EXPRESSION OF FOREIGN AND ENDOGENOUS METALLOTHIONEIN GENES IN SKIN FIBROBLASTS DERIVED FROM A TRANSGENIC SHEEP

PAUL DAVID SMITH

THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF EDINBURGH

1990
To my parents for their moral support and easy terms credit facilities.
I declare that this work is my own, except where otherwise stated.

Paul Smith
ACKNOWLEDGMENTS

Firstly I would like to thank my supervisor, John Bishop, for his sound advice and for providing the equipment and materials used during the course of this work.

My thanks also go to those with whom I had useful discussions they include Raya Al-Shawi, Paul Simmons, John Clark and most of the members of the Bishop laboratory over the past four years.

Maggie McClenaghan provided the cell-line, John Jenkins the anti-SV40 Tag antibody and Julian Mercer the gene specific probes for the sheep metallothionein gene family. Thanks.

Special thanks must go to those who worked closest to me on the first floor, I am indebted to their patience and good humour. In particular I should like to thank Jane Kinnaird for her invaluable assistance and encouragement in all aspects of the project. I am grateful to Deborah Johnson, with whom I shared a laboratory, for her resilience and cheerful, plain talking. Thanks are due to Nydia de Avila Pineda for stemming the inexorable progress of disorder in the tissue culture room, for providing good-natured and enthusiastic help at all times, but most of all for smiling a great deal. Thanks also go to Qiang Lin Qi, Cynthia Heinlein, Melville Richardson, Steven Harrison and other members of the Bishop laboratory.

The photographs were lovingly prepared, noisily by Frank Johnston and in comparative silence by Graham Brown. Dennis Toghill and George Newall from the Genetics Dept. Workshop made the electroporation device. My thanks to them all.
PhD ABSTRACT

Transgenic sheep 229 carries one copy of a mouse metallothionein-I promoter HSV-tk fusion gene (pMK). Poly A+ RNA from a liver biopsy revealed no HSV-tk message.

A skin fibroblast cell-line was established from sheep 229 and is the subject of this thesis. The cell-line is non-immortal and senesces after 30-40 population doublings in vitro.

The cell-line carries one intact copy of pMK which is not rearranged as detected by southern blot analysis of genomic DNA. The cells do not display any HSV-tk activity and no HSV-tk mRNA is detected by northern analysis.

The expression of the sheep metallothionein gene family in the cell-line was investigated using gene specific probes. All four known active members are expressed and induced in response to zinc. The two major transcripts, from the Ia and II genes, are also induced by copper. The basal level and zinc induced expression of the gene family is not significantly altered by transformation of the cells by transfecting plasmids encoding the SV40 early region.

When mouse metallothionein-I promoter fusion genes (including pMK) are introduced into the cell-line (by electroporation) both expression and zinc regulation is noticed. Based upon stable transformation efficiency the mouse metallothionein-I promoter is 3-4 times less efficient than the SV40 early region promoter. Immuno-staining was used to detect the transient expression of the SV40 large T antigen (Tag) driven by either the SV40 early region (pSV3gpt) or the mouse metallothionein-I promoter (pMTLT). In the absence of zinc the number of cells detectably expressing pMTLT is 2% of that expressing pSV3gpt, following zinc induction the number of cells expressing pMTLT increases to 20% of that expressing pSV3gpt. The mMT-I promoter is
therefore deficient in the establishment of expression compared to
the SV40 early region promoter in sheep cells.

Since both the endogenous sheep metallothionein genes and
mouse metallothionein-I promoter fusion genes are active in the
cell-line, it is concluded that an overriding, negative position
effect is the most likely explanation for the failure of the pMK
transgene to be expressed.

The ability of viral and cellular oncogenes to transform the
cell-line was investigated. Expression of SV40 Tag causes the cells
to divide more rapidly, loose contact inhibition and anchorage
dependant growth but does not lead to immortalisation. Clones
established which express SV40 Tag have a vastly increased
proliferative capacity. However, after 60-70 population doublings
all clones enter crisis and die. Failure to express SV40 Tag is not
the cause of crisis.

No stable phenotypic effect is noticed when activated Ha-ras
and c-myc (singly or in conjunction) or the bovine papilloma virus
transforming region are introduced. A transforming murine p53 clone
severely inhibits the plating efficiency of the cells.

Attempts were made to re-claim the transgene for further
analysis. Screening of a random genomic library failed to detect any
positive clones. Attempts to clone a Hind III fragment containing
the whole transgene from size fractionated DNA enriched for the
fragment, were also unsuccessful.
CONTENTS

INTRODUCTION

UNUSUAL ASPECTS OF TRANSGENE EXPRESSION
Consistent Non-Expression of Transgenes 1
Low Level Expression 3
Occasional Non-Expression Occurs with most Transgenes 4
Ectopic Expression in Transgenic Animals 6
Expression of Genes from Heterologous Species in Transgenic Animals 8
Variation between Lines 10

ASSEMBLY AND INTEGRATION OF ARRAYS OF FOREIGN DNA MOLECULES: IMPLICATIONS FOR EXPRESSION 10

THE ASSEMBLY OF FOREIGN DNA MOLECULES INTO CONCATENEMERIC ARRAYS MAY INFLUENCE EXPRESSION 12

INFLUENCE OF FLANKING CHROMOSOMAL DNA: THE POSITION EFFECT 14
P-element Mediated Gene Transfer in Drosophila 16
Insertion of Retroviral Provirus 17
The Position Effect in Transgenic Animals and Transfected Cells 20

CONSISTENT, HIGH LEVEL EXPRESSION RELATED TO COPY NUMBER IN TRANSGENIC MICE 22

METHYLATION AND GENE ACTIVITY: IMPLICATIONS FOR GENE ACTIVITY 32
Methylation of Retroviral Proviruses 36
Methylation of Microinjected Transgenes: Implications for Expression 37
Expression of Autosomal Genes Integrated into the
Skin Fibroblasts from Sheep 229 Harbour a Single, Intact Copy of pMK

FIGURE 1: RESTRICTION MAP OF pMK AND THE TRANSGENIC INSERT

FIGURE 2: SOUTHERN BLOTTING ANALYSIS OF THE TRANSGENIC INSERT IN 229 SKIN FIBROBLASTS

The Promoter and Coding Sequences are not Obviously Rearranged or Deleted

The 3' Bam HI Site has been Regenerated

FIGURE 3: TWO POSSIBLE MECHANISMS BY WHICH THE 3' BAM HI SITE IS REGENERATED

Methylation of the 229 Insert

FIGURE 4: METHYLATION OF THE 229 INSERT

CHAPTER 2: EXPRESSION OF THE pMK TRANSGENE IN 229 SKIN CELLS

There is no detectable HSV-tk Activity in 229 Skin Cells

TABLE 3: NO HSV-TK ACTIVITY IN 229 CELLS

There is no Detectable HSV-TK mRNA IN 229 CELLS

FIGURE 5: NO DETECTABLE HSV-tk mRNA IN 229 CELLS

CHAPTER 3: EXPRESSION OF THE SHEEP METALLOTHIONEIN GENE FAMILY IN SHEEP SKIN CELL-LINES

Use of RNA Probes Specific to the Members of the Sheep Metallothionein Gene Family

FIGURE 6: SPECIFICITY OF SHEEP METALLOTHIONEIN GENE SPECIFIC RIBO-PROBES

Expression of the Sheep Metallothionein Gene Family in 229 Skin Fibroblasts: Response to Zinc and Copper

FIGURE 7: EXPRESSION OF MEMBERS OF THE SHEEP METALLOTHIONEIN GENE FAMILY IN 229 SKIN FIBROBLASTS

FIGURE 8: RESPONSE TO VARIOUS COPPER INDUCTION REGIMES
Expression of the Sheep Metallothionein Gene Family in Other Cell-lines

FIGURE 9: EXPRESSION OF THE SHEEP METALLOTHIONEIN GENE FAMILY IN NORMAL AND TRANSFORMED SHEEP SKIN CELL-LINES

CHAPTER 4: THE MOUSE MT-I PROMOTER IS FUNCTIONAL IN SHEEP SKIN FIBROBLASTS

TABLE 4: STABLE TRANSFECTION OF 229 SKIN CELLS
Co-transfection Using SV40 Tag to Promote Clonal Expansion
Transfection Using Single Plasmids

CHAPTER 5: EXPRESSION AND REGULATION OF mMT-I FUSION GENES IN 229 SKIN FIBROBLASTS

FIGURE 10: ANALYSIS OF FOREIGN AND ENDOGENOUS METALLOTHIONEIN GENE EXPRESSION IN CLONAL CELL-LINES
Expression of pMTneo in Non-transformed 229 Skin Cells and Mouse L-Cells
FIGURE 11: EXPRESSION OF pMTneo IN MOUSE L-CELLS AND NON-TRANSFORMED 229 SKIN CELLS
Sheep Cells can Transcribe pMK and Produce Functional HSV-tk
FIGURE 12: REGULATION OF TRANSFECTED pMK IN SHEEP CELLS
TABLE 5: PRODUCTION OF FUNCTIONAL HSV-tk IN SHEEP CELLS

CHAPTER 6: TRANSIENT EXPRESSION VISUALIZED IN INDIVIDUAL CELLS

Validity of the Assay
Using pSV3neo and pMTLT, SV40 Tag is Transiently Expressed in the Nucleus of 229 Skin Cells
FIGURE 13: TRANSIENT EXPRESSION OF pSV3neo AND pMTLT IN 229 SKIN CELLS
The effect of Zinc upon the Transient Expression of pMTLT
FIGURE 14: REGULATION BY ZINC OF SV40 Tag EXPRESSION IN 229 SKIN CELLS STABLY TRANSFECTED WITH pMTLT

FIGURE 15: TRANSIENT AND STABLE EXPRESSION OF pMTneo in COS-7 CELLS

Enhancer Mediated Activation of Transcription: As assayed by indirect immuno-staining

Metallothionein Promoters as Inducible Enhancer Elements

CONCLUSIONS REGARDING THE INACTIVITY OF THE 229 TRANSGENIC INSERT

CHAPTER 7: ATTEMPTS TO EXTEND THE in vitro LIFESPAN OF 229 SKIN FIBROBLASTS BY TRANSFECTION OF VIRAL AND CELLULAR ONCOGENES

Identification and Characterisation of Transformed Cells

The Use of SV40 Tag to Transform Sheep Skin Cells

SV40 Tag Producing Cell-lines Grow More Rapidly to a Higher Saturation Density Than Primary 229 Skin Fibroblasts

FIGURE 16: ALTERED MORPHOLOGY AND GROWTH OF 229 SKIN CELLS FOLLOWING TRANSFORMATION BY SV40 Tag

FIGURE 17: SV40 Tag PRODUCING CELL-LINES GROW MORE RAPIDLY TO A HIGHER SATURATION DENSITY THAN PRIMARY 229 SKIN FIBROBLASTS

TABLE 6: PLATING EFFICIENCIES OF TRANSFORMED CELL-LINES ON PLASTIC CULTURE DISHES AND IN SOFT AGAR

The Plating Efficiency of the Transformed Cell-lines is Greatly Increased

The Transformed Cell-lines will form Colonies in Soft Agar

Sheep Skin Cells Expressing SV40 Tag have a Greatly Increased Proliferative Capacity

TABLE 7: EFFECT OF MURINE p53 AND SV40 Tag UPON COLONY FORMATION IN 229 SKIN FIBROBLASTS

TABLE 8: PROLIFERATIVE CAPACITIES OF TRANSFORMED AND NON-TRANSFORMED SHEEP CELL-LINES
FIGURE 18: PROLIFERATIVE HISTORIES OF NON-TRANSFORMED 229 SKIN CELLS AND SV40 Tag TRANSFORMED DAUGHTER CLONES

Cells Transformed with Plasmid pMTLT also have an Extended but Finite Lifespan in Culture

Failure to Synthesise SV40 Tag at High Levels does not Account for the Senscence of Transformed Cell-lines

FIGURE 19: pMTLT TRANSFORMED CELL-LINES SYNTHESISE SV40 Tag DURING CRISIS

Utility of SV40 Transformed Sheep Cells

DISCUSSION: SV40 MEDIATED TRANSFORMATION OF CULTURED CELLS

Immortality, the Escape From Senescence

Transformation of Human Skin Fibroblasts by SV40 Tag: Parallels with Sheep Skin Fibroblasts.

ATTEMPTS TO TRANSFORM SHEEP SKIN FIBROBLASTS WITH OTHER VIRAL AND CELLULAR ONCOGENES

TABLE 9: TRANSFECTION OF ADDITIONAL VIRAL AND CELLULAR ONCOGENES

TABLE 10: PROLIFERATIVE CAPACITIES OF CLONAL CELL-LINES

Summary: General Conclusions Regarding th Transformation of Sheep Skin Fibroblasts

APPENDIX 1: USE OF ELECTROPORATION TO INTRODUCE DNA INTO SHEEP SKIN FIBROBLASTS AND OTHER CELL-LINES

Transfection of Cultured Cells by Electroporation

FIGURE 20: CIRCUIT DIAGRAM OF THE ELECTROPORATION DEVICE

TABLE 11: EXPRESSION OF HSV-tk ACTIVITY IN BHK tk- CELLS FOLLOWING ELECTROPORATION AT VARIOUS VOLTAGES

TABLE 12: STABLE TRANSFECTION OF BHK CELLS BY ELECTROPORATION

Successful Electroporation of BHK tk- Cells

FIGURE 21: TRANSIENT EXPRESSION OF SV40 Tag IN BHK CELLS FOLLOWING ELECTROPORATION AT VARIOUS VOLTAGES
Transfection of Sheep Skin Cells by Electroporation

**FIGURE 22: TRANSIENT EXPRESSION OF SV40 Tag IN 229 SKIN CELLS**

**TABLE 13: TRANSFECTION OF 229 SKIN CELLS BY ELECTROPORATION**

Recovery of DNA from Cells after Electroporation

**FIGURE 23: RECOVERY OF FOREIGN DNA FROM CULTURED CELLS FOLLOWING ELECTROPORATION**

**APPENDIX 2: ATTEMPTS TO CLONE THE 229 TRANSGENIC INSERT**

Construction of a Random Genomic Library from 229 Skin Cells

**TABLE 14: GENOMIC LIBRARIES CONSTRUCTED FROM 229 DNA**

Attempt to Clone a Hind III Fragment Spanning the 229 insert

**MATERIALS AND METHODS**

**REFERENCES**

**ABBREVIATIONS USED IN THE TEXT**
INTRODUCTION

UNUSUAL ASPECTS OF TRANSGENE EXPRESSION

It is apparent that most transgenes are expressed appropriately in at least some of the lines generated. This is particularly true of unaltered genes (Palmiter & Brinster, 1986; Bishop & Al-Shawi, 1989). Genes comprising a hybrid of heterologous promoter and coding sequences appear to be more susceptible to low level, aberrant or non-expression (Palmiter & Brinster, 1986; Bishop & Al-Shawi, 1989).

Consistent Non-Expression of Transgenes

Correct expression is to be expected if the essential regulatory sequences are present in the transgene DNA construct. The converse case may or may not be true when consistent non-expression is the outcome with a given transgene. Experiments with the human α-globin gene are a case in point. Even when the entire α-globin locus on a 42 kb cosmid clone is introduced no expression is detected in any tissue of the 61 transgenic animals generated (Ryan et al., 1989a). However, when two erythroid specific DNAase I super-hypersensitive (SHS) sites usually found upstream of the human β-globin locus (sites II and I) (Ryan et al., 1989b) are included immediately upstream of the human α1-globin gene, high level, erythroid specific expression is observed in all mice retaining an intact copy of the transgene (Ryan et al., 1989a). These SHS sites lie within a region of the human β-globin locus termed the dominant control region (henceforth referred to as DCR), which also includes three other such sites and can confer high level erythroid specific expression on the human β-globin gene and heterologous genes in transgenic mice and cultured cells (Ryan et al., 1989b; Van Assendelft et al., 1989; Talbot et al., 1989; Grosveld et al., 1987 and later). Control transgenic mice harbouring the equivalent α1-
globin construct without the super-hypersensitive sites again failed to express human α1-globin mRNA. Clearly, the DCR from the β-globin locus can compensate for the lack of similar sequences in the α-globin construct used previously. If such sequences are associated with the α-globin locus then they must lie far 5' or 3' to the normal locus. It is not clear what role if any the α-globin gene and flanking sequences play in the generation of erythroid specificity. The α-globin sequences may determine the developmental timing of expression. Alternatively the α-globin gene may simply play the role of a passenger gene as do Thy-1 and HSV-tk neo when linked to the β-globin DCR and introduced into mouse erythroleukaemic cells (Van Assendelft et al., 1989).

Prokaryotic sequences can have an overriding, prejudicial effect on the expression of a transgene. Vector sequences have been identified as inhibitory when linked to certain genes; phage λ sequences induce low frequency and level of human β-globin expression (Townes et al., 1985), plasmid sequences inhibit expression of mMT-hGH (mouse metallothionein promoter-human growth hormone fusion gene), α-actin and α-alpha fetoprotein genes (Hammer et al., 1985; Krumlfauf et al., 1985a; Shani, 1986). It is not clear why some genes should be more sensitive to the presence of prokaryotic sequences than others. The issue is complicated when one considers that expression of the human GH gene is inhibited by plasmid sequences when linked to the mMT-I promoter but not when coupled to the rat elastase promoter (Ornitz et al., 1985a; Ornitz et al., 1985b). Nor is the inhibitory effect of prokaryotic sequences a consistent feature of mMT-I fusion genes in transgenic mice (see later). That is to say, the effect is specific to certain combinations of promoters and coding sequences which are not necessarily affected when linked to other promoters or genes.

Non-expression can be the outcome whether vector sequences are present or not. Since some non-expressed hybrid transgenes are active when introduced into tissue culture cells (Palmiter & Brinster, 1986), the non-functional joining of diverse DNA segments
per se, cannot serve as an adequate explanation. More likely, failure of such transgenes to be expressed results from the transgene being subject to a novel developmental, differentiation programme which commits it to a stable, inactive state such as can occur when traditionally autosomal genes are inserted into the X-chromosome (see later). Such an effect may underlie the inactivation of mMT-β-galactosidase genes which are highly active for several days after injection into fertilised eggs yet fail to be expressed in transgenic mice (Palmiter & Brinster, 1986)

Low Level Expression

When unaltered genes are expressed at low levels the most likely explanation is the lack of critical regulatory sequences. The expression of the hß-globin gene in transgenic mice is a case in point (Ryan et al., 1989b; Grosveld et al., 1987). The same is also true of hybrid genes and often relates to the absence of regulatory elements 3' to the promoter eg (Aronow et al., 1989). When the same promoter eg elastase or mMT-I is spliced to several different genes the level of expression can be extremely variable (Palmiter & Brinster, 1986 and see later in this introduction mMT-I Fusion genes in transgenic mice). Clearly the promoter cannot ensure consistent levels of expression and though RNA stability is certainly involved, it is likely that different structural genes can influence the rate of transcription.

