for β-galactosidase activity. **(i)** K5-LacZ^D clone 2. **(ii)** K5-LacZ^D clone 11. **(b)** Table of β-galactosidase activity in the Act-LacZ^A C5O electroporation ringclones. Transfected cells were selected in G418 for three weeks until discrete colonies were obtained. The colonies were then ringcloned and grown for a further week until sufficient numbers had been obtained to use in experimental proceedures. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. The percentage figure given above represents an estimate of the proportion of staining cells based on screening 5 x 10^5 cells in a 25cm^2 flask under an inverted light microscope.

(a)

(i) K5-Lac Z^D clone 2



(ii) K5-Lac Z^D clone 11



(b)

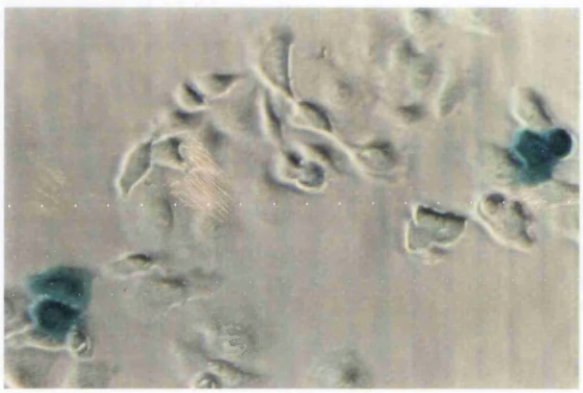
Ringclone.	Percentage of lac Z positive cells
D1	0%
D2	0%
D3	0.10%
D4	0%
D5	0.10%
D6	0.50%
D7	0%
D8	0%
D9	0%
D10	30%
D11	0%
D12	0%
D13	0%
D14	0%
D15	0%
D16	0%
D17	0%
D18	20%

Figure 21. Expression of β-galactosidase protein in 'K5-LacZ^E electroporation ringclones.

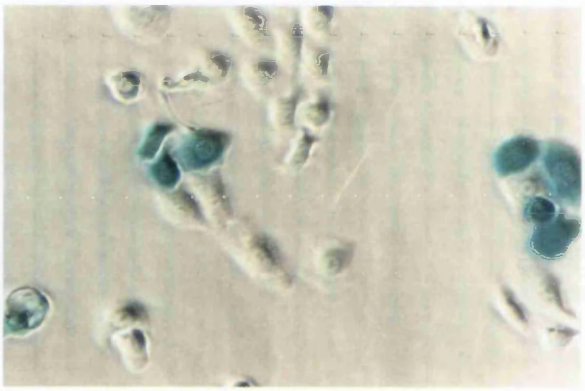
(a) Electroporated C5O cell ringclones stained with X-gal for β-galactosidase activity. **(i)** K5-LacZ^E clone 1. **(ii)** K5-LacZ^E clone 2. **(iii)** K5-LacZ^E clone 4. **(iv)** K5-LacZ^E clone 6. **(b)** Table of β-galactosidase activity in the Act-LacZ^A C5O electroporation ringclones. Transfected cells were selected in G418 for three weeks until discrete colonies were obtained. The colonies were then ringcloned and grown for a further week until sufficient numbers had been obtained to use in experimental proceedures. Cells were stained in 25cm^2 flasks according to the method outlined in Section 2.15. The percentage figure given above represents an estimate of the proportion of staining cells based on screening 5 x 10^5 cells in a 25cm^2 flask under an inverted light microscope.

(a)

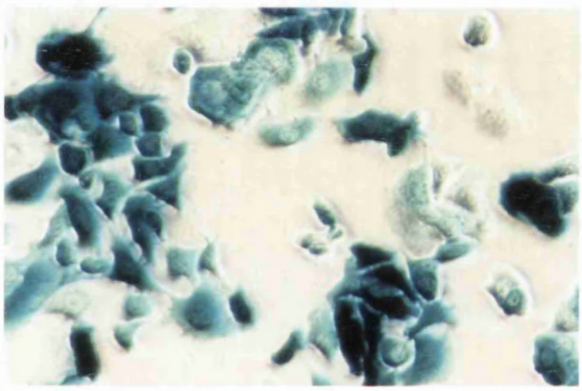
(i) K5-Lac Z^E clone 1



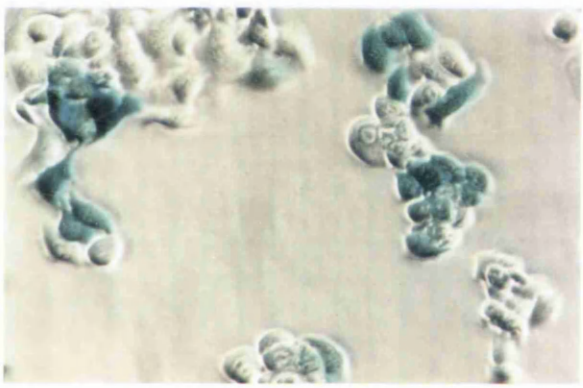
(ii) K5-Lac Z^E clone 2



(iii) K5-Lac Z^E clone 4



(iv) K5-Lac Z^E clone 6



(b)

Clone	Percentage of X-gal staining cells
E1	5.00%
E2	20.00%
E3	0.00%
E4	95.00%
E5	0.00%
E6	25.00%

lac Z gene (Figure 19). Only one clone of the transfected cells (clone 5) showed any expression of the *lac Z* gene at all. This expression was attributed to mutation arising during the transfection of the construct into the C5O cells. A similar pattern of expression was observed in the cells transfected with the K5-LacZ^{D} construct. Only 2 of 18 transfected cell ringclones showed expression of the β-galactosidase protein in a proportion of the cells greater than 1% of the total (Figure 20). In contrast most of the K5-LacZ^{E} expressed the β-galactosidase protein at high levels (Figure 21).

3.3.6 Southern blot analysis of the Act-LacZ ^{A} electroporation ringclones.

Twelve ringcloned cell lines containing the Act-LacZ^{A} introduced by electroporation (as described in Section 3.3.5) were selected and grown up for analysis. Cells from each of the ringcloned cell lines were grown to confluence in duplicate F175 tissue culture flasks. One of the flasks was used for the isolation of DNA from the individual ringclones, the other was used for the isolation of RNA from the ringclones. DNA was prepared from the cell lines by the method detailed in Section 2.17. After isolation the DNA from the different Act-LacZ^{A} cell lines was analysed by Southern blot. The results of the Southern blot analysis are shown in Figure 22 (described in Section 2.19). The blots were probed with a 1 kb fragment of the *lac Z* gene.

Twelve separate ringcloned cell lines, transfected with the Act-LacZ^{A} construct, were analysed and bands were detected which corresponded to the sequences from the *lac Z* gene. This demonstrated that the transfection had worked and that copies of the construct had integrated into the genome of the electroporated cells. Of the 12 clones analysed only 9 had detectable fragments of the *lac Z* gene and only 6 of these gave strong signals. Only four of the six electroporation cell clones with strong bands gave the correct size of fragment expected for the entire *lac Z* gene. Two of these clones (A4 and A5 in Figure 22b) contained a 9.6 kb band equivalent to the size of the whole plasmid when digested with *Hind III*. This enzyme cuts only once within the plasmid sequence and it is likely therefore that these clones contain more than one copy of the plasmid integrated in a tandem array. When this blot was repeated with the same DNA samples and probed with a fragment of the *neo* gene all the clones hybridised strongly to the probe (Figure 23). This demonstrates that both the transfection and the selection steps of the protocol had worked.

Figure 22. Southern blot analysis of Act-LacZ^A electroporation ringclones.

(a) Southern blot of *Bam HI* digests of DNA samples from the Act-LacZ^A electroporation ringclones, probed with a 1 kb fragment of the *lac Z* gene. Only 4 of the twelve ringclones contain the expected 3.1 kb band. **(b)** Southern blot of Act-LacZ^A plasmid (P) and Act-LacZ^A cell line DNA (A3-6) probed with a 1 kb fragment of the *lac Z* gene. The DNA samples have been digested with both *Bam HI* and *Hind III*. Clones A4 and A5 contain a 9.6 kb band upon digestion with Hind III which suggests that more than one copy of the gene is integrated.

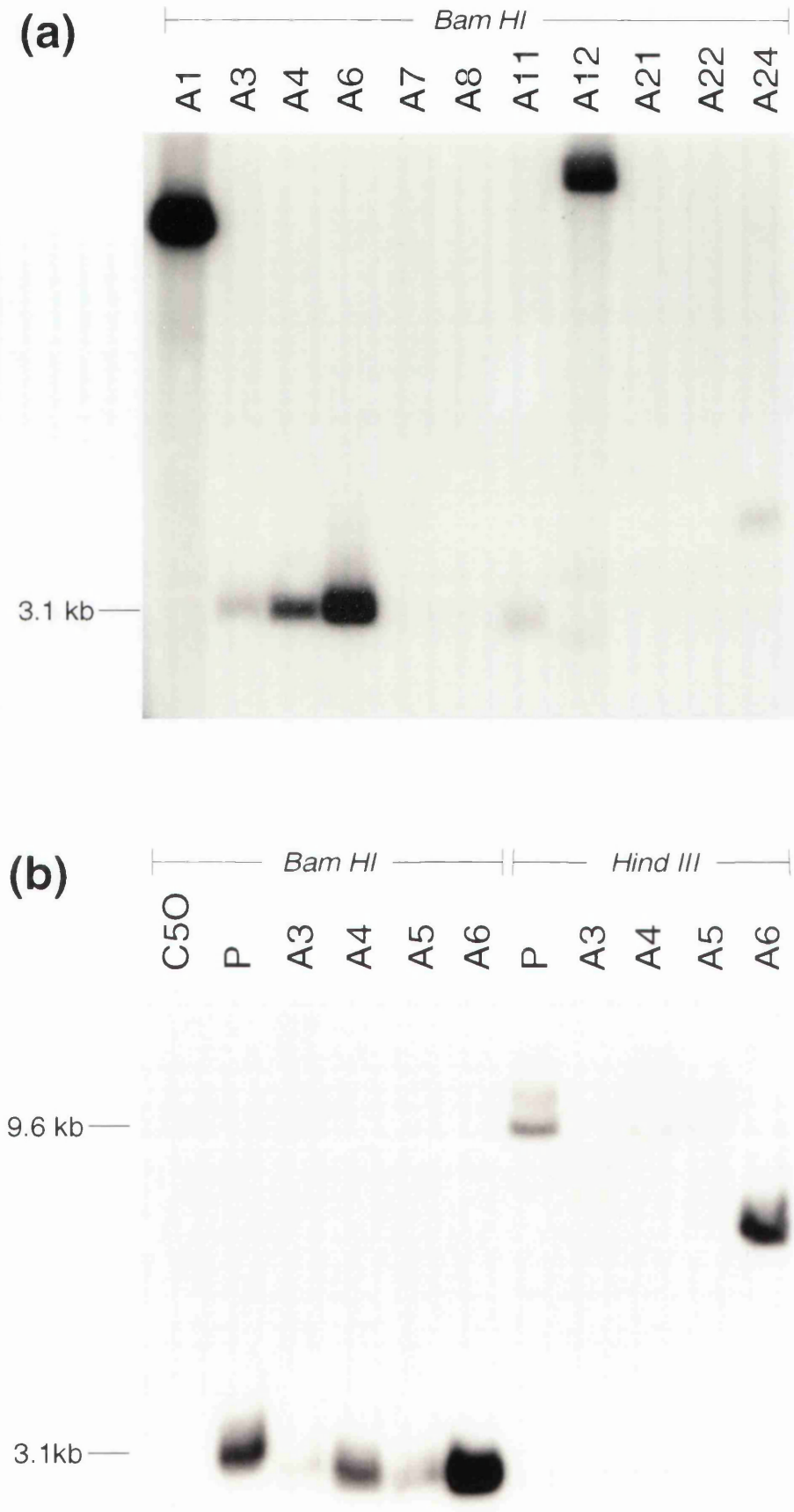
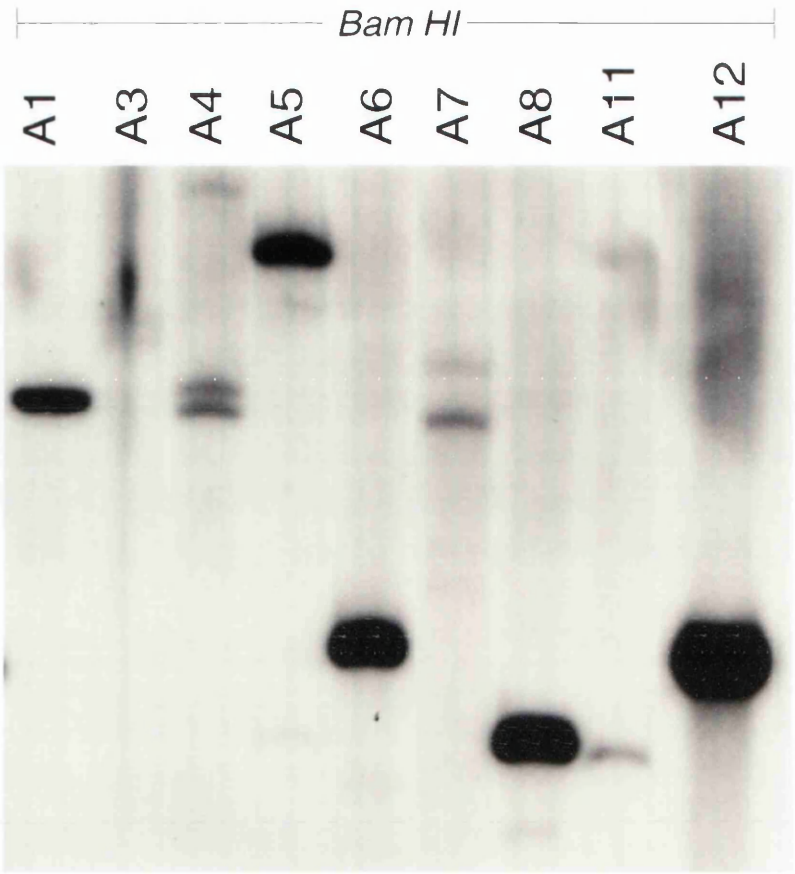

(a)
Bam HI
A1
A3
A4
A6
A7
A8
A11
A12
A21
A22
A24
3.1 kb
(b)
Bam HI
Hind III
C5O
P
A3
A4
A5
A6
P
A3
A4
A5
A6
9.6 kb
3.1kb

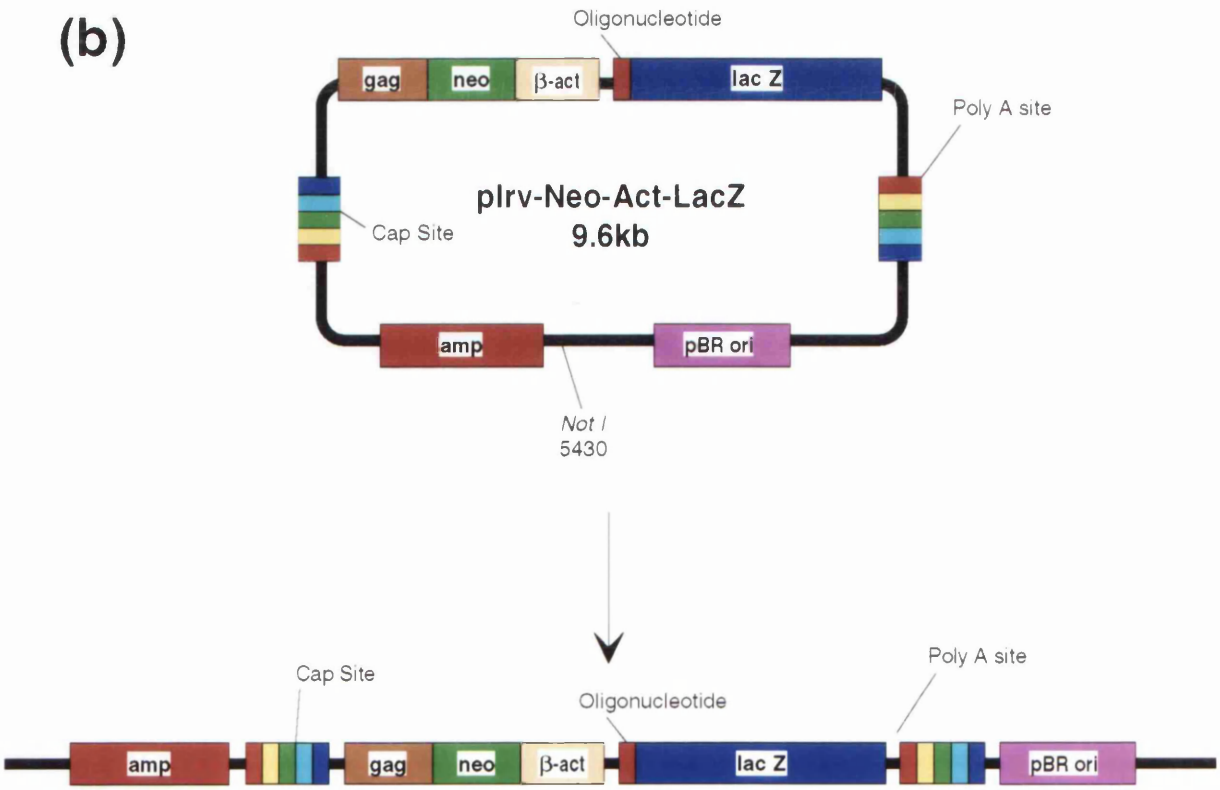
Figure 23. Presence of the *neo* gene in the Act-LacZ^{A} electroporation ringclones.

(a) Southern blot of Act-LacZ^{A} cell line DNA samples digested with *Bam HI* and probed with a 1 kb fragment of the *neo* gene. All the cell lines contain the *neo* gene, some which did not have strong bands when probed with *lac Z* have strong bands when probed with *neo* suggesting that they have lost the *lac Z* gene. **(b)** Diagram of construct used for electroporation. Plasmid was linearised by digestion at the Not I site and electroporated into C5O cells. The same strategy was used for both the Act and K5 vectors.

(a)


Bam HI
A1
A3
A4
A5
A6
A7
A8
A11
A12

(b)


Oligonucleotide
gag
neo
β-act
lac Z
Poly A site
pIrv-Neo-Act-LacZ
9.6kb
Cap Site
amp
pBR ori
Not I
5430
Cap Site
Oligonucleotide
Poly A site
amp
gag
neo
β-act
lac Z
pBR ori

It seems surprising that although that all the transfected cell lines contained the *neo* resistance gene that more of them do not hybridise strongly to the *lac Z* sequences, especially as the two gene are located on the same DNA construct. It is obvious that some clones (*e.g.* A7, A8 and A11) hybridise very strongly with the *neo* probe but not with the *lac Z* probe. This suggests that there is rearrangement of the *lac Z* gene going on after electroporation into the cells.

These results were confirmed by PCR (polymerase chain reaction) analysis of the Act-LacZ^{A} electroporation ringclone DNA samples with the LacZA and LacZB amplimers (described in Sections 2.23.1 and 2.23.3). Bands corresponding to the presence of *lac Z* DNA were identified in only six of the twelve clones. These clones were the same six which hybridised most strongly with the *lac Z* probe when the DNA samples were analysed by Southern blot. The reason for this rearrangement of the *lac Z* gene remains unknown.

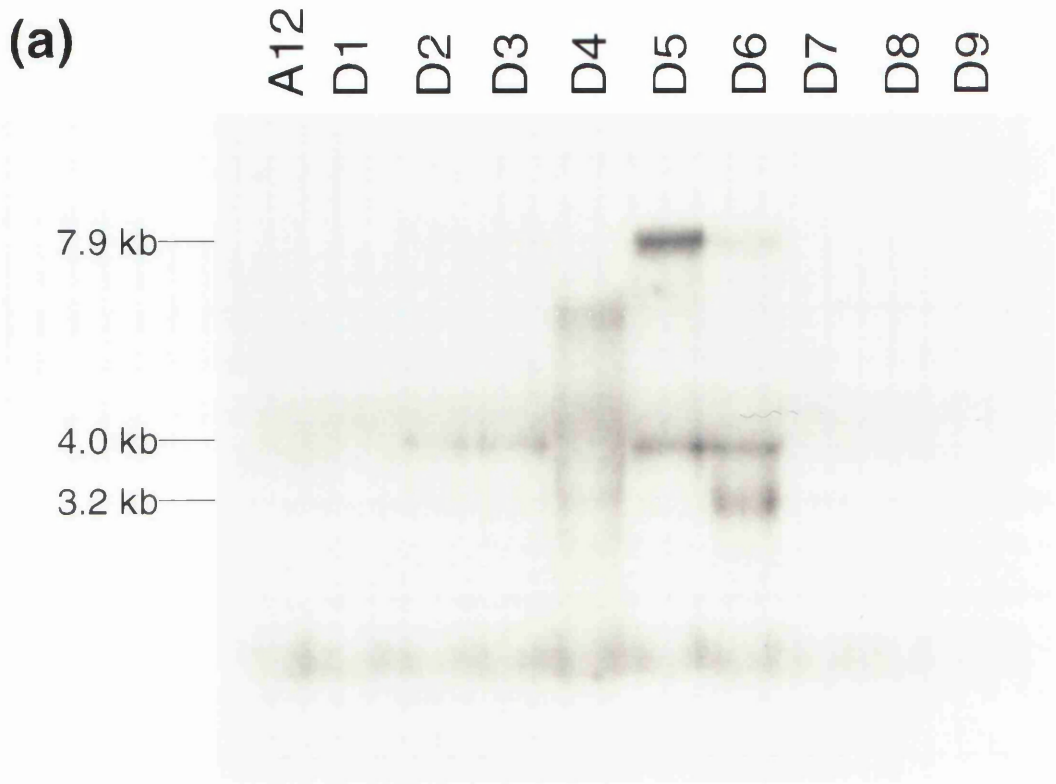
3.3.7 Measurement of expression of *lac Z* fusion gene levels in Act-Lac Z and K5-LacZ electroporation ringclones.

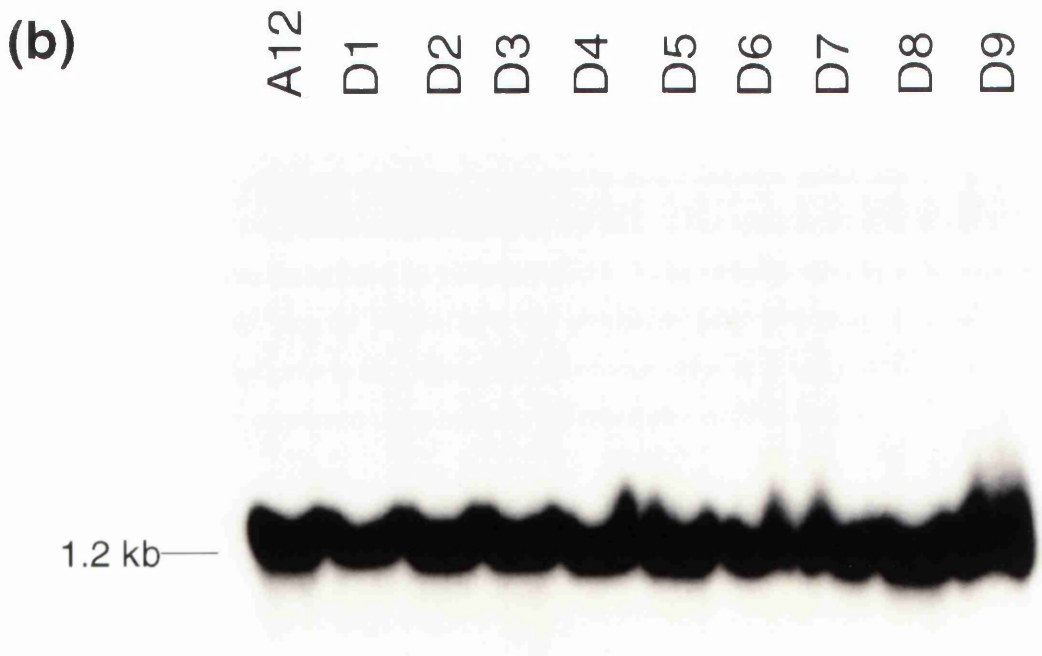
It was important to analyse the RNA and protein levels of the β-galactosidase enzyme in order to identify cell populations in which the activation of the oligo/*lac Z* fusion genes by chemical mutagens could be tested. The levels of RNA in these cell lines were analysed by three different methods: northern blots: the polymerase chain reaction and ribonuclease protection. The levels of protein were analysed by simply staining the cells with the chromogenic β-galactosidase substrate, X-gal. Cell lines transfected by electroporation with the Act-LacZ^{A}, K5-LacZ^{D} and K5-LacZ^{E} were all analysed for expression of *lac Z* RNA and protein.

The Act-LacZ^{A}, K5-LacZ^{D} and K5-LacZ^{E} cell lines which had already been analysed for the presence of the β-galactosidase protein were subsequently analysed for expression of the *lac Z* RNA. The levels of RNA were initially measured by northern blot analysis. Figure 24a shows a northern blot of RNA from 9 of the K5-LacZ^{D} cell lines. Although expression is detectable the levels of expression from the *lac Z* under the control of the K5 promoter were low. The expression of the *lac Z* gene from the β-actin promoter was even lower. The blot was reprobed with a probe for the ubiquitous housekeeping enzyme probe, *gapdh* as a loading control (Figure 24b).

Figure 24. Northern blot analysis of C5O electroporation ringclones.

(a) Northern blot of RNA derived from Act-LacZ^{A} and K5-LacZ^{D} electroporation ringclone cell lines. Bands corresponding to approximately 7.9 and 4.0 kb are visible in the cell lines D2, D3, D5 and D6. The 7.9 kb band corresponds to a transcript of the entire viral genome while the 4.0 kb band corresponds to expression of the *lac Z* gene from the internal K5 promoter. Exposed for 3 days at -70°C. **(b)** Reprobe of the above blot with the *gapdh* as a loading control. Exposed for 5 hours at - 70°C.


(a)
A12
D1
D2
D3
D4
D5
D6
D7
D8
D9
7.9 kb
4.0 kb
3.2 kb


(b)
A12
D1
D2
D3
D4
D5
D6
D7
D8
D9
1.2 kb

Only five of the nine K5-LacZ^{D} clones showed any sign of activity on the northern blot. An 8.0 kb band which probably corresponds to the whole virus is visible in lanes D2, D3, D5 and D6. In these lanes there is also a 3.9 kb band which probably correspond to expression of the *lac Z* gene from the internal K5 promoter. When this blot was reprobed with the a fragment of the *neo* gene the 8.0 kb band was visible but not the 3.9 kb band. This supports the hypothesis that the 8.0 kb band corresponds to a transcript of the whole virus.

As the northern analysis was not conclusive, as the levels of expression were extremely low it was decided to use PCR analysis of the RNA. This technique was performed according to the method detailed in Section 2.23.2. Basically this method involved initially generating single-stranded cDNA from RNA by reverse transcription. The downstream PCR oligonucleotide was used as a specific primer for this reaction. This cDNA was then amplified by PCR and produced a band if the sequence of interest was present in the cell RNA. A control to measure any DNA present in the RNA present in the sample was run at the same time. The LacZA/LacZB oligonucleotide primer pair were used to amplify the cDNA produced by reverse transcription of the RNA in the Act-LacZ^{A} C50 electroporation ringclones.

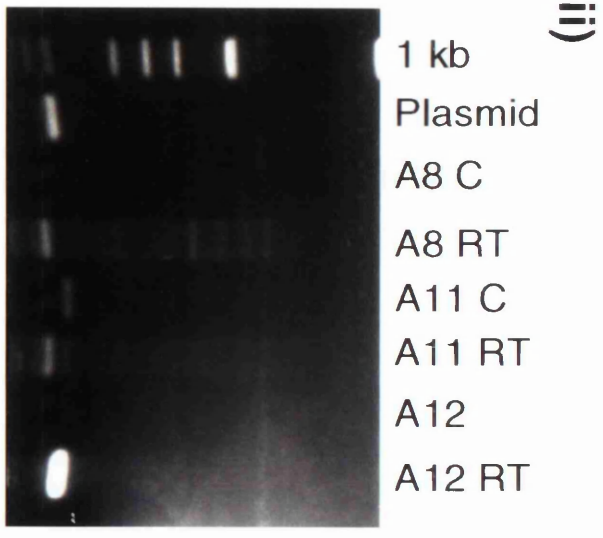
The results from this technique (shown in Figure 25) were largely in agreement with the northern blot analysis. Three clones which were shown to express the *lac Z* RNA by northern blot also were shown to express the RNA by PCR analysis as well. Two clones showed low level expression of the *lac Z* mRNA when analysed by PCR. Although expression of the lac Z RNA was detectable by PCR, the results from the PCR technique proved to be highly variable.

In order to confirm the PCR results the RNA levels were further analysed by another technique, the ribonuclease protection assay (Figure 26). This technique was carried out according to the method described in Section 2.24. The results of the ribonuclease protection assays agreed largely with both the northern blot analysis and the PCR reaction (Figure 25). High levels of expression were detected in 3 Act-LacZ^{A} clones 1,5 and 6. Lower levels of expression were seen in clones Act-LacZ^{A} 3 and 4.

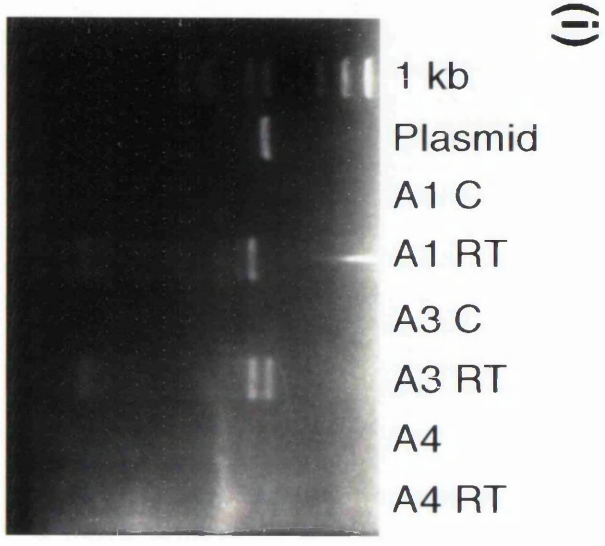
The only major difference was Act-LacZ^{A} clone 12 which had been shown to express relatively high levels of RNA by PCR but showed no expression by both the northern blot and ribonuclease protection. This clone also showed rearranged bands on the Southern blot. As the techniques analyse different parts of the *lac Z* gene it is possible that this clone only expresses part of the *lac Z* RNA. The PCR analysis covers a 217 bp

Figure 25. PCR analysis of *lac Z* RNA expression in C5O electroporation ringclones.