It has now been established that the presence of introns can increase the frequency of transgene expression and raise the mRNA levels 10-100 fold (Brinster et al., 1988). In one case, the primary effect of introns on rGH expression would appear to be on the rate of transcription rather than upon mRNA stability (Brinster et al., 1988). Since the effect is noticed in tissues which do not normally express growth hormone, the presence of tissue specific intronic enhancers cannot readily account for the effect. More puzzling and perhaps discounting the presence of an intronic general
enhancer is the failure of introns to achieve any effect in transfected cultured cells. This may be an example of a transcriptional effect only manifest during development and not when the gene is introduced into phenotypically stable cell-lines. Since the extended correlation between the presence of introns and increased rate of transcription was confined to one transgene (three mice), the generality of this type of intronic effect is open to question. The absence of an effect in cultured cells questions the validity of using immortal cultured cells to delineate specific regulatory sequences, thus, primary cells may be more informative. In contrast to the results reported by Brinster et al (1988), Buchman and Berg (1988) demonstrate clear intron dependent expression in cultured cells. The rabbit β-globin transcriptional unit failed to produce stable RNA in the absence of an intron. Even the stable RNA production of genes which were not totally dependent upon introns was increased when one was included. It is plausible that these two contradicting demonstrations of the effects of introns upon expression (Brinster et al., 1988; Buchman & Berg, 1988) may not have the same mechanistic basis. It is not clear whether the contradictory data relates to the fact that Brinster et al analysed expression of stably integrated genes whereas Buchman and Berg analysed transient expression of extra-chromosomal DNA.

Occasional Non-Expression Occurs with most Transgenes

Occasional non-expression is probably the result of stochastic phenomena relating to the assembly and integration of an array of foreign DNA molecules. In cases where transgenes are clearly functional in some lines, the inactivation probably occurs by chance. However, it is also clear that some constructs are more susceptible to silent integration than others. This is particularly true if the so called "position effect" is the cause of gene inactivity (for discussion of the position effect see later). The frequency of non-expression probably relates to the presence or absence of sequences required for full transgene promoter activity,
in other words the strength of the promoter. Absence of the entire complement of regulatory sequences may allow flanking chromosomal sequences to compromise expression. Alternatively, inclusion of all cis control sequences may insulate the transgene from such prejudicial effects. It should be noted, that interference in cis could arise from neighbouring copies of the transgene as well as from flanking chromosomal DNA. In cultured cells the frequency of stable transformation is strongly related to the potency of promoter sequences (Luciw et al., 1983; Spandidos & Wilkie, 1983). In transgenic mice some genes such as elastase and α-fetoprotein are usually appropriately expressed and at reasonable levels (Swift et al., 1984; Hammer et al., 1987). Others can exhibit a much lower frequency of expression for instance mMT-I hepatitis B surface antigen fusion genes (Chishri et al., 1985). In this case substitution of the promoter with a mouse albumin promoter resulted in high frequency expression (Palmiter and Brinster, 1986).

Arrays of foreign DNA molecules commonly found in transgenic animals are built up by a series of inter and intra-molecular reactions which involve random breakage, end-joining, homologous exchange and DNA synthesis (Bishop & Smith, 1989). Therefore point mutations, deletions and rearrangements of the transgenes must account for at least some of the observed cases of non-expression (see later for discussion).

Silencer sequences are proposed to act in a similar way to enhancers but have the converse effect. Trans-acting factors which interact with such sequences are presumed to suppress gene transcription (Maniatis et al., 1987; Levin & Manley, 1989). Examples of specific gene suppression include the negative regulation of the thyroid-stimulating hormone α gene by thyroid hormone which is thought to be achieved by receptor interaction adjacent to and downstream of the TATA box (Chatterjee et al., 1989) and the negative regulatory element within the mouse albumin enhancer has been proven to interact with a novel DNA-binding protein (Herbst et al., 1990). Integration within the sphere of
influence of a silencer type sequence might have an overriding negative effect in some cases. The role of specific silencing sequences in the inactivation of transgenes depends upon their ubiquity in the mammalian genome. We know from enhancer trap experiments in mammalian cells that the frequency of integration within functional range of an enhancer/promoter element is less than $10^{-5}$. This represents a 200 fold reduction compared to stable transformation with an enhancer containing construct (Hamada, 1986). Therefore, unless silencers are considerably more potent, common or effective at longer ranges than enhancer elements, their significance with respect to inactive transgenes is likely to be very slight.

Silencer sequences presumably mediate repression via specific, sequence dependent mechanisms. For a transgene to be trapped within this repression system it must lie proximal rather than distal to the silencer and therefore, for the reasons discussed above this must be considered a rare occurrence. Since the expression of many genes is restricted to certain cell-types, repression is both dominant and extremely common. It seems unlikely that such repression is mediated by specific mechanisms; more likely a general mechanism for the silencing of genes exists in most cases. Since inactive transgenes are also relatively common it is likely that the integration has occurred in a region of general inactivity rather than in discrete inactive locales where non-expression is established and maintained by specific trans-acting factors. This implies that repression is the norm and that expression must be positively established. For any given transgene this may largely depend upon the nature of the regulatory sequences present in the foreign DNA construct.

Ectopic Expression in Transgenic Animals

It is not uncommon, particularly when hybrid genes are used, for expression to be observed in a novel tissue with which the
transgenic sequences are not normally associated. Such ectopic expression can be common to all lines (Al-Shawi et al., 1988), arise sporadically in a small number of lines (Lacey et al., 1983). Ectopic expression may or may not be accompanied by expression in predicted tissues (Al-Shawi et al., 1988; Behringer et al., 1988). The anomalous expression of rabbit β-globin genes in mouse skeletal muscle and testes occurred in only 2 of 7 lines (Lacey et al., 1983). Sporadic ectopic expression is likely to be due to the individuality of each line, namely the site of integration and the arrangement of foreign DNA molecules at that locus. The presence of prokaryotic sequences has also been implicated in promoting ectopic expression (Palmiter & Brinster, 1986).

Occasional inappropriate expression may be more common than is immediately apparent since tissues which are sites of ectopic expression may, with good reason, not be analysed unless the line shows an unexpected phenotype such as a tumour. In this way several eccentric patterns of SV40 large T antigen expression have been detected (Behringer et al., 1988; Botteri et al., 1987; Messing et al., 1985).

For some DNA constructs ectopic expression is always detected in the same tissues. Most (if not all) examples of this arise when hybrid genes are used. It has become increasingly clear that sequences which regulate gene activity are not confined to the immediate 5' promoter regions. Essential cis-acting elements have been located far upstream (Grosveld et al., 1987), in intronic and 3' sequences (Greaves et al., 1989; Arnow et al., 1989; Bomstein et al., 1988; Karpinski et al., 1989; Behringer et al., 1987). Hybrid genes, by their nature, are likely to lack any intronic and 3' sequences which would normally complement the promoter to which they are coupled. Instead, the hybrid gene may represent the assembly of a novel combination of controlling elements which may provoke expression in tissues which are generally outwith their normal sphere of influence. One can envisage that switching on expression in novel tissues would be realised because the hybrid gene can
respond to tissue specific signals which promote transcription. The expression in ectopic tissues can be low (Lacey et al., 1983; Palmiter & Brinster, 1986) or very high (Al-Shawi et al., 1988; Low et al., 1985) when expression can exceed that in the predicted tissue. Whether low level and high level ectopic expression arise for the same reasons is not clear. When both ectopic and predicted expression are detected one may invoke a scenario whereby a tissue specific silencer locus is missing from the construct. Such cis elements might be the targets of tissue specific extinguisher loci such as that which suppresses expression of some liver specific genes in rat hepatoma cells bearing chromosome ii from mouse fibroblasts (Gourdeau et al., 1989). Clearly such an explanation cannot, by itself, account for all cases of ectopic in addition to predicted expression since, for instance, only a minority of mMT-I promoter fusion genes are expressed aberrantly in the central nervous system (see later).

Expression of genes from Heterologous Species in Transgenic Animals

For some genes, subtle differences in the tissue distribution of expression exist between species. Transgenic animals provide a good system in which to study the nature and causes of such differences. It is useful to determine whether precise patterns of expression are inherited with the gene and flanking DNA or dictated by the host environment. In the case of tissue specific genes expressed only in one cell-type the specificity is conserved between species and the transgene behaves accordingly (for review see Palmiter & Brinster, 1986).

As a general rule, the pattern of expression observed when genes from one species are introduced into the genome of another most closely resembles that of the donor rather than the host. Several examples are available (Koopman et al., 1989; Arnow et al., 1989; Kelsey et al., 1987; Gordon et al., 1987). The human α₁-antitrypsin gene largely retains a human-like pattern of expression
with the exception of the persistence of abundant transcripts in the adult small intestine (Koopman et al., 1989; Kelsey et al., 1987). The endogenous mouse transcripts are found in a limited but overlapping sub-set of those tissues expressing the human transgene.

When the human Thy-1 gene is transferred into mice, the expression also mimics that observed in humans and is not distorted by the divergent expression of the resident mouse gene (Gordon et al., 1987). In tissues where the expression does not overlap, the human gene is active when the mouse gene is inactive and vice-versa. In an equivalent experiment with the murine Thy1.1 gene, the endogenous expression is faithfully reproduced barring non-expression at some expected sites (Gordon et al., 1987). In this case, it would appear that cis-regulatory sites, or their arrangement has diverged between mice and humans. The alternate explanation invoking divergence of trans-regulatory functions cannot easily account for the convincing inactivity of the transferred mouse gene in tissues where the human gene is clearly expressed (Gordon et al., 1987).

An instance in which transgenic mice have suggested an alternate explanation for species specific gene regulation is the acquisition by human foetal globin genes of embryonic rather than foetal specific expression (Kollies et al., 1986). This might be explained by temporal variation in the expression of transregulators between mice and men. This is an example of anomalous developmental specific transgene expression. A hybrid Thy1.1 gene consisting of the 5' flanking and first 3 exons of the mouse gene coupled to the 3' region from the equivalent human gene (Kollies et al., 1987) was developmentally regulated in lymphoid tissue in a manner distinct from the endogenous gene since the expression of the transferred gene was specifically repressed during T cell maturation. With respect to lymphoid expression the hybrid transgene has characteristics of the human and mouse expression. Such a situation possibly results from a novel combination of regulatory sequences donated by the human and murine portions of the transgene.
Variation between Lines

Variation in expression between individuals of the same transgenic line has been observed (Palmiter et al., 1982; Al-Shawi et al., 1988; McGowan et al., 1989; Krumlauf et al., 1985b). In the case of a line of mice harbouring copies of pMK (mouse metallothionein promoter-HSV-tk fusion gene) extinction of expression was observed (see later). The influence of DNA methylation on "within line" expression is discussed later. To extrapolate from results in tissue culture, the expression may be unstable due to integration into specific regions of the genome, e.g., pericentromeric heterochromatin (Butner & Lo., 1986) and this instability might be due to rearrangements of the flanking sequence or methylation of the transgene. Rearrangement or mutation of the transgenes may also account for individual cases of variation. Since most transgenic mice are obtained by micro-injecting eggs collected from the F₂ offspring of a cross between two different inbred strains the G₁ offspring of the transgenic founder mouse may be considered the F₃ offspring of the original mating. At this stage independent assortment of unlinked alleles will have occurred; if two original inbred strains harbour different alleles of genetic modifiers (Belayew & Tilghman, 1982; Chandlee & Scandalio, 1987) novel combinations of transgene and modifier alleles will have been generated in the progeny of the transgenic founder mouse. It is plausible that this can account for within line variation. Indeed evidence in favour of this has recently emerged (McGowan et al., 1989) and is discussed later.

Assembly and Integration of Arrays of Foreign DNA Molecules: Implications for Expression

When cloned DNA is introduced into the mammalian pre-implantation embryo by pro-nuclear microinjection or into somatic cells by conventional transfection techniques, the DNA is sometimes incorporated into the host cell chromosomes. Most commonly,
this occurs at one or very few sites. Each integration is unique and represents a series of events which initially see some of the foreign DNA molecules assembled to produce an array or concatemer and then, usually one of these arrays or a single molecule may be integrated at random sites into the host genome. The assembly and integration of an array of foreign DNA molecules are separable and essentially stochastic phenomena; hence for a given foreign DNA construct each transgenic mouse line or stably transfected clonal cell-line, although sharing the same foreign DNA molecule, can in other respects be considered unique.

Although most transgenes are correctly regulated in at least some lines, expression is generally characteristic of the lines and not often conserved between lines. Typically the level of expression is highly variable between lines and does not relate to the number of copies of foreign DNA. Also it is common for the expression per copy of the transgene to be less than that of the resident homologue. Other aberrant phenomena are also apparent. Certain expression patterns occur sporadically and include non-expression in any tissue, non-expression in the predicted tissue, expression in a novel tissue and failure to express in a tissue which, although not a predicted site of expression, nonetheless shows consistent activity in all other lines. Where abnormal patterns of expression are peculiar to only one or a small number of lines or where the level of expression is highly variable between lines, the expression characteristics can be considered line specific and perhaps not directly caused by the presence or absence of cis regulatory sequences within the foreign DNA. This is to say, that like the structure and location of the transgenic integrant, the subsequent expression can also be considered unique to a particular line. As such, the expression is not only a characteristic of the type of DNA molecule introduced but also a heritable property of the integrant in each line.

Two features of the integrant vary between transgenic lines. Firstly the number and arrangement of foreign DNA molecules and
secondly the chromosomal location. Both may be considered as potential sources of the variation in expression between lines.

THE ASSEMBLY OF FOREIGN DNA MOLECULES INTO CONCATEMERIC ARRAYS MAY INFLUENCE EXPRESSION

When single copies are integrated they often, though probably not always, integrate via the original ends of the input molecule and will be non-rearranged and more or less intact depending on the extent of damage to the ends (Bishop & Smith, 1989). In extreme cases it can be clear that rearrangement of single copies has resulted in non-expression. For instance Ryan et al. (1989a) report the apparent deletion of the distal super-hypersensitive site from their ha-globin transgene construct and concomitant failure to express. The molecular events occurring prior to integration are discussed in detail in Bishop & Smith, (1989). The most common type of array is that composed of head to tail concatemers of the foreign DNA molecule, for instance in one study all 51 transgenic integrants harboured head-to-tail tandem arrays (Ryan et al., 1989b). Head to head and tail to tail arrangements are much less common. The head to tail arrangement is built up by repeated cycles of non-conservative homologous recombination between non-aligned linear molecules. Linear molecules can be circularised and then re-broken to produce circularly permuted linear molecules and it is these that are able to recombine with other linear molecules, including input form linear molecules, to extend an head to tail array. The recombination process involves exposure and annealing of homologous single-strands of the recombining molecules to form a duplex. Single strands not involved in heteroduplex formation, though not by definition non-homologous, are removed and the hybrid repaired. In view of recent experiments the single strands are likely to be exposed by a 5' exonuclease activity, in this case repair of the hybrid duplex involves DNA synthesis (Lin et al., 1987, 1990a and 1990b). Clearly random cleavage, ligation, exonuclease action and DNA synthesis are involved in the generation
of concatemeric arrays. Therefore since arrays can contain in excess of 100 copies of the foreign DNA molecule the inactivation of some of the copies by deletion; point mutation or rearrangement must be considered to be highly likely. Thus, the process of assembly of DNA molecules into an array could contribute to the typical lack of copy number dependent expression associated with transgenic animals (Bishop & Al-Shawi, 1989).

Previously it was suggested (Palmiter & Brinster, 1986; Bishop and Al-Shawi, 1989, and earlier in this introduction) that hybrid genes may represent the novel juxta-position of cis regulatory elements and thereby show unexpected expression characteristics. In a similar way, the copies of transgenes within an array experience a novel arrangement of functional as well as coding sequences. This may create the potential for unprecedented interactions with trans-regulatory proteins and, as a possible corollary, novel levels and patterns of expression. Interaction between copies of the transgene might be stimulatory or inhibitory. A cooperative effect between enhancer elements might increase the potential for transcription though such an effect might be copy number dependent. Conversely, interference between transcriptional units may allow an inverse correlation between copy number and expression for which there is a precedent (for example see Al-Shawi et al., 1988). Certainly, the expression per copy generally decreases with rising copy number. This implies either that at higher copy numbers an increasing number of the copies are inactive or that the expression of each copy is suppressed by the presence of additional copies. Transcriptional interference has been demonstrated between duplicated copies of the ha-globin gene in cultured cells. Substantial inhibition of the downstream gene by transcription of the upstream gene was reported. The inhibition is overcome by placing transcriptional termination signals between the two copies (Proudfoot, 1986). In a transient assay the inhibition was 3-fold whereas pools of stably transfected clones demonstrated a 5-20 fold excess of the message from the 5' copy. In this case the two copies were separated by 700 bp of 3' and 5' flanking sequence.
Down regulation of alternate promoters upstream of the \( \text{he} \)-globin gene in fully differentiated embryonic red blood cells is proposed to be due to transcriptional interference from \( \text{Alu} \) transcripts proceeding in the opposite direction between the alternate promoters and the gene (Wu et al., 1990). The possible implications of such observations for the expression of genes within arrays probably depends upon the presence or absence of cognate termination signals, these may lie hundreds to thousands of bases \( 3' \) of the polyadenylation site (Manley et al., 1989) and it seems plausible that some transgenes do not include these sites. The phenomenon of promoter suppression has been described in recombinant retroviruses bearing two selectable markers. In such a case selection for the expression of one apparently can preclude the expression of the other (Emmerman & Temin., 1984). However this clearly does not occur in the stable co-transfection of naked, cloned DNA into cultured cells.

Depletion of trans-acting factors cannot account for lack of copy number dependence since repression of endogenous genes is not detected when expression and copy number of the transgene is high (Grosveld et al., 1987; Van Assendelft et al., 1989; Aronow et al., 1989). An independant observation with similar implications is the additive effect on expression when two loci are united by mating transgenic strains (Bishop & Al-Shawi., 1989). Lastly, selective methylation of integrants or copies within an array might correlate with the expression. The possible effect of methylation with respect to the expression of transgenes is discussed in length later.

THE INFLUENCE OF FLANKING CHROMOSOMAL DNA: THE POSITION EFFECT

The host chromosomal DNA flanking the integrant may exert long or short range effects in \( \text{cis} \) upon the transgene and thereby influence expression. The effects might result in a quantitative reduction in transcription efficiency or complete silence of the gene. Alternatively, the effect might stimulate expression. The
imposition might be tissue specific and alter the expression in selected tissues, create novel sites of synthesis or suppression of the gene. All these effects have been observed in transgenic mice and where the occurrence is sporadic the position effect may be implicated. Position effect traditionally refers to chromosome rearrangements which alter or abolish developmentally regulated genes located great distances from the site of breakage. Examples of the position effect are observed after the juxtaposition of euchromatic and heterochromatic genomic regions which can occur during chromosome translocations (Spradling & Rubin, 1983). β-globin gene inactivation by DNA translocation in Dutch γβ-Thalassemia is considered to be a classical position effect phenomenon. (Kioussis et al., 1983).

It is common practice to label variable or occasional inappropriate expression of transgenes as position effect phenomena. This may be correct but formal proof is hard to achieve chiefly because each mouse bearing more than one copy of a transgene has a different arrangement of transgenes. Whilst few would dispute that integration occurs at random sites in the genome thus leaving each integrant susceptible to any of a wide variety of chromosomal environmental effects, the possibility that the expression observed is, at least, partly derived from interactions between copies within an array cannot be discounted. In the case of single copy mice the position effect may be invoked with more confidence to explain variable expression, in one study where single copy transgenic mice were obtained the transgene expression of serum growth hormone was 8, 60 and >4000 ng/ml (Morello et al., 1986). Definitive proof of the position effect might be achieved by reclaiming a transgene showing abnormal expression along with substantial flanking sequence and reintroducing the gene into mice. If the secondary transgenic mice from such an experiment all showed the same expression characteristics as the initial transgenic line then positive demonstration of the position effect might be achieved. This has not been accomplished. Alternatively, if the expression characteristic is not retained in the secondary transgenic lines this can also
serve as a proof of the position effect in that at another chromosomal location the expression of the integrant is different or returns to that observed in the other primary transgenic lines. The problem with interpreting this type of experiment are identical to those factors which might have caused the initial expression characteristic, namely the arrangement of the transgenes and the site of integration. In this respect, the most convincing demonstrations of the position effect can be achieved by re-introducing an aberrantly expressed transgene without flanking sequence and demonstrating retention of the original array but reversion to the normal pattern of expression.

P-Element Mediated Gene Transfer in *Drosophila*

Consistent integration of single copies of genes ought to allow that gene unfettered discretion to report position effect phenomena. This might be achieved during P-element mediated gene transfer into *Drosophila* (Spradling & Rubin, 1983). In such an experiment the expression of the *Drosophila* xanthine dehydrogenase gene has been proven to show only moderate, putative position effects in a minority of expressing flies (Spradling & Rubin, 1983). The line of flies which showed a large decrease in expression was found to carry an insert in the region of the chromocenter of the polytene chromosome previously associated with reduced gene activity. In this study there was positive selection for expression so the frequency of silent integrations is not known. The possibility that point mutations or rearrangement might have caused the observed effects was not discounted. Even so, if the effect were due to position effect the variation was slight compared to that observed even in single copy transgenic mice (for example of variable expression between single copy mice see Morello et al., 1986). Subsequently a more convincing demonstration of position effect was observed after transduction of the *Drosophila* white gene. This gene was expressed normally at any of at least 20 different chromosomal locations, but in three cases a mutant eye colour was
observed. The white gene from two of these mutants could be moved to other locations and this restored the wild type phenotype. This result strongly suggests that the mutant eye colour observed at low frequency was due to a chromosomal position effect and not mutation of the gene (Levis et al., 1985).

It is not clear why P element mediated gene transfer suggests, at best, the operation of moderate position effects on the level of expression. In some cases, this may reflect subtle differences between the long range regulation of Drosophila genes compared to those of higher eukaryotes. Indeed it is apparent that correct regulation of Drosophila genes can be achieved with only a few kb or less of flanking sequence (Maniatis et al., 1987). The same is clearly not always true of mammalian genes as analysis of the hB-globin gene locus in transgenic mice proves (Grosveld et al., 1987). Alternatively, the moderate position effect noticed may simply reflect the apparently non-random sites of P-element insertion (O'Hare & Rubin, 1983; Spradling & Rubin, 1983).

Insertion of Retroviral Provirus

Insertion of retroviral provirus into host cell chromosomes is part of the virus life-cycle and requires specialised viral sequences and functions. The viral RNA must first be transcribed by the virally encoded reverse transcriptase. The resulting DNA copy is then integrated into the host chromosome, a process which is catalysed by endonuclease activities encoded by the 3' end of the viral pol gene (Grindley et al., 1987). Integration is accompanied by short (4-6 bp) duplications of host DNA at the ends of the provirus (Grindley et al., 1987). However, no other rearrangements of the host DNA have been reported. Since the integration is an active property of the proviral DNA the sites of insertion in the genome are unlikely to be totally random. Preferred but multiple sites of retroviral integration have been described recently (Shih et al., 1988). This is compatible with the observation that
transcriptionally active genome regions are preferred targets for retroviral integration (Scherdin et al., 1990). It seems plausible then that insertion site preference may not be directly attributable to the sequence but to particular topological states of the chromatin. Perhaps such regions are more accessible than others. Therefore the preferred targets might include DNA replication origins or nuclear matrix attachment sites; for the latter there is some evidence (Scherdin et al., 1990).

Since retroviruses integrate as single copies and do not appear to rearrange the host locus during integration, as can occur during the integration of naked DNA (Wilkie & Palmiter, 1987; Covarrubias et al., 1987; Gordon & Ruddle., 1985; Mahon et al., 1988; Robins et al., 1981), there exists a potential to report specific chromosomal position effects without interference from other copies of the transferred DNA. The most notable problem in interpreting putative position effects arising from retroviral insertion is the possibility of a mutationally induced expression characteristic. This may or may not be significant but it is clear that retroviruses show a high mutation rate owing to the obligatory involvement of reverse transcriptase in their life-cycle (Roberts et al., 1989; Shields et al., 1978) and therefore must be regarded as a potential source of phenotypic variation.

Line specific variation is observed in retrovirally mediated transgenic mice (Harbers et al., 1981; Jaenisch et al., 1981; Jähner & Jaenisch, 1985) and in some cases this correlates with line-specific demethylation (Jähner & Jaenisch, 1985). The existence of a chromosomal position effect on integrated provirus is demonstrated by integration specific activation/expression of a recombinant retrovirus bearing the neo' gene in EC cells (embryonal carcinoma cells) (Sorge et al., 1984); EC cells are generally refractory to retroviral gene expression. This is analogous to the enhancer trap experiments with naked transfected DNA (Hamada, 1986). The possibility that the retrovirus is expressed in a sub-set of "leaky" EC cells which support viral transcription is excluded by the
observation that super-infection with a wild-type virus produced no viral transcripts. Therefore, the argument that the activation of the recombinant provirus is due to a cis effect is a strong one. One of the active proviral copies was cloned along with flanking sequence and reintroduced into EC and NIH3T3 mouse fibroblasts by transfection. No G418 resistant EC clones were obtained though the ability to produce G418 resistant NIH3T3 cells was not impaired. Though not unequivocally proving the point, this observation makes it less likely that the provirus had picked up an activating mutation.

The observations in the latter study have been confirmed and extended (Barklis et al., 1986; Peckham et al., 1989). The restricted expression of retroviruses in EC cells was shown to be overridden in two ways. Firstly, and in this instance more commonly, by integration within the 5' intron of an active transcriptional unit and secondly, in one confirmed case, by point mutation at the tRNA primer binding site (Barklis et al., 1986). It is not clear how frequently mutation accounts for the activation. The assay used to test this involved the generation of infectious virus from molecular clones of the provirus and 5' chromosomal flanking region and using these to infect EC cells. One out of the four reclaimed viruses assayed in this way was active in EC cells indicating that it had acquired an activating mutation. Five proviruses in seventeen selected cell-lines had integrated into two distinct chromosomal regions, implying that the assay exerts a strong selection with regard to the site of integration and that those sites capable of promoting expression of provirus are rare in EC cells. The number of such sites has been estimated to be around 100 (Barklis et al., 1986; Peckham et al., 1989). In all three of the proviruses analysed in detail, transcription of the viral genome appeared to initiate in the 5' chromosomal flanking region, splicing of the RNA involving a viral acceptor and a cellular donor splice site. In two of these cases the integration was in the first intron of a cellular gene and close to a very strong cellular promoter. One of these cellular
genes is likely to be the murine homolog of the yeast ribosomal L3 protein (Peckham., 1989).

In one instance a retroviral genome and flanking sequences have been recovered from the mouse genome and re-introduced by pronuclear micro-injection (Harbers et al., 1981). Only one secondary transgenic animal was obtained. The pattern of expression was significantly different from that of the primary lines and observed at high levels in an ectopic tissue. This may have been caused by a chromosomal position effect. However the events leading to the integration of the clone are not clear. The integrated viral DNA clone was found to have lost the chromosomal flanking sequences. If the insertion occurred via the expression of a viral RNA intermediate and subsequent reverse transcription then mutation may have provoked the expression pattern observed. Alternatively, if integration was achieved by illegitimate recombination between the micro-injected and chromosomal DNA then rearrangements have clearly occurred and may underpin the abnormal expression.

The Position Effect in Transgenic Animals and Transfected Cells.

In the case of transgenic lines and stably transfected clonal cell-lines variation in expression between different integrants can be considerable. The assembly and integration of arrays of foreign DNA molecules is essentially the same in both cases (Bishop & Smith, 1989). However it is clear that transfection of cell-lines and the creation of transgenic animals are not truly analogous since germ-line integrants are subject to the whole gamut of developmental and differentiation specific signals in any tissue bearing a copy. The potential for variation is therefore greater. In other words the position effect may have different consequences in different tissue or stages of development in transgenic mice. Thus expression in ectopic tissues may occur in addition to more general effects on the level of expression.
In two instances expressionally inert HSV-tk genes have been recovered from mice (Al-Shawl et al., 1990) and from cultured cells (Butnar & Lo, 1986) by molecular cloning techniques and successfully expressed when re-introduced into mice and mouse L-cell fibroblasts respectively. In cultured cells the activity of the HSV-tk gene was modulated by flanking mouse pericentromeric heterochromatin and repression was accompanied by DNA methylation and de-repression by rearrangements in the flanking DNA (Butnar & Lo, 1986). This is analogous to the classical position effect variegation observed in Drosophila when euchromatic loci are juxtaposed by translocation to heterochromatic regions (Baker, 1968). Complementing this observation, co-transfection of an HSV-tk gene with mouse satellite DNA into L-tk- cells resulted in tk+ transformants in which the HSV-tk genes were not only flanked by satellite DNA but had integrated preferentially into regions rich in repetitive sequences. Such HAT resistant clones showed an abnormally high phenotypic instability. Suppression and re-expression occurred at a frequency of greater than 3% and did not correlate with loss of transfected sequences (Talarico et al., 1988). The same effect was not observed in cells transfected with HSV-tk and main band DNA.

The transgenic mouse example involves a major urinary protein gene promoter HSV-tk supf fusion gene which is expressed in a line specific manner, correctly in the liver and ectopically in the testes and preputial gland of transgenic mice (Al-Shawl et al., 1988). However, 1 out of 5 lines shows a heritable lack of expression in the preputial gland (Al-Shawl et al., 1988). The integrant in this line comprising two tandem copies of the foreign DNA separated by a short sequence of mouse DNA was recovered with flanking mouse DNA and re-introduced into mice. Preputial expression was restored in all 6 secondary transgenic males (Al-Shawi et al., 1989). Although re-arrangement of the cloned transgene was detected in at least 2 of the lines the results suggest that the site of integration and not the tandem arrangement of the foreign DNA molecules had prevented expression in the preputial gland of the primary line. The presence of cloned mouse flanking DNA does not
have any noticeable effect since expression in the liver, testes and preputial gland are as highly variable between the secondary transgenic lines as they were in the initial experiment. This indicates that novel position effects are operative in the secondary transgenic mice, for instance two mice show no expression in the liver.

CONSISTENT, HIGH LEVEL EXPRESSION RELATED TO COPY NUMBER IN TRANSGENIC MICE

The sub-optimal, line specific and copy number independent expression observed in a large number of transgenic mouse experiments makes useful quantitation of expression and comparison of \textit{in vitro} mutated genes extremely difficult. In such experiments it is important to establish that differences in the level and pattern of expression occur because the sequences included on the transgenic construct have been altered and not because the genes have assumed a novel arrangement within an array or inserted at different chromosome locations. To some extent, these problems can be eased by comparing several lines bearing the same construct, consistent trends probably reflect the nature of the transgenes, inconsistent features are likely relate to the structure and location of the integrant.

Resident genes are correctly expressed because they lie within the cognate locus or perhaps the appropriate chromosome (eg in the case of X-chromosome inactivation). Cytological, genetic and molecular studies indicate that eukaryotic chromosomes are both structurally and functionally divided into domains of around 10-100 kb of DNA. In gene transfer experiments, a domain may be defined as a region of contiguous DNA which permits full, correctly regulated expression of the gene regardless of the site of insertion. The functional boundaries of domains have not been formally defined but
candidate regions include sites of attachment to the nuclear matrix or scaffold and DNAase I hypersensitive sites. Although the critical elements of an active locus may extend for long distances both 5' and 3' to the gene correct, if variable, expression of transgenes can be achieved by relatively short flanking sequences. For instance, only 205 bp of rat elastase 5' sequence is required for normal expression (Ornitz et al., 1985). The ability to mimic the normal locus or domain of a gene in a transgenic construct might be expected to insulate the expression against position effects and interference from neighbouring copies, thereby permitting detailed analyses such as the cooperativity between the regulatory elements associated with that gene. Similarly, homologous integration into the homologous locus should also result in appropriate expression and in one instance this has been observed in cultured cells (Smithies et al., 1985; Nandi et al., 1988). The frequency of homologous integration is low in micro-injected transgenic mice (1 in 500 mice, Brinster et al., 1989) and the only realistic route currently available is via embryonic stem cells selected for homologous integration (Capecchi, 1989).

Recently, several transgenic and cell-culture transfection experiments have achieved consistent, high level, tissue specific expression showing a varying but largely direct dependence upon copy number (Ryan et al., 1989a; Ryan et al., 1989b; Van Assendelft et al., 1989; Talbot et al., 1989; Grosveld et al., 1987; Aronow et al., 1989; Greaves et al., 1989; Collis et al., 1990; Steif et al., 1989).

The expression of the human β-globin gene in transgenic mice is well characterised and is typically subject to variable, low level or non-expression (Ryan et al., 1989b and refs therein). The entire human β-globin locus is well characterised at both the gene and the chromatin level. Two classes of DNAase I hypersensitive sites have been described for the locus. Erythroid specific sites located 5' to the individual globin genes are present only when the gene is active (Groudine et al., 1983). In addition to these
developmentally regulated sites are a set of developmentally stable, erythroid specific super-hypersensitive sites flanking the locus and located 6-18 kb upstream of the ε-globin gene and 20 kb downstream of the β-globin gene (Tuan et al., 1985; Forrester et al., 1986).

Using DNA bearing the developmentally stable DNAase 1 super-hypersensitive sites which naturally flank the entire human β-globin locus to construct a mini-locus transgene carrying just the β-globin gene and its 3' and 5' flanking regions, Grosveld et al (1987) achieved erythroid specific, position independent and copy number related expression of the human β-globin gene at levels equivalent to that of the endogenous mouse β-globin gene. Later experiments revealed that a 6-5 kb region of DNA encompassing only the upstream erythroid specific super-hypersensitive sites could act as a dominant control region (DCR) in the same way as the much larger mini-locus used earlier (Talbot et al., 1989). Most of this work was carried out in pools of transfected MEL cells (mouse erythroleukemia cells) in which the DCR behaves in an erythroid specific manner as in mice.

Subsequently the hβ-globin mini-locus has been used to direct the erythroid, differentiation specific expression of heterologous genes Thy-1 and HSV-tk-neo in transfected MEL cells (Van Assendelft et al., 1989). Following induction of *in vitro* differentiation the expression, which mimics that of the cellular β-globin gene, is at a high level and independent of position. The copy number dependence was not absolute, the expression/copy in some clones being very high and in some cases correlated with deletion at the 3' end of the construct.

The human β-globin DCR has been analysed extensively in transgenic mice (Ryan et al., 1989b). Six different DCR locus constructs were used bearing various deletions of the superhypersensitive sites (shs) and other sequences in and around the DCR. A total of 51 G. foetuses were analysed. The most stringent deletion which left only 1·9 kb of the 5' region of the DCR and one
she site was sufficient to activate high level erythroid specific expression amounting to 40% of the mouse β-globin message. The construct bearing all five upstream sites promoted expression, on average, at 109% of the endogenous transcript level. Incidentally, most of the constructs did not retain the 3' she site which appears to have little effect on expression.

In these experiments the expression was not tightly linked to the copy number even when all six she sites are included in the mini-locus. In this respect these results differ from those reported by Grosveld and colleagues (1987). Although Grosveld's initial report dealt with 9 transgenic mice most of their subsequent data is provided by work with cultured MEL cells. Therefore, only the initial report using the original, unmodified mini-locus bearing 21 kb of 5' and 12 kb of 3' β-globin locus flanking sequence shows copy number dependent expression in transgenic mice (Grosveld et al., 1987). Using a similar construct Ryan et al (1989b) did not observe strict copy number dependence. The observation of copy number dependence in pools of transfected MEL cells does not necessarily imply that the expression in individual clones within the pool follows the same, strict relation to copy number. The pool reflects only an average. This data is in accord with that provided by Ryan et al (1989b). In this case, two similar transgenic constructs containing the same number and arrangement of she sites differing only in the length of sequence 5' to the most distal she site showed indentical expression per copy when the data from each construct (13 and 9 G0 mice bearing each construct respectively) were pooled. The range of expression as a percentage of the mouse β-globin transcript was similar in both cases. It has been suggested (Collis et al., 1990) that the failure to observe copy number dependent expression in transgenic mice (Ryan et al., 1989b) results from analysis of mosaic mice selected because the foetuses were analysed at day 16 at which time fully transgenic foetuses are apparently anaemic and therefore non-viable (Talbot et al., 1989). There are, however, other explanations.
Copy number dependence, provided the quantification of RNA and DNA levels are precise, assumes that all copies are functionally intact and similarly competent in supporting transcription. This is an assumption with no factual support. Our knowledge of the assembly and integration of arrays of foreign DNA molecules implies that when multiple copies are present the criteria are unlikely to be met in all cases. Inactivation of some copies by rearrangement, deletion or mutation, which may not be noticed during copy number analysis is a plausible explanation for lack of copy number dependence. This is not to say that lack of equivalence between copies within an array can explain all cases when copy number dependence is not observed (Ryan et al., 1989b) but it is likely to be a factor in at least some of them.