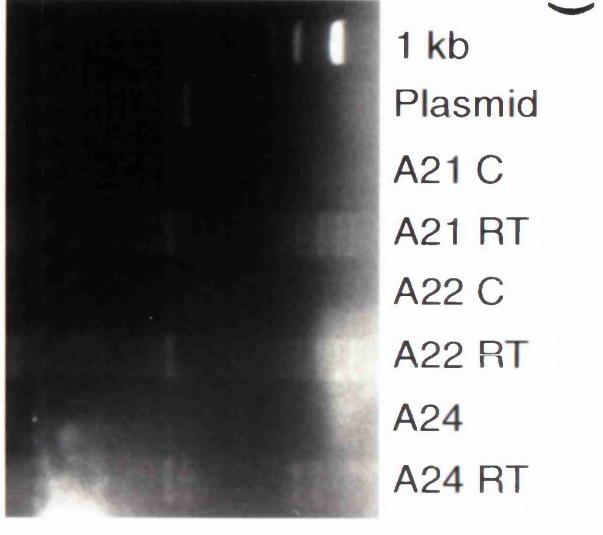
cDNA was generated from the RNA samples derived from the Act-LacZ^{A} electroporated cell lines. 50 pg of plasmid was used as a positive control for the PCR reaction. The reaction were performed with reverse transcriptase (RT) and without (C) as a control for DNA contamination. **(i)** Cell lines A1, A3 and A4. **(ii)** Cell lines A5, A6 and A7. **(iii)** Cell lines A8, A11 and A12. **(iv)**Cell lines A21, A22 and A24.


1 kb
Plasmid
A8 C
A8 RT
A11 C
A11 RT
A12
A12 RT

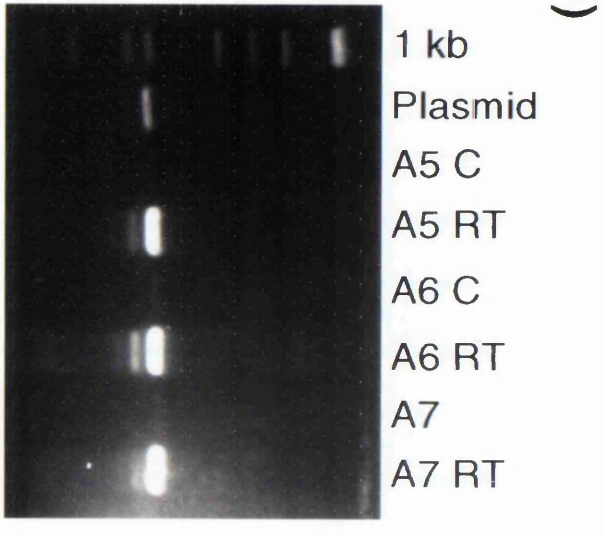
(iii)


1 kb
Plasmid
A1 C
A1 RT
A3 C
A3 RT
A4
A4 RT

(ii)


1 kb
Plasmid
A21 C
A21 RT
A22 C
A22 RT
A24
A24 RT

(iv)

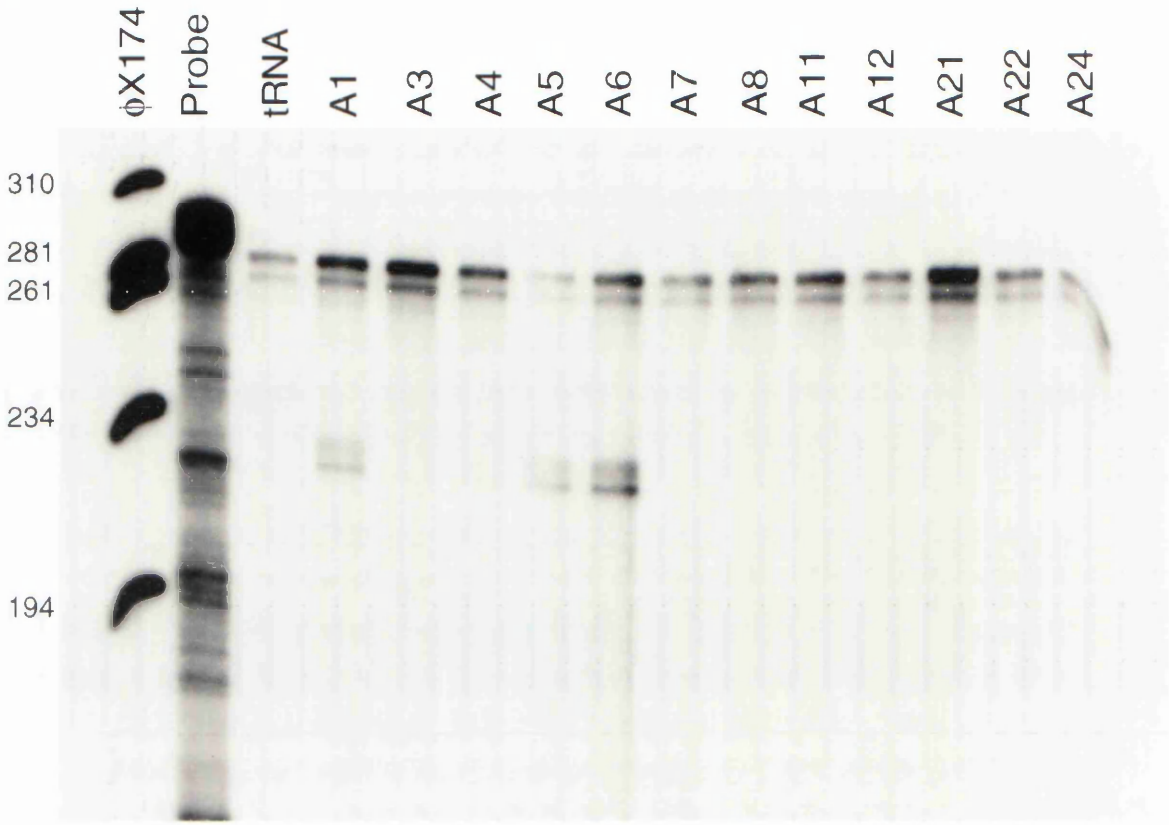

1 kb
Plasmid
A5 C
A5 RT
A6 C
A6 RT
A7
A7 RT

(i)

Figure 26. Expression levels of *lac Z* RNA in Act-LacZ^A electroporation ringclones.

(a) RNAse protection assay performed with a 227 bp fragment of the *lac Z* gene. Radiolabelled ϕX174 markers were used to size the bands on the gel. The probe is approximately 300 bp in size. tRNA was used as a negative control. The 227 bp protected fragment is visible in lanes A1, A3, A4, A5 and A6. **(b)** A graph of the relative expression levels of the *lac Z* gene in the Act-LacZ^A electroporated cell lines.

(a)


ϕX174
Probe
tRNA
A1
A3
A4
A5
A6
A7
A8
A11
A12
A21
A22
A24
310
281
261
234
194

(b)

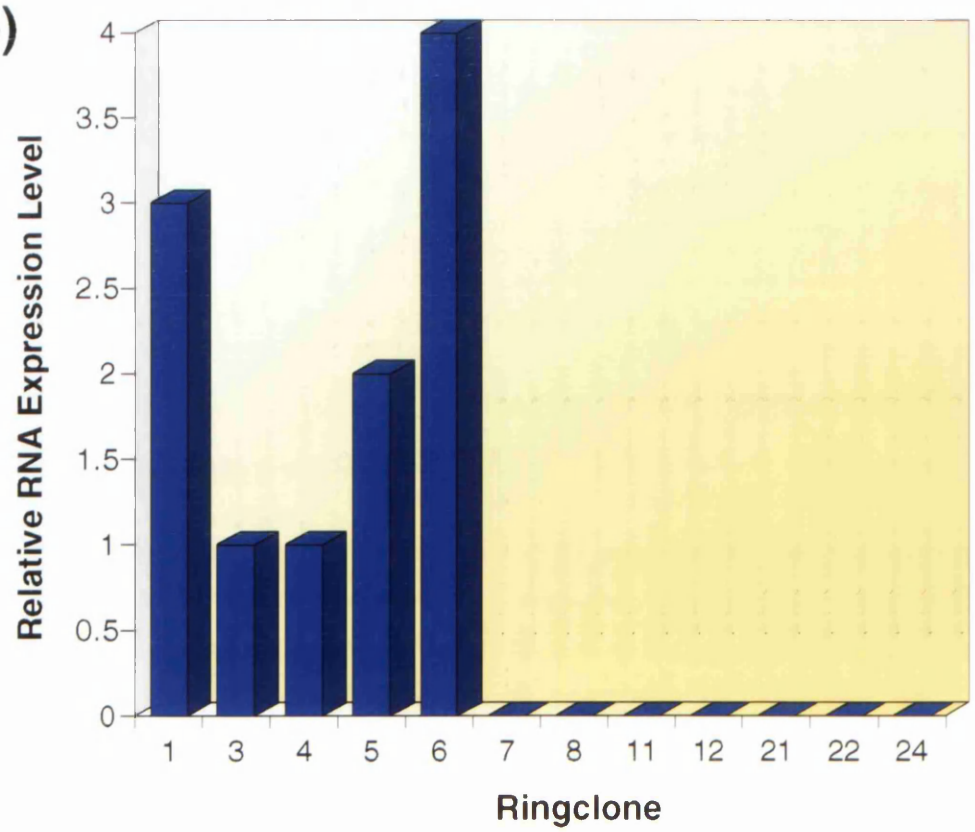
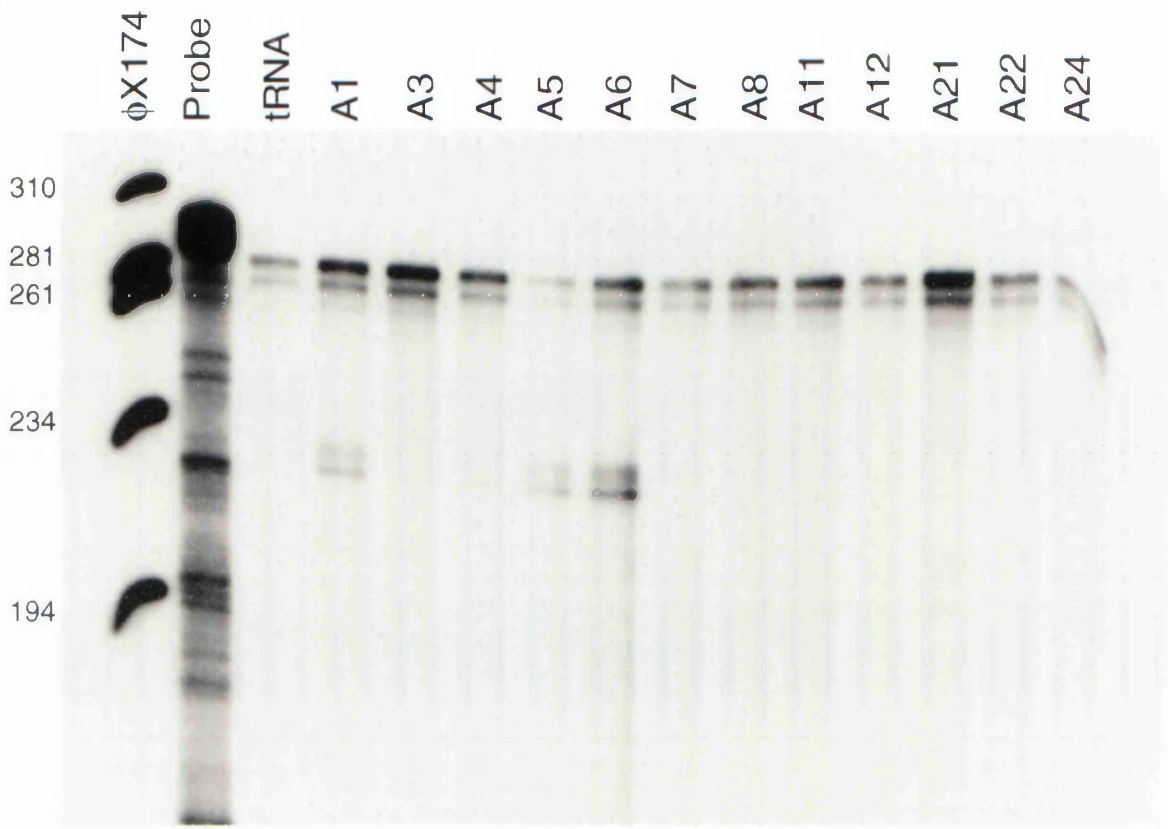

Relative RNA Expression Level
0
0.5
1
1.5
2
2.5
3
3.5
4
1
3
4
5
6
7
8
11
12
21
22
24
Ringclone

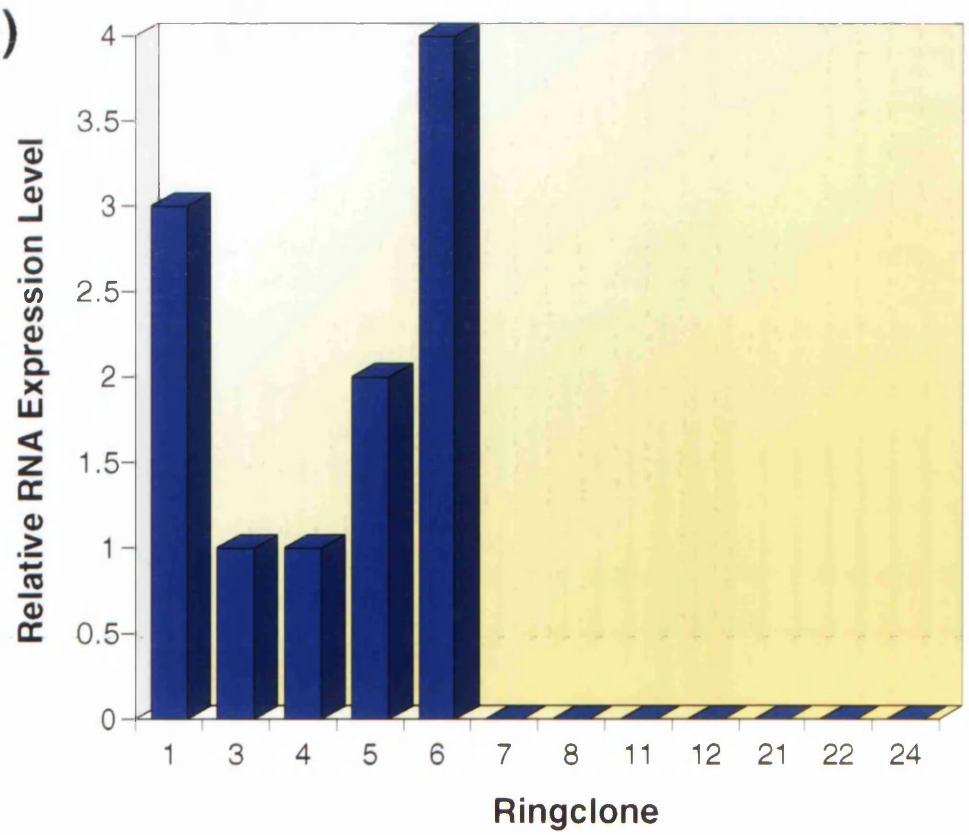
Figure 26. Expression levels of *lac Z* RNA in Act-LacZ^A electroporation ringclones.

(a) RNAse protection assay performed with a 227 bp fragment of the *lac Z* gene. Radiolabelled ϕX174 markers were used to size the bands on the gel. The probe is approximately 300 bp in size. tRNA was used as a negative control. The 227 bp protected fragment is visible in lanes A1, A3, A4, A5 and A6. **(b)** A graph of the relative expression levels of the *lac Z* gene in the Act-LacZ^A electroporated cell lines.

(a)


ϕX174
Probe
tRNA
A1
A3
A4
A5
A6
A7
A8
A11
A12
A21
A22
A24
310
281
261
234
194

(b)


Relative RNA Expression Level
4
3.5
3
2.5
2
1.5
1
0.5
0
1
3
4
5
6
7
8
11
12
21
22
24
Ringclone

fragment of the gene from base 72 to base 289 in the gene, the ribonuclease protection assay probe covered the section of the gene from base 836 to base 1063 while the probe for the northern analysis covered the bases 1063 to base 1952. Therefore it is possible that Act-LacZ^A clone 12 may only express the first part of the lac Z gene up to base 863.

Eighteen K5-LacZ^D clones were analysed for the presence of *lac Z* RNA and β-galactosidase protein. This construct contains the AAF-strategy test oligo/*lac Z* fusion gene. As the Act-LacZ^A cell lines had shown that the most reliable method of detecting RNA expression was the ribonuclease protection assay, this was the principal method used to determine the RNA expression levels in these cell lines. The level of protein expression in the cell lines had already been determined by staining cell with X-gal (Figure 20). The results of these analysis are shown in the graphs in Figure 27.

Six C5O electroporation ringclones containing the K5-LacZ^E construct were chosen for analysis. The results of these analyses are shown in Figure 28. Two of the six K5-LacZ^E clones (E3 and E5) analysed showed no expression of either the RNA (Figure 28) or the protein (Figure 22). All the cell lines which do express detectable levels of RNA also express the protein in a significant proportion of the cell population. However only in one of the four cell lines does the expression level approach 100% of the population. One of the four clones contains as little as 5% of cells which express the β-galactosidase enzyme. These results perhaps again suggest that rearrangement of the lac Z gene is occurring after transfection of the marker gene constructs.

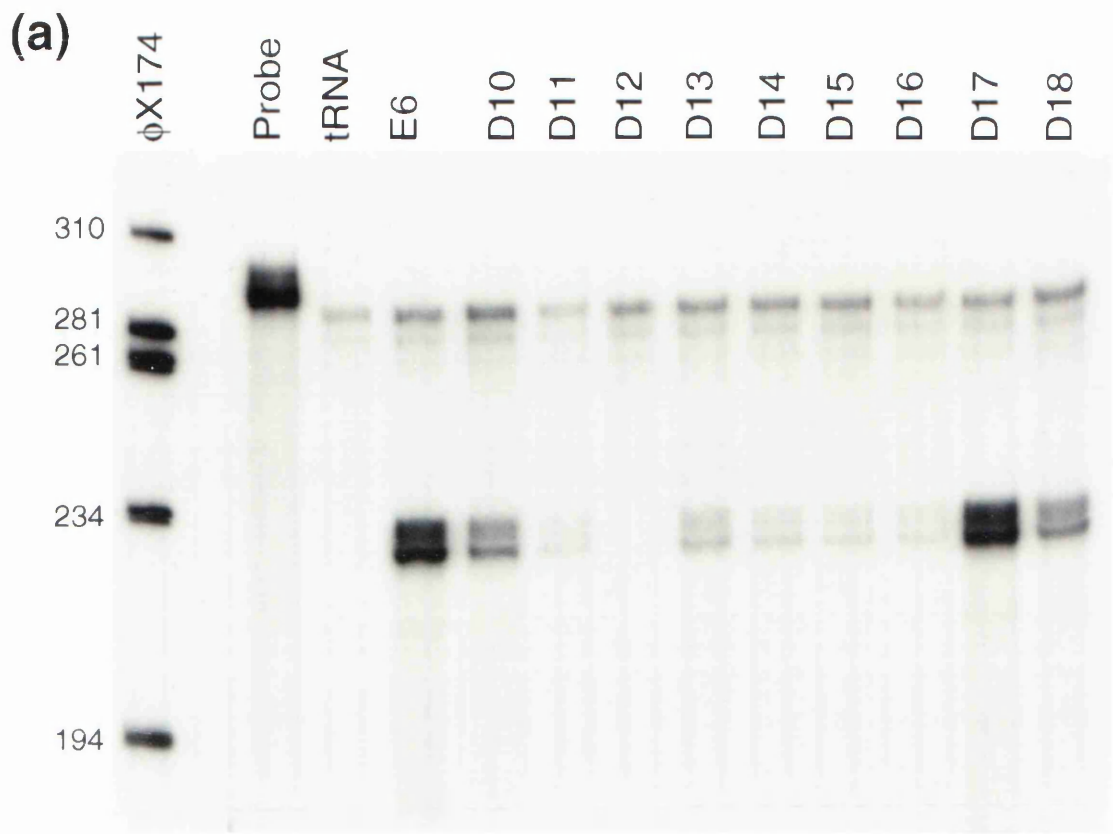
Comparison of the results of the K5-LacZ^D and K5-LacZ^E construct electroporation demonstrate that although the levels of protein expression in the Act-LacZ^D clones are detectable they are significantly less than the levels obtained in the pIrv-Neo-Act-LacZ^E cell ringclones.

3.3.8 Activation of β-galactosidase expression in cells containing Act-Lac Z A by treatment with MNNG.

The final step in testing the marker genes *in vitro* is to demonstrate that the marker genes can be activated by treatment with a chemical mutagen. This has been attempted for the MNNG-test oligo construct Act-LacZ^A . The activation experiment requires a stably transfected cell population which is known to be capable of expressing the β-galactosidase protein. The Act-LacZ^A clone 5 was selected for use in the MNNG

Figure 27. Expression levels of *lac Z* RNA in K5-LacZ^{D} electroporation ringclones.

(a) RNAse protection assay performed with a 227 bp fragment of the *lac Z* gene. Radiolabelled ϕX174 markers were used to size the bands on the gel. The probe is approximately 300 bp in size. tRNA was used as a negative control. The 227 bp *lac Z* fragment corresponding to expression of the *lac Z* gene is visible in all lanes except D12. **(b)** A graph of the relative expression levels of the *lac Z* gene in the K5-LacZ^{D} electroporated cell lines.


(a)
ϕX174
Probe
tRNA
E6
D10
D11
D12
D13
D14
D15
D16
D17
D18
310
281
261
234
194

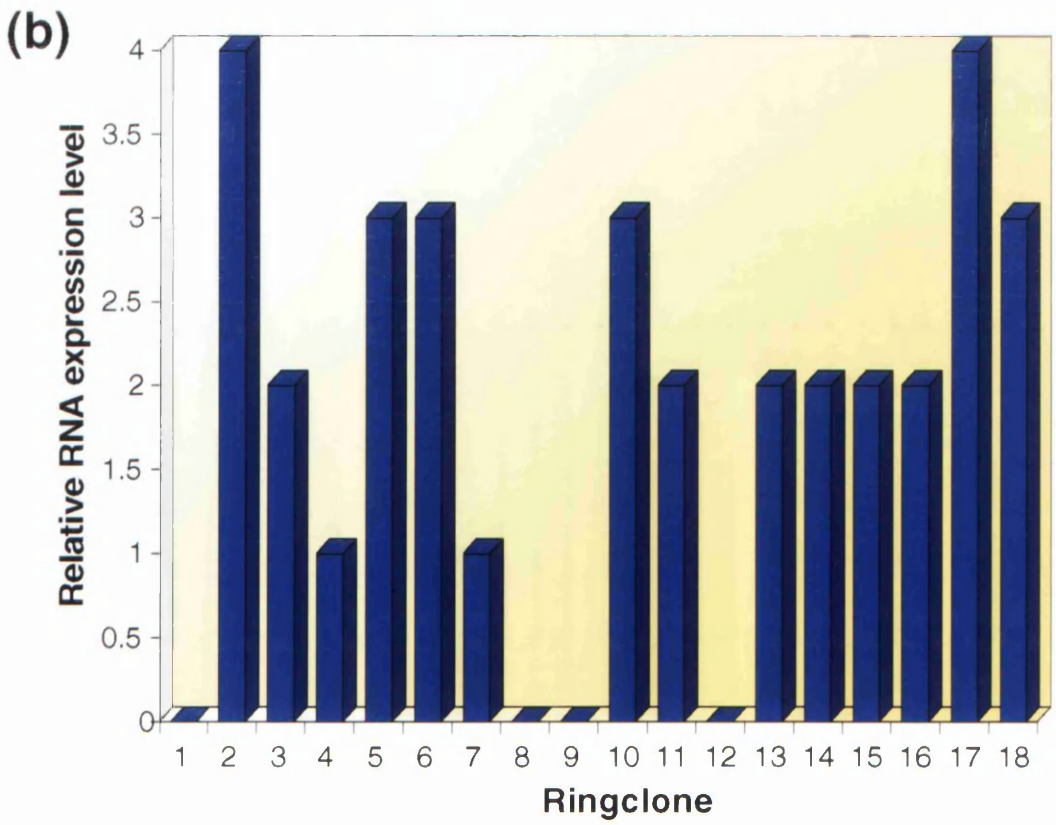
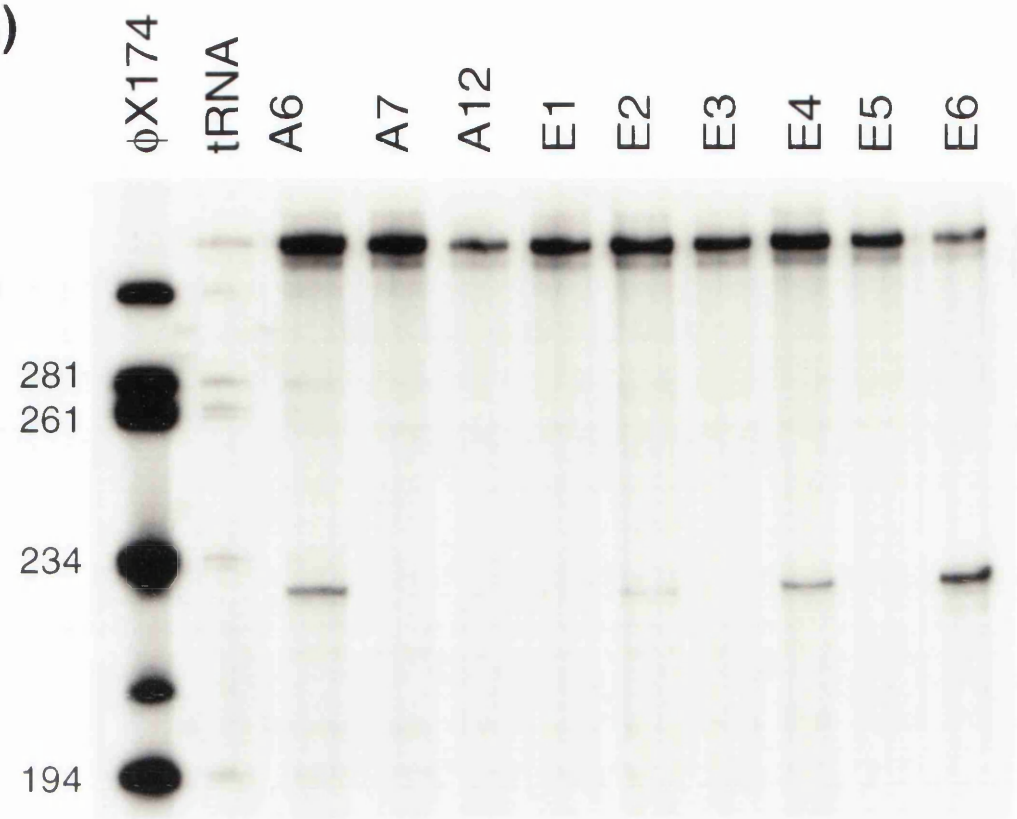

(b)
Relative RNA expression level
4
3.5
3
2.5
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Ringclone

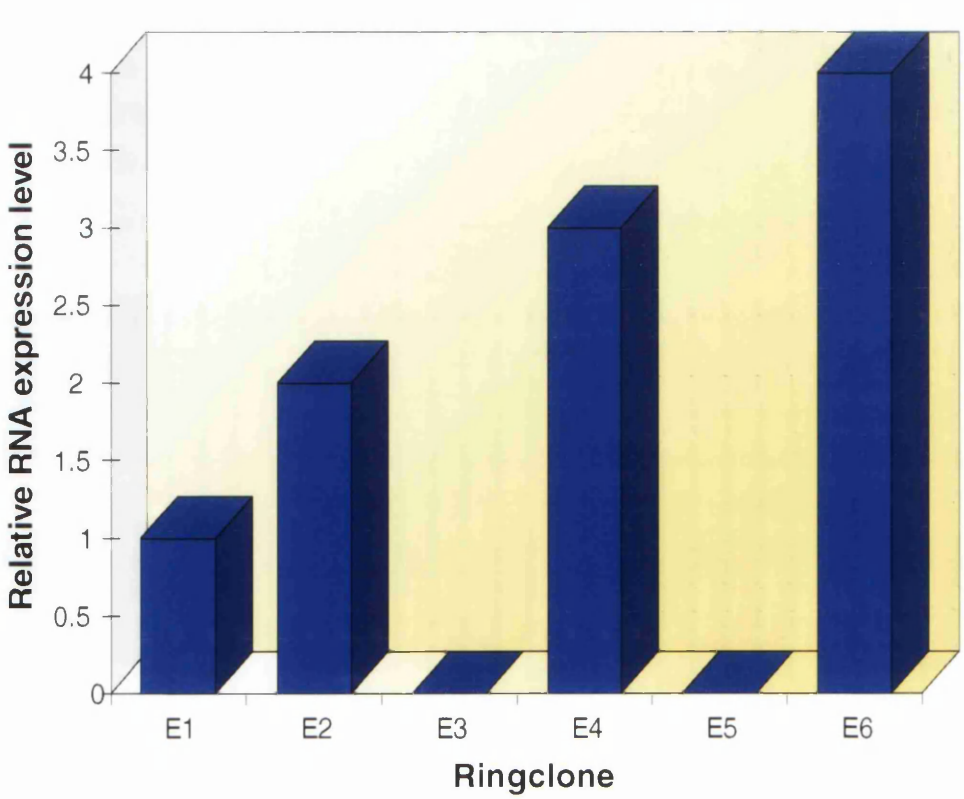
Figure 28. Expression levels of *lac Z* RNA in K5-LacZ^{E} electroporation ringclones.

(a) RNAse protection assay performed with a 227 bp fragment of the *lac Z* gene. Radiolabelled ϕX174 markers were used to size the bands on the gel. The probe is approximately 300 bp in size. tRNA was used as a negative control. The 227 bp protected *lac Z* fragment is visible in cell lines A6, E1, E2, E4 and E6. **(b)** A graph of the relative expression levels of the *lac Z* gene in the K5-LacZ^{E} electroporated cell lines.

(a)


φX174
tRNA
A6
A7
A12
E1
E2
E3
E4
E5
E6
281
261
234
194

(b)


Relative RNA expression level
4
3.5
3
2.5
2
1.5
1
0.5
0
E1
E2
E3
E4
E5
E6
Ringclone

treatment experiments. This cell line is a C5O murine epidermal ringclone transfected with the Act-LacZ^A construct by electroporation and selected on G418.

Analysis of the DNA of this cell line has shown that it contains the *lac Z* gene and that the *lac Z* gene is intact and not rearranged in any way (Figure 22). Furthermore the RNA from this cell line has been analysed by a three different techniques and shown to express a relatively high level of *lac Z* RNA (Figures 25 and 26). The staining of the cells with X-gal revealed that a very small number (<0.01%) of the cells express the protein from the *lac Z* gene. This very low expression of the β-galactosidase protein indicates that when a mutation occurs in the *lac Z* gene in this cell population it is capable of being expressed. The Act-LacZ^A clone 5 is therefore suitable for use in the MNNG-reactivation experiment.

The results of the reactivation experiment are detailed in Figure 29. After treatment with MNNG at a concentration of 1μgml^{-1}, the cells were replated and were fixed and stained with X-gal three days later. At this concentration the MNNG kills 87.5% of the clonogenic cells in the culture. The number of *lac Z* expressing cells detected in cell populations which were treated with MNNG was considerably greater than the number of *lac Z* positive cells in the control mock-treated population. No blue cells were ever observed in untransfected cells treated with MNNG.