When analyzing the effect of removing shs sites it is appropriate to consider the possible effects of neighbouring DCR's in the array. This is especially relevant considering the ability of the upstream shs's to function in either orientation and 5' and 3' to the gene (Talbot et al., 1989). This may assume a greater importance when the 3' shs site is missing. The minimal requirements of the β-globin DCR have recently been defined in MEL cells (Collie et al., 1990) and the results broadly corroborate those observed in transgenic mice (Ryan et al., 1989b). It seems clear that although two of the upstream shs sites are sufficient to direct strong, differentiation specific transcription in MEL cells, full biological expression of the human β-globin gene requires an intact DCR. There is also an absolute requirement for the second β-globin intron (Collie et al., 1990).

The DCR itself has many of the properties of a classical enhancer sequence. Namely it markedly improves the frequency of expression in transgenic mice (50/51 as opposed to 7/23 carrying the β-globin gene alone (Ryan et al., 1989b) and stimulates expression around 300 fold. The effect however, is clearly distinct from that of viral enhancers because the stimulatory effect of at least three of the shs sites within the DCR is undetectable in classical
enhancer-trap experiments (Tuan et al., 1989). Although classical viral and cellular enhancers can stimulate expression in a stable transformation assay (Luciw et al., 1983; Spandidos & Wilkie, 1983) they cannot do so in a copy number dependant, position independant fashion. The guarantee of expression is presumably achieved by assuming an open chromatin configuration regardless of the chromosomal location. This, in itself probably allows the other enhancing elements associated with the individual globin genes, which alone cannot establish high level and position independant transcription, to promote physiological levels of expression. It is interesting to note that, although the expression of the β-globin gene is erythroid specific, the expression of heterologous genes within the mini-locus is not firmly suppressed in other cell-types but that such expression does revert to a variable and low level (Van Assendelft et al., 1989; Talbot et al., 1989). This obviously implies that the globin genes themselves and local sequences are tissue specific.

In one respect MEL cells can be more informative than transgenic mice since they are required to differentiate in culture before high levels of globin expression can be observed. This pre-differentiation, non-induced state does not exist in transgenic mice. It has been reported that the mini-locus does not abolish the position-effect prior to the induction of MEL cells even though the shs sites have been established. Under such conditions both β-globin and tk-neo genes are expressed at variable and low levels. Therefore although the development of the active chromatin configuration is erythroid-specific, high level β-globin transcription is not the obligatory consequence. Since the heterologous tk-neo gene can contribute no additional erythroid-specific signals one conclusion to be drawn is that the DCR functions on two levels, primarily to establish an erythroid-specific chromosomal domain and secondly to respond to signals during the differentiation of erythroid cells (Van Assendelft et al., 1989) required for efficient expression. It has recently been observed that certain elements within the DCR can substitute for the upstream (-250 to -100 bp) erythroid-specific
region of the hβ-globin promoter, and can stimulate expression (in a non-erythroid specific manner) from a promoter containing only a TATA box (Antoniou & Grosveld, 1990). The observation that the DCR can function solely through a minimal promoter element which contains CCAAT and CAC elements (Chodosh et al., 1988; Xiao et al., 1987) and binds ubiquitous transcription factors accounts for the previous observation that non-erythroid specific genes such as HSV-tk and murine Thy-1 can be expressed in an erythroid specific manner when linked to a hβ-globin DCR (Van Assendelft et al & Talbot et al., 1989). The compatibility of the DCR and an heterologous, non-globin promoter appears to depend upon functional CCAAT and CAC elements since the murine histone H4 promoter which lacks both of these is unaffected by the DCR (Antoniou & Grosveld, 1990). Although not required in the presence of a complete DCR, the distal, erythroid-specific region of the promoter could substitute for one or other of snps 2 and 3 (Antanou & Grosveld, 1990), suggesting that synergism between proximal and distal elements is responsible for high level expression of β-globin in red cells.

The location of the DCR relative to the genes within the β-globin locus plays a role in determining the developmental specificity of expression. When the DCR is placed next to the human γ foetal globin gene, transgenic mice express the γ gene in definitive adult globin expressing erythroid cells as well as in the primitive embryonic globin expressing cells (Enver et al., 1989). Indeed, recent observations have established that correct temporal regulation of γ (foetal) to β (adult) globin gene switching is determined by sequences associated with the individual genes and that developmental regulation is invalidated when either of the genes is placed next to the DCR (Enver et al., 1990; Behringer et al., 1990). Developmental regulation of globin gene switching within the context of the DCR requires that both genes are linked to the same DCR. This suggests that switching derives from competition between the individual globin genes for the attentions of the DCR, the outcome of this competition being determined by the availability
of developmentally regulated, promoter specific trans-acting factors (Enver et al., 1990; Behringer et al., 1990).

The human CD2 glycoprotein is expressed in most lymphocytes and probably all peripheral T-cells. In transgenic mice expression of the human CD2 gene, with 4.5 kb of 5' and 9 kb of 3' flanking sequence, is confined to the T-cells, independent of position and at a high level which positively correlates with the transgenic copy number (Greaves et al., 1989). This dominant control activity is located in the 3' flanking region and is able to exert a similar effect on another T-cell specific gene, Thy-1 and also direct expression of the human β-globin gene to the T-cells. Interestingly the human β-globin gene was expressed in an integration specific manner in the liver as opposed to the position independent nature of its expression in T-cells. Since both heterologous genes are expressed in their cognate tissues, brain and foetal liver respectively, the CD2 DCR does not establish T-cell specific expression by exerting a dominant negative effect in a non-T-cell environment. Therefore in the cognate cell-type the CD2 and β-globin DCRs are functionally equivalent.

The existence of tissue specific dominant control regions which are capable of re-directing expression of heterologous genes to the tissue in which the DCRs are active suggests, for instance, that provided the β-globin gene is within a T-cell specific DCR or a thy-1 gene is within an erythroid specific DCR then transcription can be directed by regulatory factors present in those cells. The question arises, to what extent is there conflict between the tissue specificity of the DCR and that of elements closely associated with the heterologous gene. What is the role of local cis acting elements? In such a situation they may exert little specific influence. In this respect it would be of interest to study the expression of a heterologous gene whose promoter directs a characteristic other than tissue specificity for instance induction by an external agent. Persistence of the induction after linkage to
a heterologous DCR would indicate that gene specific mechanisms are not excluded by the DCR.

The DCRs described for \( \beta \)-globin and CD-2 are extremely potent. It is possible that other genes possess similar regions which do not function quite as convincingly. Where transgenes are expressed correctly at high frequency, DCRs of varying strength may be operative. The expression of an ADA-CAT fusion gene (ADA-adenosine deaminase) in transgenic mice may be a case in point (Aronow et al., 1989). High level expression of the fusion gene in the thymus of transgenic mice was only achieved when 12.8 kb of sequence from intron 1 was placed 3' to the CAT gene. This intron exhibited a tissue specific array of DNAase 1 hypersensitive sites and displayed enhancer-like properties in a transient cell transfection assay. Like the human ADA gene, the ADA-CAT fusion gene is expressed in several tissues. Only in the thymus did the expression correlate tolerably well with copy number.

The dominant control regions described for the CD-2 and \( \beta \)-globin genes are known to contain tissue specific DNAase 1 hypersensitive sites and when these are restored the genes adjacent to them can be expressed at high levels and independently of chromosomal location. Thus a relationship between sites in chromatin and appropriate expression is established. Although it is presumed that the sequences at the sites are recognised by tissue specific factors, the means by which active chromatin and expression is established remain obscure. Nor is it clear that only those regions at or over-lapping the sites are required.

The high level, position independent, copy number dependent activity of a chicken lysozyme promoter CAT fusion gene in chicken promacrophage cells is dependent upon the presence of a specific enhancer element and a nuclear DNA scaffold attachment site (A element) derived from the 5' boundary of the active chromatin domain of the chicken lysozyme gene (Steif et al., 1989). The same nuclear DNA scaffold attachment site (also known as the matrix attachment
region) has been shown to exert a similar effect upon an heterologous promoter (HSV-tk) in heterologous cells (rat-2 cells) (Phi-Van et al., 1990). Thus, a DNA element of known function can act in an analogous manner to the DCRs described above, even though the attachment site does not overlap a known hypersensitive site. The enhancer alone was able to stimulate expression but not in a copy number dependent fashion. The A element, when flanking the enhancerless gene also stimulated expression in the stable transformation assay, though not in a transient assay unlike the enhancer element. When both enhancer and A element were used true copy number dependence was observed. Indeed, the effect of increasing copy number was more than additive. Any mechanism proposed to explain the action of the A element and enhancer should offer some comment regarding this cooperative effect. One conclusion to be drawn from this observation is that individual A elements within an array of foreign DNA molecules can influence the transcriptional efficiency of neighbouring copies. This observation has important implications regarding the chief function of dominant control regions where more than one copy of the foreign DNA molecule is present. In such cases it is fair to say that the effects of the DCR may be two-fold. Firstly to establish an active chromatin domain (ie the potential for full transcriptional activity) regardless of the site of integration and secondly to allow each unit within the array to reach its transcriptional potential regardless of the number and arrangement of other units within the array. The data presented concerning DCRs cannot discount the second possibility as an explanation for the copy number dependent expression observed (Grosveld et al., 1987; Steif et al., 1989; Greaves et al., 1989). The issue will be hard to clarify since both the site of integration and the structure of the integrant vary between lines, and position independence cannot be observed in the absence of copy number dependence.
At one level of resolution genes can be distinguished by their level and pattern of expression. At this level, the distinction between active and inactive genes is huge. Inactive genes are suppressed by a factor of approximately $10^{9}$ to $10^{10}$ relative to the fully active state (Weintraub, 1985; Ivarie et al., 1983). Active genes can often be distinguished from inactive genes by their sequestration into chromosomal domains which are more sensitive to DNAase I relative to inactive genes and the bulk of the genome. Active chromosomal domains are further characterised by small regions of DNAase I hypersensitivity which probably represent local perturbations of the chromatin structure. The same genes in different tissues or at alternate stages of development may be distinguished in this way. Genes within the $\beta$-globin cluster provide a good example of this (Groudine et al., 1983). Genes can retain this identity through many rounds of cell division. Therefore the assembly and maintenance of active and inactive chromosomal domains is a central issue in our understanding of the mechanisms which underpin the regulation of gene expression with respect to tissue and stage of development.

The imposition of an expression programme upon a given gene is almost certainly regulated by interaction with DNA binding proteins of which some will be sequence specific. Interaction with proteins can stably alter the configuration of the chromatin and thereby determine accessibility to additional cellular factors which can promote transcription. Another mechanism by which genes may be stably altered is by covalent modification, notably the methylation of cytosine residues. Methylation of DNA can be maintained through rounds of DNA replication by the action of a hemi-methylase enzyme. Possibly all cytosine methylation occurs at the CpG dinucleotide of which 60-90% are methylated. This dinucleotide is represented in the DNA of vertebrates at around 20% of the expected frequency and non-methylated CpG's are non-randomly distributed in the genome (Bird,
1986). The non-random clustering of non-methylated CpG's lead to the
development, except in the case of the inactive X-chromosome where
CpG's are heavily methylated (Monk, 1986). Tissue specific genes,
which in general lack a true MFI character, are virtually fully
methylated in the male and possibly the female germ-line. This, like
the stably repressed state is maintained during development except
in the tissues where the gene is active and the DNA is under-
methylated. Thus a link, though not necessarily a causal one, is
established between under-methylation and gene activity. (reviewed
Cedar, 1988; Bird, 1986).

When DNA is in vitro methylated prior to transfection the
outcome is often suppression of gene activity (Cedar, 1988). In some
cases, this effect has been linked to a chromatin structure
different from that assumed by non-methylated, expressed,
transfected DNA. Therefore it seems likely that some cellular
factors can distinguish between methylated and non-methylated DNA
(Keshet et al., 1986; Buschausen et al., 1985; Buschausen et al.,
1987).

In vivo, de novo methylation has been associated with
transcriptional suppression. The methylation of CpG islands
following X chromosome inactivation (Lock et al., 1987; Monk, 1986)
and the coincidental methylation and silencing of retroviral genomes
are cases in point (Jähner & Jaenisch, 1985). Current evidence
points to the suppression of transcription preceding methylation
(Lock et al., 1987; Mechan et al., 1989; Grosschedl & Marx, 1988).
Even though methylation may, at least in some cases, be a
consequence of gene inactivity, its role in the stable maintenance
of the repressed state is demonstrated by the reactivation of genes
following treatment with inhibitors of DNA methylation, most notably by 5-Azacytidine. Genes on the inactive X-chromosome (Venolia et al., 1982;), silent retroviral genomes (Jaenisch et al., 1985), cellular house-keeping self-help genes (eg metallothioneins) and tissue specific genes can be reactivated by 5 Azacytidine treatment (Heguy et al., 1986; Compere & Palmiter, 1981; Konieczny & Emerson, 1984). Perhaps one of the most potent examples of methylation being involved in the suppression of a tissue specific gene in inappropriate cell-types is the in vitro myogenesis observed when 10T1/2 fibroblasts are briefly treated with 5 Aza-cytidine (Konieczny & Emerson, 1984). The frequency of myogenic induction is consistent with the activation of a very small number (possibly one) of loci. A myogenic determination locus (Myo D1) was isolated by subtractive cDNA hybridisation (Davis et al., 1987) and this clone can induce myogenesis when reintroduced into the same cells.

Analysis of the Myo D1 locus in aza-cytidine treated 10T1/2 cells revealed that selective demethylation accompanied the treatment but did not necessarily lead to active transcription of the gene (Michalowsky & Jones, 1989). However, the extent of the demethylation did correlate with Myo D1 expression and the extent of myogenesis. When nuclei from 5 aza-cytidine treated and control cells were treated with the methylation insensitive restriction enzyme MspI the sites in the Myo D1 locus were offered a greater degree of protection in the control cells. Not only does this illustrate a change in chromatin structure following treatment but also, and perhaps more importantly, implies that methylated sites are bound by nuclear proteins and that these proteins are no longer bound when the DNA is de-methylated. This is consistent with other data presented later (Mechan et al., 1989; Antequera et al., 1989).

Theories regarding the role of DNA methylation in the regulation of gene activity invoke the modification of trans-regulator binding to methylated DNA. In practice, this could be achieved in two ways. Firstly by the inability of trans-activator proteins to bind methylated DNA and secondly by the recognition of \( ^{m}\text{CpG} \) by a cellular factor which then interferes with the action of
transcription factors either by steric hindrance or by perturbing the local configuration of the gene. Recently, the trans-activator protein SP1 for which the consensus sequence features 1 CpG dinucleotide, showed no binding preference for methylated or unmethylated sites. More significantly SP1 mediated transcription was not compromised by methylation of the binding site (Hoelller et al., 1988; Harrington et al., 1988). In contrast, a number of recent papers report that certain trans-activator proteins are less effective at methylated binding sites (Dynan, 1989). In particular, there is evidence that methylation of a target site for a trans-acting factor upstream of the tyrosine amino-transferase gene regulates the cell-type specific factor binding that can be detected by genomic footprinting (Becker et al., 1987). This data perhaps implicates direct exclusion of trans-acting factors by specific \textsuperscript{m}CpGs. This is plausible since the inclusion of a methyl group can provide an additional hydrophobic contact. The effect may be subtle and conditional upon the type of protein interaction prevalent at that site. This might explain the variable effect of cytosine methylation on different trans-acting factors (for review see Dynan, 1989). Whether the data garnered from \textit{in vitro} methylation studies truly reflects the natural influence of cytosine methylation is not totally clear. This is largely because the sites methylated may have no \textit{in vivo} effect either because the methylation at that site is irrelevant or because the methylation is not extensive enough to mediate assembly into inactive chromatin or prevent active transcription (Murray & Grosveld, 1987). The latter case may be particularly true when only one or two sites are methylated.

Some of the current data are consistent with the hypothesis that sequence specific interaction at methylated DNA may be blocked by the action of a \textsuperscript{m}CpG specific binding protein such as that recently identified and shown to bind a series of methylated but otherwise unrelated sequences (Mechan et al., 1989). It seems likely that this protein (MeCP) can bind several \textsuperscript{m}CpGs and in this light a possible role has been proposed in the assembly of a higher order chromatin structure which is incompatible with transcription. Such a
structure might be the nucleosomal solenoid stabilised by interaction of MeCP with several CpGs on internucleosomal linkers (Mechan et al., 1989; Antequera et al., 1989). Recently, specific protection of CpGs in mammalian nuclei against digestion by CpG insensitive enzymes such as MspI has been demonstrated (Antequera et al., 1989; Michalowsky & Jones, 1989). Indirect evidence that this type of phenomenon might occur in vivo is available. A methylated HSV-tk gene is transcribed for 40 hours after microinjection prior to suppression (Buschausen et al., 1985 & 1987). The implication here being that direct exclusion of trans-acting factors by the methylated DNA does not account for the suppression, rather, the DNA is eventually recruited into a configuration which is susceptible to cytosine dependent repression (Antequera et al., 1989). In contrast, recent data demonstrate that non-methylated CpG islands assume an alternate chromatin structure characterised by histone H1 depletion, histone H3 and H4 acetylation and the presence of nucleosome free regions (Tazi & Bird, 1990). Therefore, evidence now exists that hyper-methylation and hypo-methylation correlate with distinct chromatin configurations thereby providing a functional insight into the relationship between methylation and gene activity.

In summary, there may be two mechanisms by which cytosine methylation might interfere with gene expression. Firstly by maintaining and perhaps aiding in the assembly of inactive chromatin domains and secondly by directly reducing the efficacy of transcription factors whose binding site contains a CpG dinucleotide. In both cases, removal of the methyl groups may release the repression and leave transcription to the discretion of the available trans-acting factors, in other words de-methylation and transcription may not be obligatorily linked.

Methylation of Retroviral Proviruses

Some retroviral proviruses can be inactivated in the mouse embryo and this repression can be lifted by treatment with 5 Aza-
cytidine (Jaenisch et al., 1985). Recently, the methylation at a single SmaI site in four separate proviral insertions in the genome of mice has been analysed (Mcgowan et al., 1989). At one of these loci the proviral copy was not equivalently methylated between cells. This was observed in a variety of somatic tissues whose last common ancestor was the primitive ectoderm.

**Methylation of Microinjected Transgenes: Implications for Expression**

Since methylation is, at a minimum associated with the maintenance of gene repression, selective methylation of individual copies of foreign DNA molecules may relate to the general absence of copy number dependent expression in transgenic mice. Complete silence of a transgene or tissue specific suppression might be marked by methylation. Since most transgenic arrays are composed of multiple copies of foreign DNA molecules and therefore multiple sites for methylation sensitive enzymes, correlation between expression and methylation is problematical. However, it is clear that de novo methylation, of uncertain stability, can be achieved (McGowan et al., 1989). Since methylation is detected in G0 mice, it can occur independently of passage through the germ-line (Palmiter et al., 1982). Palmiter et al (1982) showed a correlation between partial demethylation at some sites and an increase in expression in two progeny of the same mother. Siblings of these two mice showed complete demethylation at the same sites but no increase in expression relative to the mother. Clearly, in this case, neither the expression nor the methylation status of the transgene is stably inherited. The transgene in this case was pMK containing the 5' promoter region from the mouse metallothionein-I gene which harbours an MFI like region (Bird, 1986). This line of mice contained around 100 copies of the transgene per cell and therefore it is not clear what conclusions can be drawn regarding the relationship between demethylation and gene expression in this case.
In one instance the maintenance of a MFI has been demonstrated in transgenic mice. A mouse thy 1 5' : human thy 1 3' hybrid gene maintained its 1-0 kb MFI in 4/4 transgenic mice irrespective of expression at various developmental stages even though the flanking DNA within the construct was de novo methylated (Kolsto et al., 1986). This indicates that the exclusion of methyl cytosine in MFI like sequences can be determined by cis-acting sequences associated with the gene or perhaps by the MFI itself and that this property is not necessarily sensitive to chromosomal location. In this case (Kolsto et al., 1986), the MFI is retained in the absence of transcription suggesting that under methylation is necessary but not sufficient for expression.

Recently, a more convincing correlation between methylation of a transgene and its expression has been demonstrated (McGowen et al., 1989). This concerns the expression of an heat shock 68: lacZ fusion gene in the floor plate of the neural tube during days 10-14 of mouse development (Kothary et al., 1988). Within a transgenic line carrying 15-20 copies some transgenic foetuses showed no expression. Southern analysis of head DNA from non-expressing and expressing foetuses, revealed that hypomethylation was correlated with expression. Furthermore, intermediary methylation states were detected in some foetuses. When the corresponding neural tubes were analysed the number of expressing cells was reduced compared to that in the foetuses whose DNA was unmethylated. This was interpreted as a mosaic population of expressing and non-expressing cells. An inverse correlation between the extent of methylation of the array and the number of expressing cells was noticed. The intensity of staining in expressing cells was roughly equivalent between mosaic and non-mosaic mice. The methylation and expression was demonstrated to be under genetic control since matings of the same transgenic male to females from different inbred strains resulted in different methylation and expression patterns in G0 progeny. Since the maternal effect appears after fertilisation, the strain specific regulation of expression and methylation at this locus is controlled in trans. The integrant in this strain of mice actually lies within
or adjacent to the endogenous $dt$ locus (Kothary et al., 1988) and therefore might represent a special case rather than a more widespread phenomenon. However, cellular mosaicism of expression of hemizygous transgenes has been observed in other instances (Katsuki et al., 1988; Sweetser et al., 1988a & 1988b), as has methylation mosaicism (McGowan et al., 1989).

One question posed by these observations is how the methylation phenotype is represented within individual cells. Consider the case where the same locus can display three methylation phenotypes high, low and intermediate (McGowan et al., 1989). This might be achieved by selective methylation of the copies within the locus. Alternatively, one of two types of methylation may exist within each cell, high and low, the observed methylation phenotype being determined by the extent of the mosaicism (McGowan et al., 1989). The recent data demonstrating that a proviral genome was not equivalently methylated between cells and the cellular mosaicism of expression (Mcgowan et al., 1989) favour the latter explanation though it must be noted that formal proof of either will be necessarily hard to achieve.

Expression of Autosomal Genes Integrated into the X-Chromosome

During mammalian X-chromosome inactivation physically linked genes are repressed and this is followed by CpG methylation (Lock et al., 1987; Monk, 1986). The repression is considered to start at a specific site or sites and spread. If such sites are associated with X-linked genes then an autosomal transgene inserted into an X-chromosome might prove a useful probe for the mechanism of X-chromosome inactivation. At least two instances of this have been reported. They show that some X-linked autosomal transgenes can evade inactivation whilst others may be correctly inactivated (Goldman et al., 1987; Krumlauf et al., 1987). The latter case involving the inactivation of an alpha feto-protein transgene in foetal liver but continued expression in the extra-embryonic
visceral yolk sac, suggests that X-linked inactivation may be conditional. This example is perhaps more informative since it shows that autosomal genes can be subject to X-linked inactivation and that this probably does not require X-specific sequences to be intimately associated with X-linked genes. The first case whereby a chicken transferrin gene escaped X-linked inactivation differed from the second in two respects, the array was large (187 kb) and secondly there is no X-chromosome inactivation in chickens.

Parental Imprinting in Transgenic Mice

Parental imprinting refers to the phenomenon whereby the expression of a gene is determined by the parent from which it is inherited. The methylation of a transgene has been linked to parental legacy in several reports (Swain et al., 1987; Hadouchouel et al., 1987; Sapienza et al., 1987; Reik et al., 1987; Sasaki et al., 1989). In at least two instances, the link is extended to expression of the transgene. In these cases, the correlation is absolute (Hadouchouel et al., 1987; Swain et al., 1987). In only one of these studies (Sapienza et al., 1987) was the parental imprinting detected in more than one line (4/5), in this case expression of the transgene was not mentioned, therefore, the reported imprinting is at the level of DNA methylation only. One strain showed no effect, 3 demonstrated methylation via the female germ-line and 1 via the male germ-line. This high incidence suggests a pre-disposition on the part of the transgene for imprinted methylation. In the other cases the methylation occurred during passage through the female germ-line (Swain et al., 1987; Reik et al., 1987; Sasaki et al., 1989; Hadouchouel et al., 1987). In one case (Hadouchouel et al., 1987) the maternal inhibition of expression could not be reversed by passage through either the male or female germ-line. Expression in this and the other line bearing the same foreign DNA, could be linked to hypomethylation at specific sites and DNAase I sensitivity (Pourcel et al., 1990). In most cases, the imprinting was confined to one line, 1/7 (Reik et al., 1987), 1/2 (Hadouchouel et al.,
This suggests that the foreign sequences were not constitutively liable to stable methylation and so, by implication, that integration in the vicinity of an imprinted domain was responsible for the effect if the imprinting may reflect a position effect phenomenon. In one of the studies showing a strain with an expression imprinting, another line of mice resulting from microinjection of the same DNA construct showed some degree of parental dependent methylation but with no apparent phenotypic effect. It was suggested that the imprinting may be an unstable property of the transgene. The methylation thumb-print might not be the sole determinant of the extent of the phenotypic effect of the imprinting, additional factors may be required which could depend upon the arrangement of foreign genes and the site of integration.

In the past, the frequency of imprinting of endogenous loci has been estimated to be very low. This data is based upon experiments using Robertsonian translocation chromosomes, at the level of gross morphological or behavioural phenotypes; these experiments suggested that imprinted loci are present on six mouse autosomes (Sapienza et al., 1989a & 1989b for reviews). Although differential methylation can fulfill the criteria demanded of an imprinting mechanism, the frequency at which this has been observed at hemizygous transgenic loci is rather high. Around one third of transgenic loci are imprinted in some way, though it must be noted that this data is derived only from experiments where some form of imprinting was found, i.e. it is derived from a small number of experiments. For the vast majority of transgenic experiments, the data is not available. It is therefore appropriate to ask whether the methylation imprinting observed at hemizygous transgenic loci reflects true genome imprinting as defined by pronuclear transplantation or genetic criteria (Sapienza et al., 1989a & 1989b). In some cases the methylation and expression are apparently unrelated (Sasaki et al., 1989; Reik et al., 1987) and even when the imprinting is related to expression, the methylation was identical in all tissues tested (Swain et al., 1987). Allele specific methylation differences have been detected in cultured human cells.
and in human tissue (Silva & White, 1988; Chandler et al., 1987), though, in the latter case such "allelic blueprints" did not show differential parental imprinting. Detailed discussion of mammalian genome imprinting, and models to explain the imprinting of hemizygous transgenic alleles are available (Sapienza et al., 1989; Monk, 1990).

SUMMARY

Transgenes can be de novo methylated independently of passage through either germ-line (Palmiter et al., 1982), the pattern of methylation can vary between lines (Palmiter et al., 1982; Sasaki et al., 1989) and within lines (Palmiter et al., 1982; McGowen et al., 1989). Therefore methylation may be conditional and regarded as a position effect phenomenon in some cases. There is a serious lack of data regarding the variation in methylation between transgenic loci and its correlation with expression. In some cases the methylation may depend upon parental legacy (Swain et al., 1987; Hadouchouel et al., 1987; Reik et al., 1987; Sapienza et al., 1987; Sasaki et al., 1989) and this may correlate with expression (Swain et al., 1987, Hadouchouel et al., 1987). Within lines, variable expression and methylation may co-segregate (McGowen et al., 1989) and there is also evidence that this reflects cellular mosaicism in both expression and methylation (McGowen et al., 1989). Therefore the failure to inherit methylation patterns may allow variable expression between offspring of the same parent. It seems likely that methylation of the same locus is under some form of genetic control since at least two loci show different patterns of methylation when transgenic males were mated with females of different inbred lines (McGowen et al., 1989). Such trans-acting modifier loci have been termed imprinting alleles (Sapienza et al., 1989). Polymorphism and segregation of such alleles may offer some explanation regarding the variation of expression sometimes observed within and between transgenic lines (Bishop & Al-Shawl, 1988).
Metallothioneins (henceforth referred to as MTs) are small cysteine rich proteins which characteristically bind heavy metals. All vertebrates examined contain two or more distinct MT isoforms, categorised as MT-I and MT-II by virtue of their elution position from DEAE-cellulose columns. MTs are not exclusive to vertebrates, they have been found in invertebrates, plants and eukaryotic micro-organisms. Small, metal binding proteins have also been detected in prokaryotes. (for review of metallothioneins refer to Hamer, 1986).

The synthesis of MTs is homeostatically regulated by heavy metals and this is reflected by an increase in the rate of transcription and dependent upon cis-regulatory sequences located 5' to the MT genes. The expression of MTs also responds to several other physiological cues such as glucocorticoid hormones, interferon, interleukin 1, tumour promoters, inhibitors of protein and DNA synthesis, ionizing radiation and stress conditions (reviewed Hamer, 1986; Karin, 1985). Regulation by heavy metals is the best characterised induction and has been noticed in virtually all tissues, cell-lines and species so far studied (Hamer, 1986). Those heavy metals known to invoke a response include cadmium, zinc, copper, nickel, silver, mercury and bismuth; Cd, Zn and Cu are commonly used when investigating MT gene expression.

Although no precise physiological roles have been assigned to MTs, their binding to and induction by heavy metals implies that they are almost certainly involved in the de-toxification of heavy metals and the homeostatic regulation of zinc and copper levels. For a more detailed discussion regarding the physiological and toxicological aspects of MTs refer to (Hamer, 1986 and refs therein). MTs are of interest to molecular biologists chiefly because they provide a model system in which to study the mechanisms underlying the control of gene expression.
Metallothionein Gene Families in Mice, Humans and Sheep

Most studies involve mouse or human genes though the genes from other species have been cloned (Hamer, 1986). The mouse gene family consists of two closely linked functional genes MT-I and MT-II. The MT-II gene lies around 6 kb upstream of MT-I. The human gene family comprises at least 12 genes, the functional members, of which there are at least 7, are located on chromosome 16. The sheep metallothionein gene family comprises 4 functional genes and a pseudogene containing only one exon (Peterson et al., 1986 & 1988b). Consistent with observations made in other mammals there are two isoforms, I and II, sMT-I consisting of three variants Ia, Ib and Ic. In this respect the sheep family bears stronger relation to that found in humans than in mice. Structurally the genes are similar to those isolated from other mammals. The coding region homology of the sMT-Ia gene to the hMT-IIa and mMT-I gene is 86% and 89% respectively (Peterson et al., 1984).

The transcription of MT genes is subject to two distinct types of regulation, a low basal level and a high level induced by various external factors. Both modes of regulation are dependent upon the cis-regulatory sequences associated with each gene. The two mouse MT genes are coordinately regulated by heavy metals and glucocorticoids (Searle et al., 1984; Yagle & Palmiter, 1985). In species where several MT genes are active (e.g., humans and sheep) the genes may show characteristic expression (Heguy et al., 1986; Schmidt & Hamer., 1986; Peterson & Mercer, 1988a). The hMT-IIa and Ia genes are differentially regulated by heavy metals and glucocorticoids (Hamer, 1986; Richards et al., 1984). This pattern of regulation is dependent upon the promoter regions which have diverged so that they are no longer functionally equivalent as the mouse genes appear to be (Searle et al., 1984; Yagle & Palmiter, 1985). In particular hMTIa and sMT-II genes have a very high basal level of expression (Karin et al., 1987; Peterson et al., 1988a).
Tissue Distribution of Expression

Although MTs are considered housekeeping proteins by virtue of their expression in many different cell-types, tissue specific and developmental specific variations in the level and type of MT synthesis exist (Hamer, 1986). The data regarding mouse MT expression is the most comprehensive. Two studies report the number of mMT-I RNA molecules per cell in various mouse tissues in the presence and absence of induction (Durnham & Palmiter, 1981; Palmiter et al 1983), only in the testes was the level unaffected by cadmium. The hierarchy of expression changes after cadmium induction (Durnham & Palmiter, 1981) but the liver remains the primary site of expression, the kidney and intestine show high levels after Cd induction. The heart and testes are intermediately expressing tissues, the brain, spleen, lung and muscle express at lower but still Cd inducible levels. Zinc, copper and mercury all induce MT-I expression in the liver and kidney though quantitative and kinetic variation is apparent between tissues and inducer (Durnham & Palmiter, 1981). Where heavy metals have been administered to whole animals, it is not obvious whether variation in the response to different metals reflects the ability of the tissues to accumulate or respond to the metal ion. The two mouse genes I and II are coordinately regulated by Cd, lipopolysaccharides and dexamethasone in the liver, kidney, heart and brain (Searle et al., 1984). It is possible that detailed analysis of mMT-II expression in more tissues and stages of development may reveal a departure from the mMT-I pattern thereby providing an explanation for the presence of two genes owing to their non-overlapping function.

In sheep, the sMT-Ia gene is principally expressed in the liver and ovary, with lower levels in the kidney and brain (Shanahan et al., 1989). Any differences in tissue distribution of expression between species may derive from differential regulation of either gene expression or heavy metal metabolism.
Differential expression of the human MT gene family has been demonstrated (Hamer, 1986; Richards et al., 1984; Heguy et al., 1986; Schmidt & Hamer., 1986). The MT-IIa and Ia genes are expressed in all cell-types examined but MT-Ib is expressed in a sub-set of these cells (Heguy et al., 1986). The expression of hMT-Ie, If and IIa have been compared in various cell-lines (Schmidt & Hamer., 1986), the II-a gene was expressed ubiquitously but the Ie and If genes were expressed in a highly specific and often reciprocal fashion which related to the embryonic germ-layer origin of the cells.

Regulation of Metallothionein Gene Expression by Heavy Metals

Several early gene transfer experiments into cultured cells, mouse eggs and transgenic mice revealed that the cis acting DNA sequences which mediate the response to heavy metals lie 5' to the start of transcription, and that the 5' region could confer metal inducibility upon heterologous genes (Mayo et al., 1982; Karin et al., 1984b; Brinster et al., 1982; Brinster et al., 1981; Palmiter et al., 1982). Both the mMT-I and the hMT-IIa promoter regions behave like a bona fide inducible enhancer element that can activate transcription over long distances (Serfling et al., 1985; Maniatis et al., 1987; Westin & Schaffner., 1988). Later, discrete metal regulatory elements (hereafter referred to as MREs) within the promoter regions were located by classical reverse genetics and those derived from the mMT-I promoter have been shown to confer metal regulation in the absence of other MT promoter sequences (Stuart et al., 1984; Culotta & Hamer., 1989; Stuart et al., 1985; Searle et al., 1985; Carter et al., 1985).

The MREs from the mMT-I promoter are best characterised; it contains five (and possibly a sixth, Mueller et al., 1988) MREs which lie between 40 and 175 bp upstream of the transcription initiation site. The elements are 12-17 bp long and are highly conserved between flies and man (Culotta & Hamer., 1989). All
consist of a core heptanucleotide TGCpuCXG partially over-lapping a less conserved G-C rich region. The MREs appear to function cooperatively since duplication of an element permits a more than linear increase in the heavy metal induction ratio (Stuart et al., 1984 & 1985; Carter et al., 1985; Westin & Schaffner., 1988). MREd is able to confer the strongest induction to an heterologous promoter in vivo (Culotta et al., 1989; Stuart et al 1984). Under similar assay conditions MREE was non-functional (Stuart et al., 1985). MREs alone only mediates a significant induction when duplicated (Searle et al., 1985). MREd on the other hand, can function in isolation, but duplication increases the induction ratio 3-5 fold. It also responds to the same spectrum of heavy metals as does the complete promoter region (Culotta & Hamer, 1989). Fine mapping of MREd reveals that the nucleotides within the conserved core region are critical to the induction whereas those outside the core influence the response to a varying but lesser degree (Culotta & Hamer, 1989). Full response requires the wild-type MRE (Culotta & Hamer., 1989; Westin & Schaffner, 1988), in the same way that the full metal response requires an intact promoter region bearing all the MREs (Carter et al., 1985). The MREa confers a stronger heavy metal response to a promoter bearing the cognate TATA box than it does to a promoter bearing the corresponding region from the HSV-tk gene (Carter et al., 1985). Therefore it appears that the mMT-I promoter consists of several discrete cis-acting elements which provide full regulation of transcriptional activity in a cooperative manner.