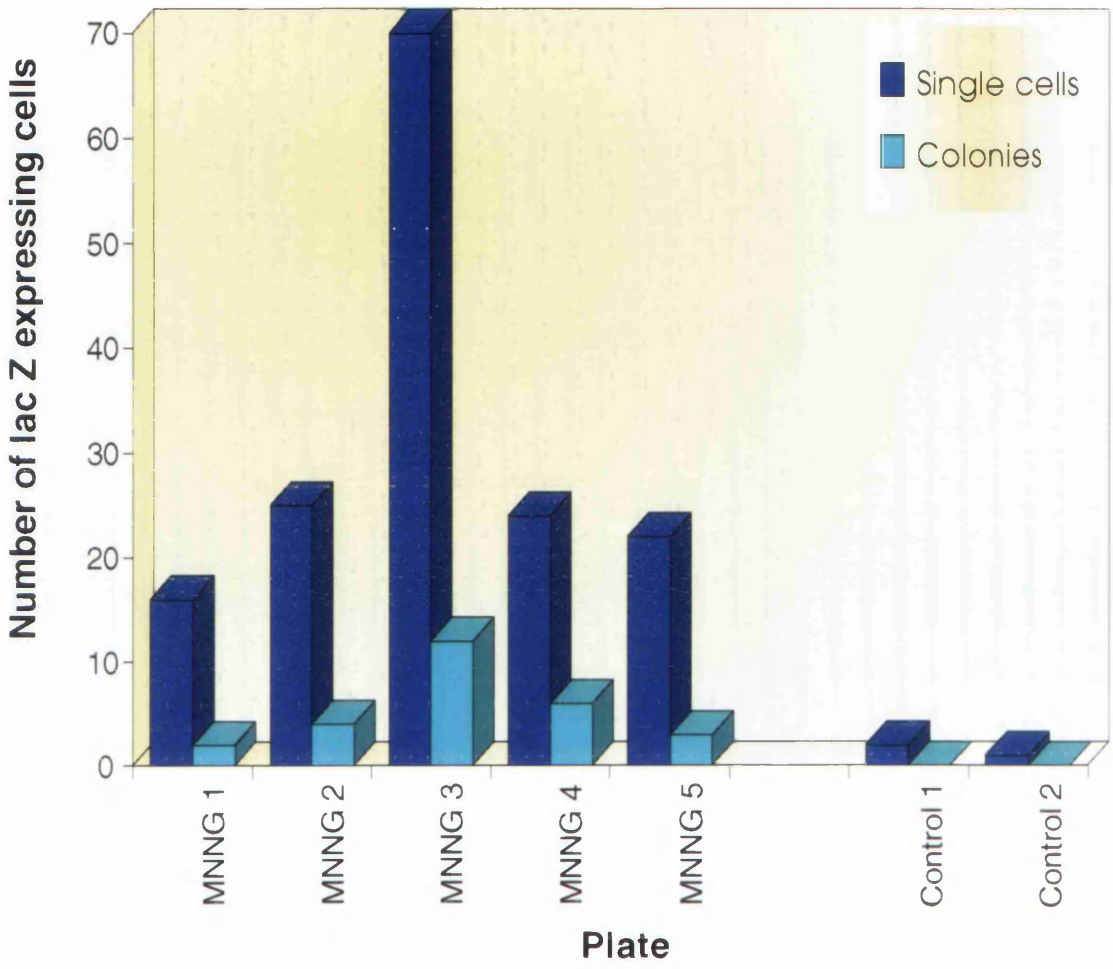
The frequency of activation was calculated in two separate ways. The total numbers of blue cells on each plate and the mutation frequency was calculated as the number of activated cells per 5 x 10^5 cells added to each plate. The average number of blue cells on each plate is 32 and hence the mutation frequency is determined to be 1.5 x 10^{-4}. The other method depends on the calculation of the number of clonogenic cells cell in the population of 5 x 10^5. This was determined by plating out cells in small numbers into F25 flasks and looking at colony formation after 2 weeks. The number of clonogenic cells in the untreated population was 20% and therefore the surviving clonogenic cells in 5 x 10^5 treated cells was determined to be 1.25 x 10^4. The number of blue cell colonies (groups of two or more adjacent expressing cells) on each plate was determined by examining the plates under an inverted microscope. The average number of blue colonies on each plate was 5.3 and hence the mutation frequency was determined to be 4.32 x 10^{-3} per clonogenic cell.

The accuracy of both figures is debatable. The number of constructs present in Act-LacZ^A clone 5 was never accurately determined. Even if it had been determined the

Figure 29. Activation of β-galactosidase expression in Act-LacZ^A clone 5 cells by treatment with MNNG.

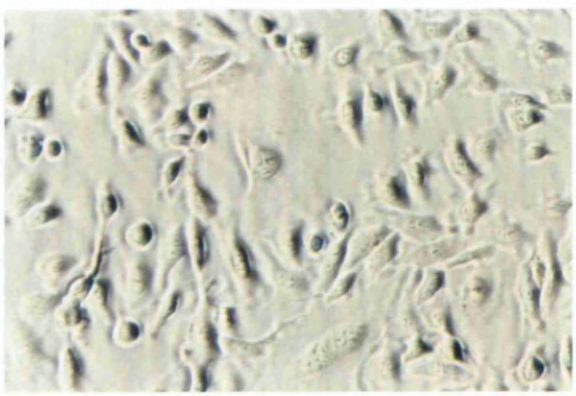
(a) Graph showing the numbers of *lac Z* expressing single cells and clones (groups of two or more cells) expressing the *lac Z* gene in plates containing 5 x 10^5 MNNG-treated or control Act-LacZ^A clone 5 cells. **(b)** Control **(i)** and MNNG-treated **(ii)** cells stained with X-gal. A single β-galactosidase expressing cell is visible in the treated cell population.

(a)

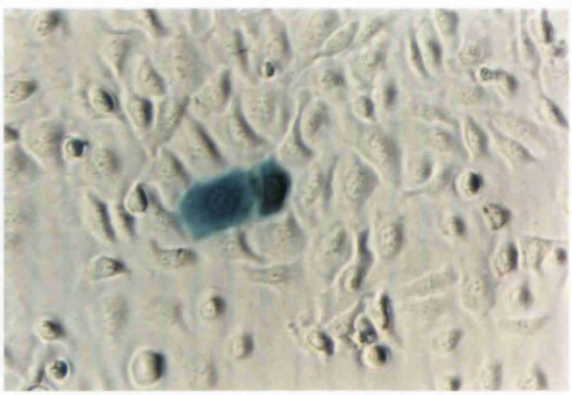


(b)

(i) Control Act-Lac Z ^{A} cells



(ii) MNNG-treated Act-Lac Z ^{A} cells



number of cells in the population which could express the gene would then have to be determined. In some of the K5-LacZ^{E} ringclone cell lines as little as 5% of the total population actually expressed the gene when stained with the X-gal stain. Of the two figures obtained the second figure is possibly more accurate. The first figure may count a lot of dead and dying cells which are incapable of expressing the gene.

These reservations do not detract from the main conclusions of the experiment which shows that the LacZ^{A} marker gene is activated in cells by exposure of the cell to the chemical mutagen MNNG *in vitro*. The establishment of an exact frequency is not all that important for the aims of this project. Any mutation frequency determined would not be relevant to the conditions *in vivo* when the cell environment is totally different to the environment of the cells in tissue culture. Work with the AAF marking strategy (LacZ^{D}) was discontinued at this time due to time constraints on the project and the positive results with the MNNG strategy (LacZ^{A})marker gene.

3.4 Testing of the cell marking system *in vivo.*

3.4.1 Generation of transgenic mice.

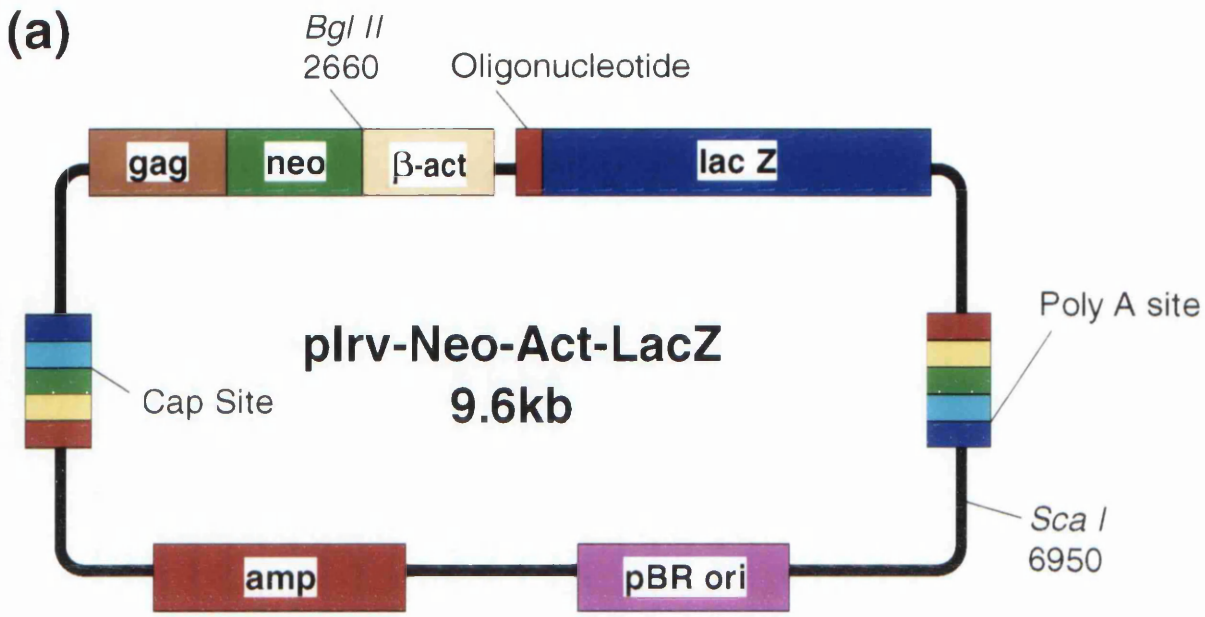
Having demonstrated that the marker genes worked *in vitro* the next step was to test the marker genes *in vivo*, by introducing the constructs into the germline of mice. All transgenic mice generated in the course of this project were generated by pronuclear microinjection of fertilised embryos (this method is discussed in detail in Section 2.26).

The first marker gene construct used to generate transgenic mice was the *tg*Act-LacZ^{A} MNNG-marker gene construct. The first step in generating transgenic mice is the purification of the marker gene construct. As vector sequences are known to inhibit the expression of transgenes the majority of the vector sequences were removed by restriction enzyme digestion. The *tg*Act-LacZ^{A} insert used for the production of transgenic mice are shown in Figure 30.

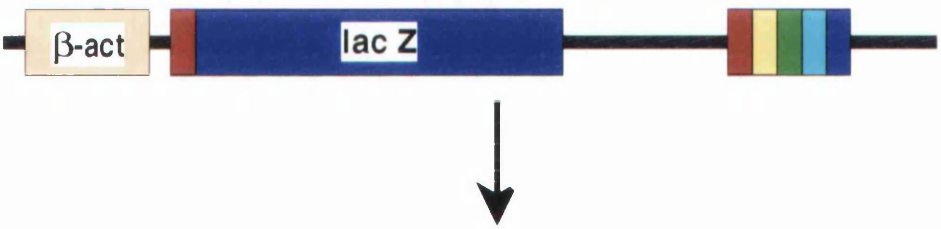
The Act-LacZ^{A} vector was digested simultaneously with the restriction endonucleases *Bgl II* and *Sca I*. This removes a 4.3 kb fragment from the plasmid containing the β-actin promoter, the LacZ^{A} marker gene and the poly A sequence located in the 3' viral LTR sequences. This DNA fragment will subsequently be referred to as *tg*Act-LacZ^{A}.

Figure 30. Structure of *tg*Act-LacZ^A construct used to generate transgenic mice.

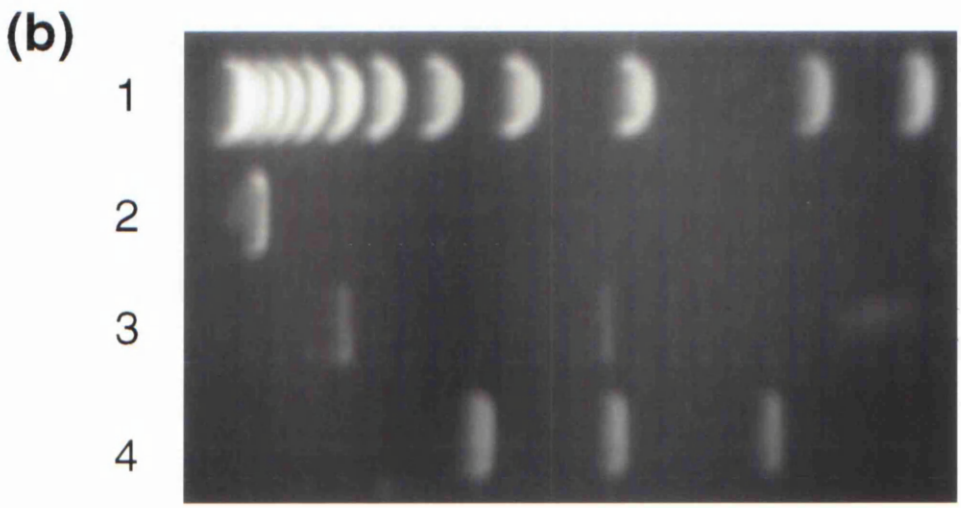
(a) Diagram of the structure of the *tg*Act-LacZ^A construct used for microinjection. **(b)** Restriction enzyme analysis of the construct. **1)** 1 kb ladder **2)** Act-LacZ^A plasmid linearised by digestion with *Not I*. **3)** 3.1 kb *lac Z* band excised from Act-LacZ^A plasmid by digestion with Sal I. **4)** 4.3 kb *tg*Act-LacZ^A insert excised by digestion with *Bgl II*/*Sca I*.



Insert excised by digestion
with *Bgl II*/ *Sca I* enzymes



4.3kb insert purified and used
for pronuclear microinjection



The prospective transgene sequences were the separated on an agarose gel and purified as described in Section 2.25.

The DNA samples were then introduced into mouse embryos by injection of the pronucleus of fertilised eggs. The fertilised mouse embryos were harvested from superovulated three week old (C57BL/6J x CBA/J) F1 female mice. The superovulation protocol is described in section 2.26.2. After injection of the embryos, they were cultured overnight and transferred into pseudopregnant recipients the following day. Following the transfer the recipient mice were left alone in cages for a period of three weeks. If the mice became pregnant and produced offspring the offspring were analysed for the presence of transgenes (Figure 31 shows a table detailing the results of the microinjection experiments).

The *tg*Act-LacZ^A transgene was injected into embryos at a range of concentrations. The concentrations used were 1, 2 and 5 μgml^{-1}. A total of 22 offspring were obtained from injection of *tg*Act-LacZ^A DNA at a concentration of 1 μgml^{-1}. Genomic DNA was isolated from the mice derived from injected embryos by tail tip biopsies according to the method outlined in Section 2.27. The DNA from the offspring were then analysed for the presence of *tg*Act-LacZ^A DNA by Southern blot.

One of these 22 mice analysed contained a *tg*Act-LacZ^A transgene. These results were confirmed by analysis of the genomic DNA by PCR. The low rate of incorporation of the transgene may be due to the slightly low DNA concentration. Figure 32a shows a Southern blot of DNA isolated from tail tip biopsies of offspring from embryos injected with 1 μgml^{-1}. Only one mouse (number 1591) contained a transgene band. The DNA on the blot is digested by the enzyme *Kpn I* which cuts only once within the sequence of the transgene. This should help determine the organisation of the transgene insertions. Mouse 1591 seems to have a single copy of the gene inserted into the genome as digestion with *Kpn I* results in the production of only 1 transgene band. Although it is possible that the integration site may contain a couple of constructs arranged in a head to head sequence, although this is an unusual arrangement of transgenes. Transgenes present in multiple copies are usually arranged in head to tail concatamers. If this was the case the extra copies of the transgene would be present as a 4.3 kb band equivalent to the size of the transgene.

The concentration of *tg*Act-LacZ^A DNA injected was then raised to 2 μgml^{-1}. A total of 47 offspring were derived from embryos injected with the *tg*Act-LacZ^A DNA injected at this concentration. Southern blot analysis determined that thirteen of these 47 offspring

Figure 31. Table of results from injection of *lac Z* fusion gene constructs into mouse embryos.

In total 22 transgenic founder animals containing three separate constructs were generated from 147 offspring derived from embryos microinjected with DNA. Injected offspring represents the total number of live newborn mice derived from embryos which had been microinjected with DNA before transfer into pseudopregnant female mice. The transgenic offspring represent the total number of live newborn mice which were transgene positive when screened by Southern blot.

Transgene	Concentration	Injected offspring	Transgenic offspring
Act-LacZ[A]	1	22	1
	2	47	13
	5	0	0
K5-LacZ[D]	0.5	21	4
K5-LacZ[E]	0.5	22	4
	Total	112	22

Figure 32. Southern blot analysis of founder mice containing the *tg*Act-LacZ^A transgene.

Southern blots of DNA derived from tail tip biopsies of mice injected with *tg*Act-LacZ^A DNA. The DNA samples are digested with *Kpn I* which cuts once in the 3' end of the DNA construct. **(a)** One mouse (1591) is transgene positive. **(b)** Blot of Act-LacZ^A cell line and *tg*Act-LacZ^A injected mouse DNA. Untransfected C5O cell line DNA is used as a negative control. The Act-LacZ^A cell line show a 6.4 kb band corresponding to the size of the *Kpn I* fragment of the construct. Two positive mice contain a 4.3 kb band suggesting that the construct is integrated in a head to tail array.

(a)
C5O
1586
1587
1588
1589
1590
1591
1592
1593
1594
8.5 kb
(b)
C5O
Act-Lac ZA6
1301
1302
1303
1304
1305
1306
1307
1308
10.3 kb
8.6 kb
6.4 kb
5.9 kb
4.3 kb

produced contained *tg*Act-LacZ^A transgenes. The organisation of the transgene bands in these mice demonstrated that most of the mice have a number of copies of the transgenes. Figure 32b shows a Southern blot of DNA samples derived from tail tip biopsies of offspring from embryos injected with *tg*Act-LacZ^A DNA at 2 μgml^{-1}. The blot is probed for the presence of the *lac Z* gene.

Two positive mice are detected on this blot of genomic DNA digested with the *Kpn I* restriction enzyme. Both the positive DNA samples exhibit bands at 4.3 kb which indicates a number of copies of the transgene inserted in a head to tail concatamer. The positive control lane shows cell line Act-LacZ^A DNA which contains the entire DNA construct. The *Kpn I* enzyme cuts twice in the Act-LacZ^A construct to produce a 6.4 kb band containing the *lac Z* gene sequence. The presence of a 4.3 kb band in the positive samples rules out the possibility of plasmid contamination, which would result in the presence of a 6.4 kb band instead of the 4.3 kb band.

In addition to the 4.3 kb band there are another two bands present in each of the positive DNA samples on the blot. One of these probably represents the integration site of the construct. The other band corresponds to a size roughly equivalent to two constructs. This could be produced by head to head concatamers of the construct, instead of head to tail. The *Kpn I* enzyme site is located near one end of the construct. If the end of the construct including the *Kpn I* site was deleted upon integration of the gene this might result in the two adjacent constructs becoming fused together. When the DNA is digested by *Kpn I* the two constructs would not be separated as the site is lost and this would result in the production of an approximately 8.0 kb band. The other alternative explanation of this band is incomplete digestion of the DNA, although this pattern is maintained in DNA samples from offspring generated from these mice.

While the mice from the previous injections were being analysed the *tg*Act-LacZ^A DNA injection concentration was then increased again to 5 μgml^{-1}. No positive offspring were obtained from injection of the *tg*Act-LacZ^A construct at this concentration. Although the injected embryos continued to progress to the two-cell stage the resultant litter sizes obtained from the transfer of embryos into pseudopregnant recipient females was very low. A few offspring were born but these were in very small litters which were subsequently killed or abandoned by their mother. Presumably this reduction in litter size was due to a harmful effect of either the increased DNA concentration or the increased concentration of contaminants in the DNA sample.

Injection of the *tg*Act-LacZ^A construct was discontinued when 14 transgenic founders had been produced by injection of this construct at various concentrations. The next set of injections were performed with the *tg*K5-LacZ^D construct which contained the AAF-strategy control oligo/*lac Z* fusion gene. The prospective transgene was removed from the K5-LacZ^D construct by digestion with the restriction enzymes *Hind III* and *Sca I* (this is shown in Figure 33). The *tg*K5-LacZ^D insert was injected at a concentration similar to that of the *tg*Act-LacZ^A 2 μgml^{-1} sample which according to the optical density measurement on the spectrophotometer was 0.5 μgml^{-1}. The concentrations of the two DNA samples were compared by running them side by side on an agarose gel and judging the intensity of the bands by visual inspection. The *tg*K5-LacZ^D prospective transgene DNA was purified by the same method as described above for the purification of the *tg*Act-LacZ^A DNA (this method is also outlined in Section 2.25).

DNA isolated from tail tip biopsies of mice derived from embryos injected with the *tg*K5-LacZ^D construct was analysed by Southern blot. DNA samples from four of the twenty-two offspring were positive for the *tg*K5-LacZ^D transgene. Figure 34a shows a Southern blot with two control and eight *tg*K5-LacZ^D-injected mouse DNA samples digested with the restriction enzyme *Kpn I*. This blot was hybridised to a probe specific for the *lac Z* gene sequences. The negative control lane shows no hybridisation to the *lac Z* probe. The positive control is DNA from mouse number 1298 previously shown to contain the Act-LacZ^A transgene. The *lac Z* probe hybridises to the characteristic 4.3 kb band produced by this transgene. There is one positive *tg*K5-LacZ^D sample on the blot in mouse number 2178. The *BK III* promoter in the *tg*K5-LacZ^D construct is 1 kb larger than the β-actin promoter in the *tg*Act-LacZ^A transgene. The size of this transgene is therefore 5.3 kb. This 5.3 kb band is observed in the DNA sample isolated from the tail tip biopsy of mouse number 2178.

The final construct which was injected into mouse embryos in order to generate transgenic mice was the *tg*K5-LacZ^E construct, which contains the AAF-strategy control oligo/*lac Z* fusion gene. The prospective transgene sequences were isolated in the same fashion as the *tg*K5-LacZ^D transgene by digestion of the K5-LacZ^E plasmid with the restriction enzymes (Figure 33). The *tg*K5-LacZ^E DNA was purified (as described in Section 2.25) and also injected at a concentration of 0.5 μgml^{-1}.

Figure 34b shows a Southern blot of one positive control and five *tg*K5-LacZ^E DNA samples digested with the restriction enzyme *Kpn I*. This Southern blot has been hybridised to a probe for the *lac Z* gene sequence. The overall structure of the *tg*K5-LacZ^E transgene is the same as that of the *tg*K5-LacZ^D and it is therefore the same size

Figure 33. Structure of *tg*K5-LacZ^{D} and *tg*K5-LacZ^{E} constructs used to generate transgenic mice.

(a) Diagram of the structure of the *tg*K5-LacZ$^{D/E}$ construct used for microinjection. **(b)** Restriction enzyme digestion of the K5-LacZ^{D} plasmid. **1)** 1 kb ladder. **2)** K5-LacZ^{D} plasmid linearised by digestion with *Not I*. **3)** 3.1 kb *lac Z* gene excised from the K5-LacZ^{D} plasmid by digestion with *Sal I*. **4)** 5.3 kb *tg*K5-LacZ^{D} insert excised from plasmid by digestion with *Hind III/Sca I*.

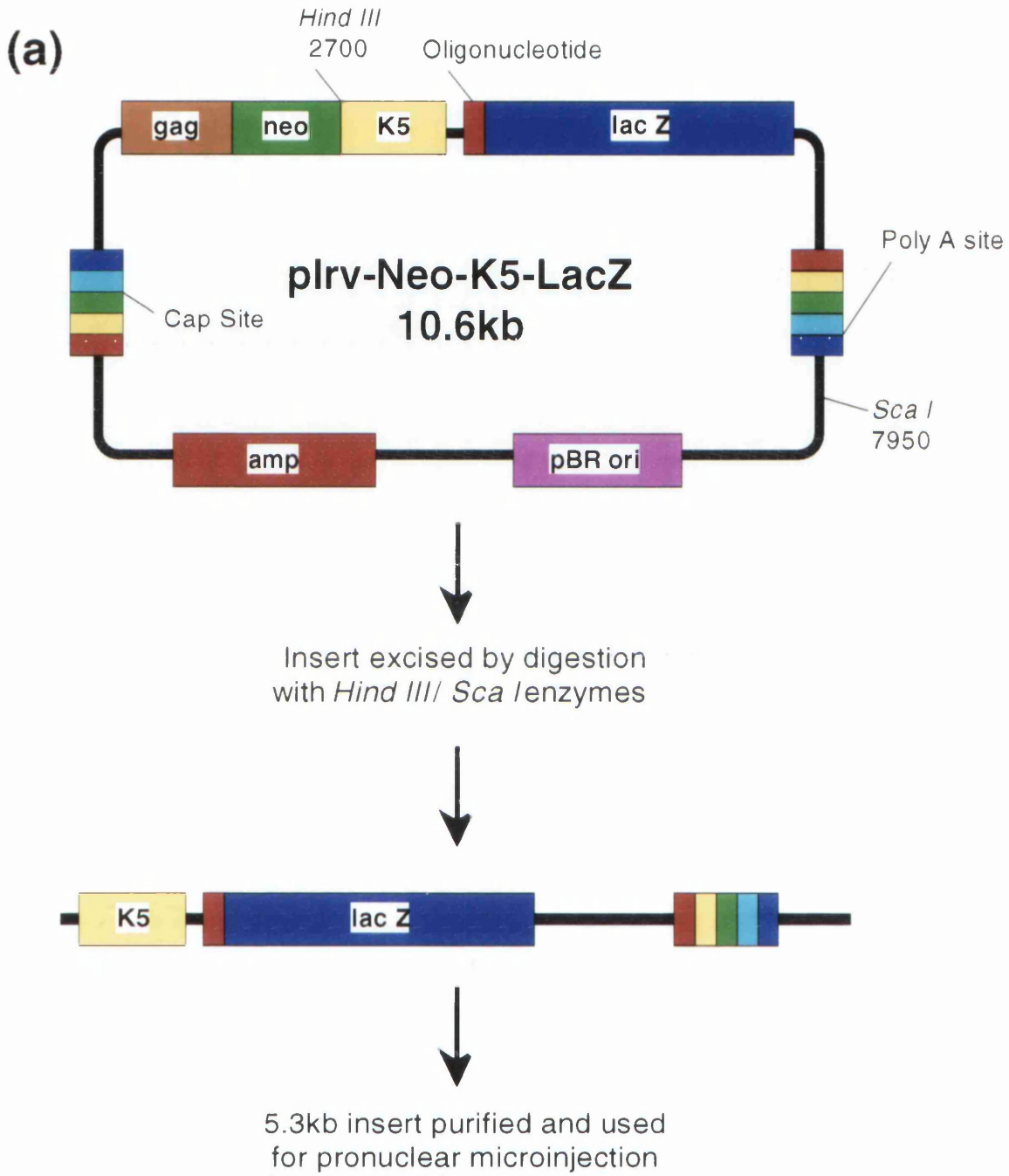
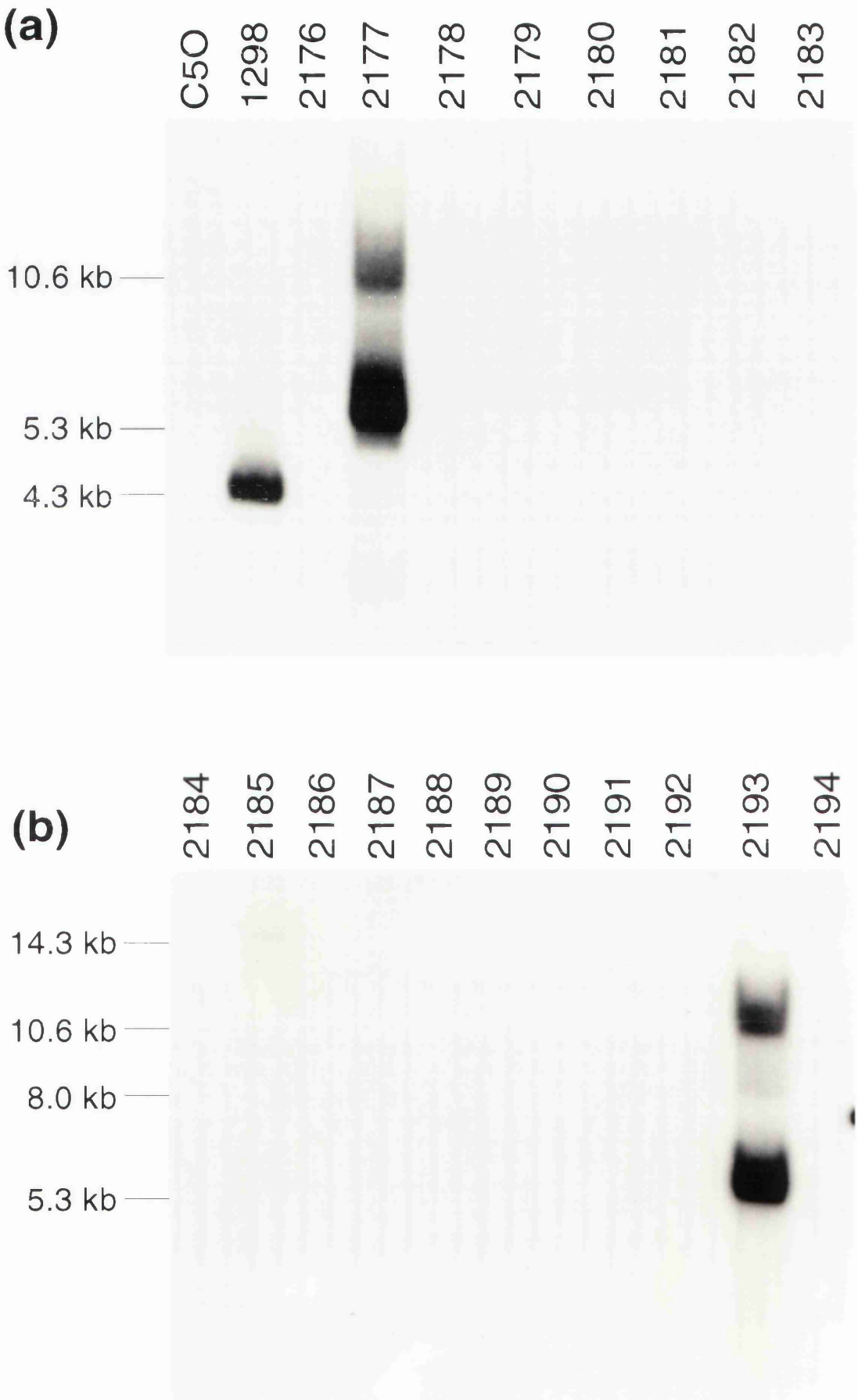

(a)
Hind III
2700
Oligonucleotide
gag
neo
K5
lac Z
Poly A site
Cap Site
pIrv-Neo-K5-LacZ
10.6kb
Sca I
7950
amp
pBR ori
Insert excised by digestion
with Hind III/ Sca I enzymes
K5
lac Z
5.3kb insert purified and used
for pronuclear microinjection
(b)
1
2
3
4

Figure 34. Southern blot analysis of founder mice containing the *tg*K5-LacZ^{D} and *tg*K5-LacZ^{E} transgenes.