The MREs are Sites of Interaction with Trans-acting Factors.

Although it is natural to propose that the MREs shown to be required for metal dependant regulation represent the binding sites of metal responsive trans-acting factors, formal demonstration of this has only been achieved quite recently. In vivo competition studies provided indirect evidence that metal regulation was positively controlled by limiting factors (Seguin et al., 1984).
Later, in vivo foot-printing on amplified mMT-I promoters (using the DHFR gene amplification system) revealed a metal dependent occupancy of MREs (Mueller et al., 1988) in addition to other constitutive footprints, of which two overlap the SP1 (Kadonaga et al., 1986) and adenovirus major late transcription factor (MLTF) (Carthew et al., 1987) consensus motifs. The MLTF had previously been shown to bind this sequence and stimulate transcription from the mMT-I promoter in an in vitro reconstituted system (Carthew et al., 1987). SP1 can bind to a GC element in the hMT-IIa promoter and stimulate RNA synthesis in vitro (Lee et al., 1987). No similar effect has been reported for the mouse promoter. Only MREd shows any protection in the absence of zinc. The level of protection of all the MREs is in rough correlation with their respective potency revealed by previous transfection assays (Stuart et al., 1985). A sixth metal dependent in vivo footprint was also detected and assigned MREF, sequence homology and deletion analysis had previously suggested the presence of a sixth MRE (Stuart et al., 1984 & 1985; Searle et al., 1985). The in vivo foot-printing experiments clearly do not reveal all DNA:protein interactions, for instance no TATA footprint is visible. Therefore it is probable that other proteins may bind the region but have not, as yet, been detected.

As the foot-print stands, there are sites which may represent regions of competition between factors, most notably at the MREd/SP1 overlap site (Mueller et al., 1988). The occurrence and significance of such competition in vivo remains to be demonstrated but the notion of a fine transcriptional control determined by the relative abundance of trans-acting factors is an attractive one. It is known that the MREs are necessary for the metal response but that other sequences determine the absolute level of expression in a cooperative fashion (Westin & Schaffner, 1988; Stuart et al., 1985; Carter et al., 1985; Sequin et al., 1984).

Multiple MREs interdigitated by putative basal and other enhancer elements is a common theme among various MT regulatory arrays. The putative factors binding to these sites are not specific to metallothioneins and include SP1 (Kadonaga et al., 1986), MLTF
(Carthew et al., 1987; Mueller et al., 1988), API (Angel et al., 1987), AP2 (Imagawa et al., 1987), AP4 (Mermod et al.), glucocorticoid receptor (Karin et al., 1984b), nuclear factor 1 (Imbert et al., 1989) and the α interferon-inducible factor (Friedman & Stark, 1985). Neither binding to or a role in MT gene regulation is established for all these factors, but it is tempting to speculate that conflict and cooperation between these and metal regulatory factors can determine the fine tuning of expression to suit the requirements of various cell-types.

In agreement with the data presented by Mueller et al (1987), the MREd has been shown to form a Zn dependent complex with a protein or proteins from a Hela cell extract (Westin & Schaffner, 1988). As predicted from its sequence the same element can bind the transcription factor SP1. The ability of mutant MREds to form complexes, as assayed by band shift assays, correlates with their in vivo potency as metal regulatory elements. Incidentally the effect was exclusive to Zn, since cadmium was ineffective. In puzzling contrast to this Seguin et al (1987) detected only a Cd dependent in vitro complex between the same element and proteins from a mouse L-cell extract. Whether this reflects a species variation whereby the same element can respond preferentially to different metals is not known. In this case the in vitro binding data do not agree with the in vivo transcription assays which demonstrate that MREd can respond independently to Zn, Cd and Cu (Culotta & Hamer, 1989).

Recently an MRE binding factor has been purified from mouse L-cells by using a trout MT-B gene MRE as an affinity reagent (Imbert et al., 1989). An MRE from such a divergent species as the trout was used because it is A-T rich and devoid of other recognisable and potentially misleading promoter motifs associated with mammalian genes (Zafarullah et al., 1988). This factor (MBF-1) gave a footprint over the trout MRE which could be inhibited by other functional murine MRE sequences. However, the only footprint produced by MBF-1 on the mMT-I promoter was over mMREe which, by some criteria, offers, at best, a weak response to heavy metals.
(Stuart et al., 1985). Nevertheless, MBF-1 was able to stimulate expression in response to Zn from the trout promoter in an in vitro assay using mouse cell extracts. The effect was not detected when Zn was replaced by Cd or Cu. It is not clear whether MBF-1 functions in vivo, indeed it lacks certain characteristics predicted for a metal responsive trans-acting protein, most significantly no metal dependent MRE binding has been detected. Possibly MBF-1 was artefactually saturated with metal ions during extraction and purification or perhaps cannot bind metals at all and lies downstream of the metal sensitive stage of the induction pathway. It is not known at what stage zinc is required in order to stimulate the in vitro transcription. Such data, if available might predict whether MBF-I is bound to DNA but inactive in the absence of zinc. The in vivo footprinting data do imply that protein binding to MREs is dependent upon the heavy metal challenge (Mueller et al., 1988).

The metallothionein gene in yeast (Sacchromyces Cerevisiae) is specifically induced by Cu (1) or Ag (1) ions. These ions induce transcription by altering the conformation of a trans-regulator protein encoded by the ACE1 locus (Furst et al., 1989). Cu and the ACE1 regulatory protein can reversibly activate yeast MT gene transcription in a mouse L-cell extract. Order of addition experiments indicate that Cu is not solely required for DNA binding, ie initiation of the induction, but is also required to sustain the response (Culotta et al., 1989). Although the ACE1 protein has the characteristics required of a mammalian metal responsive factor it does not influence the in vitro transcription from the mMT-I promoter (Culotta et al., 1989) probably because the ACE1 target sequence is not closely related to known mammalian MREs. Therefore it is not clear whether analogous proteins and mechanisms operate in mammalian cells.

Clearly there is some conflict between the in vitro binding and in vivo transfection data relating to the mMT-I promoter. In this case the in vivo footprinting data (Mueller et al., 1988) are
perhaps the most informative and logical since all the functional MREs are occupied in a metal dependant fashion. Also, the degree of protection correlates with the \textit{in vivo} potency of the respective MREs. Apart from the putative MRE binding proteins, at least two other known transcription factors bind the mMT-I promoter \textit{in vivo} and \textit{in vitro}, these are SPI and MLTF (Muellar et al., 1988; Carthew et al., 1987), the latter being able to enhance transcription in an \textit{in vitro} assay (Carthew et al., 1987). It is speculated that these and other factors which are ubiquitous rather than cell or gene specific may promote or stabilise the basal level of transcription thereby allowing maximal response to heavy metals without themselves playing a direct role in mediating the induction. The metal regulatory elements are known to act positively (Seguin et al., 1984) and be present in non-induced cells because the induction does not require protein synthesis. Further characterisation of MRE binding activities should allow a more accurate prediction of their mode of action.

**Promoter motifs are conserved between mice, humans and sheep**

The conserved function of metallothioneins predicts that the mechanisms which govern their expression will be similar between species. Analysis of the promoter regions of the mouse, human and sheep genes confirms this. The promoters from mammalian genes bear several metal regulatory elements for which a consensus can be derived. The sheep genes each have between 5 and 7 MRE-like sequences. The 6 MREs upstream of the sMT-Ia gene have between 8/15 and 13/15 nucleotides in common with the consensus and a perfect match with the heptanucleotide core sequence (Peterson et al., 1986).

Sequence with identity to the glucocorticoid response consensus such as found in the hMT genes are found in sMT-Ia, sMT-Ib and sMT-II (8, 9, and 9 out of 14 respectively) but not in sMT-Ic (Peterson et al., 1988b). The level of identity with the consensus
is low and subsequent expression analysis suggests that only sMT-II responds to glucocorticoids (Peterson et al., 1988a).

Both the mouse and human genes have consensus binding sites for the transcription factor SP1 (Kadonaga et al., 1986). The upstream regions of the sheep genes have from 2-4 potential SP1 binding sites, several of which overlap MREs. This is reminiscent of the murine MREd from the mMT-I promoter (Culotta & Hamer, 1989; Mueller et al., 1988; Westin & Schaffner, 1988). The SP1 sites in the mMT-I promoter have not been assigned a role in mediating the metal response, but may be required for full transcriptional activation (Culotta & Hamer, 1989; Mueller et al., 1988; Westin & Schaffner, 1988). Deletion of the SP1 site in the hMT-IIa promoter reduces both the constitutive and induced activity of the promoter (Karin et al., 1987; Lee et al., 1987).

Deletion analysis of the hMT-IIa gene (Karin & Holtgreave, 1984) and the mMT-I gene (Stuart et al., 1984; Carter et al., 1985) revealed an element which determined the basal level of expression, in the absence of induction. Only a poorly matched sequence is located in the Ia (8/13) 5' region whereas Ic and Ib have 10/13 and II 12/13. This is consistent with the observed levels of basal expression (Peterson et al., 1988a), only II is substantially expressed in the absence of any inducer (see later).

It is now clear that the basal expression element of the human MT-IIa gene includes sequences which mediate a response to serum factors and activators of protein kinase C. The human MT-IIa gene and several others including the SV40 early region promoter and the human collagenase gene (Angel et al., 1987), are inducible by the tumour promoter, TPA (phorbol ester). The induction is mediated by the transcription factor AP1 which binds to a 9 bp element (TRE) found conserved between TPA inducible promoters (Angel et al., 1987). A perfect match for the hMT-IIa TRE in exactly the same location is found upstream of the sMT-II gene (Peterson et al., 1988b). Indeed, the region between -112 and -60 base pairs upstream...
of the CAP site in the hMT-IIa gene is repeated with only 6 mismatches in the sMT-II promoter (Karin et al., 1987; Peterson et al., 1988b). This region includes sequences which are critical for the basal and induced expression of the hMT-IIa gene (Karin et al., 1987).

There are sequences sharing some identity with a proposed α-interferon inducible element (Friedman & Stark, 1985). However, in view of the relatively poor match, the significance to the expression of the sMT genes remains in doubt. Another conserved element termed "new" is common to all the sheep genes and comparison of mouse and human sequences revealed a consensus, though the function of the sequence remains obscure.

The strong identity of critical promoter elements between mouse, human and sheep MT genes suggests that species differences will not present a barrier to the expression metallothionein fusion genes. Species specific differences in the relative level of expression between tissues may arise because of variation in heavy metal metabolism between species.

UNUSUAL ASPECTS OF METALLOTHIONEIN GENE EXPRESSION

Transfected mMT-I genes are not responsive to glucocorticoids

Both the mMT-I and mMT-II genes are coordinately induced by glucocorticoids when in their natural chromosome residence (Yagle & Palmiter, 1985; Searle et al., 1984; Hager & Palmiter, 1981). The major effect is upon the level of transcription initiation (Hager & Palmiter, 1981), mRNA stability plays a lesser role. In gene transfer experiments the response of both genes is consistently lost (Hamer, 1986; Mayo & Palmiter, 1981; Durnham et al., 1984; Palmiter et al., 1982). Glucocorticoid regulation of expression is also abolished during amplification of the mMT-I gene in cadmium resistant mouse cells (Mayo & Palmiter, 1981). One possible reason
for the transfected mMT-I gene losing the glucocorticoid response is that the immediate 5' flanking sequence does not contain a recognizable glucocorticoid response element (Searle et al., 1984). If loss of response in transfected and amplified genes can both be accounted for by separation from essential cis-regulatory elements then it is necessary to suppose that such sequences are located at least 25 kb distal to the gene since this is the minimum size of the amplification unit (Mayo & Palmiter, 1981). Alternate explanations would predict a consistent change in chromatin structure or perhaps methylation which occurs as an obligatory accompaniment to gene transfer or amplification. Since the mMT-I fusion genes in transgenic mice also fail to respond to glucocorticoids, passage through the germ-line and exposure to full developmental cues cannot restore the response (Palmiter et al., 1982; Durnham et al., 1984). Whatever the explanation, it seems that at least some aspects of mMT gene expression may be dependent upon residence at the natural chromosomal locus.

The hMT-IIa gene does not behave in this way after gene transfer (Karin et al., 1984b). An element showing strong identity with the glucocorticoid responsive element consensus is found upstream of the gene (Karin et al., 1984c). Therefore loss of response to glucocorticoids following gene transfer is not a feature of MT genes in general.

**Hypermethylation is associated with Metallothionein gene inactivity**

Several reports have linked hypermethylation with the silencing of MT genes in certain cell types (Heguy et al., 1986; Compere & Palmiter, 1981; Gouhari et al., 1987; Lieberman et al., 1983; Schmidt & Hamer, 1986).

Some mouse T-cell-lines, W7 thymoma (Compere & Palmiter, 1981) and S49 lymphoma (Lieberman et al., 1983), fail to synthesise MTs and are, as a consequence, exquisitely sensitive to cadmium.
In neither case are the MT genes deleted or non-functionally rearranged. In both cell-lines expression of the mMT-I gene can be reactivated by treatment with 5-azacytidine (an inhibitor of DNA methylation). Expression in the S29 cell line (Lieberman et al., 1983) can also be activated by exposure to ultra-violet light. In both cases, the activation of MT gene expression can be correlated with extensive de-methylation over the entire gene. Assuming the mMT-II gene can produce a functional protein, this gene must also be inactive in the two cell-lines and may also be reactivated by the treatments described. The mechanism by which u/v light promotes expression of the mMT-I gene is unclear but appears to relate to the stimulation of de-methylation possibly by reducing the methyl-transferase activity or perhaps by some quirk of the DNA repair mechanism which is induced following exposure to u/v.

Similar observations have been made in chinese hamster ovary cells. Metallothionein gene reactivation has been achieved in MT- tk CHO cells by inhibitors of DNA methylation (5-azacytidine and its deoxy-counterpart) and correlated to specific de-methylation (Gouhari et al., 1987). Activation was also noted when in vitro mutagenised CHO cells were screened for temperature sensitive reversion to the tk phenotype. A single clone was derived which showed a temperature sensitive reversion to tk and MT phenotypes of $7.2 \times 10^{-4}$ and $6 \times 10^{-4}$ respectively. Reactivants showed specific demethylation of the MT and tk sequences. These cells were presumed to represent mutants in the regulation of DNA methylation (Gouhari et al., 1987).

The human MT-Ib gene, unlike the hMT-IIa and hMT-Ia genes shows cell type specific expression (Heguy et al., 1986). In Hela cells the gene is not normally expressed but treatment with 5-azacytidine results in a substantial increase in mRNA levels. In Hela cells the MT-Ib gene is methylated whereas in HepG2 cells, an expressing cell-type, the gene is hypomethylated. When non-methylated hMT-Ib fusion genes are introduced into either Hela or
HepG2 cells, expression is similar in both cases. This further indicates that silence of the endogenous hMT-Ib gene in HeLa cells is under epigenetic control. This observation differs from the methylated, suppressed mMT genes described earlier where presumably both the I and the closely linked II genes are inactive. In this respect, the natural cell type specificity of expression of human genes appears to correlate with differential methylation of closely linked genes. Expression data relating to the expression of hMT-Ie and hMT-IIf to that of hMT-IIa gene in several cell-types reveals that the cell-specific expression observed for these two other type I genes might also be under epigenetic control (Schmidt & Hamer, 1986).

The data correlating specific de-methylation with expression of MT genes suggests that MT-fusion transgenes might behave in a similar way if the methylation observed is sequence dependant and not reliant upon residence at the MT chromosomal locus. Clearly, methylation of MT genes is also cell type specific when the genes are resident at their natural locus. Therefore methylation of a transgene in a sequence dependant manner may only occur in one or a few cell types or tissues. Exhaustive analysis of the methylation and expression of MT-fusion genes in various tissues of transgenic animals would be instructive in this respect.

Methylation of transgenes as a consequence of their site of integration, perhaps by juxta-position to a methylated or imprinted domain (Palmiter et al., 1982; Sapienza et al., 1989a & b) might be expected to occur sporadically regardless of the sequence of the transgene and this may represent the same series of events which lead to the cell-specific repression of endogenous genes. Concurrent suppression of transgene and cellular homologues is unlikely, partly because of a presumed selective pressure against failure to synthesise metallothioneins. There is a serious lack of data relating transgene expression and methylation (see elsewhere in this thesis) however, the data discussed in this section indicate a more than casual relationship between methylation and expression of
metallothioneins. Therefore, if an mMT fusion gene is methylated following transfer then non-expression may be the outcome. This may be particularly relevant to tissues which are not generally considered primary sites of MT synthesis.

**EXPRESSION OF METALLOTHIONEIN FUSION GENES IN TRANSGENIC ANIMALS**

**Expression of pMK in transgenic mice**

One of the first genes to be introduced into the germ-line of mice and expressed in somatic tissues was pMK (mouse metallothionein-I promoter HSV-tk fusion gene) (Brinster et al., 1981; Palmer et al., 1982). The same plasmid had been demonstrated to function when introduced into cultured cells and mouse eggs and to respond to heavy metals but not glucocorticoids (Mayo et al., 1982; Brinster et al., 1982). Production and analysis of these mice was therefore significant with regard both to the new transgenic technology and the regulation of gene expression. In the initial report (Brinster et al., 1982) 4 transgenic mice were obtained when circular pMK molecules were used. One of these mice expressed HSV-tk activity in the liver and kidney when assayed following cadmium induction. All the transgenic mice harboured at least one functionally intact pMK molecule and the reasons for non-expression are obscure. At the mRNA level, the expression of the fusion gene was approximately 100 fold less than the resident mMT-I gene.