Southern blots of DNA derived from tail tip biopsies of mice injected with *tg*Act-LacZ^{D} and *tg*Act-LacZ^{E} DNA. The DNA samples are digested with Kpn I which cuts once in the 3' end of the DNA construct. **(a)** Southern blot of DNA samples derived from mice injected with the *tg*Act-LacZ^{D} insert. DNA from mouse 1298 which contains the *tg*Act-LacZ^{A} transgene was used as a positive control. One mouse contains the *tg*Act-LacZ^{D} transgene and contains a 5.3 kb band suggesting the transgene is integrated in a head to tail array. **(b)** Blot of *tg*Act-LacZ^{E} injected mouse DNA. Two positive mice contain a transgene band.


(a)
C5O
1298
2176
2177
2178
2179
2180
2181
2182
2183
10.6 kb
5.3 kb
4.3 kb
(b)
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
14.3 kb
10.6 kb
8.0 kb
5.3 kb

at 5.3 kb (Figure 33). Two of the eleven K5-LacZ^{E} samples were positive and one (2193) contained the 5.3 kb *tg*K5-LacZ^{E} transgene band suggesting tandem array of the transgenes. The other mouse (2185) contained a single integration site band at 14.3 kb. Twenty-two mice were derived from embryos injected with the K5-LacZ^{E} transgene construct at 0.5 μgml^{-1} and four of these proved to be transgenic.

In total 22 founder mice were generated with three different oligonucleotide/*lac Z* fusion constructs. These 22 founder mice were produced from a total of 112 mice derived from injected embryos and hence the average frequency of transgenic mouse generation was very nearly 1 transgenic mouse in every five offspring obtained from injected embryos. The construct which had the highest frequency of transgene positive animals obtained from injected embryos was the Act-LacZ^{A} construct. Of 47 animals born from Act-LacZ^{A} injected embryos 13 were transgene positive which represent a frequency of almost 30%.

3.4.2 Breeding mice and establishing transgenic lines.

After a number of transgenic founder animals had been produced they were bred with non-transgenic animals in order to establish permanent transgenic lines. All the transgenic animals (male or female) were initially bred with NIH mice. NIH mice were selected to breed with the transgenic mice as they were commonly used in the laboratory and breeding the transgenic mice with this strain would help to make this project compatible with the others in the laboratory at areas of overlap.

Ten of the fourteen *tg*Act-LacZ^{A} founder mice were selected for breeding with the NIH founder mice. Five male founders were bred with NIH female mice and five female founder mice were bred with five NIH male mice. DNA isolated from tail tip biopsies of offspring from the breeding were screened by Southern blot analysis, slot blot analysis and the polymerase chain reaction (PCR).

All five of the male mice bred with the NIH female mice gave rise to transgene-positive offspring. Only three of the female mice bred with NIH mice gave rise to positive offspring. One of these female mice did not breed with the NIH male and produced no offspring. One of the female *tg*Act-LacZ^{A} founder mice bred with the NIH male mouse and produced offspring but did not pass on the gene to its offspring. Figure 35 shows a Southern blot of DNA samples digested with the *Kpn I* restriction enzyme and

Figure 35. Analysis of the offspring of transgenic founder animals by Southern blot.

Southern blots of genomic DNA samples prepared from mouse tail tip biopsies and digested with Kpn I. **(a)** Blot of DNA from founder mouse 1591 and 9 offspring derived from breeding the founder mouse with an NIH male. Six of the nine mice are positive. **(b)** Small southern blot used to screen offspring. C5O cell line DNA is used as a negative control. Founder mice 1301 and 1313 and their offspring are screened. 5 of the 11 offspring derived from founder 1301 are positive. Mouse 1313 does not pass the transgene to any of its offspring.

(a)
C5O
1591
2020
2021
2022
2023
2024
2025
2026
2027
2028
8.5 kb
(b)
C5O
1301
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
4.3 kb
4.3 kb
C5O
1313
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071

Figure 36. Table of transgenic lines generated from founder animals.

This table describes the breeding of eleven lines of transgenic mice containing three different DNA constructs from founder mice.

Construct	Transgene	Line	Founder
pIrv-Neo-Act-LacZ[A]	tgAct-LacZ[A]	A	1591
		B	1296
		C	1301
		D	1303
		E	1316
		F	1320
		G	1322
		H	1323
		I	2004
pIrv-Neo-K5-LacZ[D]	tgK5-LacZ[D]	A	2178
pIrv-Neo-K5-LacZ[E]	tgK5-LacZ[E]	A	2193

hybridised to a radioactive probe. This was carried out according to the methods detailed in Sections 2.19, 2.21 and 2.22. The DNA samples included a negative control, a positive control which was the transgenic founder animal number 1591 DNA previously shown to contain the *tg*Act-LacZ^{A} transgene. The other DNA samples were isolated from tail tip biopsies of offspring generated by breeding the transgenic founder mouse 1591 with a male NIH mouse. A positive transgene band is seen in the DNA sample from the parent founder mouse 1591. The same band can be observed in six of the nine offspring DNA samples indicating that mouse 1591 passed on the gene to its offspring. Mouse 1313 another *tg*Act-LacZ^{A} founder mouse does not pass the integrated transgene on to any of its offspring.

The founder mice for the four *tg*K5-LacZ^{D} and *tg*K5-LacZ^{E} transgenic mice were also bred (Figure 36 shows a table of the transgenic lines bred). Only 1 out of 4 *tg*K5-LacZ^{D} founder mice passed the gene on to its offspring. One female founder did not breed and two produced offspring negative for the presence of the transgene. Three of the four *tg*K5-LacZ^{E} founder mice passed the genes onto their offspring but strangely only one of the lines showed expression in the mice bred from the founder although all the founder mice expressed the gene. PCR has now been developed to routinely analyse the offspring of transgenic animals for the presence of the *lac Z* transgene sequences (Figure 37).

3.4.3 Measurement of transgene expression in mice.

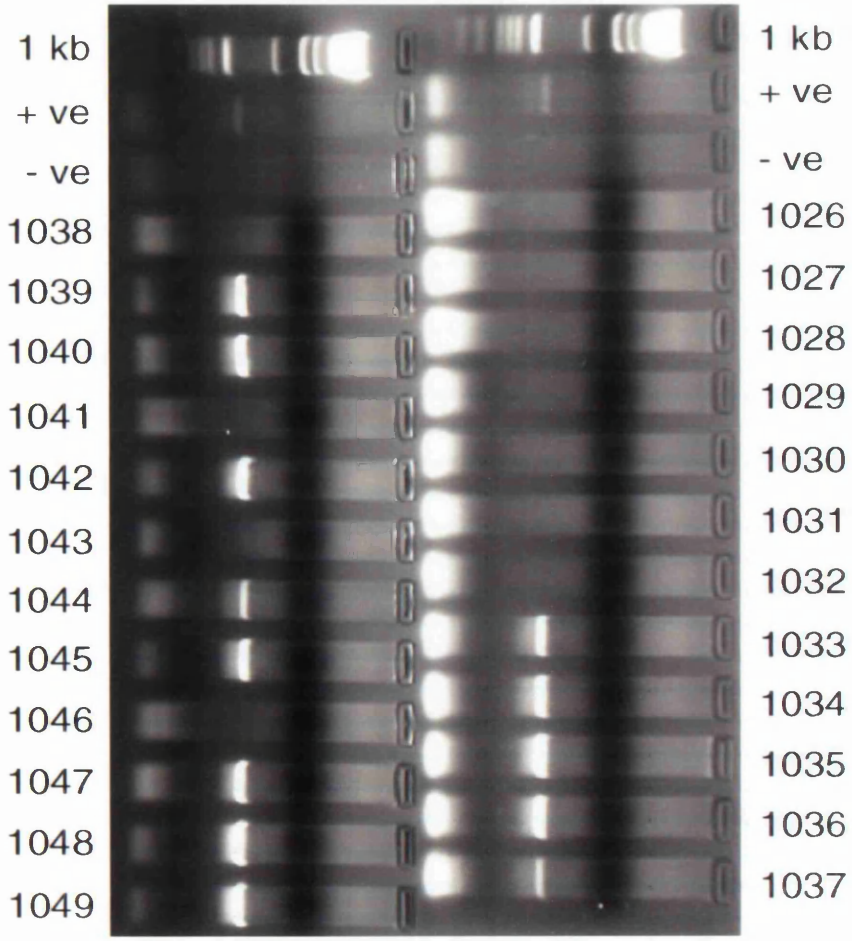
Before the activation of the marker genes could be attempted *in vivo* it is essential to check the expression levels of the transgenes introduced into the mice. The expression levels of two of the transgene constructs have been checked. The *tg*K5-LacZ^{E} construct contains the AAF-strategy control gene as has been shown in tissue culture to express high levels of β-galactosidase protein in epidermal cell lines maintained in tissue culture. The *tg*Act-LacZ^{A} transgene construct on the other hand contains the MNNG-strategy test oligo and therefore expresses no β-galactosidase protein. The expression levels of this construct had to be checked by measuring the *lac Z* RNA levels.

The RNA levels produced by the *tg*Act-LacZ^{A} transgene in eight lines produced by different founder animals were checked by ribonuclease protection. Figure 38 shows the results of this technique. The first analysis performed was to screen different tissues of an adult mouse of *tg*Act-LacZ^{A} line D. RNA samples were prepared from a variety of

Figure 37. Analysis of the offspring of transgenic founder animals by PCR.

PCR of genomic DNA samples obtained from tail tip biopsies of mice. PCR was performed with LacZ U2/LacZ L2 amplimers. The positive and negative controls are genomic DNA samples **(a)** Analysis of mouse samples 2490-2500 and 1001-1011. **(b)** Analysis of mouse samples 1026-1049.

(b)


1 kb
+ ve
- ve
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1 kb
+ ve
- ve
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037

(a)

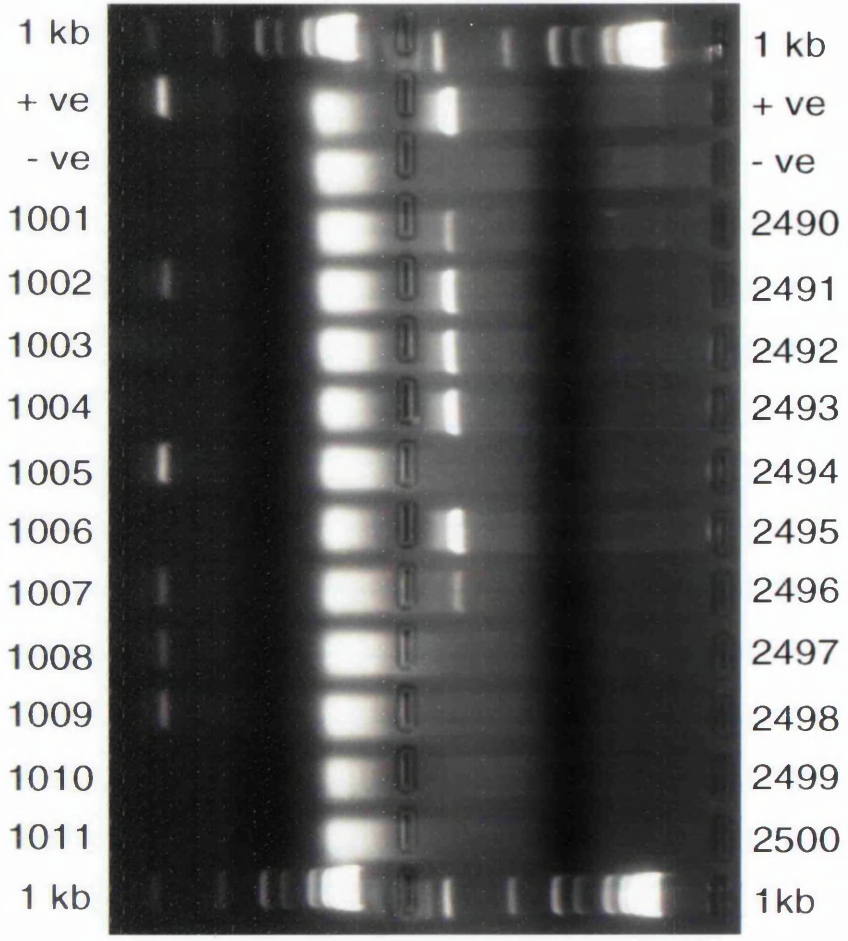

1 kb
+ ve
- ve
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1 kb
1 kb
+ ve
- ve
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
1kb

Figure 38. Expression of the *tg*Act-LacZ^A transgene in transgenic mouse lines.

RNAse protection assay performed with a 227 bp fragment of the *lac Z* gene. Radiolabelled ϕX174 markers were used to size the bands on the gel. The probe is approximately 300 bp in size. tRNA was used as a negative control. E6 cell line RNA, previously shown to express the *lac Z* RNA at high levels was used as a positive control. **(a)** RNA samples prepared from an adult mouse of *tg*Act-LacZ^A line D. No expression is detected in any of the tissues. **(b)** RNA samples from new-born mice from eight separate *tg*Act-LacZ^A transgenic mouse lines (A-H). Expression is detected in lines A, C, E, G and H.

(a)
φX174
Probe
tRNA
K5-Lac Z^E 6
Brain
Intestine
Kidney
Liver
Lung
Spleen
310
281
261
234
194
(b)
φX174
tRNA
K5-Lac Z^E 6
A3
B3
C2
C5
D5
E3
F2
G4
H5
310
281
261
234
194

the mouse tissues and analysed for *lac Z* expression by ribonuclease protection. No expression was observed in any of the tissues of this mouse which were analysed. RNA samples were then prepared from newborn mouse tissue from each of eight separate *tg*Act-LacZ^{A} transgenic mouse lines and analysed for RNA expression by ribonuclease protection with the 227 bp probe covering nucleotides 837 to 1063 of the *lac Z* gene. Three of the tested lines (B, D and F) show no expression of the *lac Z* RNA. Another two of the lines (A and C) show very little expression of the *lac Z* RNA. Three lines (E, G and H) show relatively high expression but this is still significantly lower than the expression of the control construct in the cell lines.

The levels of expression of the control construct *tg*K5-LacZ^{E} were measured by staining the skins of transgenic animals with the chromogenic β-galactosidase substrate, X-gal. The expression of this gene was confined mainly to the hair follicles in the transgenic animals. Figure 39 shows a picture showing the typical expression pattern of the gene in the epidermis of the mouse. Mice were routinely analysed for expression of the gene by analysis of ear punch biopsies stained with X-gal. Of the four founders generated with the *tg*K5-LacZ^{E} construct only three of them passed the gene onto the their offspring. Only one of these three passed on the expression of the gene to its offspring. The expression pattern of the gene in all four founder mice was restricted to the hair follicle. Treatment of TPA upregulated the expression level of the gene but did not widen the population of cells expressing the gene the gene expression was still confined to the hair follicle.

3.4.4 Activation of β-galactosidase expression in skin and brain tissue of Act-LacZ A transgenic mouse line by treatment with MNNG.

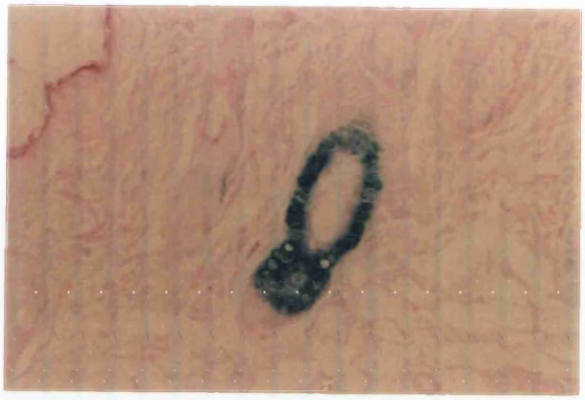
The activation of the *tg*Act-LacZ^{A} fusion gene construct by MNNG has already been demonstrated in a stably transfected epidermal cell line *in vitro*. The *tg*Act-LacZ^{A} transgene construct contains an oligonucleotide/*lac Z* fusion gene which is specifically designed to be activated by a mutation in the sequence of the oligonucleotide induced by the chemical mutagen N-nitro-N'-methyl-N-nitrosoguanidine (MNNG). The design and specificity of this reaction are discussed above in Sections 1.16 and 3.1.

Transgenic mice have been generated with the *tg*Act-LacZ^{A} gene in the hope it can be used as a marker to analyse the lineage of cell populations in the tissues of mice *in vivo*. The mice have been shown to express RNA from the gene construct. The mice should

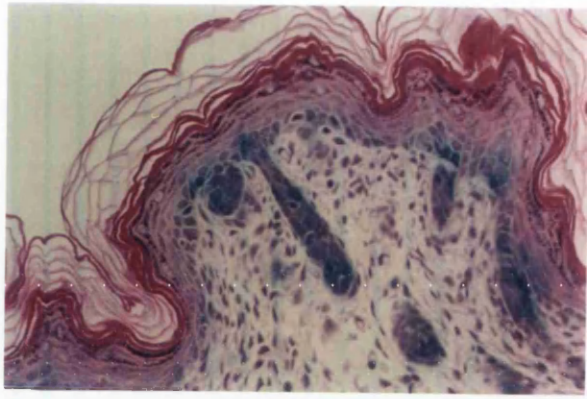
Figure 39. Expression of the *tg*K5-LacZ^{E} transgene in the skin of transgenic animals.

Skin from tgK5-LacZ^{E} transgenic mice. Stained with X-gal and counterstained with eosin. **(a)** The animal is six weeks old and activity is restricted to the hair follicles. Magnification 200×. β-galactosidase (blue-staining cells) clearly visible in the hair follicle of the transgenic mouse. **(b)** Skin from a newborn mice stained with X-gal, haemotoxylin and eosin. Magnification 200×. **(c)** Newborn mice stained with X-gal and eosin only. Magnification 100×. **(d)** Newborn mouse skin stained with X-gal and eosin only. β-galactosidase activity is clearly visible in the basal layer of the epidermis and the follicle. Magnification 200×. **(e)** Separate newborn mouse skin again displaying high activity. Magnification 200×. **(f)** High magnification of newborn mouse skin.. Magnification 400×.

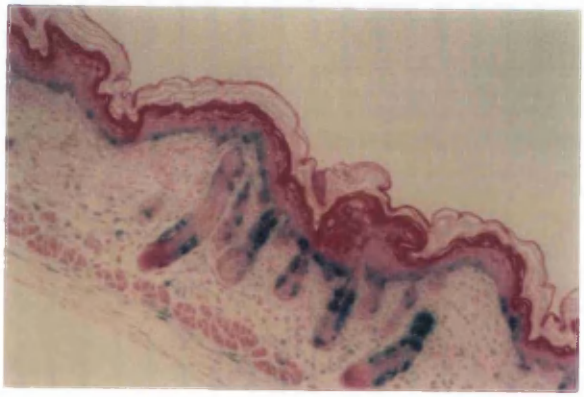
(a)



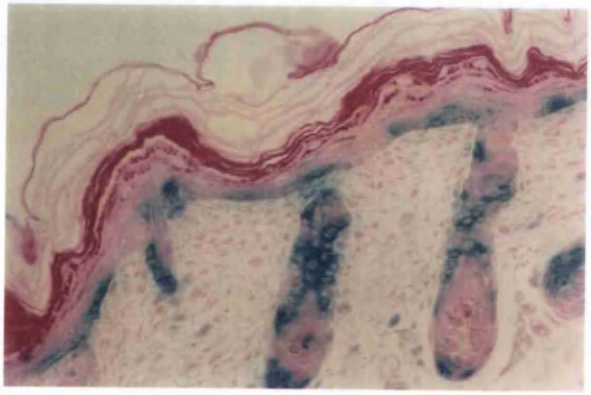
(b)



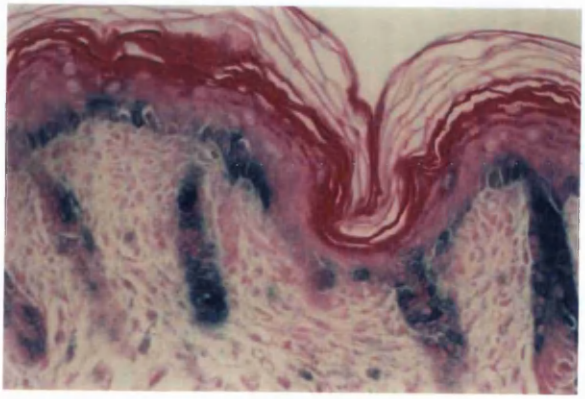
(c)



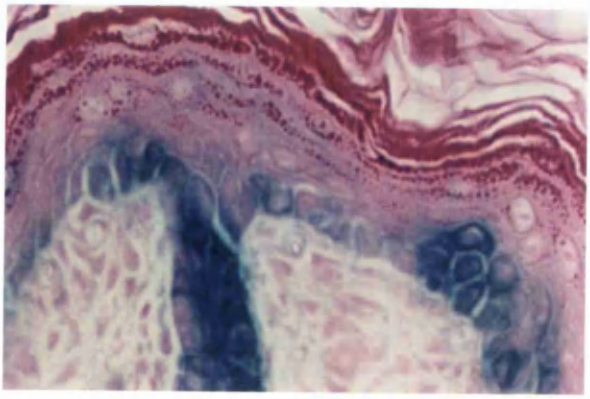
(d)



(e)



(f)



not produce any protein from this gene as it contains no ATG initiation sequence. *In vitro* the *tg*Act-LacZ^A gene produces RNA for the *lac Z* gene but no β-galactosidase protein. Exposing the mouse tissues to the chemical mutagen MNNG should produce mutations in the target sequences in the oligonucleotide. The specific mutations caused by MNNG should result in the production of an in-frame ATG sequence which will initiate translation of the mRNA and the production of the β-galactosidase protein.

Initial attempts to look for the activation of the marker gene in mouse embryos proved unsuccessful. Heterozygote female mice from the *tg*Act-LacZ^A were mated to heterozygote male mice from the same transgenic mouse line. Each female mouse was inspected twice each day for signs of a coital plug. When the female mice were plugged this was taken to be day 0 of pregnancy. On day 8 of pregnancy the female mice were weighed and injected with MNNG solution such that they received a dose equivalent to 25 mgkg^{-1}. Four days later the pregnant mice were sacrificed. The embryos were removed and stained for β-galactosidase activity with X-gal. Since the mouse litter was produced by mating two heterozygote mice the litter will be composed of homozygote, heterozygote and non-transgenic mice. The non-transgenic mice act as controls for the staining experiment as they should demonstrate no β-galactosidase activity. However no significant staining was observed in the tissues of the mouse embryos stained when inspected under a stereomicroscope. In addition 50 newborn mice treated by MNNG in the same way have been examined for activation of the marker gene in epidermis and intestine and no evidence of β-galactosidase expression has been discovered.

Subsequent experiments provided some good evidence to suggest that the marker gene system was working. Newborn mice from the *tg*Act-LacZ^A line H were inspected for spontaneous mutational activation of the marker gene by staining a range of tissues with the chromogenic β-galactosidase substrate X-gal. Brain tissue was stained from both control non-transgenic mice and transgenic mice was stained and compared. No staining was observed in the brain tissue of the non-transgenic control mice. Small patches of staining tissue were observed in the brains of the *tg*Act-LacZ^A transgenic mice. The staining cells could not be identified by inspection of the tissue under a stereomicroscope.

These results were confirmed by staining frozen sections of brain tissue from six week old mice from the transgenic mouse strain *tg*Act-LacZ^A line H. Newborn transgenic mice and control non-transgenic mice were treated by topical application of MNNG twice weekly for four weeks starting at three days of age (described in Section 2.30). After the MNNG treatment was stopped the mice were then left for two weeks in cages.

When six weeks old the mice were sacrificed. Frozen sections of brain tissue were cut and stained according to the methods described in Section 2.29. Although five transgeic and three control mice had been treated, resources were only available to get frozen sections cut from one transgenic and one control mouse brain.

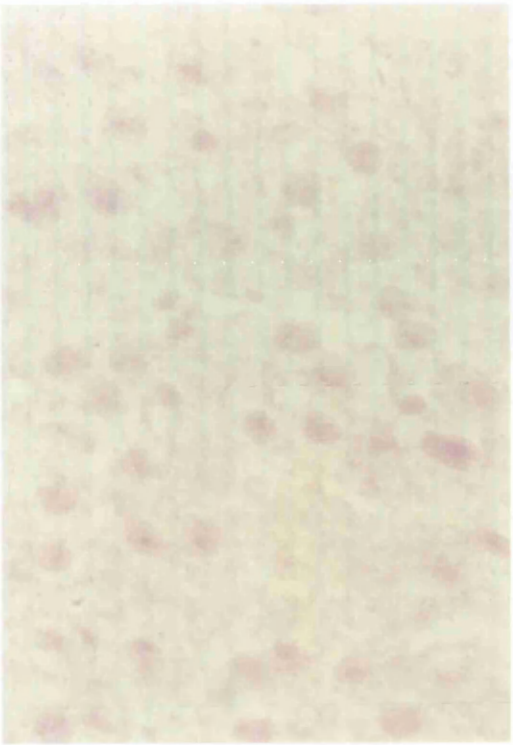
A total of 624 frozen serial sections were cut from the brain of one transgenic mouse which had been treated with MNNG. These frozen sections were all stained with X-gal for β-galactosidase activity before counterstaining with eosin and mounting. Upon inspection of the frozen sections under a light microscope a total of three potential colonies of cells were identified each containing a number of blue staining cells. Each of these colonies covered a number of frozen sections (22,12 and 14 respectively) and the colonies were separated by a number of serial sections in which no X-gal staining cells could be seen. Figure 40 shows the largest colony of β-galactosidase expressing cells located in the forebrain of a transgenic mouse, in a region of the brain overlying the olfactory bulb. This colony is visible in 22 consecutive serial sections.

No such colonies are visible in the brain of the non-transgenic control mouse. A total of 168 frozen serial sections of control mouse brain were checked and no sign of β-galactosidase activity was discovered. These sections were stained for β-galactosidase activity under exactly the same conditions as the control mouse.

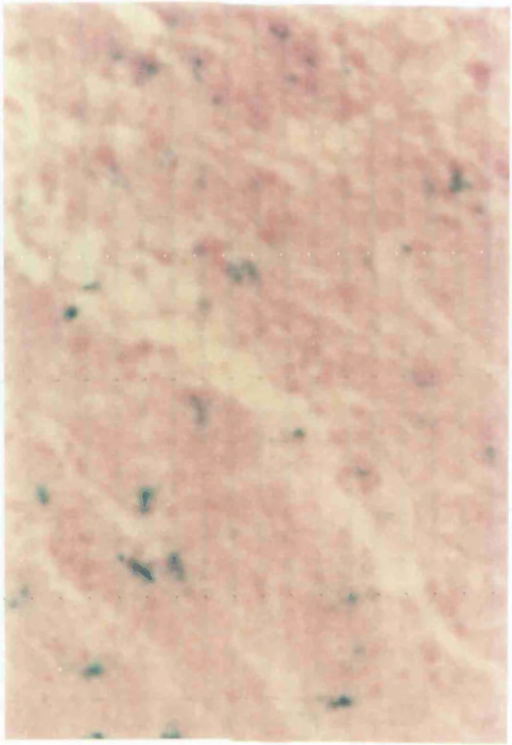
Figure 40. Activation of β-galactosidase activity in *tg*Act-LacZ^A transgenic animals following treatment with MNNG.

Photographs of mouse brain tissue sections stained with X-gal and counterstained with eosin. Magnification 400x. **(i)** Section of control mouse brain tissue. **(ii)** β-galactosidase positive cells in the brain of a six week-old *tg*Act-LacZ^A line H transgenic mouse. Magnification 200×. **(iii)** View of same colony in subsequent serial section. Magnification 200×. **(iv)** High magnification of lac Z positive cells in the brain. Magnification 400×.

(i) control mouse

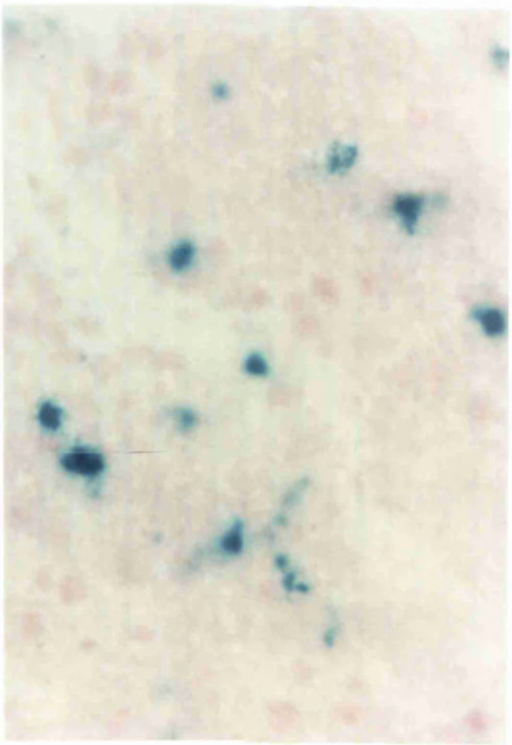


(ii) MNNG-treated mouse 1



(iii) MNNG-treated mouse 2

(iv) MNNG-treated mouse 3



CHAPTER 4.

Discussion.

4. Discussion.

4.1 Design of a transgenic marker system.

The requirements of a good cell marking system have already been discussed in Sections 1.12 and 3.1. The basis for the design of the system has been discussed in these sections so the discussion here is confined to a few highly speculative considerations the influence of which could not be predicted. The design of this project was extremely important for a number of reasons. The time taken to develop, clone and test the system *in vitro* and *in vivo* was likely to quite considerable. In order to ensure that the chances of a successful outcome great care was taken in the design of the project to maximise the chances of success.

The oligonucleotide system was conceived in order to attempt to maximise the chances of gene activation. There could be some degree of debate about how much of an increase in the rate of mutation the oligonucleotide would produce. Although there are eight potential target sites in the MNNG-strategy test oligo and nine target sites in the AAF-strategy oligonucleotide the frequency of mutation may not increase proportionately to the number of mutagen target sites. After all the target sites are all grouped in a very small part of the gene. Of course the opposite argument can also be made. That is that the grouping of a large number of mutagen target sites may act synergistically and attract the chemical mutagen more strongly increasing the mutagenic frequency by more than the number of target sites.

Although the mutagens that were used in the course of this project were chosen because they acted on defined target sequences very little is known about the effect of surrounding sequences on mutagens binding to DNA. As the oligonucleotide designs used in the course of this project were based on highly repetitive sequences it is highly likely that the oligonucleotides may take up unusual conformations *in vivo*. Indeed, it is known that stretches of purine pyrimidine repeats, which form the basis of the AAF-strategy oligonucleotide, are known to from Z-DNA. The influence of these potentially unusual DNA structures on the binding of chemical mutagens was largely unpredictable but is worth considering when interpreting the results of the activation experiments. It would be possible to test these ideas but it is not really worthwhile for the purposes of this project.

One other advantage of the design of this marking system is the flexibility. Care was taken to ensure that the marking system developed would be a general system which would be applicable to tissues other than the epidermis which is the main focus of the project. The use of a transgene as a cell marker instead of selecting a tissue specific gene already in place ensures that a multitude of tissues are able to be analysed by this marker gene approach.

Another aspect of the flexibility of the system is the use of the oligonucleotide as a target for gene activation. The oligonucleotide strategy has already been used to generate *lac Z* marker genes which can be activated by two separate types of chemical mutagen; MNNG is a point mutagen whereas AAF is a frameshift mutagen. If these chemicals had not worked to produce a functional marker gene the oligonucleotide could have been redesigned to take account of the mutational specificity of another chemical mutagen. This mutagen may have worked more effectively to activate the gene.

4.2 Construction of the *lac Z* fusion gene vectors.

The first step in this project was to clone a set of oligonucleotides in front of a *lac Z* gene in the vector pMC1871. During the cloning steps the oligonucleotides proved to be very unstable. The initial attempts to clone the oligonucleotides in front of the *lac Z* gene in the pMC1871 vector were unsuccessful because mutations in the sequence of the oligonucleotides. If the oligonucleotides were extremely mutable when they were cloned then they may be extremely mutable in the tissues of transgenic mice. This may result in the activation rate of the oligonucleotide by chemical mutagens being quite high.

This problem was eventually circumvented by cloning the oligonucleotides into a strain of bacteria which lacked recombination enzymes. This implied that the recombination enzymes must have been responsible for mutations occurring in the sequence of the oligonucleotides. The problem was not completely solved. There was still a high rate of deletions within the construct. The reason for these deletions remains to be adequately explained. Perhaps the most likely explanation is that recombination was taking place between some bacterial sequences and the initial part of the *lac Z* gene. Recombination taking place between endogenous bacterial sequences and the plasmid may have resulted in the 200 base pair deletion. It is possible that the cloning site around the beginning of the *lac Z* gene may be a frequently used sequence in cloning and may have caused the

recombination. If this is this case then is a tremendous coincidence that the deletions occur in the area surrounding the oligonucleotide elsewhere.

Perhaps the most likely explanation for the instability of the oligonucleotide is that the oligonucleotide sequences mismatch when annealed. When they are subsequently transformed into bacteria a rearrangement occurs which results in mutations in the sequence of the oligonucleotide or perhaps the short deletions in the plasmid. All the oligonucleotides are composed of highly repetitive sequences. It is possible that when annealed that when the oligonucleotides anneal they are slightly offset. The mismatch annealing of the oligonucleotides may result in short deletions produced in the oligonucleotide sequences when transformed into bacteria. This explanation does not account for the apparently specific 200 bp deletions which resulted when the oligonucleotides were cloned into pBRLacZ in the SURE™ bacteria. It would explain the deletions observed in the oligonucleotide sequences when they are ligated into the pMC1871 plasmid. It also accounts for the fact that, when the oligonucleotides already forming the fusion genes were cloned from the pBRLacZ vector into the pIrv-Neo-Act plasmid, the oligonucleotide sequences remained intact and no further mutations were observed in the sequence of the oligonucleotide.

Apart from those cloning steps which involved cloning the oligonucleotides or the *lac Z* fusion genes containing the oligonucleotides all the cloning reactions performed in the course of the project were achieved easily with high efficiency.

4.3 Transfection of the fusion genes into cell lines.

4.3.1 Expression of β-galactosidase from the LacZ^D fusion gene.

The transfection of the various oligonucleotide *lac Z* fusion genes also produced a few unexpected results. The first of these was the expression obtained from the transfection of the Act-LacZ^D vector. This vector contains the Act-LacZ^D oligonucleotide fusion gene under the control of the β-actin promoter. This fusion gene contains oligonucleotide D, the AAF-strategy test oligo, which contains an ATG sequence which is two base pairs out of frame with the coding sequence of the *lac Z* gene.

When this construct is transfected into fibroblast or epidermal cells the level of expression is very similar to that of the positive control constructs Act-LacZ$^{wild\text{-}type}$ and Act-LacZ^{E}. Both the positive control constructs contain functional *lac Z* genes and produce β-galactosidase protein at high levels in transfected cells. It seems likely that the Act-LacZ^{D} vector expresses β-galactosidase protein aberrantly from an ATG sequence subsequently located in the promoter sequence which is eighty base pairs upstream of the *lac Z* gene and which is in-frame with the coding sequence of the gene.

When the Act-LacZ^{D} construct was transiently or stably transfected into epidermal cells *in vitro* the β-galactosidase protein was expressed at the same levels and in the same proportion of the cell population. When the K5-LacZ^{D} construct was stably transfected into the epidermal cells the levels of expression was considerably lower than the control. This supports the hypothesis that the upstream ATG is supporting the expression of the β-galactosidase protein from the Act-LacZ^{D} construct.

The K5-LacZ^{E} construct which contains the LacZ^{E} fusion gene with the AAF-strategy control oligonucleotide E was stably transfected in C5O cells and the expression of the protein was analysed in ringclones of selected cell colonies. Four out of six selected cell clones expressed the β-galactosidase protein at reasonably high levels. Of these four expressing cell lines all contained detectable levels of mRNA from the *lac Z* gene. No RNA was detectable in the two clones which showed no detectable expression of the β-galactosidase protein.

Approximately two thirds of the K5-LacZ^{D} C5O epidermal cell ringclones which expressed the RNA in the did not express the protein. Of the remaining one-third which showed a degree of expression most of the clones (3 of 5) showed extremely low level. Only 2 of the fourteen clones which expressed the *lac Z* mRNA showed a significant expression of the protein (> 1% of the total cell population express the β-galactosidase gene). This means that of the K5-LacZ^{D} transfected cell ringlones, which express the *lac Z* RNA, approximately 14% expressed a significant level of β-galactosidase activity. In contrast, 100% of the K5-LacZ^{E} transfected ringclones which express the *lac Z* RNA also express significant levels of the β-galactosidase protein.

This level of expression obtained from the K5-LacZ^{D} transfected construct is still quite high but in contrast to the transfections with the Act-LacZ^{D} construct is considerably lower than the control. It is possible that the lower level of expression obtained during transfection of the K5-LacZ^{D} construct is caused by mutation or rearrangement of the gene or by an effect of the integration site. One other aspect of the expression of the β-

galactosidase expression that supports this conclusion is the level of expression in individual cells.

When β-galactosidase protein is expressed from the K5-LacZ^{D} it is expressed at the same level as the protein in the control constructs *i.e.* each individual cell which expresses the protein stains with similar intensity to the cells in the control population. If the oligonucleotide in the LacZ^{D} fusion gene supported inefficient translation of the gene then the expected expression pattern would be a low level of expression in most of the transfected cells. Most of the transfected cells would therefore be expected to stain more faintly with X-gal than the control construct transfected cells. This is not the observed pattern. The pattern in K5-LacZ^{D} transfected cells is that a lower number of cells express the gene than in control transfections but that the cells which do express the protein express a similar amount of protein to the control construct transfected cells.

As the expression of the β-galactosidase in the K5-LacZ^{D} in individual transfected cells is similar to that obtained in individual control transfected cells there is an implication that the expression in both cell populations is sustained by the same method. This common method could be that both cell contain a *lac Z* construct with an in-frame ATG codon. If this is the case there may have been a mutation in the K5-LacZ^{D} construct in the cell which brings the ATG in the oligonucleotide D into frame with the coding sequence of the *lac Z* gene.

If mutation within oligonucleotide D explains the reasonably high level of β-galactosidase expression from the K5-LacZ^{D} construct the construct may prove to be very useful. It suggests that the mutation rate of this construct is quite high. The levels of expression obtained from this construct are considerably higher than the expression levels detected with the transfection of the Act-LacZ^{A} construct. It is quite possible therefore that spontaneous mutations may occur *in vivo* in transgenic mice at a high frequency. If cells can be spontaneously marked then the relationships of cells in a tissue can be examined. For instance, this mouse could be used to determine whether cells in the hair follicle can divide to produce cells in the interfollicular epidermis.

A high spontaneous mutation frequency which activates a marker gene would have some potential advantage over a marker gene which is induced. To induce a marker gene by the system described in this thesis it would be necessary to treat the tissue of interest with a chemical mutagen. Other marker systems are even more disruptive. In order to mark epidermal cell effectively with retroviruses it would be necessary to scarify the tissue. Scarification of the tissue is extremely damaging to the tissue and the picture of

lineage relationships built up by cell marking in the damaged tissue by this technique could differ markedly from the normal relationships formed in untreated tissue. There is a possibility that this may disturb the behaviour of cell in the tissue and the picture of cell lineage obtained may be slightly distorted, although the chances of this being a significant problem are probably remote.

The one drawback of using spontaneous mutation is that it is impossible to determine when the cells became marked. It would therefore be impossible to do any work on the progressive restriction of lineages that occurs throughout development. If a treatment is used to induce cell marking the time of gene activation can be more or less assumed to be at the time of treatment. The half life of the chemical mutagen MNNG is only 30 minutes and therefore the time at which mutation occurs is likely to be very shortly after application of the mutagen(Bartsch et al., 1982).

A high spontaneous mutation rate also implies that the activation frequency induced by mutation may also be quite high. If the explanation of the high expression observed from the K5-LacZ^{D} construct is spontaneous mutation then the rate of mutation induced by treatment with the chemical mutagen N-acetoxy-AAF *in vivo* in the skins of transgenic mice may also be extremely high.

One way of attempting to determine if mutation is the cause of the expression of β-galactosidase in some cells in the ringclones electroporated with the K5-LacZ^{D} construct would be to isolate and sequence the gene from the expressing cells to look for mutations. This problem could have been approached by isolating cell transfected with the K5-LacZ^{D} construct. In one of the K5-LacZ^{D} electroporated ringclones approximately 30% of cells express the β-galactosidase protein. If these cell were plated out at low density so the cell grew as individual colonies it may be possible to isolate a pure clone of cells in which 100% of the cell population expresses the protein. The transfected construct, which would presumably be identical in these cells, could then be retrieved from these cells, cloned and sequenced or even sequenced directly.

This has not been done because any conclusions reached from such an experiment would not really generate any information useful to the progress of the project. One potential problem is that the cells may contain several constructs only one of which is expressing the β-galactosidase protein. This would confuse the analysis of the sequence. Even if the cells only contained the one copy of the gene and a mutation was identified in the sequence of the oligonucleotide the results would still be open to question. There is a

possibility that the mutation in the sequence may have occurred when the construct was grown in bacteria, during transfection or due to contamination with control sequences.

The mutation frequency of the AAF-strategy test oligonucleotide fusion gene $LacZ^D$ remains to be determined. This work with this construct was halted as there was not time to pursue both strategies during the course of the project and successful results had already been acheived with the MNNG-strategy marker gene.

4.3.2 Determination of the MNNG-activation frequency of the $LacZ^A$ gene.

No unexpected results were obtained during the transfection of the construct Act-$LacZ^A$ into cells *in vitro*. No significant levels of expression of the β-galactosidase protein were observed in any of the cells which had been stably or transiently transfected with the Act-$LacZ^A$ construct. This construct contains the MNNG-test strategy oligonucleotide A which consist of a run of $^{5'}GTG^{3'}$ trinucleotides with no ATG initiation codon. It was designed as a silent marker gene construct to be activated by the chemical mutagen MNNG and it was hoped that this construct would not produce any expression of the *lac Z* gene when transfected into cells. The control constructs Act-$LacZ^B$ and Act-$LacZ^C$ which contained marker genes, in which one of the run of $^{5'}GTG^{3'}$ had been replaced by an $^{5'}ATG^{3'}$ in order to initiate translation of the RNA, both produced high levels of β-galactosidase activity. This indicated that if the specific mutation in the oligonucleotide, designed to be caused by MNNG, was induced then the silent marker gene would become activated and mark the cell.

This hypothesis was tested in ringclones of cells which had been electroporated with the Act-$LacZ^A$ construct and selected on G418. These cells were treated with MNNG (described in Sections 2.16 and 3.3.8). MNNG was shown to activate the marker gene in these cells at a frequency of ~ 4×10^{-3}. The accuracy of this figure is open to dispute although it is probably approximately correct. Even if the figure was accurately determined it would probably have very little relevance to the treatment of mice *in vivo*. The concentration of MNNG used to treat cells *in vitro* is usually in the region of 1 μgml^{-1} whereas the concentration used for the topical treatment of mice *in vivo* is approximately 3 $mgml^{-1}$. This is a difference somewhere in the region of 3000 times and this fact alone is sufficient to indicate how the surrounding environment of cell influences their reaction to MNNG.

The mutation frequency that was established for this construct is open to question is mainly due to the difficulty of determining the proportion of the transfected cell population which express the *lac Z* RNA and are available to act as a targets for the gene reactivation. This is further complicated by the possibility that the cells may contain more than one construct. The calculation of an exact mutation frequency was not thought necessary. The figure obtained was likely to be approximately correct and the work that would have been required to calculate a more exact figure was impossible to do in the time scale of this project. Any mutation rate obtained *in vitro* may not be of much relevance to the situation *in vivo* where the environment of the cells is significantly different. The calculation of an exact figure for the activation rate *in vitro* was not critical for the progress of the project and would probably have had little bearing on the eventual aim of the project which is to apply the technology to the analysis of cell lineage on transgenic mice.

Since this project was started two studies have shown that genes can be activated by reverse mutation in mouse and human cells (Eccles et al., 1989; Schaff *et al.*, 1990). One of these studies demonstrated the reversion of an amber mutation in the mouse *H-2K^b* gene in response to ultraviolet irradiation. In this system revertant were induced at a frequency of 3×10^{-6}. The revertant clones were sorted by FACS analysis and sequence to show that the reactivating mutation was the expected mutation.

The other study was more comparable to the this project. The human cell line HT 1080 *APRT*$^-$ was transfected with the mouse *APRT* gene inactivated by an A to G transition. This was shown to be reverted by exposure to the chemical mutagen ethyl methanesulphonate (EMS) which had previously been shown to cause G to A transitions in both bacteria (Coulondre and Miller, 1977) and mammalian cells (Ashman and Davidson, 1987; Greenspan *et al.*, 1988; Lebkowski *et al.*, 1986). This reversion was induced at a frequency of $4 - 7.5 \times 10^{-4}$ in the cell lines transfected with the mutant construct. The construct was not reverted by treatment by 2-aminopurine (2AP) which has been shown to cause primarily A to G transition mutations (Caras *et al.*, 1982).

Both the previous study and this project have used chemical mutagens to specifically revert a mutation in a gene. The reactivation frequency obtained in this study (4×10^{-3}) is slightly higher than the mutation frequency obtained in the *APRT* gene in response to EMS (7.5×10^{-4}). The *APRT* gene only contained one target site for the chemical mutation whereas the *lac Z* gene used in this study contained nine potential target sites. Eight of the sites had been engineered to take account of the mutagenic sequence specificity of MNNG and had a guanine residue situated immediately 5' to the target

guanine residue (Burns *et al.*, 1987). The mutation induced in the APRT gene took no account of any sequence specificity of EMS-induced mutation. It is therefore not surprising that the mutation rate obtained in this study is slightly higher than that obtained in the APRT gene.

Both these studies (Eccles et al., 1989; Schaff *et al.*, 1990) were performed in such a way that the reverted cells could be selected either by FACS analysis or by drug selection and the gene sequences retrieved and sequenced. An ideal method of testing gene reactivation with the oligonucleotide in this system would have been to use the oligo to activate a silent drug resistance gene, such as *APRT*, in the same fashion as it has been used to activate *lac Z*. This would have allowed large numbers of cells to be treated with MNNG and screened for reversion. It would also enabled the recovery of the selected genes and these could then be analysed for sequence specific changes. This though would have required the construction of another series of constructs which would not have been possible in the time frame of this project. The gene used in this study *lac Z* does not easily allow selection or FACS analysis, although there is a fluorescent β-galactosidase substrate available (Nolan *et al.*, 1988; Karttunen and Shastri, 1991). The usual assay for the β-galactosidase protein is a histochemical stain which first involves fixing the cells before they can be stained.

4.3.3 Instability of the transfected *lac Z* sequences.

The analysis of the Act-LacZ^A constructs also produced another interesting result. Again the *lac Z* fusion genes seemed to be unstable. Southern blot analysis of the Act-LacZ^A electroporated ringclone cells suggested that the construct may be unstable. Of the 12 electroporation ringclones analysed only four of these gave the appropriate *lac Z* bands when DNA samples were analysed by Southern blot. Three of the clones had no detectable *lac Z* bands at all and a further five of the clones contained bands which were of an inappropriate size. All these clones contained the *neo* resistance gene which confers resistance to G418 when essentially the same blot was probed with a fragment of the *neo* gene.

Most of the transfected constructs were expected to remain intact and it was surprising that only four out of twelve did. The other ringclones exhibited rearranged *lac Z* bands or no bands at all. There are a number of possible explanations for the apparent rearrangement of the *lac Z* genes in these ringclones. The deletions of the *lac Z* gene

from the constructs could have occurred upon integration of the constructs into the cellular genome. The selection is based solely on the resistance of the cells to the antibiotic G418 and hence on the expression of the *neo* gene. The loss of the *lac Z* gene upon integration would confer no selective disadvantage on the cells and there is no necessity to keep the *lac Z* sequences. It is also possible that the oligonucleotides were responsible for the construct instability.

Another potential explanation of the deletions could be that the selection in G418 placed too strong a pressure on the cells. Although the concentration of G418 (200 μgml^{-1}) had been used previously to select C5O cell populations *in vitro* it is possible that the selection was too strong. This may have resulted in the deletion of the *lac Z* sequences in an attempt to maximise the expression of the *neo* gene. A much higher selection level was placed upon the transfected fibroblasts which survived in concentrations of up to 800 μgml^{-1} to 1 $mgml^{-1}$ of G418. The fibroblasts seemed to cope perfectly well in the higher antibiotic concentrations although the level of expression of the construct as evidenced by staining for β-galactosidase activity in these cells was higher than the expression levels obtained in the C5O epidermal cell line. This suggests that the construct is quite capable of generating resistance in cells to higher levels of G418 than used to select. This hypothesis could be tested by lowering the concentration of G418 used to select the C5O electroporation ringclones. One of the cell lines was useful for testing the activation of the fusion gene with MNNG and hence this experiment would have been of little more than purely technical interest. Any subsequent transfections performed with these cells may benefit from a titration of the G418 concentration.

Despite these problems a Act-LacZ^{A} electroporation cell line has been used to demonstrate that the oligonucleotide/*lac Z* fusion gene cell marking strategy does work *in vitro*. This is certainly true with the MNNG-strategy construct and is also likely to be true with the AAF-strategy construct. Even if the AAF construct was no successful the system designed is sufficiently flexible to allow new oligonucleotides to be designed and used instead of the AAF construct. The general system has been demonstrated to work with MNNG and there is no reason why new oligonucleotides could not be engineered to take account of the specificity of other chemical mutagens and so be used to generate novel marker fusion genes.

4.4 Transgenic mice.

4.4.1 Demonstration of activation of marker genes *in vivo*.

The results obtained with the transgenic mouse lines are very encouraging although the project has not progressed far enough for the lineage analysis on mouse skin to begin. The most important result so far achieved in the course of this project is the demonstration of colonies of *lac Z* positive cells in the brain of MNNG-treated transgenic animals. This demonstrates that the marker genes can be activated *in vivo* as well as *in vitro*. Whether these colonies were induced by the exposure to MNNG is admittedly open to question. Even so this result still indicates that the basic design of the marker gene system is sound and should be useful in cell lineage analysis in transgenic mouse tissues.

Figure 40 shows a colony of β-galactosidase staining cells visible in the forebrain of a six week old Act-LacZ^A transgenic mouse which had been treated with MNNG. The transgenic mouse line used in this experiment, *tg*Act-LacZ^A line H, had previously been shown to express RNA for the *lac Z* gene by ribonuclease protection.

The β-galactosidase positive cells in Figure 40 have been proposed to be a colony of neurons in the white matter tract above the olfactory bulb in the mouse forebrain (J. Price, personal communication). No such colonies were visible in the brains of non-transgenic animals which had been treated with MNNG. A total of 168 frozen sections from a control mouse brain were examined for *lac Z* activity. Neuronal cells have a very low intrinsic activity towards the chromogenic β-galactosidase substrate, X-gal and therefore show very little background staining. The colony is also visible in the same spot fourteen consecutive serial sections. These three pieces of evidence suggest that the colony could not be a random effect of background staining. This suggests that the β-galactosidase staining colony is likely to have been caused by a mutation which has arisen in the progenitor cell of the cells in this colony.

The division of neuronal cell in the brain is largely completed by the birth of the animal and it may be unlikely that a large colony of neuronal cells could have been marked by a mutation after birth. This suggests that the MNNG treatment may not have been responsible for the activating mutation which switched on the transgene. The mutation may have arisen spontaneously during the development of the mouse. This is not

entirely unexpected as the treatment of MNNG was applied topically and the amount of MNNG which reached the brain having first to pass through the skin may have been minimal. Furthermore only three such colonies were seen in the brain of the with the one pictured being by far the largest. This alone suggests that the mutation rate must have been very low. Intraperitoneal injection of MNNG may be a more effective route to expose the brain cells to higher concentrations of the chemical mutagen.

Disappointingly no β-galactosidase positive cells were observed in the skin of the same transgenic animals which had been treated with MNNG. The likely reason for this is expression levels of the transgene may be very low in the skin. The expression of the Act-LacZ^{A} construct is quite low *in vitro* and this may also be the case for the *tg*Act-LacZ^{A} mice in vivo. It is quite possible therefore that the transgene underwent the appropriate mutation but went undetected because the promoter did not support expression of the RNA for the gene.

This β-actin promoter has already been used to make transgenic mice and is capable of giving widespread expression of a *lac Z* reporter gene during embryogenesis (Beddington et al., 1990). This expression quickly declines after birth and is very patchy and often totally switched off in adult mice (R. Beddington, personal communication). It is therefore possible that the mice express sufficiently highly during embryogenesis for the cell marking to work. The one experiment carried out on embryonic mice did not, however, produce any marked cells. This experiment was not repeated as it was more important to acheive cell marking in adult mice for lineage analysis and because there was no suitable control available.

There is still good evidence from the generation of colonies in mouse brain that marker gene system works *in vivo* and should be useful for the analysis of cell lineages in the tissues of transgenic mice, *in vivo*.

4.4.2 RNA expression levels from the β-actin and K5 promoters.

Although the constructs generated so far may prove useful both sets of constructs (the *tg*Act-LacZ^{A} and the *tg*K5-LacZ^{D}) demonstrate the same potential problem - the level of expression from the promoter is not very high. When the level of expression from these promoters was analysed in tissue culture, the levels of expression were barely detectable by northern blot analysis. The expression level from the BK III (K5) promoter was

slightly higher in the epidermal cells than the expression from the β-actin promoter in the same cells.

The low expression of the silent marker genes may have been predictable because evidence exists to suggest that RNA molecules which are not translated, are quickly degraded (Atwater et al., 1990; Peltz and Jacobson, 1992). As the marker gene constructs have no ATG sequences they may be unlikely to bind to ribosome and certainly should not be translated.

There was very little evidence for this occurring in cells in transfected cells in tissue culture. The control and test construct mRNA levels appear to be reasonably similar, although the control construct expression levels may be slightly higher. This suggests that the silent marker gene RNA is not subject to extremely fast turnover which reduces the apparent expression levels. The low level expression observed from the promoters is probably therefore intrinsically low. This correlates with the low levels of expression detected in the tissues of the *tg*Act-LacZ^{A} transgenic mouse lines. No expression was detected in three of these lines with a further two expressing extremely low levels (Figure 38). Three of the lines assayed did express detectable amount but these did not appear to be very high in comparison to the levels expressed by the cell lines.

The β-actin and K5 promoters used for the generation of the first transgenes used in this project may not prove very useful in the long term, for lineage analysis in the mouse. The levels of expression obtained from these promoters seem extremely low in comparison to the endogenous gene expression levels in the cell lines transfected with the Act-LacZ and K5-LacZ constructs. The expression of the K5 constructs in mouse skin also seem more or less limited to the hair follicle with only the occasional interfollicular cell expressing the β-galactosidase protein. It may therefore be necessary to use alternate promoters to express the marker genes at high levels for cell marking in different tissues.

4.5 Stem cells and Carcinogenesis.

During the course of this project a number of advances have been made towards the understanding of the organisation of the proliferative hierarchy in the skin. These studies have concentrated mainly on the organisation of cell proliferation and proliferation in the hair follicles.

A study by Cotsarelis *et al.* (1990) demonstrated that there were label-retaining (*i.e.* slowly cycling) cells located in the bulge area of the murine hair follicle. These authors concluded that this area may be the location of the stem cells for the hair follicle. The conclusions of this study were reinforced by another recent report (Koyabashi et al., 1993) in which the cells with the highest *in vitro* clonogenic potential in rat vibrissae were also localised to the bulge area, also a site of label retaining cells (Lavker *et al.*, 1991). Earlier work on the rat vibrissa had demonstrated that the lower part of the hair follicle (the matrix cells of the hair bulb) was not important for whisker growth in response to stimulus from the dermal papilla (Oliver, 1966; Oliver, 1967a, b).

More recently two conflicting reports have localised cells with a high proliferative potential to the outer root sheath of human hair follicles (Yang *et al.*, 1993; Rochat *et al.*, 1994). The first of these studies localised the highly clonogenic cells to the upper part of the follicle at or around the bulge area. Human hair follicles have a prominent bulge area during morphogenesis of the follicle (Holbrook, 1979; Holbrook *et al.*, 1993) but this area diminishes soon after birth (Holecek and Ackerman, 1993). The latter of the studies localised the cells to an area of the outer root slightly lower than the bulge area at the insertion site of the arrector pilli muscle. Despite this conflict both results show that highly proliferative cells are located in the outer root sheath of the follicles and very few proliferative cells are located in the hair bulb.

The hair follicles have also been widely implicated as the origin of skin tumours in carcinogenesis studies. A series of grafting experiments performed by Billingham *et al.* (1951) suggested that only when carcinogen-treated skin grafts included the hair follicle would they give rise to tumours. This experiment was subsequently repeated using two strains of mice with different histocompatibility genes (Steinmuller, 1971). A thin skin graft from one strain was grafted onto carcinogen treated dermis of the other strain. Tissue typing of the resultant skin tumours demonstrated they had arisen from the recipient strain, implying that the tumours were derived from the follicles. When tumours are induced in mouse skin with ionizing radiation few tumours were induced when the radiation penetrated less than 0.3 mm, when the penetration of the radiation was deeper (0.4 - 1.4 mm in depth) numerous tumours were produced (Albert *et al.*, 1967a,b,1969). The number of tumours produced was directly proportional to the amount of damage inflicted upon the hair follicle.

One interesting anomaly in the hypothesis, that hair follicle cells are involved in the generation of a large proportion of skin tumours, has recently been challenged. It has previously been shown that dividing cells are most susceptible to tumour initiation

(Bowden and Boutwell, 1974; Hennings et *al.*, 1978). However, it has been generally accepted that skin tumour initiation is best performed when the animals are in the telogen (resting) phase of the hair cycle to maximise the tumour yield (Stenbeck *et al.*, 1981). In the telogen phase the proposed stem cells in the bulge region are not dividing, these cells only come into cycle briefly during the anagen phase of the cell cycle (Lavker et al., 1993). However a recent report (Miller *et al.*, 1993) contradicts the hypothesis that telogen initiation is more effective than anagen initiation. The authors showed that the highest yield of tumours was produced when animals were initiated in early anagen (growing) phase by either one or two-stage chemical carcinogenesis treatment protocols.

Also recent results obtained in this laboratory suggest a role for hair follicle cells in the generation of malignant skin tumours. The work performed with the K5-LacZ^{E} positive control mouse suggests that this 1.4 kb shortened version of the K5 promoter targets expression of the gene specifically to the outer root sheath of the hair follicles in adult mice(Figure 39). When the same promoter was used to drive expression of a mutant H-*ras* gene it was capable of producing spontaneous malignant tumours at a high rate (Brown *et al.*, in preparation). This did not happen when the H-*ras* was expressed in the suprabasal layers of the interfollicular epidermis under the control of the K10 promoter (Ballieul *et al.*, 1990). Although these animals developed papillomas at sites of wounding the tumours rarely progressed to malignancy. Taken together these results suggest that when the mutant H-*ras* is expressed in the stem cells, the resultant tumours are more likely to progress to malignancy.

The recent data described above demonstrate the importance of hair follicle cells in the maintenance of the follicle and as target cells for the generation of carcinogenesis. While this population of cells has been implicated in the maintenance of the follicle there is still relatively little information on the relationship between the stem cells of the follicle and other epidermal structures such as the interfollicular epidermis and the sebaceous gland. It is well known that after the destruction of the epidermis by abrasion (Argyris, 1976) or severe burning (Lavker et al., 1993) that replacement of the epidermal cells can occur from the hair follicles. No label retaining cells have ever been seen in the sebaceous glands suggesting that these structures do not contain a unique stem cell population. In addition there is evidence that glabrous cell can be recruited to form hair by the dermal papilla. This suggests that the differentiation state of the epidermis can be altered.

All these facts suggest the existence of a multipotential stem cell population in the epidermis. This is contrast to the usual proposed model of the epidermis which dictates that the epidermal stem cell is unipotential (*e.g.* Sell and Pierce, 1994). The lineage

analysis system described in this thesis should be ideal for investigating the relationship between the different cell types in the epidermis. The extent to which the cells of the hair follicle can contribute to the interfollicular epidermis and the sebaceous gland could be determined. These structures may be descended from common stem cells or different sets of lineage restricted cells. The relationship between these different cell types is one of the most fundamental and important questions yet to be addressed in the study of the epidermal cell biology. The answers to these questions could have important implications both for the biology of the skin and in the development of disease states such as cancer.

4.6 The Way Forward.

4.6.1 The immediate prospects.

The β-actin promoter used in the Act-LacZ^{A} construct has previously been used to generate transgenic mice. The expression of the wild-type *lac Z* transgene under the control of this promoter is widespread during mouse development but that expression falls off during post-natal growth. Mice carrying the Act-LacZ^{A} transgene should be very useful for the study of lineage divergence during murine development.

The immediate priorities with the AAF-strategy oligonucleotide *lac Z* fusion gene is to test the activation of the gene in the cell lines and in the transgenic mice. Expression of the gene in the epidermal cell lines could be tested by exposing these cell lines to the chemical N-acetoxy-AAF. Characterised cell lines are already available for this work.

Alternatively the construct K5-LacZ^{D} could be electroporated into liver cell lines. Oval cells, the proposed liver stem cells, can now be grown in culture and would represent an ideal system to test this construct (Brill *et al.*, 1993). Unlike epidermal cells, liver cells are able to activate the AAF chemical and -AAF adducts will be produced on the DNA of these cells. These -AAF adducts have already been shown to cause a two base pair deletion in a series of GC base pairs, the sequence of the oligonucleotide in the K5-LacZ^{D} marker gene and should therefore activate translation of the β-galactosidase protein. Stably transfected cell lines could be used for testing of the AAF-strategy

fusion gene in the same way that the C5O epidermal cell ringclones allowed the testing of the MNNG-strategy fusion marker gene.