Subsequently, the study of pMK expression in transgenic mice was extended (Palmiter et al., 1982). In this report, transgenic mice were derived after injection of 4 different DNA conformation, supercoiled pMK, pMK linearised at the Bam H1 site, Bam H1 linear pMK ligated to a 6 fold molar excess of Bam H1 cleaved mouse genomic DNA and a blunt-ended Bst EII fragment of pMK. A majority of the transgenic mice (6/9) expressed HSV-tk activity in the liver. The three transgenic mice which failed to show expression in the liver...
were obtained by injection of supercoiled pMK. No correlation between expression and transgene copy number was observed. This is now an expected and recurring phenomenon in transgenic mice (Brinster & Palmiter, 1986; Bishop & Al-Shawi, 1989 and see earlier in this introduction). The level of expression varied up to 15-fold between lines. Cadmium but not dexamethasone regulation of HSV-tk activity was demonstrated in the livers of four transgenic mice. This confirms data obtained in cultured cells that mMT-I fusion genes respond to heavy metals but not glucocorticoids. Two of the lines generated in this study provided the first example of variable expression among offspring of the same parent. In the extreme case some offspring from an expressing parent failed to express at all. This extinction of expression was passed on to subsequent generations. An investigation of methylation of the transgenes in these lines revealed that foreign DNA can be subject to de novo methylation but, particularly in the line showing extinction of activity, there was no striking correspondence between DNA methylation and HSV-tk activity.

The results obtained with pMK in transgenic mice are significant to the work presented in this thesis because they indicate that the construct itself is functional in the majority of cases when introduced into the germ-line of mice. Aberrant, or non-expression is likely to be related to the nature of the individual integration events which occur during the generation of each transgenic mouse. Since the pMK sequences injected included plasmid sequences, it is important to know that they do not have an overriding, prejudicial effect on the subsequent expression of the transgene. Prokaryotic sequences have been identified as inhibitory when in combination with certain genes (see earlier in this introduction). Expression of the human growth hormone gene in transgenic mice is inhibited by plasmid sequences when linked to the mMT-I promoter but not when coupled to the rat elastase promoter (Ornitz et al., 1985a & b). However the inhibitory effect of prokaryotic sequences is not a consistent feature of mMT-I fusion genes in transgenic mice (Palmiter & Brinster, 1986). Non-expression
Expression of other mouse metallothionein-I promoter fusion genes in transgenic mice

The 5' and promoter region of the mouse metallothionein-I gene has been coupled to the coding region from several other genes and the resulting hybrid genes used to create transgenic mice; the outcome of such experiments is summarised in Table 1. Most of the hybrid genes are expressed in at least some lines. The frequency of expression can be high, 11/14 hGHRF (Hammer et al., 1985b) or low, HPRT, 4/17 (Stout et al., 1985). Some of the variation may be due to the fact that a number of the injected genes still retained prokaryotic vector sequences, but it is also clear that some hybrid genes are expressed more efficiently than others. Thus, the experiments with mMT-1 promoter fusion genes indicated that sequences other than the promoter can stimulate or interfere with expression. Indeed, it appears that the presence of introns can increase the transcriptional efficiency in transgenic mice (Brinster et al., 1988 and earlier in this introduction). A good example of different mMT-1 fusion genes being expressed at disparate levels is provided by comparing the expression of HSV-tk (pMK) and growth hormone fusion genes to the expression of the resident metallothionein gene in the livers of transgenic mice. Whereas the metallothionein and growth hormone mRNAs were present in comparable amounts (Palmiter et al., 1983), the expression of HSV-tk was around 1% of the endogenous gene (Brinster et al., 1981; Palmiter et al., 1982). Later by nuclear "run on " assays the rate of transcription of pMK was shown to be less than 5% of the endogenous gene (Palmiter et al., 1984).

For a small number of the hybrid genes, no expression was detected (see Table 1 eg mMT-1 β-Galactosidase). Since some of these genes are functional when introduced into cultured cells and even
TABLE I

MOUSE METALLOTHIONEIN-I PROMOTER FUSION GENES IN TRANSGENEIC MICE.

<table>
<thead>
<tr>
<th>TRANSGENE</th>
<th>EXPRESSION</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>mMT-I HSV-tk (pMK)</td>
<td>6/9</td>
<td>low mRNA, HSV-tk activity in liver, Cadmium induction.</td>
</tr>
<tr>
<td>mMT-I rGH</td>
<td>6/7</td>
<td>mRNA and GH expression in various tissues, ectopic expression brain</td>
</tr>
<tr>
<td>mMT-I hGH</td>
<td>23/33</td>
<td>expression in various tissues, response to Zn &amp; Cd, expression pituitary, gonadotrophs</td>
</tr>
<tr>
<td>mMT-I bGH (cDNA)</td>
<td>7/10</td>
<td>various tissues, induction Zn</td>
</tr>
<tr>
<td>mMT-I pGH (cDNA)</td>
<td>2/11</td>
<td>various tissues</td>
</tr>
<tr>
<td>mMT-I hGHRF</td>
<td>11/14</td>
<td>similar to endogenous gene but low in kidney</td>
</tr>
<tr>
<td>mMT-I rSomatm</td>
<td>8/11</td>
<td>various tissues, v. high in ant. pituitary</td>
</tr>
<tr>
<td>mMT-I HBsAg</td>
<td>2/6</td>
<td>various tissues, Zn response</td>
</tr>
<tr>
<td>mMT-I hHPRT</td>
<td>4/17</td>
<td>various tissues, high in brain</td>
</tr>
<tr>
<td>mMT-I B-Gal</td>
<td>0/11</td>
<td></td>
</tr>
<tr>
<td>mMT-I ho-globin</td>
<td>0/9</td>
<td></td>
</tr>
<tr>
<td>mMT-I hPLα</td>
<td>2/5</td>
<td>high expression</td>
</tr>
<tr>
<td>mMT-I hFactor IX</td>
<td>0/21</td>
<td>vector sequences included</td>
</tr>
<tr>
<td>mMT-I h factor IX</td>
<td>0/10</td>
<td>no vector sequences</td>
</tr>
<tr>
<td>mMT-I mPOMC</td>
<td>0/6</td>
<td>vector sequences included</td>
</tr>
<tr>
<td>mMT-I mPOMC</td>
<td>0/6</td>
<td>no vector sequences</td>
</tr>
<tr>
<td>mMT-I mMyc</td>
<td>ND/23</td>
<td>no tumours</td>
</tr>
<tr>
<td>mMT-I vSrc</td>
<td>ND/4</td>
<td>no tumours</td>
</tr>
<tr>
<td>mMT-I hIGF-I (cDNA)</td>
<td>5/21</td>
<td>various tissues</td>
</tr>
<tr>
<td>TRANSGENE</td>
<td>EXPRESSION</td>
<td>COMMENTS</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>mMT-I hIGF-II (cDNA)</td>
<td>2/17</td>
<td>various tissues</td>
</tr>
<tr>
<td>mMT-I SV40 T-Ag</td>
<td>ND/9</td>
<td>5</td>
</tr>
<tr>
<td>mMT-I Ad E1A (13S cDNA)</td>
<td>2/2</td>
<td>5</td>
</tr>
<tr>
<td>mMT-I INF-β</td>
<td>6/10</td>
<td>various tissues, Cd induction, Male sterility</td>
</tr>
<tr>
<td>mMT-I TTR</td>
<td>Various, Zn induction</td>
<td>11</td>
</tr>
<tr>
<td>mMT-I-LDL</td>
<td>8/20</td>
<td>various, Cd induction</td>
</tr>
</tbody>
</table>

**ABBREVIATIONS:**

m: - murine,  
R: - rat,  
h: - human  
p: - porcine,  
b: - bovine,  
GH: - growth hormone,  
GHRF: - growth hormone releasing factor,  
HBsAg: - hepatitis B surface antigen,  
HPRT: - hypoxanthine phosphoribosyl transferase,  
β-Gal: - galactosidase,  
hPRL: - placental lactogen,  
factor IX: - clotting factor,  
POMC: - proopiomelanocortin,  
IGF: - insulin-like growth factor,  
SV40 T-Ag: - simian virus 40 large T antigen,  
Ad: - adenovirus,  
INF: - interferon,  
TTR: - transthyretin,  
LDL: - low density lipoprotein,  
Somatin: - Somatostatin.

**REFERENCES:**

1) Brinster et al., 1981.  
2) Palmiter et al., 1982a.  
3) Palmiter et al., 1982b.  
4) Palmiter et al., 1983.  
6) Hamer et al., 1985b.  
7) Low et al., 1985.  
9) Stout et al., 1985.  
10) Iwakura et al., 1988.  
fertilized mouse eggs (mMT-1 β-Galactosidase, Stuart et al., 1984), the non-functional joining of diverse DNA segments per se cannot serve as an adequate explanation. More likely, the consistent non-expression observed with some constructs in transgenic mice is due to some event occurring during development that is specific to those constructs and probably not a position effect phenomenon. The so called position effect more readily explains occasional aberrant or failure of expression, a feature of many gene constructs in transgenic mice. In view of the many attempts to use the mMT-1 promoter to direct heterologous gene expression in transgenic mice, variable frequency of expression is not suprising and it is entirely possible that other promoter regions would show similar characteristics when linked to such a diverse selection of gene coding regions. It is also worthy of note that all the MT based transgenic experiments have, naturally, been with hybrid genes which are more prone to unusual frequency, levels and distribution of expression than unaltered genes (Brinster & Palmiter, 1986; Bishop & Al-Shawi., 1989 and see earlier in the introduction).

The resident mouse metallothionein gene is often considered to be a "housekeeping gene" since expression is detected in several cell-types and tissues. There exists an heirarchy of expression, liver > kidney > intestine > heart > brain/testes > spleen/lung (Palmiter et al., 1983; Durnham & Palmiter, 1981). The hybrid genes are typically expressed in the liver, though this is not always the major site. Hybrid genes containing rGH (Swanson et al., 1985), hGH (Low et al., 1986), rSomatostatin (Low et al., 1985) and hHPRT (Stout et al., 1985) genes are all maximally expressed in various regions of the brain. Somatostatin and human growth hormone expression is directed exclusively to the gonadotrophic cells of the anterior pituitary (Low et al., 1986) and rGrowth hormone to a restricted subset of neurones in the hypothalamus. One transgenic mouse expressed hHPRT only in the brain and kidney. In another mouse, the hierarchy of MT expression was effectively reversed with the brain showing the highest hHPRT expression and liver the lowest (Stout et al., 1985). In the latter case, HPRT expression in the

- 62 -
brain was very poorly induced by cadmium (1.4 fold), thereby excluding the possibility that the mMT-I promoter alone was the cause of the high expression in this tissue which is traditionally bottom of the mMT-I expression hierarchy. More likely, this high constitutive level expression is directed by sequences within or 3' to the HPRT cDNA or by the generation of a novel combination of regulatory sequences contributed by both the promoter and the gene sequences (see earlier).

As mentioned earlier in this introduction, metallothionein genes are characteristically inducible by various physiological cues most notably heavy metal ions and glucocorticoid hormones. Whilst glucocorticoid hormones never regulate mMT-1 promoters following gene transfer, heavy metals remain potent inducers of mMT-1 fusion gene transcription. The magnitude of the induction, like the basal and absolute (induced) levels of expression, vary between lines.

Expression of mMT-I fusion genes in transgenic livestock

Transgenic pigs and sheep bearing the mMT-I promoter fused to growth regulating genes have been produced (for review see Pursel et al., 1989). The frequencies of transgene expression compared to similar constructs in transgenic mice are illustrated in Table 2. In general the frequencies of expression are similar to the same fusion genes in transgenic mice (Purcel et al., 1989; Rexroad et al., 1989). Since the frequency of production of transgenic pigs and sheep is up to ten-fold lower than that achieved in mice, the absolute numbers of transgenic animals is low. This obviously makes conclusions difficult to draw, however, the hGHRF fusion gene does appear to express rather less well in sheep and pigs than it is in mice (Table 2). The tissue distribution of transgene expression has been reported for the bGH fusion genes in transgenic pigs and sheep (Purcel et al., 1989; Rexroad et al., 1989). In the two transgenic pigs analysed the tissues could be divided into two groups regarding level of expression with the liver, kidney and pancreas in the
<table>
<thead>
<tr>
<th></th>
<th>MOUSE</th>
<th>SHEEP</th>
<th>PIG</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mMT-I hGH</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. transgenic</td>
<td>33</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>No. expressing</td>
<td>22</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>% expressing</td>
<td>66%</td>
<td>100%</td>
<td>55%</td>
</tr>
<tr>
<td><strong>mMT-I hGHRF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. transgenic</td>
<td>14</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>No. expressing</td>
<td>11</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>% expressing</td>
<td>79%</td>
<td>11%</td>
<td>28%</td>
</tr>
<tr>
<td><strong>mMT-I bGH</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. transgenic</td>
<td>10</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>No. expressing</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>% expressing</td>
<td>70%</td>
<td>89%</td>
<td></td>
</tr>
<tr>
<td><strong>mMT-I hIGF-I</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. transgenic</td>
<td>21</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>No. expressing</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>% expressing</td>
<td>24%</td>
<td>25%</td>
<td></td>
</tr>
</tbody>
</table>

**References:**
1. Palmiter et al., 1983.

**Abbreviations:**
- hGH: human growth hormone.
- hGHRH: human growth hormone releasing factor.
- bGH: bovine growth hormone.
- hIGF-I: human insulin-like growth factor.
highest and the lung in the lowest group, expression in other tissues eg gonads, adrenal glands being in either group (Purcell et al., 1989). Expression of the same gene in two transgenic sheep showed very little variation between tissues except for exceptionally high expression in the kidney and cerebrum of one of the animals (Rexroad et al., 1989). Broadly speaking expression of the transgenes in pigs and sheep occurred in tissues which are sites of synthesis for similar genes in transgenic mice (Hammer et al., 1985a; Palmiter et al., 1983). In conclusion, the use of mMT-I fusion genes in livestock is successful in that at least some expressing animals are derived. Overall, the frequency of expression is around 50%. When this is allied to the general difficulty in generating transgenic livestock, this means that for every 1000 injected ova transferred less than 1% will result in expressing transgenic animals.

Expression of a sheep metallothionein-Ia: sheep growth hormone fusion gene in transgenic mice and sheep

Several lines of transgenic mice have recently been established bearing a construct comprising an 861 bp sMT-Ia promoter element fused to the coding and 3' regions of the sheep growth hormone gene (Shanahan et al., 1989). The sheep promoter was used because it lacked a sequence related to the basal level expression motif found in other metallothionein genes (Peterson et al., 1986; 1988a & b), and had exhibited very low levels of expression in sheep skin fibroblasts in the absence of zinc (Peterson et al., 1986 & 1988a). The possible advantages gained by maintaining a low level of basal growth hormone are: avoidance of deleterious side effects previously observed using mMT-I GH fusion genes and also to allow greater control over the growth rate of the animal. Basal levels of expression were detected in the brain and testes of all the lines examined and occasionally in the lung and ovary. Following zinc induction, expression, (in order of decreasing magnitude) could be detected in the small intestine, kidney and liver. The degree of
induction by zinc ranged from no response to 30 fold in the small intestine whilst the brain and testes failed to respond significantly. These patterns of expression represent a change in tissue specificity compared to the expression of the sMT-Ia gene in sheep. The resident sMT-Ia gene is primarily transcribed in the liver and ovary and to a lesser degree in the kidney and brain (Shanahan et al., 1989). The reason for these differences may relate to the use of a fusion gene or perhaps to differences in heavy metal metabolism between ruminants and rodents.

The same fusion gene has also been used to produce three transgenic sheep which all had elevated blood levels of ovine growth hormone (Murray et al., 1989)

Expression of a human metallothionein-IIa:human growth hormone fusion gene in transgenic mice

Six out of nine transgenic mice showed detectable fusion gene transcripts which in all cases were inducible by zinc. The induction was variable (less than two fold to over ten fold (Hammer et al., 1985a)). The aim of this study was to demonstrate glucocorticoid regulation of the hMT IIa promoter in transgenic mice. One of the two mice studied responded five fold to the steroid induction.

Summary.

The mMT-I promoter region has been used successfully to direct expression of heterologous genes to those tissues which traditionally express large amounts of metallothionein. Expression was regulated by heavy metals but not by glucocorticoids. Occasionally aberrant patterns of expression are observed particulary in various regions of the brain. The ratio of fusion gene to endogenous gene expression varied between construct, lines and tissues. The outcome of a transgenic experiment in terms of
expression is not predictable. Usually some lines do not express at all though some behave as predicted. The level of expression can vary between lines and even within lines and also aberrant patterns of expression may be observed. All these phenomena have been observed with mMT-I fusion genes in transgenic mice, and as such provide a good indication of the advantages and possible problems inherent in transgenic experiments. Although the level of expression of pMK in transgenic mice is low the frequency of expression, 6/9, is above average for mMT-I fusion genes in transgenic mice several of which are not expressed in any mice generated (Palmiter & Brinster, 1986).
CHAPTER 1

ANALYSIS OF THE FOREIGN DNA INTEGRANT IN TRANSGENIC SHEEP 229

Transgenic sheep 229 was produced by microinjecting Bam HI linearised plasmid pMK (mMT-I:HSV-tk fusion gene) into one of the pronuclei of a one cell stage sheep embryo. The transgenic sheep was identified by Southern analysis of DNA extracted from a blood sample. Liver poly A+ RNA was obtained from a biopsy sample. No HSV-tk mRNA was detected by northern analysis (Simons et al., 1988). RNA samples from other tissues also failed to reveal HSV-tk mRNA. A fibroblast cell line was prepared from a skin biopsy (by M. McClenaghan IAPGR). The work presented in this thesis refers to these cells.

Skin Fibroblasts from Sheep 229 Harbour a Single, Intact Copy of pMK

By the technique of Southern, using a variety of restriction enzymes to digest genomic DNA and probing the blots with a selection of pMK derived fragments, a restriction map of the integrant present in 229 skin cells can be constructed (see Figure 1). The results indicate that the skin cells harbour a single, structurally intact pMK molecule which may have integrated via the ends of a single Bam HI linearised input molecule. Enzymes such as Hind III and Bam HI which fail to cut the linear input molecule reveal a single band larger than linear pMK itself (see Figure 2 panel A & B, linear pMK is 8.4 kb in length). This suggests that the pMK DNA is flanked by sheep DNA bearing sites for these two enzymes.