There is considerable interest in the organisation and location of potential stem cells in the liver (discussed in Section 1.5). This construct would be extremely useful for lineage analysis in this tissue. This would require transgenic mice to be generated with the $LacZ^D$ oligonucleotide/*lac Z* fusion gene under the control of a promoter which is highly expressed in liver cells. These mice could then be exposed to the AAF chemical *in vivo* both during development of the liver in embryos and maintenance of the liver in adult mice.

Transgenic mice at different ages and stages of development would be exposed to AAF at varying concentrations and left for times varying from a few days to a few months. Frozen sections of the liver could then be stained with the chromogenic β-galactosidase substrate X-gal. The clones of *lac Z* cell could be identified and a picture of lineage of cells in the liver could be built up piece by piece through the analysis of clones from a large number of individuals. The response of liver stem cells to injury of the tissue could also be assessed.

It will be possible to assess the capacity of the developmental stem cells to contribute to the formation of different tissues and the of different cell types. The developmental stem cell may work in a different fashion to the stem cell in renewing tissues. In order to study the stem cells of renewing tissues it will be necessary to place the marker genes under the control of a promoter which can sustain high expression of a transgene in adult animals. It would be useful if such a promoter induced a high level of expression in the epidermis and other tissues of the mouse to enable the spectrum of lineage analyses to be broadened.

4.6.2 Cell lineage analysis and mutagenicity testing.

The control *lac Z* fusion genes should be placed under the control of a wide variety of promoters which may induce widespread expression of the transgenes in the tissues of the transgenic mouse. A keratin fusion gene promoter which includes elements of both the K5 and K6 gene is available and could be used to increase the expression of the marker genes in the epidermis of transgenic mice. Other possible promoters would include the H-ras gene promoter which has already been characterised in the laboratory

and shown to be widely expressed in a number of different tissues. Promoters of housekeeping enzymes such as 3-phosphoglycerate kinase (*Pgk-1*), may also be used to promote expression of the marker gene. Recently a promoter trapping experiment in embryonic stem (ES) cells (described in Section 1.2) has identified a gene which is widely expressed throughout adult mouse tissues. The promoter of this gene could be cloned in order to generate high expression of the marker genes in the tissues of transgenic mice.

Furthermore the mice generated in the course of this promoter trapping experiment may be used in a complementary approach to lineage analysis, along with the transgenic marker gene system developed in this laboratory. If the expression of the *lac Z* gene is truly ubiquitous mutagenesis of the *lac Z* gene would mark cells. A mutation occurring in the sequence of the gene would prevent expression in single cells and clones of cells derived from the original marked cells. These could be detected amongst the background of cell in the tissue should also express the gene (see Section 3.1). This does depend critically on expression of the gene being maintained at a constant level throughout the population. The promoter is maintained in its normal location, expression of the transgene may be maintained more efficiently than expression of a transgene under the control of a promoter located at a random position throughout the genome.

If the marker genes are expressed at high levels in the tissues of transgenic mouse the analysis of cell lineage both during development and the maintenance of tissues in adult mice will be able to progress. The mutation frequency of the transgene will be established. Obviously this will be restricted in a similar fashion to the work with the cell lines. Both the number of expressed marker gene constructs and the percentage of cell in the population that express the marker transgene will have to be determined to allow calculation of the mutation frequency. This work can be performed on mice containing the test oligos for both the MNNG and AAF oligo/*lac Z* fusion gene strategies. The mice containing these fusion genes will be treated with different concentrations of the appropriate chemical mutagen and will then be left for times varying from a few days to months. The number of resultant clones can then be determined by staining the tissues with the *lac Z* stain, X-gal. Approximate reactivation frequencies could be calculated by comparison of the clone numbers to the number of cells in a tissue which have be analysed for the presence of the stain.

It would be of interest to breed these marker gene mice with mice in which the *p53* gene has been inactivated by homologous recombination. These mice have already been used

in the laboratory in studies of tumour formation. These studies demonstrated that *p53* was involved in the benign to malignant conversion in chemically induced tumours of mouse skin (Kemp *et al.*, 1993). There is a possibility that the *p53* gene influences the development of tumours by increasing the rate at which the cells which do not express this protein acquire novel mutations. Whether this phenomenon occurs by increasing the rate at which the cell will acquire point mutations or whether only the rate of gross gene rearrangements is open to question. As the marker gene constructs developed during the course of this project should allow the estimation of mutation frequencies, the breeding of the marker gene transgenic mice with the *p53* knockout mice should allow the influence of the *p53* gene of the rate of point mutation induced by chemical mutagens to be established. In fact breeding of the *p53* knockout mice with the Act-LacZ^{A} transgenic mouse has already started.

The mutant gene constructs could also be used form the basis of a mutagenicity-testing system *in vivo*. As it requires a specific mutation to activate the gene, if activation of the gene is observed in the tissues of a mouse the predicted specific mutation must have occurred. Substances could therefore be tested to see if they are capable of inducing different types of mutation in the tissues of transgenic mice. A similar system already exists and utilises transgenic mice carrying the *lac I* gene in phage λ sequences (Provost et al., 1993; Shephard et al., 1993). After exposing the mice to the substance to be tested, the phage λ sequences must then be extracted from the tissue of interest and transformed into *E. coli*. Any mutations in the *lac I* gene resulting from this must then be sequenced.

There would be two significant advantages of using the oligonucleotide fusion gene constructs described in this thesis. Firstly as the marker genes are activated by a defined specific mutation, the mutation caused by the substance tested is determined without sequencing the gene. Secondly the pattern of mutations caused by the substance can be identified as the assay for mutant cells involves examining intact tissues for β-galactosidase activity. The tissue can be sectioned and stained with X-gal or stained as a wholemount. The *lac I* system requires the tissue to be disintegrated before analysis of the results and any significant pattern of mutagenesis is therefore lost. The creation of this system would require the design of a few new oligonucleotide/*Lac Z* fusion genes to cover the range of potential mutations which could occur.

There would be a couple of minor drawbacks of using this system. The system would require the expression of the gene *in vivo*. It would therefore require the use of reliable promoters in transgenic mice which allowed expression in the tissue of interest at a high

level. This may require the creation of several lines of transgenic mice with the marker genes under the control of different promoters. The *lac I* gene in the system described above is silent and expression of the gene in transgenic mice is irrelevant. It is debatable whether or not the frequency of mutation would be easily calculable in the tissues of transgenic mice *in vivo*. This would require the estimation of the number of mutant cell arising within the population of cells tested. The estimation of mutation frequency in this way may be laborious and inaccurate. Despite these reservations the development of this system is probably still a worthwhile project as it reduces the work involved in isolating phage λ constructs and sequencing large parts of the *lac I* gene.

4.6.3 The overall goals of the project.

A number of other potential uses of these mice in the longer term exist. Once the normal patterns of cell renewal in the epidermis have been established it will be possible by the use of the techniques described above to look to see how these patterns are disturbed by the expression of an oncogene or growth factor in the same tissues. One possibility would be to breed these mice with mice expressing the H-*ras* oncogene under the control of the short 1.4 kb K5 promoter. These mice show an interesting pattern of abnormalities in the epidermis and hair follicles and develop epidermal tumours spontaneously as adults. The alteration of the pattern of stem cell division could be assessed in such doubly transgenic animals (containing a marker gene and H-*ras* oncogene). This could be done by treating the mice with chemical mutagens to activate the marker gene and look at the how the clones produced are altered when compared to animals containing the marker gene construct alone and which are essentially normal.

This type of analysis could performed not only with the H-*ras* oncogene but also with transgenic mice containing other oncogenes or growth factor. One of the most interesting aspects of this analysis would be that it allows the effect of oncogenes and growth factors on the behaviour of single cells as opposed to the gross effects on the whole tissue. These cells are obviously not excluded from the influence of the surrounding cells in the tissues but still the analysis of the behaviour of single cells may be expected to reveal some interesting aspects of the influence of an oncogene or growth factor on a particular tissue.

The construction of better transgenic models for oncogenesis could be another long-term aim of this project. One of the drawbacks of all present transgenic models of

tumour development is that the mutant oncogene or growth factor is switched on in many or all cells of the tissue at the same time, even when under the control of a highly tissue-specific promoter. The system developed for the activation of marker genes could be used instead to activate an oncogene such as H-*ras*. This could be achieved by substituting an oligonucleotide in place of the first few amino acids of the oncogene protein. The oligonucleotide could then be used to activate translation of the gene in the same fashion as it has been used to activate the *lac Z* marker gene for the purposes of this project (described in Section 3.1). The oncogene will not be expressed until a mutation in the sequence of the oligonucleotide occurs either spontaneously or by application of a chemical mutagen. The AAF-mutation strategy oligonucleotide could be used to specifically activate oncogene introduced into the skin. A mutant *ras* oncogene cloned behind the AAF-strategy oligonucleotide D would not be expressed as the ATG in the oligonucleotide would be out of frame with the coding sequence of the *ras* gene. It could then be specifically activated by treatment with AAF which will not activate the endogenous H-*ras* gene. This should activate the oncogene in individual cells which may help create a more accurate model of the true situation of oncogenesis *in vivo* where oncogenes are activated in discrete individual cells.

Another potentially intriguing use of this system would be to combine the cell marking aspect of the project with other genes which would influence the behaviour of the marked cells. This type of analysis has been used to great effect in *Drosophila* to establish some of the most fundamental rules about the development of this organism. Cell markers genes were combined with mutations at the *Minute* locus which influence the rate of growth of individual cells to establish the existence of developmental compartments cells in both the ectoderm and mesoderm of the fly. The combination of the marker gene approach with the activation of other genes which affect cell behaviour in the same single cell at the same time could be of enormous benefit to the study of developmental biology in the mouse.

The approaches which have been described in this thesis are mainly aimed towards the development of cell lineage analysis in the skin. Although this approach has been developed for use in the skin there are a wide variety of murine tissues which have been proposed to contain stem cells. The study and understanding of the function of stem cells in these tissues would benefit by the application of cell marking to these tissue. One of the most important aspects of the system developed for the purposes of this project is that it can be applied to any tissue in the mouse. This may require the use of a specific promoter to target expression of the genes to these tissues. The central nervous system and the liver are two tissues to which the application of this cell marking

technique may be extremely effective. Brain cells express a low level of methyl transferase activity and may be extremely susceptible to mutagenesis with MNNG (Walter *et al.*, 1993; Buecheler and Kleihues, 1977; Craddock and Henderson, 1984). The liver on the other hand is extremely susceptible to mutagenesis with AAF and this strategy may prove extremely effective for liver cell marking (Tiwawech *et al.*, 1991; Poirier *et al.*, 1991; Glauert *et al.*, 1991). It will probably be necessary to collaborate with other labs with experience in working with other tissues to obtain the full benefit of the application of this technique to other tissues.

This thesis describes the design and development of a marker system for the analysis of cell lineage. The system has now been tested and shown to work *in vitro* in cell lines and *in vivo* in the tissues of transgenic mice. The marker system will now be applied to the analysis of skin cell lineage in the tissues of developing and adult transgenic mice which carry the marker genes described. In the future it is hoped to apply the same marker system to analysis of cell lineage in other tissues such as intestine, liver, and central nervous system. It is hoped that this system may also be applied to the activation of other genes leading to the creation of good transgenic models of oncogenesis and the study of the effect of the activation of genes in individual cells. The combination of these approaches may provide a good insight into the role of stem cells in the development and maintenance of tissues in the mouse.

CHAPTER 5

References

5. References

ABRAMSON, S., MILLER, R. G. & PHILLIPS, R. A. (1977) The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems. *J. Exp. Med.* **145**, 1567-1579.

ARGYRIS, T. S. (1976) Kinetics of epidermal production during epidermal regeneration following abrasion in mice. *Am. J. Pathol.* **83**, 329-340.

AL-BAWARI, S. E. & POTTEN, C. S. (1976) Regeneration and dose response characteristics of irradiated mouse dorsal epidermal cells. *Int. J. Rad. Biol.* **30**, 201-216.

ALBERS, K.M., SETZER, R. W. & TAICHMAN, L. B. (1986) Heterogeneity in the replicating population of cultured human epidermal keratinocytes. *Differentiation* **31**, 134-140.

ALBERT, R. E., BURNS, F. J. & HEIMBACH, R. D. (1967a) The effect of penetration depth of electron radiation on skin tumour formation in the rat. *Radiat. Res.* **30**, 515-524.

ALBERT, R. E., BURNS, F. J. & HEIMBACH, R. D. (1967b) The association between chronic radiation damage of the hair follicles and tumour formation in the rat. *Radiat. Res.* **30**, 590-599.

ALBERT, R. E., PHILLIPS, M. E., BENNET, P. & HEIMBACH, R. D. (1969) The morphology and growth characteristics of radiation induced epithelial skin tumours in the rat. *Cancer Res.* **29**, 658-668.

AKAM, M. (1987) The molecular basis for metameric pattern in the Drosophila embryo. *Development* **101**, 1-22.

AMBROS, V. & MOSS, E. G. (1994) Heterochronic genes and the temporal control of C. elegans development. *Trends Genet.* **10**, 123-127.

ANDERSEN, B., ZIERZ, S. & JUNGERMANN, K. (1983) Perinatal development of the distribution of phosphoenolpyruvate carboxykinase and succinate dehydrogenase in rat liver parenchyma. *Eur. J. Cell Biol.* **30**,126-131.

ANGEVINE, J. B. & SIDMAN, R. L. (1961) Autoradiographic study of cell migration during histogenesis of the cerebral cortex in the mouse. *Nature* **192**, 766-768.

ANSELL, J. D. & MICKLEM, H. S. (1986) Genetic markers for following cell populations, in *Handbook of Experimental Immunology Vol. 2.* (Weir, D. M., Herzenberg, L. A. & Blackwell, C. eds.) p 56. Blackwell, New York.

ARBER, N., ZAJICEK, G. & ARIEL, I. (1988) The streaming liver II: hepatocyte life history. *Liver* **8**, 80-87.

ASSELINEAU, D., BERNARD, B. A., BAILLY, C., DARMON, M. & PRUNIÉRAS, M. (1986) Human epidermis reconstructed by culture: Is it normal? *J. Invest. Dermatol.* **86**, 181-185.

ATWATER, J. A. WISDOM R. & VERMA, I. M. (1990) Regulated mRNA stability. *Annu. Rev. Genet.* **24**, 519-541.

AUSTIN, C. P. & CEPKO C. L. (1990) Cellular migration patterns in the developing mouse cerebral cortex. *Development* **110**, 713-732.

BAILLEUL, B., SURANI, M. A., WHITE, S., BARTON, S. C., BROWN, K., BLESSING, M., JORCANO, J. & BALMAIN, A. (1990) Skin hyperkeratosis and papilloma formation in transgenic mice expressing a ras oncogene from a suprabasal keration promoter. *Cell* **62**, 697-708.

BALABAN, C.,TEILLET, M.-A. & LEDUORAIN, N. (1988) Application of a quailchick chimaera system to the study of brain development and behavior. *Science* **241**, 1139-1142.

BARRACLOUGH, R. & RUDLAND, P. S. (1990) Differentiation of mammary stem cells in vivo and in vitro. *Environ. Health Perspect.* **80**, 39-48.

BARRANDON, Y. & GREEN, H. (1987) Three types of keratinocyte with different capacities for multiplication. *Proc. Natl. Acad. Sci. USA.* **84**, 2302-2306.

BARTEK, J., IGGO, I., GANNON, J. & LANE, D. P. (1990) Immunochemical analysis of mutant p53 in human breast cancer cell lines. *Oncogene* **5**, 893-899.

BARTSCH, H., MALAVEILLE, C., TOMATIS, L., BRUN, G., DODET, B. & TERRACINI, B. (1982) Quantitative comparisons between carcinogenicity, mutagenicity and electrophilicity of direct-acting N-nitroso compounds and other alkylating agents. *I. A. R. C. Sci. Pub.* **41**, 525-532.

BASLER, K. & STRUHL, G. (1994) Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein. *Nature* **368**, 208-214.

BEDDINGTON, R. S. P. (1982) An autoradiographic analysis of tissue potency in different regions of the embryonic ectoderm during gasrtrulation in the mouse. *J. Embryol. Exp. Morphol.* **69**, 269-285.

BEDDINGTON, R. S. P. (1986) Analysis of tissue fate and prospective potency in the egg cylinder, in *Experimental Approaches to Mammalian Embryonic Development* (Rossant, J. & Persden, R. A. eds.) pp. 121-147. Cambridge University Press, New York.

BEDDINGTON, R. S. P., MORGERNSTERN, J., LAND, H. & HOGAN A. (1989) An in situ transgenic enzyme marker for the midgestation mouse embryo and the visualisation of inner cell mass clones during early organogenesis. *Development* **106**, 37-46.

BEDDINGTON, R. S. P. & ROBERTSON, E. (1989) An assessment of the developmental potential of embryonic stem cells midgestation mouse embryo. *Development* **105**, 733-737.

BELAND, F. A., DOOLEY, K. L. & JACKSON, C. D. (1982) Persistance of DNA adducts in Rat Liver and Kidney after multiple doses of the carcinogen *N*-hydroxy-2-acetylaminofluorene. *Cancer Res.* **42**, 1348-1354.

BENNET, A. L., PAULSON, K. E, MILLER, R. E. & DARNELL JR., J. E. (1987) Aquisition of antigens of adult and pericentral hepatocytes by differentiating fetal hepatoblasts in vitro. *J. Cell Biol.* **105**, 1073-1085.

BERNSTINE, E. G., HOOPER, M. L., GRANDCHAMP, S. & EPHRUSSI, B. (1973) Alkaline phosphatase activity in mouse teratoma cells. *Proc. Natl. Acad. Sci. USA.* **70**, 3899-3903.

BIACHARA, M. & FUCHS, R. P. P. (1985) DNA binding and mutation spectra of the carcinogen N-2-acetylaminofluorene in Escherichia coli: A correlation between the conformation of the premutatgenic lesion and the mutation specificity. *J. Mol. Biol.* **183**, 341-351.

BICKENBACH, J. R. (1988) Identification of label-retaining cells in oral mucosa and skin. *J. Dent. Res.* **60**, 1611-1620.

BILLINGAM, R. E., ORR, J. W. AND WOODHOUSE, D. L. (1951) Transplantation of skin componenets during chemical carcinigenesis with 20-methylcholanthrene. *Br. J. Cancer.* **5**, 417-432.

BOEGER-BROWN, U. (1993) Analysis of multistage skin carcinogenesis by retroviral mediated gene transfer. Ph.D. thesis, University of Glasgow.

BOGGS, D. R., BOGGS, S. S., SAXE, D. F., GRESS, L. A. & CARFIELD, D. R. (1982) Haemopoietic stem cells with high proliferation potential. Assay of their concentration by the frequency and duration of cure of W/W^{v} mice. *J. Clin. Invest.* **70**, 243-247.

BÖGLER, O., WREN, D., BARNETT, S. C., LAND, H. & NOBLE, M. (1990) Cooperation between two growth factors promotes extended self-renewal and inhibits differentiation of the O-2A progenitor cells. *Proc. Natl. Acad. Sci. USA.* **87**, 6368-6372.

BOWDEN, G. T. & BOUTWELL, R. K. (1974) Studies on the role of stimulated epidermal DNA synthesis in the initiation of skin tumours in mice by N-methyl-N'-nitro-N-nitrosoguanidine. *Cancer Res.* **34**, 1552-1563.

BRADLEY, A., EVANS, M. J., KAUFMAN, M. H. & ROBERTSON, E. (1984) Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-256.

BRADLEY, T. R. & METCALF, D. (1966) The growth of mouse bone marrow cells in vitro. *Aust. J. Exp. Biol. Med. Sci.* **44**, 287.

BRILL, S., HOLST, P., SIGAL, S., ZVIBEL, I., FIORINO, A., OCHS, A., SOMASUNDARAN, U. & REID, L. M. (1993) Hepatic progenitor populations in embryonic, neonatal and adult liver. *Proc. Soc. Exp. Biol. And Med.* **204**, 261-269.

BRINSTER, R. L. (1974) The effect of cells transferred into the mouse blastocyst on subsequent mouse development. *J. Exp. Med.* **140**, 1049-1056.

BRYANT, P. J. (1970) Cell lineage relationships in the imaginal wing disc of Drosophila melanogaster. *Dev. Biol.* **22**, 389-411.

BURNOUF, D. & FUCHS, R. P. P. (1985) Construction of frameshift mutation hotspots within the tetracycline resistance gene of pBR322. *Biochimie* **67**, 385-389.

BURNOUF, D., KOEHL, P. & FUCHS, R. P. P. (1989) Single adduct mutagenesis: Strong effect of the position of a single acetylaminofluorene adduct within a mutation hotspot. *Proc. Natl. Acad. Sci. USA.* **86**, 4147-4151.

BUECHELER, J. & KLEIHUES, P. (1977) Excision of O6-methylguanine from DNA of various mouse tissues following a single injection of N-methyl-N-nitrosourea. *Chem.-Biol. Interactions* **16**, 325-333.

BURNS, P. A., GORDON, A. J. E. & GLICKMAN, B. W. (1987) Influence of neighbouring base sequence on N-methyl-N'-nitro-N-nitrosoguanidine mutagenesis in the *lac I* gene of *Escherichia coli*. *J. Mol. Biol.* **194**, 385-390.

BYRNE, C., TAINSKY, M. & FUCHS, E. (1994) Programming gene expression in the developing epidermis. *Development* **120**, 2369-2383.

CAMPARA, J. & REYNOLDS, R. (1989) The hepatic circulation, in *The Liver: Biology and Pathobiology, 2nd edn.* (Arias, I. M., Jakoby, W. B., Popper, H., Schacter, D. & Schafritz, D. A. eds.) pp. 911-930. Raven, New York.

CAPEL B., HAWLEY, R., COVARRUBIAS, L., HAWLEY, T & MINTZ B. (1989) Clonal contributions of small numbers of retrovirally marked haemopoietic stem cells engrafted in unirradiayed neonatal W\W^{v} mice. *Proc. Natl. Acad. Sci. USA.* **86**, 4564.

CAPPECCI, M. R. (1989) Altering the genome by homologous recombination. *Science* **244**, 1288-1292.

CARAS, I. W., MACINNES, M. A., PERSING, D. H., COFFINO, P. & MARTIN, D. W. (1982) Mechanism of 2-aminopurine mutagenesis in mouse T-lymphosarcoma cells. *Mol. Cell Biol.* **2**, 1096-1103.

CAVINESS, V. S. (1976) Patterns of cell and fibre distribution in the neocortex of the reeler mutant mouse. *J. Comp. Neurol.* **170**, 435-448.

CAVINESS, V. S. (1982) Neocortical histogenesis in normal and reeler mutant mice: a developmental study based upon ^{3}H-thymidine autoradiography. *Dev. Brain Res.* **4**, 293-302.

CAVINESS, V. S. & SIDMAN, R. L. (1973) Time of origin of corresponding cell classes in the cerebral cortex of normal and reller mutant mice: an autoradiographic analysis. *J. Comp. Neurol.* **148**, 141-152.

CEPKO, C. L. (1988) Retrovirus vectors and their applications in neurobiology. *Neuron* **1**, 345-353.

CEPKO, C. L., ROBERTS, B. E. & MULLIGAN, R. C. (1984) Construction and application of a highly transmissable murine retrovirus shuttle vector. *Cell* **37**, 1053-1062.

CHENG, H. & LEBLOND, C. P. (1974a) Origin, differentiation and renewal of the for main epithelial cell types in the mouse small intestine: IV. Paneth cells. *Am. J. Anat.* **141**, 521-536.

CHENG, H. & LEBLOND, C. P. (1974b) Origin, differentiation and renewal of the for main epithelial cell types in the mouse small intestine: V. Unitarian theory of the origin of the four epitheial cell types. *Am. J. Anat.* **141**, 537-562.

CHISHOLM, J. C., JOHNSON, M. H., WARREN, P. D., FLEMING, T. P. & PICKERING, S. J. (1985) Developmental variability within and between mouse expanding blastocysts and their ICMs. *J. Embryol. Exp. Morphol.* **86**, 311-336.

COHN, S. M., ROTH, K. A., BIRKENMEIER, E. H. & GORDON, J. I. (1991) Temporal and spatial patterns of transgene expression in againg adult mice provide isights about the origins, organisation and differentiation of the intestinal epithelium. *Proc. Natl. Acad. Sci. USA.* **88**, 1034-1038.

COTSARELIS, G., CHENG, S.-Z., DONG, G., SUN, T.-T. & LAVKER, R. M. (1989) Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: Implications on epithelial stem cells. *Cell* **57**, 201-209.

COTSARELIS, G., SUN, T.-T. & LAVKER, R. M. (1990) Label-retaining cells reside in the bulge area of the pilosebaceous unit: Implications for follicular stem cells, hair cycle and skin carcinogenesis. *Cell* **61**, 1329-1337.

COULOMB, B., SAIAG, P., BELL, E., BREITBURD, F., LEBRETON, C., HESLON, M. & DUBERTRET, L. (1986) A new method for studying epidermalisation in vitro. *Brit. J. Dermatol.* **114**, 91-101.

COULONDRE, C. & MILLER, J. H. (1977) Genetic studies of the lac repressor. III. Additional correlation of mutational sites with specific amino acid residues. *J. Mol. Biol.* **117**, 577-606.

CRADDOCK, V. M. & HENDERSON, A. R. (1984) Repair and replication of DNA in rat and mouse tissues in relation to cancer induction by N-nitoso-N-alkylureas. *Chem.-Biol. Interactions* **52**, 223-231.

CRICK, F. H. C. & LAWRENCE P. A. (1975) Compartments and polyclones in insect development. *Science* **189**, 340-347.

CUPPLES, C. G. & MILLER, J. H. (1989) A set of lac Z mutations in Escherichia coli that allow rapid detection of each of the six base substitutions. *Proc. Natl. Acad. Sci. USA.* **86**, 5345-5349.

DAMJANOV, I., SOLTER, D. & SKREB, N. (1971) Teratocarcinogenesis as related to the age of embryos grafted under the kindey capsule. *Roux Arch. Dev. Biol.* **173**, 228-234.

DE ROOIJ, D. G. (1988) Regulation of the proliferation of spermatagonial stem cells. *J. Cell Sci. Suppl.* **10**, 181-194.

DEXTER, T. M. (1982) Stromal cell associated haemopoiesis. *J. Cell Physiol. Suppl.* **1**, 87-94.

DEXTER, T. M., ALLEN T. D. & LAJTHA, L. G. (1977) Conditions controlling the proliferation of haemopoietic stem cells *in vitro*. *J. Cell Physiol.* **91**, 335-344.

DEXTER, T. M., HEYWORTH, C. & WHETTON, A. D. (1986) The role of growth factors in haemopoiesis. *Bioessays* **2**, 154-158.

DEXTER, T. M. & SPOONCER E. (1987) Growth and differentiation of the haemopoietic system. *Ann. Rev. Cell Biol.* **3**, 423-441.

DEXTER, T. M., SPOONCER, E., SIMMONS, P. & ALLEN, P. D. (1984) Long-term bone marrow cultures: An overview of techniques and experience, in *Long-Term Bone Marrow Culture* (Wright, D.G. & Greenberger, J. S. eds.) pp 57-96. Liss, New York.

DICK, J. E., MAGLI, M. C., HUSZAR, D., PHILLIPS, R. A. & BERNSTEIN, A. (1985) Intoduction of a selectable gene into primitive cells capable of long-term reconstitutuion of the haemopoietic system of W/W^v mice. *Cell* **42**, 71.

DINARDO, S., HEEMSKERK, J., DOUGAN, S. & O'FARRELL, P. H. (1994) The making of a maggot: patterning of the Drosophila embryonic epidermis. *Curr. Opinion. Gen. Dev.* **4**, 529-534.

DOETSCHMAN, T. C., EISETTER, H., KATZ, M., SCHMIDT, W. & KEMLER, R. (1985) The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol.* **87**, 27-45.

DONIGER, J., DAY, R. S. & DIPAOLO J. A. (1985) Quantitative assessment of the role of O^6-methylguanine in the initiation of carcinogenesis by methylating agents. *Proc. Natl. Acad. Sci. USA*. **82**, 421-425.

DOVER, R. & POTTEN C. S. (1983) Cell cycle kinetics of cultures human epidermal keratinocytes. *J. Invest. Dermatol.* **80**, 423-429.

DRAGER, U. (1981) Observations on the organisation of the visual cortex in the reeler mouse. *J. Comp. Neurol.* **201**, 555-570.

DUBRIDGE, R. B., TANG, P., HSIA, H. C., LEONG, P.-M., MILLER, J. H. & CALOS, M. P. (1987) Analysis of mutation in human cells by using a Epstein-Barr virus shuttle system. *Mol. Cell. Biol.* **7**, 379-387.

DUFFILL, M. B., APPLETON, D. R., DYSON, P., SHUSTER, S. & WRIGHT, N.A. (1977) The measurement of cell cycle time in squamous epithelium using the metaphase arrest method with vincristine. *Brit. J. Dermatol.* **96**, 493-502.

ECCLES, S. J., VIDAL, M. A., WRIGHTON, C. J., GROSVELD, F. G. & BURKE, J. F. (1989) Rapid detection of ultraviolet-induced reversion of an amber mutation in mouse L cells. *Proc. Natl. Acad. Sci. USA.* **86**, 1297-1301.

ELLIS, R. E. & HORVITZ, H. R. (1986) Genetic control of programmed cell death in the nematode C. elegans. *Cell* **44**, 817-829.

EPSTEIN W. & MAIBACH A. (1965) Cell renewal in human epidermis. *Arch. Dermatol.* **92**, 462-468.

EVANS, M J. (1972) The isolation and properties of a clonal tissue culture strain of pluripotential mouse teratoma cells. *J. Embryol. Exp. Morphol.* **28**, 163-173.

EVANS, M. J. & KAUFMAN, M. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.

EVARTS, R. P., NAGY, P., MARSDEN, E. & THORGEIRSSON, S. S. (1987) A precursor relationship exists between oval cells and hepatocytes in rat liver. *Carcinogenesis* **8**, 1737-1740.

EVARTS, R. P., NAGY, P., NAKATSUKASA, H., MARSDEN, E. R. & THORGEIRSSON, S. S. (1989) In vivo differentiation of rat liver oval cells into hepatocytes. *Cancer Res.* **49**, 1541-1547.

FARBER, E. (1984) Cellular biochemistry of the stepwise development of cancer with chemicals. *Cancer Res.* **44**, 5463-5474.

FAUSTMAN, E. M., KIRBY, Z., GAGE, D. & VARNUM M. (1989) In vitro developmental toxicity of five direct acting alkylating agents in rodent embryos: Structure activity patterns. *Teratology* **40**, 199-210.

FEINBERG, A. P. & VOGELSTEIN, B. (1983) Hypomethylation of ras oncogenes in primary human carcinomas. *Biochem. Biophys. Res. Commun.* **111**, 47-54.

FEINBERG, A. P. & VOGELSTEIN, B. (1984) A technique for radiolabelling DNA restriction fragments to high specific activity. *Anal. Biochem.* **137**, 226-227.

FFRENCH-CONSTANT, C. & RAFF M. C. (1986) Proliferating bi-potential glial progenitor cells in adult rat optic nerve. *Nature* **319**, 499-502.

FIALKOW, P. J. (1973) Primordial cell pool size and lineage relationships of five human cell types. *Ann. Hum. Genet.* **37**, 39.

FUSENIG, N. E. (1986) Mammalian epidermal cells in culture, in *Biology Of The Integument, in Vertebrates Volume 2* (Bereiter-Hahn, J., Matoltsy, A.G., Richards, K. S. eds.), pp 409-442. Springer-Verlag, Berlin.

GAASBEEK-JANZEN, J. W., GEBHARDT, R., TEN VOORDE, G. H. J., LAMERS, W. H., CHARLES, R. & MOORMAN, A. F. (1987) Heterogeneous distribution of glutamine sythetase during rat liver development. *J. Histochem. Cytochem.* **35**, 49-54.

GALILEO, D. S., GRAY, D. E., OWENS, G., MAJORS, J. & SANES J. (1990) Neurons and glia arise from a common progenitor in chick optic tectum: demonstration with two retroviruses and cell type specific antibodies. *Proc. Natl. Acad. Sci. USA.* **87**, 458-462.

GARCIA-BELLIDO, A., RIPOLL, P. & MORATA G. (1973) Developmental compartmentalisation of the wing disc of Drosophila. *Nature New Biol.* **245**, 251-253.

GARCIA-BELLIDO, A., MORATA G. & LAWRENCE, P. A. (1979) Compartments in animal development. *Sci. Am.* **241**, 90-98.

GARDNER, R. L. (1983) Origin and differentiation of extra-embryonic tissues in the mouse. *Int. Rev. Exp. Path.* **24**, 63-133.

GARDNER, R. L. (1985) Regeneration of endoderm from primitive ectoderm in the mouse: fact or artifact? *J. Embryol. Exp. Morphol.* **88**, 303-326.

GARDNER, R. L. & BEDDINGTON, R. S. P. (1988) Multi-lineage stem cells in the mammalian embryo. *J. Cell Sci. Suppl.* **10**, 11-27.

GEBHARDT, R. (1988) Different proliferative activity in vitro of periportal and perivenous hepatocytes. *Scand. J. Gastroenterol. Suppl.* **151**, 8-18.

GEISSLER, E. N., RYAN, M. A. & HOUSMAN, D. E. (1988) The dominant-white spotting (W) locus of the mouse encodes the c-kit proto-oncogene. *Cell* **55**, 185.

GERMAIN, L., BLOUIN, M. J. & MARCEAU, N. (1988b) Biliary epithelial and hepatocytic cell lineage relationships in embryonic rat liver as determined by the differential

expression of cytokeratins, alpha-fetoprotein albumin and cell surface-exposed components. *Cancer Res.* **48**, 4909-4918.

GERMAIN, L., GOYETTE, R., & MARCEAU, N. (1985) Differential cytokeratin and alpha-fetoprotein expression in morphologically distinct epithelial cells emerging at the early stages of rat hepatocarcinogenesis. *Cancer Res.* **45**, 673-681.

GERMAIN, L., NOEL, M., GOUDEAU, H. & MARCEAU, N. (1988a) Promotion of growth and differentiation of rat ductular oval cells in primary culture. *Cancer Res.* **48**, 368-378.

GILBERT S. F. (1994) *Developmental Biology (Fourth Edition)* pp510-518 Sinauer Associates Inc., Sunderland, Massachusetts.

GLICKMAN, B. W., HORSFALL, M. J., GORDON, A. J. E. & BURNS, P. A. (1987) Nearest neighbour affects G:C to A:T transitions induced by alkylating agents. *Environ. Health Persp.* **76**, 29-32.

GLAUERT, H. P., LAY, L. T., KENNAN, W. S. & PITOT, H. C. (1991) Effect of dietary fat on the inititation of hepatocarcinogenesis by diethylnitrosamine or 2-acetylaminofluorene in rats. *Carcinogenesis* **12**, 991-995.

GOLDOWITZ, D. (1989) Cell allocation in mammalian CNS: evidence from murine interspecies aggregation. *Neuron* **3**, 705-713.

GORDON, J. I., SCHMIDT, G. H. & ROTH, K. A. (1992) Studies of intestinal stem cells using normal, chimaeric and transgenic mice. *FASEB J.* **6**, 3039-3050.

GORDON, M. Y. (1991) Haemopoietic growth factors and receptors: Bound and free. *Cancer Cells* **3**, 127-133.

GORDON, M. Y., RILEY, G. P., WATT, S. M. & GREAVES, M. F. (1987) Compartmentalization of a haemopoietic growth factor (GM-CSF) by glucosaminoglycans in the bone marrow microenvironment. *Nature* **326**, 403-405.

GOUGH, N. M., GOUGH, J., METCALF, D., KELSO, A., GRAIL, D., NICOLA, N. A., BURGESS, A. W. & DUNN, A. R. (1984) Molecular cloning of a cDNA encoding a murine haemopoietic growth regulator granulocyte/macrophage colony stimulating factor. *Nature* **309**, 763-767.

GRAHAM, G. J. & PRAGNELL, I. B. (1992) The haemopoietic stem cell: properties and control mechanisms. *Sem. Cell Biol.* **3**, 423-434.

GRAY, G. E., GLOVER, J. C., MAJORS, J. & SANES, J. (1988) Radial arrangement of clonally related cells in the chick optic tectum: lineage analysis with a recombinanat retrovirus. *Proc. Natl. Acad. Sci. USA.* **85**, 7356-7360.

GREENSPAN, J. A., XU, F. & DAVIDSON, R. L. (1988) Molecular analysis of ethyl methansulphonate-induced reversion of a chromosomally integrated mutant shuttle vector gene in mammalian cells. *Mol. Cell Biol.* **8**, 4185-4189.

GRIFFITHS, D. F. R., DAVIES, S. J., WILLIAMS, D., WILLIAMS, G. T. & WILLIAMS, E. D. (1988) Demonstration of somatic mutation and colonic crypt clonality by X-linked enzyme histochemistry. *Nature* **333**, 461-463.

GROBSTEIN, C. (1952) Intra-ocular growth and differentiation of clusters of mouse embryonic shields cultured with and without primitive endoderm and in the presence of possible inductors. *J. Exp. Zool.* **119**, 355-380.

GU, H.., MARTH, J. D., ORBAN, P. C., MOSSMANN, H. & RAJEWSKY K. (1994) Deletion of a DNA polymerase β gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103-106.

GURDON, J. B. (1992) The generation of diversity and pattern in animal development. *Cell* **68**, 185-199.

HAUSER, J., LEVINE, A. S. & DIXON, K.(1987) Unique pattern of point mutation arising after gene transfer ito mammalian cells. *E.M.B.O.* ***6***, 63-67.

HALL, P. A. & WATT, F. M. (1989) Stem cells: the generation and maintenance of cellular diversity. *Development* **106**, 619-633.

HAMILTON, E. & POTTEN, C. S. (1972) Influence of hair plucking on the turnover time time in the epidermal basal layer. *Cell Tiss. Kinet.* **5**, 505-513.

HARRISON, D. A. & PERRIMON, N. (1993) Simple and efficient generation of marked clones in Drosophila. *Curr. Biol.* **3**, 424-433.

HARRISON, D. E., ASTLE, C. M. & LERNER, C. (1988) Number and continuous proliferative pattern of transplanted primitive immunohaematopoietic stem cells. *Proc. Natl. Acad. Sci. USA.* **86**, 822-826.

HARVAN, W. L. & ALLISON, J. P. (1988) Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors. *Nature* **335**, 443-447.

HARVAN, W. L. & ALLISON, J. P. (1990) Origin of Thy-1+ dendritic epidermaln cells of adult mice from fetal thymic precursors. *Nature* **344**, 68.

HAYNER, N. T., BRAUN, L., YASWEN, P., BROOKS, M. & FAUSTO, N. (1984) Isozyme prolfiles of oval cells, parenchymal cells and biliary cells isolated by centrifugal elutriation from normal and preneoplastic livers. *Cancer Res.* **44**, 332-338.

HEEMSKERK, J., DINARDO, S., KOSTRIKEN, R. & O'FARRELL, P. H. (1991) Multiple modes of engrailed regulation in the progression towards cell fate determination. *Nature* **352**, 404-410.

HEGAZY, M. A. H. & FOWLER, J. F. (1973) Cell population kinetics of plucked and unplucked mouse skin: 1. Unirradiated skin. *Cell Tiss. Kinet.* **6**, 17-33.

HENDRY, J. H., POTTEN, C. S., GHAFFOR, A., MOORE, J. V., ROBERTS, S. A. & WILLIAMS, P. C. (1989) The response of murine intestinal crypts to short-range promethium -147 irradiation: deductions concerning clonogenic cell numbers and positions. *Radiation Research* **118**, 364 374.

HENNINGS, H., BOWDEN, G. T. & BOUTWELL, R. K. (1968) The effect of croton oil pretreatment of skin tumour initiation in mice. *Cancer Res.* **29**, 1773-1780.

HERRUP, K. (1986) Cell lineage relationships in the development of the mammalian CNS: role of cell lineage in control of Purkinje cell number. *Dev. Biol.* **115**, 148-154.

HERZENBERG, L. A. & SWEET, R. G. (1976) Fluorescence-activated cell sorting. *Sci. Am.* **234**, 108.

HERZLINGER, D., KOSEKI, C., MIKAWA, T. & AL-AWQATI, Q. (1992) Metanephric mesenchyme contains multipotent stem cells whose fate is restricted after induction. *Development* **114**, 565-572.

HOFFMAN, J. W., STEFFEN, D., GUSELLA, J., TABIN, C., BIRD, S., COWING D. & WEINBERG, R. A. (1982) DNA methylation affecting the expression of murine leukaemia proviruses. *J. Virol.* **44**, 144-157.

HOGAN, B., CONSTANTINI, F. & LACY, E. (1986) *Manipulating the Mouse Embryo - A Laboratory Manual.* pp 332. Cold Spring Harbour Laboratory, New York.

HOLBROOK, K. A. (1979) Human epidermal embryogenesis. *Int. J. Dermatol.* **18**, 329-356.

HOLBROOK, K. A., SMITH, L. T., KAPLAN, E. D., MINAMI, S. A., HERBERT, G. P. & UNDERWOOD, R. A. (1993) Expression of morphogens during human follicle development in vivi and a model for studying follicle morphogenesis in vitro. *J. Invest. Dermatol.* **101**, 39s-49s.

HOLOCEK, B.-U. & ACKERMAN, A. B. (1993) Bulge-activation hypothesis: Is in valid? *Am. J. Dermopathol.* **15**, 235-247.

HOLT, C. E., BERTSCH, T. W., ELLIS, H. M. & HARRIS, W. A. (1988) Cellular determination in the Xenopus retina is independent of lineage and birtdate. *Neuron* **1**, 15-26.

HORNSEY, S. (1985) The macro-colony assay in small intestine, in *Cell Clones* (Potten, C. S. & Hendry, J. H. eds.) pp 44-49. Churchill Livingstone, Edinburgh.

HUANG,E., NOCKA, K., BEIER, D. R., CHU, Y.-T., BUCK, J., LAHM, H.-W., WELLNER, D., LEDER, P. & BESMER, P. (1990) The haemopoietic growth factor KL is encoded

by the Sl locus and is the ligand of the c-kit receptor, the gene product of the W locus. *Cell* **63**, 235.

IJIRI, K. & POTTEN C. S. (1988) Circadian rhythms in the incidence of apoptotic cells and number of clonogenic cells after irradiation using normal and reversed light conditions. *Int. J. Rad. Biol.* **53**, 717-727.

IKUTA, K., KINA, T., MACNEIL, I., UCHIDA, N., PEAULT, B., CHIEN, Y. H. & WIESSMAN, I. L. (1990) A developmental switch in thymic lymphocyte maturation potential occurs at the level of haemopoietic stem cells. *Cell* **62**, 863.

INGHAM, P. W. (1988) The molecular genetics of embryonic pattern formation in Drosophila. *Nature* **335**, 25-34.

INGHAM, P. W. & MARTINEZ ARIAS A. (1992) Boundaries and fields in early embryos. *Cell* **68**, 221-235.

INOUE, M., FUKUSHIMA, Y., MATSUURA, N., SHIOZAKI, H., MORI, T., KITAMURA, Y. & FUJITA, H. (1988) Macroscopic intestinal colonies of mice as a tool for studying differentiation of multipotential intestinal stem cells. *Am. J. Path.* **132**, 49-58.

INOUYE, M. & MURAKAMI, U. (1978) Teratogenic effect of N-methyl-N'-nitro-N-nitrosoguanidine in mice. *Teratology* **18**, 263-268.

IVERSEN O. H., BJERKNES, R. & DEVIK, F. (1968) Kinetics of cell renewal, cell migration and cell loss in the hairless mouse dorsal epidermis. *Cell Tiss. Kinet.* **1**, 365-379.

JACOBSEN, K., MIYAKE, K., KINCADE, P. W. & OSMOND, D. G. (1992) Highly restricted expression of a stromal cell determinant in mouse bone marrow in vivo. *J. Exp. Med.* **176**, 927-935.

JACOBSON, M. (1985) Clonal analysis and cell lineage of the vertebrate central nervous system. *A. Rev. Neurosci.* **8**, 71-102.

JAHODA C. A. B. (1992) Induction of follicle formation and hair growth by vibrissa dermal papillae implanted into rat ear wounds: vibrissa-type fibres are specified. *Development* **115**, 1103-1109.

JAHODA, C. A. B. & REYNOLDS, A. J. (1993) Dermal-epidermal interactions - follicle derived cell populations in the study of hair growth mechanisms. *J. Invest. Dermatol.* **101**, 33s-38s.

JANKOWSKI, J. A. & WRIGHT, N. A. (1992) Epithelial stem cells in gastrointestinal morphogenesis, adaption and carcinogenesis. *Semin. Cell Biol.* **3**, 445-456.

JENSEN, P. K. A., PEDERSEN, S. & BOLUND, L. (1985) Basal-cell sub-populations and cell-cycle kinetics in human epidermal explant cultures. *Cell Tiss. Kinet.* **18**, 201-215.

JOHNSON, M. H. & ZIOMEK C. A. (1981) The foundation of two distinct cell lineages within the mouse morula. *Cell* **24**, 71-80.

JONES, A. & SPRING-MILLS, E. (1983) The liver and gallbladder, in *Histology, Cell and Tissue Biology, 5th edn.* (Weiss, L. ed.) pp. 707-748. Elsevier, New York.

JONES, P. H. & WATT F. M. (1993) Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. *Cell* **73**, 713-724.

JORDAN, C. T. & LEMISCHKA, I. R. (1990) Clonal and systemic analysis of long-term haematopoiesis in the mouse. *Genes Dev.* **4**, 230-232.

JOYNER, A., KELLER, G., PHILLIPS, R. A. & BERNSTEIN A. (1983) Retrovirus transfer of a bacterial gene into mouse haematopoietic progenitor cells. *Nature* **305**, 556-558.

JUNGERMANN, K. (1986) Functional heterogeneity of periportal and prevenous hepatocyctes. *Enzyme* **35**, 161-180.

KANAMURA, S. & ASADA-KUBOTA, M. (1980) The heterogeneity of hepatocytes during the post-natal development of the mouse. *Anat. Embryol.* **158**, 151-159.

KARTTUNEN, J. & SHASTRI N. (1991) Measurement of ligand-induced activation in single viable T-cells using the lac Z reporter gene. *Proc. Natl. Acad. Sci. USA.* **88**, 3972-3976.

KELLER, G., PAIGE, C., GILBOA, E. & WAGNER, E. F. (1985) Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haemopoietic precursors. *Nature* **318**, 149.

KELLER, G. & SNODGRASS, R. (1990) Lifespan of multipotential haemopoietic stem cells in vivo. *J. Exp. Med.* **171**, 1407.

KEMP, C. J., DONEHOWER, L. A., BRADLEY, A. & BALMAIN, A. (1993) Reduction of p53 gene dosage does not increase initiation or promotion but enhances malignanat progression of chemically induced skin tumours. *Cell* **74**, 813-822.

KEMP, C. J., FEE, F. & BALMAIN A. (1993a) Allelotype analysis of mouse skin tumours using polymorphic microsatellites: Sequential genetic alterations on chromosomes 6, 7 and 11. *Cancer Res.* **53**, 6022-6027.

KENYON, C. (1994) If birds can fly, why can't we? Homeotic genes and evolution. *Cell* **78**, 175-180.

KIMBLE J. & HIRSCH, D. (1979) The post-embryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Dev. Biol.* **70**, 396-417.

KIMMEL, C. B & WARGA, R. M. (1987) Indeterminate cell lineage of the zebrafish embryo. *Dev. Biol.* **124**, 269-280.

KIRKLAND S. C. (1988) Clonal origin of columnar, mucous and endocrine cell lineages in human colorectal epithelium. *Cancer* **61**, 1359-1363.

KORNBERG, T. (1981) Engrailed: a gene controlling compartment and sgement formation in Drosophila. *Proc. Natl. Acad. Sci. USA.* **78**, 1095-1099.

KOYIBASHI, K., ROCHAT, A. & BARRANDON, Y. (1993) Segregation of keratinocyte colony-forming cells in the bulge of the rat vibrissa. *Proc. Natl. Acad. Sci. USA.* **90**, 7391-7395.

KOZAK, M. (1983) Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbiol. Rev.* **47**, 1-45.

KREIK, A. (1974) Carcinogenesis by aromatic amines. *Biochem. Biophys. Acta* **355**, 177-203.

LAJTHA, L. A. (1979) Stem cell concepts. *Differentiation* **14**, 23-34.

LALLEMAND, Y. & BRÛLET, P. (1990) An in situ assessment of the routes and extents of colonisation of the mouse embryo by embryonic stem cells and their descendants. *Development* **110**, 1241-1248.

LAMBIE, E. J. & KIMBLE, J. (1991) Genetic control of cell interactions in nematode development. *Ann. Rev. Genet.* **25**, 411-436.

LAUFER, J. S., BAZZICALUPO, P. & WOOD W. B. (1980) Segregation of developmental potential in early embryos of Caenorhabditis elegans. *Cell* **19**, 569-577.

LAVKER, R. M., COTSARELIS, G., WEI, Z.-G. & SUN T.-T. (1991) Stem cells of the pelage vibrissa and eyelash follicles: the hair cycle and tumour formation. *Ann. NY. Acad. Sci.* **642**, 214-225.

LAVKER, R. M., MILLER, S., WILSON, C., COTSARELIS, G., WEI, Z.-G., YANG, J.-S. & SUN, T.-T. (1993) Hair follicle stem cells: Their location, role in hair cycle and involvement in skin tumour formation. *J. Invest. Dermatol.* **101**, 16s-26s.

LAVKER, R. M. & SUN, T.-T. (1983) Epidermal stem cells. *J. Invest. Dermatol.* **81**, 121s-127s.

LAWRENCE, P. A. (1981) The cellular basis of segmentation in insects. *Cell* **26**, 3-10.

LAWRENCE, P. A. (1988) The present status of the parasegment. *Development* **104**, 61-65.

LAWRENCE, P. A. & MORATA G. (1976) Compartments in the wing of Drosophila: a study of the engrailed gene. *Dev. Biol.* **50**, 321-337.

LAWRENCE, P. A., JOHNSTON, P. & MORATA, G. (1986) Methods of marking cells, in *Drosophila: A Practical Approach* (Roberts, D. B. ed.) pp 229-242. I. R. L. Press, Oxford.

LEBER, S. M., BREEDLOVE, S. M. & SANES, J. R. (1990) Lineage arrangement and death of clonally related motorneurons in chick spinal chord. *J. Neurosci.* **10**, 2451-2462.

LEBKOWSKI, J. S., MILLER, J. H. & CALOS, M. P. (1986) Determination of DNA sequence changes induced by ethyl methansulphonate in human cells using a shuttle vector system. *Mol. Cell Biol.* **6**, 1838-1842.

LEMISCHKA, I. R. (1992) What we have learned from retroviral marking of haematopoietic stem cells. *Curr. Topics Microbiol. Immunol.* **177**,

LEMISCHKA, I. R., RAULET, D. H., & MULLIGAN R. C. (1986) Developmental potential and dynamic behaviour of haematopoietic stem cells. *Cell* **45**, 917-927.

LEMMON, V. & PEARLMAN, A. L. (1981) Does laminar position determine the receptive field properties of cortical neurons? A study of corticotectal cells in area 17 of the normal mouse and the reeler mutant. *J. Neurosci.* **1**, 83-93.

LENOIR, M. C., BERNARD, B. A., PAUTRAT, G., DARMON, M. & SHROOT, B. (1988) Outer root sheath cells of the human hair follicle are able to regenerate a fully differentiated epidermis *in vitro*. *Dev. Biol.* **130**, 610-620.

LIMAT, A., BREITKREUTZ, D., HUNZIKER, T., BOILATT, C., WIESMANN, U., KLEIN, E., NOSER, F. & FUSENIG N. E. (1991) Restoration of the epidermal phemotype by follicular outer root sheath cells in recombinant culture with dermal fibroblasts. *Exp. Cell Res.* **194**, 218-227.

LIU, K. C. & WRIGHT, N. A. (1992) The migration pathway of epithelial cells an human duodenal villi: The origin and fate of 'Gastric Metaplastic' cells in duodenal mucosa. *Epith. Cell Biol.* 1, 53-58.

LOEFFLER, M. & GROSSMAN, B. (1990) A stochastic branching model with formation of submits applied to the growth of intestinal crypts. *J. Theor. Biol.*

LOUD, A. V. (1968) Quantitative stereological descriptions of the ultrastructure of normal rat parenchymal cells. *J. Cell Biol.* **37**, 27-46.

LUSKIN, M. B., PEARLMAN, A. L. & SANES, J. R. (1988) Cell lineage in the cerebral cortex of the mouse studied in vivo and in vitro with a recombinanat retrovirus. *Neuron* **1**, 635-647.

MACKENZIE, I. C. (1969) Ordered structure of the stratum corneum of the mammalian skin. *Nature* **222**, 881-883.

MACKENZIE, I. C. (1970) Relationship between mitosis and the ordered structure of the stratum corneum in mouse epidermis. *Nature* **226**, 653-655.

MACLAREN, A. (1976) *Mammalian chimaeras*. Cambridge University Press.

MACMAHON, A., FOSTEN, M. & MONK, M. (1990) X-chromosome inactivation mosaicism in the three germ layers and the germline of the mouse embryo. *J. Embryol. Exp. Morphol.* **74**, 207-220.

MAGANTO, P., TRABER, P. G., RUSSELL, C., DOBBINS III, W. O., KEREN, K. & GUMUCIO, J. (1990) Long term maintenance of the adult pattern of liver-specific expression for P-450b, P-450e, albumin and alpha-fetoprotein genes in intrasplenically transplanted hepatocyes. *Hepatology* **11**, 585-593.

MANDEL, M. & HIGA, A. J. (1970) Calcium dependent bacteriophage DNA interaction. *J. Mol. Biol.* **53**, 159-162.

MANN R., MULLIGAN, R. C. & BALTIMORE D. (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**, 153-159.

MARTIN, G. R. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma cells. *Proc. Natl. Acad. Sci. USA.* **78**, 7634-7638.

MARTIN, G. R. & EVANS, M. J. (1975) The formation of embroid bodies in vitro by homogeneous embryonal carcinoma cell cultures derived from single isolated cells, in *Teratomas and Differentiation.* (Sherman, M. I. & Solter, D. eds.) pp. 169-187. Academic Press, New York.

MARVIN. M, & MCKAY, R. (1992) Multipotential stem cells of the vertebrate CNS. *Semin. Cell Biol.* **3**, 401-411.

MATOLTSY, A. G. (1986) Structure and function of the mammalian epidermis, in *Biology Of The Integument in Vertebrates Volume 2* (Bereiter-Hahn, J., Matoltsy, A.G., Richards, K. S. eds.), pp 255-271. Springer-Verlag, Berlin.

MCBURNEY, M. W., SUTHERLAND, L. C., ADRA, C. N., LECLAIR, B., RUDNICKI, M. A. & JARDINE K. (1992) The mouse Pgk-1 promoter contains an upstream activator sequence. *Nucl. Acids Res.* **19**, 5755-5761.

MCCONNELL, S. K. (1988) Fates of visual cortical neurons in the ferret after isochronic and heterochronic transplantation. *J. Neurosci.* **8**, 945-974.

MCCONNELL, S. K. (1991) The generation of neuronal diversity in the central nervous system. *Ann. Rev. Neurosci.* **14**, 269-300.

MCCONNELL, S. K. & KAZNOWSKI, C. (1991) Cell cycle dependence of laminar determination in the developing neocortex. *Science* **254**, 282-285.

MCCULLOCH, E. A., SIMINOVITCH, L. & TILL, J. E. (1964) Sleen colony formation in anaemic mice of the genotype W/W^{v}. *Science* **144**, 844.

McCULLOCH, E. A., SIMINOVITCH, L., TILL, J. E., RUSSELL, E. S. & BERNSTEIN, S. E. (1965) The cellular basis of the genetically determined haemopoietic defect in anaemic mice of geneotype Sl/Sl^{d}. *Blood* **26**, 399.

McKAY, R. D. G. (1989) The origins of cellular diversity in the mammalian central nervous system. *Cell* **58**, 815-821.

MESSIER, B., & LEBLOND, C. P. (1960) Cell proliferation and migration as revealed by radioautography after injection of thymidine-H3 into male rats and mice. *Am J. Anat.* **106**, 247-265.

METCALF, D. (1977) *Haemopoietic colonies*. 227 pp. Springer-Verlag, Berlin.

MICKLEM, H. S., ANSELL, J. D., WAYMAN, J. E. & FORRESTER L. (1983) The clonal organisation of haematopoiesis in the mouse, in *Progress in Immunology V* (Yamamina, Y. & Tada, T. eds.) p 633. Academic, San Diego.

MILLER S. J., SUN, T.-T. & LAVKER R. M. (1993) Hair follicles, stem cells and skin cancer. *J. Invest. Dermatol.* **100**, 288s-294s.

MILLER, S. J., WEI, Z.-G., WILSON, C., DZUBOW, L., SUN, T.-T. & LAVKER, R. M. (1993) Mouse skin is particularly susceptibile to tumour initiation during early anagen of the hair cycle: possible involvement of hair follicle stem cells. *J. Invest. Dermatol.* **101**, 591-594.

MINTZ, B., ANTHONY, K. & LITWIN, S. (1984) Monoclonal derivation of mouse myeloid and lymphoid lineages from totipotent haemopoietic stem cells experimentally engrafted in fetal hosts. *Proc. Natl. Acad. Sci. USA.* **81**, 7835.

MINTZ, B. & ILLMENSEE, K. (1975) Normal genetically mosaic mice produced from malignant teratcarcinoma cells. *Proc. Natl. Acad. Sci. USA.* **72**, 3585-3589.

MORATA G., KORNBERG, T. & LAWRENCE, P. A. (1983) The phenotype of engrailed mutations in the antenna of Drosophila. *Cell* **26**, 3-10.

MORATA G. & LAWRENCE, P. A. (1975) Control of compartment development by the engrailed gene in Drosophila. *Nature* **255**, 614-617.

MORATA G. & LAWRENCE, P. A. (1977) Homeotic genes, compartments and cell determination in Drosophila. *Nature* **265**, 211-216.

MORATA G. & LAWRENCE, P. A. (1979) Development of the eye antenna imaginal disc of Drosophila. *Dev. Biol.* **70**, 355-371.

MORIYA, M., TAKESHITA, M., JOHNSON, F., PEDEN, K., WILL, S. & GROLLMAN, A. P. (1988) Targetted mutations induced by a single acetylaminofluorene DNA adduct in mammalian cells and bacteria. *Proc. Natl. Acad. Sci. USA.* **85**, 1586-1589.

MORRIS, R. J., FISCHER, S. M. & SLAGA, T. J. (1985) Evidence that centrally and peripherally located cells in the murine epidermal proliferative unit are two different populations. *J. Invest. Dermatol.* **84**, 277-281.

MORRIS, R. J., FISCHER, S. M. & SLAGA, T. J. (1986) Evidence that a slowly cycling subpopulation of adult murine epidermal cells retains carcinogen. *Cancer Res.* **46**, 3061-3066.

MOSER, A. R., PITOT, H. C. & DOVE, W. P. (1990) A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* **247**, 322-324.

MOSER A. R., DOVE, W. F., ROTH, K. A. & GORDON J. I. (1992) The Min (multiple intestinal neoplasia) mutation: its effect on gut epithelial cell differentiation and interaction with a modifier system. *J. Cell Biol.* **116**, 1517-1526.

NOBLE, M., MURRAY, K., STROOBANT, P., WATERFIELD, M. & RIDDLE, P. (1988) Platelet-derived growth factor promotes division and motility, and inhibits premature differentiation of the oligodendrocyte-type 2 astrocyte progenitor cell. *Nature* **330**, 560-562.

NOBLE, M., WREN, D. & WOLSWIJK, G. (1992) The O-2A^{adult} progenitor cell: a glial stem cell of the adult central nervous system. *Semin. Cell Biol.* **3**, 413-422.

NOLAN, G. P., FIERING, S., NICOLAS, J.-F. & HERZENBERG, L. A. (1988) Fluorescence-activated cell analysis and sorting of viable mammalian cells based on β-D-galactosidase activity after transduction of Escherichia coli lac Z. *Proc. Natl. Acad. Sci. USA.* **85**, 2603-2607.

NUDEL, U., ZAKUT, R., SHANI, M., NEUMAN, S., LEVY, Z. & YAFFE, D. (1983) The nucleotide sequence of the rat b-actin gene. *Nucl. Acids Res.* **11**, 1759-1771.

OGAWA, M. (1993) Differentiation and proliferation of haematopoietic stem cells. *Blood* **81**, 2844-2853.

O'GORMAN, S., FOX, D. T. & WAHL, G. M. (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* **251**, 1351-1355.

OLIVER, R. F. (1966) Whisker growth after removal of the dermal papilla and lengths of follicle in the hooded rat. *J. Embryol. Exp. Morphol.* **15**, 331-347.

OLIVER, R. F. (1967) Ectopic regeneration of whisker in the hooded rat from implanted lengths of vibrissa folllicle wall. *J. Embryol. Exp. Morphol.* **17**, 27-34.

OLIVER, R. F. (1967) The expeimental induction of whisker in the hooded rat by implantation of the dermal papilla. *J. Embryol. Exp. Morphol.* **18**, 43-51.

O'SULLIVAN, J. F., SCHMIDT, G. H. & PAUL, D. (1988) Spontaneous mutation rate. *Nature* **352**, 220.

PADYKULA, H. A. (1991) Regeneration in the primate uterus: the role of stem cells. *Ann. NY. Acad. Sci.* **622**, 47-56.

PAI, S. B., STEELE, V. E. & NETTESHEIM, P. (1983) Neoplastic transformation of primary tracheal epithelial cell cultures. *Carcinogenesis* **4**, 369-74.

PALMITER, R. D. & BRINSTER, R. L. (1986) Germ-line transformation of mice. *Ann. Rev. Genet.* **20**, 465-499.

PAPAIOANNOU, V. E., MCBURNEY, M. W., GARDNER, R. L. & EVANS, M. J. (1975) Fate of teratocarcinoma cells injected into early mouse embryos. *Nature* **258**, 70-73.

PARKINSON, E. K. (1992) Epidermal keratinocyte stem cells: their maintenance and regulation. *Sem. Cell Biol.* **3**, 435-444.

PARKINSON, E. K., GRAHAM, G. J., DAUBERSIES, P., BURNS, J. E., HEUFLER, C., PLUMB, M., SCHULER, G. & PRAGNELL, I. B. (1993) Haemopoietic stem cell inhibitor (SCI / MIP-1α) also inhibits clonogenic epidermal keratinocyte proliferation. *J. Invest Dermatol.* **101**, 113-117.

PELTZ, S. W. & JACOBSON, A. (1992) mRNA stability: in trans-it. *Curr. Opin. Cell. Biol.* **4**, 979-983.

PETROPOULOS, C. J., YASWEN, P., PANZICA, M. & FAUSTO, N. (1985) Cell lineagaes in liver carcinogenesis: possible clues from studies of the distribution of alphafetoprotein RNA sequences in cell populations isolated from normal regenerating and preneoplastic rat livers. *Cancer Res.* **45**, 5762-5768.

PLACZEK, M., TESSIER-LAVIGNE, A., YAMADA,T., JESSEL, T. & DODD, J. (1990) Mesodermal control of neural cell identity: floor plate induction by the notochord. *Science* **250**, 985-988.

POIRIER, M. C., DUBIN, M. A. & YUSPA, S. H. (1979) Formation and removal of specific acetylaminofluorene-DNA adducts in mouse and human cells measured by radioimmunoassay. *Cancer Res.* **39**, 1377-1381

POIRIER, M. C., FULLERTON, N. F., KINOUCHI, T., SMITH, B. A. & BELAND, F. A. (1982) Comparison between DNA adduct formation and tumourigenesis in livers and bladders of mice chronically fed 2-acetylaminofluorene. *Cancer Res.* **42**, 1371-1321.

POIRIER, M. C., FULLERTON, N. F., KINOUCHI, T., SMITH, B. A. & BELAND, F. A. (1991) Comparison between DNA adduct formation and tumorigenesis in livers and bladders and bladders of mice chronically fed 2-acetylaminofluorene. *Carcinogenesis* **12**, 895-900.

POIRIER, M. C., TRUE, B.-A. & LAISHES, B. A. (1982) Formation and removal of (guan-8-yl)-DNA-2-acetylaminofluorene adducts in liver and kidney of male rats given dietary 2-acetylaminofluorene. *Cancer Res.* **42**, 1317-1321.

POIRIER, M. C., WILLIAMS, G. M. & YUSPA, S. H. (1980) Effect of culture condidtions, cell type and species of origin on the distribution of the of acetylated and deacetylated deoxyguanosine C-8 adducts of N-acetoxy-2-acetylaminofluorene. *Mol. Pharmacol.* **18**, 581-587.

PONDER, B. A. J. (1987) Cell marking techniques and their application, in *Mammalian Development: A Practical Approach* (I.R.L.) pp 115-138.

PONDER, B. A. J., FESTING, M. F. W. & WILKINSON, M M. (1985b) An allelic difference determines reciprocal patterns of expression of binding sites for Dolichos bifluorus lectin in inbred strains of mice. *J. Embryol. Exp. Morphol.* **87**, 229-239.

PONDER, B. A. J., SCHMIDT, G. H. & WILKINSON, M.M. (1986) Immunohistochemistry in the analysis of mouse aggregation chimaeras. *Histochem J.* **18**, 217-227.

PONDER, B. A. J., SCHMIDT, G. H., WILKINSON, M.M. & WOOD, M. (1983) H2 antigens as markers of cellular genotype in chimaeric mice. *J. Embryol. Exp. Morphol.* **76**, 83-93.

PONDER, B. A. J., SCHMIDT, G. H., WILKINSON, M.M., WOOD M.J., MONK, M. & REID, A. (1985a) Derivation of mouse intestinal crypts from single progenitor cells. *Nature* **313**, 689-691.

PONDER, B. A. J. & WILKINSON, M M. (1983) Organ related differences in binding of Dolichos bifluorus agglutinin to vascular endothelium. *Dev. Biol.* **96**, 535-541.

POTTEN, C. S. (1974) The epidermal proliferative unit: the possible role of the central basal cell. *Cell Tiss. Kinet.* **7**, 77-88.

POTTEN, C. S. (1976) Identification of clonogenic cells in the epidermis and the structural arrangement of the epidermal proliferative unit (EPU), in *Stem Cells Of Renewing Populations* (Academic Press), pp 91-102.

POTTEN, C. S. (1977) Extreme sensitvity of some intestinal crypt cells to X and γ irradiation. *Nature 269*, 518-521.

POTTEN, C. S. (1981) Cell replacement in the epidermis (keratopoiesis) via discrete units of proliferation. *Int. Rev. Cytol.* **69**, 271-318.

POTTEN, C. S. (1986) Cell cycles in cell hierarchies. *Int. J. Radiat. Biol.* **49**, 257-278.

POTTEN, C. S. & HENDRY, J. H. (1973) Clonogenic cells and stem cells in epidermis. *Int. J. Radiat. Biol.* **24**, 537-40.

POTTEN, C. S. & HENDRY, J. H. (1985) The micro-colony assay in mouse small intestine, in *Cell Clones* (Potten, C. S. & Hendry, J. H. eds.) pp 50-60. Churchill Livingstone, Edinburgh.

POTTEN, C. S., HENDRY, J. H. & MOORE, J. V. (1987) Estimates of the number of clonogenic cells in the crypts of mouse small intestine. *Virchows Archiv* B **53**, 227-234.

POTTEN, C. S., HUME, W. J., REID, P. & CAIRNS, J. (1978) The segregation of DNA in epithelial stem cells. *Cell* **15**, 899-906.

POTTEN, C. S. & LOEFFLER, M. (1990) Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the Crypt. *Development* **110**, 1001-1020.

POTTEN, C. S. & MORRIS, R. J. (1988) Epithelial stem cells in vivo. *J. Cell Sci. Suppl.* **10**, 45-62.

PRICE, J. (1987) Retroviruses and the study of cell lineage. *Development* **101**, 409-419.

PRICE, J. & THURLOW, L. (1988) Cell lineage in the rat cerebral cortex: a study using retroviral-mediated gene transfer. *Development* **104**, 473-482.

PRICE, J., TURNER, D. & CEPKO, C. (1987) Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. *Proc. Natl. Acad. Sci. USA.* **84**, 156-160.

PROVOST, G. S., KRETZ, P. L., HAMNER, R. T., MATTHEWS, C. D., ROGERS, B. J., LUNDBERG, K. S., DYCAICO, M. J. & SHORT, J. M. (1993) Transgenic systems for in vivo mutation analysis. *Mutat. Res.* **288**, 133-149.

RAFF, M. C., ABNEY, E. R., COHEN, J., LINDSAY, R. & NOBLE, M. (1983a) Two types of astrocytes in cultures of developing rat white matter: differences in morphology, surface gangliosides and growth characteristics. *J. Neurosci.* **3**, 1289-1300.

RAFF, M. C. & LILLIEN, L. E. (1988) Differentiation of a bipotential glial progenitor cell: what controls the timing and the choice of developmental pathway? *J. Cell Sci. Suppl.* **10**, 77-83.

RAFF, M. C., MILLER, R. H. & NOBLE, M. (1983b) A glial progenitor cell that develops into an astrocyte or an oligodendrocyte depending on the constituents of the culture medium. *Nature* **303**, 390-396.

RAKIC, P. (1988) Specification of cerebral cortical areas. *Science* **241**, 170-176.

RAO, M. S., YELDANDI, A. V. & REDDY, J. K. (1990) Stem cell potential of ductular and periductular cells in the adult rat pancreas. *Cell Differen. Dev.* **29**, 155-163.

REDDY, J. K., RAO, M. S., QURESHAI, S. A., REDDY, M. K., SCARPELLI, D. G. & LALWANI, N. D. (1984) Induction and origin of hepatocytes in the pancreas. *J. Cell Biol.* **98**, 2082-2090.

REYNOLDS, A. J. & JAHODA, C. A. B. (1991a) Hair follicle stem cells ? A distinct germinative epidermal population is activated *in vitro* by the presence of hair dermal papilla cells. *J. Cell Sci.* **99**, 373-385.

REYNOLDS, A. J. & JAHODA, C. A. B. (1991b) Inductive properties of hair follicle cells. *Ann. N. Y. Acad. Sci.* **642**, 226-242.

REYNOLDS, A. J. & JAHODA, C. A. B. (1992) Cultured dermal papilla cells induce follicle formation and hair growth by transdifferentiation of adult epidermis. *Development* **115**, 587-593.

RHIM, J. S., FUJITA, J., ARNSTEIN, P. & AARONSON, S. A. (1986) Neoplastic conversion of human keratinocytes by Adevovirus 12-SV40 virus and chemical carcinogens. *Science* **232**, 385-388.

RIGAUD, G., GRANGE, T. & PICTET, R. (1987) The use of NaOH as a transfer solution of DNA onto nylon membrane decreases the hybridisation efficiency. *Nucl. Acids Res.* **15**, 857.

ROBERTSON, E. J. (1986) Pluripotential stem cell lines as a route to the mouse germline. *Trends Genet.* **2**, 9-13.

ROCHAT, A., KOBAYASHI, K. & BARRANDON, Y. (1994) Location of stem cells of human hair follicles by clonal analysis. *Cell* **76**, 1063-1073.

ROTH, K. A., HERMISTON, M. L. & GORDON, J. I. (1991b) Use of transgenic mice to infer the biological properties of small intestinal stem cells and to examine the lineage relationships of their descendants. *Proc. Natl. Acad. Sci. USA.* **88**, 9407-9411.

ROTH, K. A., RUBIN, E. H., BIRKENMEIER, E. H. & GORDON, J. I. (1991a) Expression of liver fatty acid binding protein/human growth hormone fusion genes within the enterocyte and enteroendocrine cell populations of fetal transgenic mice. *J. Biol. Chem.* **226**, 5949-5954.

ROTH, K. A., HERTZ, J. M. & GORDON, J. I. (1990) Mapping enteroendocrine cell populations in transgenic mice reveals an unexpected degree of complexity in cellular differentiation within the gastrointestinal tract. *J. Cell Biol.* **110**, 1791-1801.

RUBIN, D. C., ROTH, K. A., BIRKENMEIER, E. H. & GORDON J. I. (1991) Epithelial cell differentiation in normal and transgenic mouse intestinal isografts. *J. Biol. Chem.* **113**, 1183-1192.

RUBIN, D. C., SWIETLICKI, E., ROTH, K. A. & GORDON, J. I. (1992) Using fetal intestinal isografts from normal and transgenic mice to study the programming of positional information along the duodenal to colonic axis of the intestinal epithelium. *J. Biol. Chem.* **267**, 15122-15133.

RUDLAND, P. S. & BARRACLOUGH, R. (1988) Stem cells in mammary gland differentiation and cancer. *J. Cell Sci. Suppl.* **10**, 95-114.]

SACHS, L. (1987) The molecular control of blood cell development. *Science* **238**, 1374-1379.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & EHRLICH, H. A. (1988) Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-481.

SANES, J. F. (1989) Analysing cell lineage with a recombinant retrovirus. *Trends Neurosci.* **12**, 21-28.

SANES, J. R., RUBENSTEIN, J. L. R. & NICOLAS, J.-F. (1987) Use of a recombinant retrovirus to study post-implantation cell lineage in mouse embryos. *E.M.B.O.* **5**, 3133-3142.

SAWYER, T. W., GILL, R. D., SMITH-OLIVER, T., BUTTERWORTH, B. E. & DIGIOVANNI, J. (1988) Measurement of unscheduled DNA synthesis in primary cultures of adult mouse epidermal keratinocytes. *Carcinogenesis* **9**, 1197-1202.

SCHAFF, D. A., JARRETT, R. A., DLOUHY, S. R., PONNIAH, S., STOCKLEMAN, M., STAMBROOK, P. J. & TISCHFIELD, J. A. (1990) Mouse transgenes in human cells detect specific base substitutions. *Proc. Natl. Acad. Sci. USA.* **87**, 8675-8679.

SCHIERENBERG, E. (1989) Cytoplasmic determination and the distribution of developmental potential in the embryo of Ceanorhabditis elegans. *Bioessays* **10**, 99-104.

SCHMIDT, G. H., BLOUNT M. A. & PONDER, B. A. J. (1987) Immunochemical demonstration of the clonal organisation of chimaeric mouse epidermis. *Development* **100**, 533-541.

SCHMIDT, G. H., GARBUTT, D. J., WILKINSON, M.M., & PONDER, B. A. J. (1985b) Clonal analysis of intestinal crypt populations in mouse aggregation chimaeras. *J. Embryol. Exp. Morphol.* **85**, 121-130.

SCHMIDT, G. H., O'SULLIVAN, J. F. & PAUL D. (1990) Ethylnitrosourea-induced mutations in vivo involving Dolichos bifluorus agglutinin receptor in mouse intestinal epithelium. *Mutat. Res.* **228**, 149-155.

SCHMIDT, G. H., WILKINSON, M. M. & PONDER, B. A. J. (1985a) Cell migration pathway in the intestinal epithelium: An in situ marker system using mouse aggregation chimaeras. *Cell* **40**, 425-429.

SCHMIDT, G. H., WINTON, D. J. & PONDER, B. A. J. (1988) Development of the pattern of cell renewal in the crypt-villus unit of the chimaeric mouse small intestine. *Development* **103**, 785-790.

SCHOFIELD, R., DEXTER, T. M., LORD, B. I. & TESTA, N. G. (1986) Comparison of haemopoiesis in young and old mice. *Mech Ageing Dev.* **34**, 1-12.

SEEBERG, E. & FUCHS, R. P. P. (1990) Acetylaminofluorene bound to different guanines of the sequence -GGCGCC- is excised with different efficiencies by the UvrABC excision nuclease in a pattern not correlated to the potency of mutation induction. *Proc. Natl. Acad. Sci. USA.* **87**, 191-194.

SELL, S. & PIERCE, G. B. (1994) Maturation arrest of stem cell differentiation is a common pathway for the cellular origin of teratocarcinomas and epithelial cancers. *Laborat. Invest.* **70**, 6-22.

SENGEL, P. (1986) Epidermal-dermal interactions, in *Biology Of The Integument in Vertebrates, Volume 2* (Bereiter-Hahn, J., Matoltsy, A.G., Richards, K. S. eds.), pp 374-408. Springer-Verlag, Berlin.

SHEPHARD, S. E., SENGSTAG, C., LUTZ, W. K. & SCHLATTER, C. (1993) Mutations in liver DNA of lacI transgenic mice (big blue) following subchronic exposure to 2-actylaminofluorene. *Mutat. Res.* **302**, 91-96.

SHIBUTANI, S., GENTLES, R., JOHNSON, F. & GROLLMAN, A. P. (1991) Isolation and characterisation of oligodeoxynucleotides containing dG-N^2-AAF and oxidation products of dG-C8-AF. *Carcinogenesis* **12**, 813-818.

SCHIERENBERG, E. (1989) Cytoplasmic determination and distribution of developmental potential in the embryo of Caenorhabditis elegans. *Bioessays* **10**, 99-104.

SHIOJIRI, N., LEMIRE, J. M. & FAUSTO, N. (1991) Cell lineages and oval progenitors in rat liver development. *Cancer Res.* **51**, 2611-2620.

SHIRAI, A., HOLMES, K. & KLINMAN, D. (1993) Detection and quantitation of cells secreting IL-6 under physiologic conditions in BALB/c mice. *J. Immunol.* **150**, 793-799.

SICZKOWSKI, M., CLARKE, D. & GORDON, M. Y. (1992) Binding of primitive haematopoietic progenitor cells to marrow stromal cells involves heparan sulphate. *Blood* **80**, 912-919.

SIGAL, S. H., BRILL, S., FIORINO, A. S. & REID, L. M. (1992) The liver as a stem cell and lineage system. *Am. J. Physiol.* **263**, G139-G148.

SMALL, R. K., RIDDLE, P. & NOBLE, M. (1987) Evidence for migration of oligodendrocyte-type 2 astrocyte progenitor cells into the developing rat optic nerve. *Nature* **329**, 155-157.

SMITH, A. G. (1992) Mouse embryo stem cells: their identification, propagation and manipulation. *Semin. Cell Biol.* **3**, 385-399.

SMITH, L. G., WEISSMAN, I. L. & HEIMFELD, S. (1991) Clonal analysis of haemopoietic stem cell differentiation in vivo. *Proc. Natl. Acad. Sci. USA.* **88**, 2788.

SNELL, G. D. & STEVENS, L. C. (1966) Early embryology, in *Biology of the Laboratory Mouse Second Edition* (Green, L. E. ed.) pp. 205-245 McGraw-Hill, New York.

SNELL, S. (1983) Comparison of oval cells induced in rat liver by feeding N-2-flourenacetylamide in a choline devoid diet and bile duct cells induced by feeding 4,4'-diaminodiphenylmethane. *Cancer Res.* **43**, 1761-1767.

SNOW, M. H. L. & TAM, P. P. L. (1979) Is compensatory growth a complicating factor in mouse teratology. *Nature* **279**, 554-557.

SOLTER, D., SKREB, N. & DAMJANOV, I. (1970) Extrauterine growth of mouse egg cylinders results in malignant teratoma. *Nature* **227**, 503-504.

SOLURSH, M. (1989) Cartilage stem cells: regulation of differentiation. *Connect. Tiss. Res.* **20**, 81-89.

SORIANO, P. & JAENISCH, R. (1986) Retroviruses as probes for mammalian development: allocation of cells to the somatic and germ cell lineages. *Cell* **45**, 19-29.

SPANGRUDE, G. J. (1989) Enrichment of mouse haemopoietic stem cells: Diverging roads. *Immunol. Today* **10**, 344.

SPANGRUDE, G. J., HIEMFIELD, S. & WEISSMAN, I. L. (1988) Purification and characterisation of mouse haematopoietic stem cells. *Science* **241**, 58.

SPANGRUDE, G. J. & JOHNSON, G. R. (1990) Resting and activated subsets of mouse multipotent haematopoietic stem cells. *Proc. Natl. Acad. Sci. USA.* **87**, 7443-7447.

SPANGRUDE, G. J. & SCOLLAY, R. (1990) A simplified method for enrichment of mouse haemopoietic stem cells. *Exp. Haematol.* **18**, 920.

SPANGRUDE, G. J., SMITH, L., UCHIDA, N., IKUTA, K., HEIMFELD, S., FRIEDMAN, J. & WIESSMAN, I.L. (1991) Mouse haematopoietic stem cells. *Blood* **78**, 1395-1492.

STEINMULLER, D. (1971) A reinvestigation of epidermal transplantation during chemical carcinogenesis. *Cancer. Res.* **31**, 2080-2084.

STEVENS, L. C. (1970) The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev. Biol.* **21**, 364-382.

STOCKDALE, F. E. (1992) Myogenic cell lineages. *Dev. Biol.* **154**, 284-298.

STEVENS, L. C. (1983) The origin and development of testicular, ovarian and embryo-derived teratomas, in *Teratocarcinoma Stem Cells* (Silver, L. M., Martin, G. R. & Strickland, S. eds.) pp 26-36. Cold Spring Harbour, New York.

STROME, S. & WOOD, W. B. (1983) Generation of asymmetry and segregation of germ-line granules in early Caenorhabditis elegans embryos. *Cell* **35** 15-25.

STRUHL, G. & BASLER, K. (1993) Organizing activity of Wingless protein in Drosophila. *Cell* **72**, 527-540.

SULSTON, J. E. & HORVITZ, H. R. (1977) Post-embryonic cell lineages of the nematode, Ceanorhabditis elegans. *Dev. Biol.* **56**, 110-156.

SULSTON, J. E., SCHIERENBERG, E., WHITE, J. G. & THOMSON, J. N. (1983) The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* **100**, 64-119.

SUN, T.-T., COTSARELIS, G. & LAVKER, R. M. (1991) Hair follicle stem cells: the bulge activation hypothesis. *J. Invest. Dermatol.* **96**, 77s-78s.

SWEETSER, D. A. BIRKENMEIER, E. H., HOPPE, P. C., MCKEEL, D. W. & GORDON, J. I. (1990) Mechanisms underlying generation of gradients in gene expression within the intestine: an analysis using transgenic mice containing fatty acid binding protein/human growth hormone fusion genes. *Genes Dev.* **2**, 1318-1332.

SWEETSER, D. A., LOWE, J. B. & GORDON, J. I. (1986) The nucleotide sequence of the rat liver fatty acid binding protein: Evidence that exon 1 encodes an oligopeptide domain shared by a family of proteins which binds hydrophobic ligands. *J. Biol Chem.* **261**, 5553-5561.

TEMELES, D, S., MCGRATH, H. E., KITTLER, E, L., SHADDUCK, R. K., KISTER, V. K., CRITTENDEN, R. B., TURNER, B. L. & QUESENBERRY, P. J. (1993) Cytokine expression from bone marrow derived macrophages. *Exp. Haematol.* **21**, 388-393.

TEMPLE, S. & RAFF, M. C. (1986) Clonal analysis of oligodendrocyte development in culture: evidence for a developmental clock which counts divisions. *Cell* **44**, 773-779.

THOFT, R. A., WILEY L. A. & SUNDARRAJ, N. (1989) The multipotential stem cells of the limbus. *Eye* **3**, 109-113.

THOMPSON, E. M., FLEMING, K. A., EVANS, D. J., FUNDELE, R., SURANI, M. A. & WRIGHT, N. A. (1990) Gastric endocrine cells share a clonal origin with other gut cell lineages. *Development* **110**, 477-481.

THORGEIRSSON, S. S. (1993) Hepatic stem cells. *Am. J. Pathol.* **142**, 1331-1333.

THURMAN, R. G. & KAUFFMAN, F. C. (1985) Sublobular compartmentation of pharmacologic events (SCOPE): metabolic fluxes in periportal and pericentral regions of the liver lobule. *Hepatology* **5**, 144-151.

TILL, J. E. & MCCULLOCH, E. A. (1961) A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res.* **14**, 215-222.

TIWAWECH, D., HASEGAWA, R., KURATA, Y., TATEMATSU, M., SHIBATA, M.-A., THAMAVIT, W. & ITO, N. (1991) Dose dependent effects of 2-acetylaminofluorene on hepatic foci development and cell proliferation in rats. *Carcinogenesis* **12**, 985-990.

TRABER, P. G., CHIANALE, J., & GUMUCIO, J. (1988) Physiologic significance and regulation of hepatocellular heterogeneity. *Gastroenterology* **95**, 1130-1143.

TSENG S. C. (1989) Concept and application of limbal stem cells. *Eye* **3**, 141-157.

TURNER, D. L. & CEPKO, C. L. (1987) A common progenitor for neurons and glia persists in rat retina late in development. *Nature* **328**, 131-136.

TURNER, D. L., SNYDER, E. Y. & CEPKO, C. L. (1990) Lineage independent determination of cell type in embryonic mouse retina. *Neuron* **4**, 833-845.

UITERDIJK, H. G., PONDER, B. A. J., FESTING, M. F. W., HILGERS, J., SKOW, L. & VAN NIE, R. (1986) The gene controlling the binding site of Dolichos bifluorus agglutinin is on chromosome 11 of the mouse. *Genet. Res.* **47**, 125-129.

VINCENT, J. P. & O'FARRELL, P. H. (1992) The state of engrailed expression is not clonally transmitted during early Drosophila development. *Cell* **68**, 923-931.

WALSH, C. & CEPKO, C. L. (1988) Clonally related cortical cells show several migration patterns. *Science* **241**, 1342-1345.

WALSH, C. & CEPKO, C. L. (1992) Widespread dispersion of neuronal clones accross functianal regions of the cerebral cortex. *Science* **255**, 434-440.

WALSH, C. & CEPKO, C. L. (1990) Cell lineage and cell migration in the developing cerebral cortex. *Experientia* 46, 940-947.

WALTER, C. A., LU, J., BHAKTA, M., MITRA, S., DUNN, W., HERBERT, D. C., WEAKER, F. J., HOOG, T., GARZA, P., ADRIAN, G. S. & KAMOLVARIN, N. (1993) Brain and liver targeted overexpression of O6-methylguanine DNA transferase in transgenic mice. *Carcinogenesis* **14**, 1537-1543.

WATT, F. M. (1988) Epidermal stem cells in culture. *J. Cell Sci. Suppl.* **10**, 85-94.

WEINBERG, W. C., MORGAN, D. L., GEORGE, C. & YUSPA, S. H. (1992) A comparison of interfollicular and hair follicle derived cells as targets fo the v-rasHa oncogene in mouse skin carcinogenesis. *Carcinogeneisis* **13**, 1119-1124.

WEINSTEIN, G. D., MCCULLOUGH, J. L. & ROSS, P. (1984) Cell proliferation in normal epidermis. *J. Invest. Dermatol.* **82**, 623-628.

WEISSMAN, B. E., ARBOLEDA, M. J., KIMOTO, K. A., MACDONALD, M. J., OCHIAI, R. & DER, C. J. (1989) Activation of ras oncogenes in chemically transformed BALB/MK2 mouse keratinocytes. *Molec. Carcinogenesis* **2**, 150-158.

WESTPHAL, H. & GRUSS, P. (1989) Molecular genetics of development studied in the transgenic mouse. *Ann. Rev. Cell Biol.* **5**, 181-196.

WETTS, R. & FRASER, S. E. (1988) Multipotent precursors can give rise to all major cell types of the frog retina. *Science* **239**, 1142-1145.

WHETTON, A. D. & DEXTER, T. M. (1993) Influence of growth factors and substrates on differentiation of haemopoietic stem cells. *Cur. Opin. Cell Biol.* **5**, 1044-1049.

WHITING, J., MARSHALL, H., COOK, M., KRUMLAUF, R., RIGBY, P. W. J., STOTT, D. & ALLEMANN, R. K. (1991) Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. *Genes Dev.* **5**, 2048-2059.

WILLIAMS, B. P., READ J. & PRICE, J. (1991) The generation of oligodendrocytes and neurons from a common precursor cell. *Neuron* **7**, 685-693.

WILLIAMS, D. A., LEMISCHKA, I. R., NATHAN, D. G. & MULLIGAN, R. C. (1984) Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* **310**, 476-480.

WILLIAMS, D. E., DE VRIES, P., NAMEN, A. E., WIDMER, M. B. & LYMAN, S. D. (1992) The steel factor. *Dev. Biol.* **151**, 368-376.

WILSON, J. W. & LEDUC, E. (1948) The occurence and formation of binucleate cells and polyploid nuclei in the rat liver. *Am. J. Anat.* **82**, 353-391.

WINTON, D. J., BLOUNT, M. A. & PONDER, B. A. J. (1988) A clonal marker induced by mutation in mouse intestinal epithelium. *Nature* **333**, 463-466.

WINTON, D. J., BLOUNT, M. A. & PONDER, B. A. J. (1991) Spontaneous mutation rate. *Nature* **352**, 201.

WINTON, D. J., HOWARD, L. & PONDER, B. A. J. (1992) Possible non-functional crypts in small intestine defined using mouse aggregation chimaeras. *Epith. Cell Biol.* **1**, 59-64.

WINTON, D. J., PEACOCK, J. H. & PONDER, B. A. J. (1989) Effect of gamma radiation at high and low-dose rate on a novel in vivo mutation assay in mouse intestine. *Mutagenesis* **4**, 404-406.

WINTON, D. J. & PONDER, B. A. J. (1991) Stem cell organisation in mouse small intestine. *Proc. R. Soc. Lond. Ser. B.* **241**, 13-18.

WITHERS, H. R. & ELKIND, M. M. (1969) Radiation sensitvity and fractionation response of crypt cells of mouse jeujunum. *Rad. Res.* **38**, 598-613.

WITHERS, H. R. & ELKIND, M. M. (1970) Microcolony survival assay for cells of mouse intestinal mucosa exposed to radiation. *Int. J. Rad. Biol.* **17**, 261-268.

WOLPERT, L. (1988) Stem cells: a problem in asymmetry. *J. Cell Sci. Suppl.* **10**, 1-9.

WOLSWIJK, G. & NOBLE, M. (1989) Identification of an adult-specific glial progenitor cell. *Development* **105**, 387-400.

WOLSWIJK, G., RIDDLE, P. & NOBLE, M. (1990) Co-existence of perinatal and adult forms of a glial progenitor cell during development of the rat optic nerve. *Development* **106**, 691-698.

WOLSWIJK, G., RIDDLE, P. & NOBLE, M. (1991) Platelet-derived growth factor is mitogenic for O-2A^{adult} progenitor cells. *Glia* **4**, 495-503.

WORTON, R. G., MCCULLOCH, E. A. & TILL, J. E. (1969) Physical separation of haemopoietic stem cells differing in their capacity for self-renewal. *J. Exp. Med.* **130**, 91.

WREN, D., WOLSWIJK, G. & NOBLE, M. (1990) In vitro analysis of the origin and maintenance of O-2Aadult progenitor cells. *J. Cell Biol.* **116**, 167-176.

WRIGHT, N. A. (1983) The cell proliferation kinetics of the epidermis, in *Biochemistry And Physiology Of The Skin*. (Goldsmith L. A., *ed.*), pp 203-229. Oxford Univ. Press, Oxford.

WRIGHT, N. A. & ALISON, M. R. (1984) *The Biology of Epithelial Cell Populations. Vol 1*. pp. 536. Clarendon, Oxford.

WRIGHT, N. A. & APPLETON, D. R. (1980) The metaphase arrest method - a critical review. *Cell Tiss. Kinet.* **13**, 643-663.

WRIGHT, N. A. & IRWIN, M. (1982) The kinetics of villus cell populations in the mouse small intestine: Normal villi - the steady state requirement. *Cell Tissue Kinet.* **15**, 595-609.

WU, A. M., TILL, J. E., SIMINOVICH, L. & MCCULLOCH, E. A. (1968) Cytological evidence for a relationship between normal haematopoietic colony-forming cells and cells of the immune system. *J. Exp. Med.* **127**, 455-463.

YAMADA, T., PLACZEK, M., TANAKA, H., DODD, J. & JESSELL, T. M. (1991) Control of cell pattern in the developing nervous system: polarising activity of the floor plate and notochord. *Cell* **64**, 635-647.

YANG, J.-S., LAVKER, R. M. & SUN, T.-T. (1993) Upper human hair follicle contains a subpopulation of keratinocytes with superior in vitro proliferative potential. *J. Invest. Dermatol.* **101**, 652-659.

ZSEBO, K. M., WILLIAMS, D. A., GIESSLER, E. N., BROUDY, V. C., MARTIN F. H., ATKINS, H. L., HSU, R.-Y., BIRKETT, N. C., OKINO, K. H., MURDOCK, D. C., JACOBSEN, F. W., LANGLEY, K. E., SMITH, K. A., TAKEISHI, T., CATTENACH, B. M., GALLI, S. J. & SUGGS, S. V. (1990) Stem cell afctor is encoded at the Sl locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. *Cell* **63**, 213-224.