Several lines of evidence demonstrate that a single copy of pMK is present at this locus. Firstly, the intensity of bands in genomic DNA when compared to those of copy number markers is consistent with the presence of a single copy (see Figure 2 panel
FIGURE 1

RESTRICTION MAP OF pMK AND THE 229 TRANSGENIC INSERT

The restriction map of the transgenic insert in sheep 229 was based upon DNA blotting data. A co-linear map of the pMK input molecule (Bam HI linear) is provided above the map of the integrant. Restriction fragments and probes to which they hybridise are illustrated below the map.

SYMBOLS.

Restriction enzyme sites:  ● Eco RI
                          ○ Hind III
                          ◆ Bam HI
                          △ Pst I
                          □ Pvu II
                          ■ Bgl II
                          ◊ Sst I

pMK MAP:  ______ pBR322.
          ______ mouse metallothionein I promoter and 5' sequence.
          ______ HSV-tk gene coding and 3' sequences.

PROBES:  1. HSV-tk.
         2. pBR322-mMT-I.
         3. whole pMK plasmid.
         4. mMT-I-HSV-tk.
FIGURE 2

SOUTHERN BLOTTING ANALYSIS OF THE TRANSGENIC INSERT IN 229 SKIN FIBROBLASTS

DNA blotting, probe labelling and hybridisation procedures were carried out as outlined in materials and methods. 10μg of genomic DNA was loaded in each track.

PANEL A.

Lane 1: 229 DNA Hind III digest.
Lane 2: 229 DNA Bgl II digest.
Lane 3: 229 DNA Hind III/Bgl II digest.
Lane 4: XEMBL 301 Bam HI digest.
Probe: Bam HI/Bgl II fragment containing pBR322 and mMT-I sequences.

PANEL B.

Lane 1: 229 DNA Pvu II digest.
Lane 2: 229 DNA Bgl II digest.
Lane 3: 229 DNA Bam HI digest.
Lane 4: 229 DNA Eco RI digest.
Lane 5: 2 copy number markers.
Probe: Pvu II fragment of the HSV-tk gene from plasmid pTK1.

PANEL C.

Lane 1: 229 DNA Bgl II digest.
Lane 2: 229 DNA Bgl II/Bam HI digest.
Lane 3: Plasmid pMK Bgl II/Bam HI digest.
Lane 4: 12 kb ladder (Bethesda Research Laboratories).
Probe: Pvu II fragment of the HSV-tk gene.
B). Secondly, when genomic DNA is restricted with a single cut enzyme such as Bgl II and probed with sequences from pMK lying either side of the Bgl II site, a single putative junction fragment is observed in each case (see Figure 2 panel A, B & C). No bands consistent with the presence of additional Bgl II sites derived from \( \frac{1}{2} \) pMK are seen; a head to tail, head to head or tail to tail arrangement of pMK molecules should reveal bands of a predicted size. A head to tail array, which is the most common arrangement following stable gene transfer (Bishop & Smith, 1989), would produce three Bgl II fragments capable of hybridizing to pMK probes, two fragments spanning the join between foreign and native DNA (junction fragments) of unpredictable size and a fragment of unit pMK length. When other enzymes are used the fragments generated, other than the expected internal fragments, are of an unpredicted size, do not suggest the presence of additional pMK sequences and are therefore considered to be junction fragments.

The Promoter and Coding Sequences are not Obviously Rearranged or Deleted.

A trivial explanation for the non-expression of pMK in 229 skin cells would be the rearrangement or deletion of parts of the foreign DNA prior to or at integration. Figure 1 shows the presence of predicted internal fragments from pMK in the genome of 229. There is no evidence of rearrangement or deletion of the mMT-1 promoter or HSV-tk sequences. Restriction sites spanning the whole of this region have been located in 229 DNA. It is possible that plasmid pBR322 sequences 5' of the Pvu II site in pBR322 (Figure 1) have been lost since sites beyond this point have not been mapped. However, there is no reason to suspect that the loss of such sequences could compromise the subsequent expression of pMK.
The 3' Bam HI Site has been Regenerated.

When Bam HI is used in conjunction with either Bgl II (see Figure 2 panel C) or Pat I (data not shown) both genomic DNA and pMK share at least one tk positive band indicating that the DNA between the 3' Bgl II or Pat I site and the 3' Bam HI site is both tk positive and of identical length in each case. This is consistent with the re-generation of the 3' Bam HI site. There are several routes by which this could arise. Since we believe that integration of foreign DNA occurs by a process of opportunistic repair-ligation, it is possible that the Bam HI site was regenerated fortuitously during the integration of an intact input molecule. However there are other explanations.

Secondly, since we know that linear molecules can be circularised and circular molecules can be linearised when introduced into mammalian cells, a round of circularisation followed by random breakage generates a circularly permuted linear molecule (for review see Bishop & Smith, 1989). The integrant in 229 skin cells is consistent with the following scenario. Assume that a linear pMK molecule ligated via the Bam HI sticky ends, subsequently broke close to the Bam HI site within pBR322 sequence and this molecule then integrated (see Figure 3, route 1). The third possibility is that the Bam HI site could have been generated by homologous recombination either between an input Bam HI linear molecule and a circularly permuted linear molecule which overlapped extensively from the 3' end of the input molecule or alternatively by homologous recombination between two, extensively overlapping, circularly permuted linear molecules (see Figure 3, route 2, process c). The distance between the random break and the Bam HI site in the second explanation and region of non-overlap between the recombining molecules in the third mechanism would have to be small enough to avoid detection on a southern blot and not include any of the
FIGURE 3.

TWO POSSIBLE MECHANISMS BY WHICH THE 3' BAM HI SITE IS REGENERATED.

Figure 3 shows that the 3' Bam HI site may be regenerated extra-chromosomally rather than during recombination with chromosomal DNA.

Route 1

a: - the input molecule re-circularises via the Bam HI cohesive termini, the Bam HI site is thus regenerated.

b: - the re-circularised molecule breaks randomly to generate a circularly permuted linear molecule which retains the Bam HI site.

Route 2

c: - homologous recombination between the circularly permuted linear molecule generated by process "b" and either an input linear molecule (shown) or another circularly permuted linear molecule (not shown) generates a linear molecule which retains the Bam HI site and has a duplication of 5' sequence 3' of this Bam HI site.

↑ Bam HI cohesive terminus

↑ Bam HI site
input molecule

route 1

route 2

integrate

integrate
restriction enzymes used in the mapping procedures. With the data presented here it is not possible to distinguish between these explanations. This would require analysis of the 3' end of the insert.

**Methylation of the 229 insert**

As mentioned in the introduction, methylation of DNA at the cytosine of a CpG dinucleotide is often associated with silent genes. It is not clear whether the methylation is the silencing agent or merely a consequence of inactivity. What is clear however is that de-methylating agents have been shown to activate silent genes which are then found to have become hypomethylated. In particular, metallothionein genes have been shown to behave in this way (also see introduction). In previous transgenic mice experiments with pMK, *de novo* methylation of the construct has been demonstrated though little firm correlation with gene activity was reported (Palmiter et al., 1982a & 1984).

The Eco RI fragment covering the mMT-1 promoter and most of the HSV sequences in pMK contains at least 15 CCGG sequences. This tetra-nucleotide is recognised by the restriction enzymes HpaII and MspI. MspI cleavage is prevented by methylation of the outer C but not the inner C, HpaII on the other hand will only cleave when both C's are unmethylated. Therefore the two enzymes can be used to estimate the level of CpG methylation within a given sequence covered by a hybridisation probe, in this case the Eco RI fragment. Genomic DNA and plasmid pMK was digested with Eco RI and then with either HpaII or MspI. The results of southern analysis of these digests using the Eco RI fragment as a probe is shown in Figure 4. When the genomic DNA is digested with either HpaII or MspI in conjunction with Eco RI the pattern of bands is the same and indistinguishable from the pattern visualised when purified plasmid DNA is digested with the same enzymes. There is a faint band at around 1.8 kb visible on longer exposure in the HpaII genomic track which is not apparent in the MspI digest. However, given its low
FIGURE 4

METHYLATION OF THE 229 INTEGRANT

PANEL A.

Lane 1: 15µg 229 DNA Msp I/Eco RI digest.
Lane 2: 15µg 229 DNA Hpa II/Eco RI digest.
Lane 3: 15µg 229 DNA Eco RI digest.
Lane 4: 100pg pMK Hpa II digest.
Lane 5: 100pg pMK Hpa II/Eco RI digest.
Lane 6: 100pg pMK Eco RI digest.
Lane 7: 12 kb ladder (Bethesda Research Laboratories).

PANEL B.

Restriction map of the mMT-I promoter and HSV-tk gene regions from pMK. Only the Hpa II sites in the promoter and proximal portion of the HSV-tk gene are illustrated; there are many other sites within the HSV-tk gene.
intensity, its significance is questionable. It may be due to incomplete HpaII digestion rather than inhibition of cleavage by methylation. If this does reflect true cytosine methylation at a particular site then it is unlikely that this site is methylated in all copies of the insert. This data fails to reveal a correlation between the gene inactivity and methylation of the internal C in the tetranucleotide CCGG. Indeed, appreciable methylation at this site is not apparent. This tetranucleotide probably accounts for a tenth of the CpG's in the region (Palmiter et al., 1984), some of the others being present at the recognition site for other methylation sensitive enzymes HhaI, SmaI, SstII, AvaI, NciI and some not in any site at all. Since these sites were not analysed it is not possible to discount a correlation between cytosine methylation and the observed gene activity. The data do, however, show that the region is not heavily methylated.

Summary.

The foreign DNA present in the 229 skin fibroblasts has been analysed by restriction mapping and southern blotting. There appears to be one more or less intact non re-arranged copy of pMK inserted at a single locus in the genome. The data do not suggest any explanation for the inactivity of the gene. Analysis of some of the CpG dinucleotides in the promoter and coding regions did not reveal appreciable levels of methylation.
CHAPTER 2

EXPRESSION OF THE pMK TRANSGENE IN 229 SKIN CELLS

THERE IS NO DETECTABLE HSV-tk ACTIVITY IN 229 SKIN CELLS.

Three sheep skin fibroblast cell-lines 229 (transgenic), 257 and 255 (non-transgenic) were assayed for HSV-tk activity as described in materials and methods. Zinc inductions were performed on 229 cells (16 hours growth in medium supplemented with 100μM ZnSO₄). Cell extracts were prepared from induced and non-induced cells and the assay performed in the presence and absence of TTP (thymidine triphosphate) which is a specific inhibitor of cellular thymidine kinase (Jamieson et al., 1974). The results of the assay are presented in Table 3a. Enzyme activities are expressed as pmol thymidine converted to thymidine monophosphate per minute per mg of protein.

The activities in the presence of TTP represent HSV-tk activity and the values observed in the absence of TTP represent the activity of the endogenous cellular tk plus any HSV-tk activity. BHK tk- cells possess neither endogenous or HSV-tk activity. The results obtained with extracts from 229 cells, either induced or uninduced, are not significantly above those recorded for either the non-transgenic cells extracts or those from BHK tk- cells.

Subsequent to these assays, 229 cell extracts have been analysed several times during the course of this work. In no instance was there a suggestion of HSV-tk activity. The results of some of these assays are also presented in Table 3. Table 3b shows a comparison of induced and non-induced 229 cells and BHK tk- cells. The cellular tk activity is very high and is strongly inhibited in the presence of TTP. The residual activity though
### TABLE 3

**NO HSV-tk ACTIVITY IN 229 CELLS**

A  
**Activity**: pmol/min/mg protein.

<table>
<thead>
<tr>
<th>Cell Extract</th>
<th>30 minutes</th>
<th>60 minutes</th>
<th>120 minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>229 -TTP</td>
<td>23.9</td>
<td>25.7</td>
<td>23</td>
</tr>
<tr>
<td>229 +TTP</td>
<td>1.125</td>
<td>0.7</td>
<td>0.165</td>
</tr>
<tr>
<td>229 +Zn -TTP</td>
<td>19</td>
<td>17</td>
<td>14.6</td>
</tr>
<tr>
<td>229 +Zn +TTP</td>
<td>0.87</td>
<td>1.2</td>
<td>1.325</td>
</tr>
<tr>
<td>255 -TTP</td>
<td>28.4</td>
<td>26.25</td>
<td>20</td>
</tr>
<tr>
<td>255 +TTP</td>
<td>3.175</td>
<td>5.45</td>
<td>0.34</td>
</tr>
<tr>
<td>257 -TTP</td>
<td>13</td>
<td>10.71</td>
<td>8.55</td>
</tr>
<tr>
<td>257 +TTP</td>
<td>0.7</td>
<td>1.575</td>
<td>0.4</td>
</tr>
<tr>
<td>tk− -TTP</td>
<td>71</td>
<td>54.5</td>
<td>30</td>
</tr>
<tr>
<td>tk+ +TTP</td>
<td>39.4</td>
<td>29</td>
<td>20.6</td>
</tr>
<tr>
<td>tk− -TTP</td>
<td>2.4</td>
<td>0.24</td>
<td>0.46</td>
</tr>
</tbody>
</table>

The values represent the average of two independent assays.

B

| 229 -TTP           | 49         |
| 229 +TTP           | 0.8        |
| 229 +Zn -TTP       | 62         |
| 229 +Zn +TTP       | 0.512      |
| tk− +TTP           | 0.482      |

Each sample assayed in duplicate.

C

| 229 +TTP           | 0.237      |
| 229 +Zn +TTP       | 0.4        |
| 229: pMK pool +TTP | 11.1       |
| 229: pMK pool +Zn +TTP | 74.7   |

| 229: pMK pool +TTP | 0.18       |
| 229: pMK pool +Zn +TTP | 0.24    |
| 229: pMK pool +Zn +TTP | 13.3   |
| 229: pMK pool +Zn +TTP | 69.3    |
| 229: pMK pool +Zn +TTP | 0.15    |
| 229: pMK pool +Zn +TTP | 0.23    |
| 229: pMK pool +Zn +TTP | 12.5    |
| 229: pMK pool +Zn +TTP | 63      |

229: pMK pool represents a pool of clones obtained following transfection of 229 cells with plasmids pMK and pSV3neo (see chapter 5).
slightly above that observed for BHK tk- cells is not increased by zinc and may well represent incomplete inhibition of the endogenous activity. To indicate the level of HSV-tk activity attainable in sheep cells expressing pMK, 229 cells were compared with a pool of 229 transfectants expressing SV40 large tumour antigen (Tag) and transfected pMK. The results are shown in Table 3c. Uninduced, the transfected cells show an 80 fold increase in HSV-tk activity over the control 229 cells, in the presence of zinc the difference is at least 200 fold. By this benchmark, the HSV-tk activity in normal 229 cells is insignificant. It should be added that expression of Tag cannot account for the increase in HSV-tk activity since sheep cells expressing SV40 Tag alone display similar activities to those observed in normal 229 cells (data not shown).

THERE IS NO DETECTABLE HSV-tk mRNA IN 229 CELLS

HSV-tk assays are generally accepted to be more sensitive than analysis of mRNA levels as an assay for gene expression (Palmiter et al., 1982a and Raya Al-Shawi personal communication). However, the absence of HSV-tk activity in 229 cells might be due to inactivating mutations in the coding region of the gene. In such a case HSV-tk mRNA may be present but the translation product (if any) might show little or no HSV-tk activity.

To explore the possibility that HSV-tk mRNA may be present but is never translated into active protein, total cellular RNA was prepared from control 229 cells and from those induced by zinc (8 hours growth in medium supplemented with 100μM ZnSO₄). The RNA was subjected to northern analysis and hybridised to a random primed HSV-tk probe. The results are shown in Figure 5. Panel A shows two separate preparations of RNA from both control and zinc treated 229 cells compared to RNA prepared from zinc treated BHK tk- cells stably transfected with pMK. No HSV-tk mRNA is visible in any of the tracks bearing sheep RNA. Slight cross-hybridisation to ribosomal
FIGURE 5

NO DETECTABLE HSV-tk mRNA IN 229 CELLS.

Total cellular RNA was prepared and subjected to northern analysis using an HSV-tk specific DNA probe. Zinc inductions were performed for 8 hours by supplementing the culture medium with 100 μM ZnSO₄. Copper inductions were performed for 8 hours by supplementing growth medium with 200μM CuSO₄.

PANEL A

Lane 1: 15 μg RNA from BHK tk⁻ cells expressing pMK. Zn induction.
2: 5 μg RNA as lane 1.
4: 5 μg RNA as lane 3.
6: 5 μg RNA as lane 5.
8: 5 μg RNA as lane 7.
10: 5 μg RNA as above.

The X-ray film was exposed for 10 days.

PANEL B

Lane 1: 20 μg RNA from 229 cells. No induction.
2: 20 μg RNA from 229 cells. Zn induction.
4: 20 μg RNA from 229 cells expressing transfected pMK. No induction.
5: 20 μg RNA from 229 cells expressing transfected pMK. Zn induction
6: 20 μg RNA from 229 cells expressing transfected pMK. Cu induction.

The X-ray film was exposed for 10 days.
RNA is visible in the tracks with 25μg of RNA. Panel B shows a comparison between 229 cells and a pool of 229 transfectants expressing pMK and SV40 Tag. This data was obtained much later than that presented in panel A owing to difficulties encountered during transfection of sheep skin fibroblasts. The HSV-tk activities of the same cells are presented in Table 3C.

The northern analysis has been repeated on other occasions with similar results.

SUMMARY

The results presented in this chapter establish that skin cells from sheep 229 though possessing an intact copy of pMK do not appear to express the transgene as assayed by HSV-tk assays and northern analysis.
CHAPTER 3

EXPRESSION OF THE SHEEP METALLOTHIONEIN GENE FAMILY IN SHEEP SKIN CELL-LINES

Failure of a gene to be expressed may relate to specific characteristics of the cell-type in which it is resident. The range of trans-acting factors available in a certain cell may not be sufficient to establish transcription of the gene in question. Where this is the case non-expression is likely to be the unavoidable outcome in all experiments. Whether the non-expression of pMK in the skin cells represents such a phenomenon is unclear since at present only one transgenic sheep bearing pMK has been produced. Transcriptional regulation almost certainly requires the cooperative action of a number of trans-acting factors at several sites present in and around the gene, most typically 5' to the gene. The transcriptional regulation of MT genes provides a good example of this (see introduction). Some of these factors are likely to be non-gene specific, general trans-acting factors which function at several genes (SP1, AP1, CTF ref Jones et al., 1988). Loss or mal-functioning of these might be expected to have serious consequences for the cell. However, factors which are specific to a small number of genes and fail to operate might simply cause very few genes to be inactive and therefore may be tolerated by the cell.

The transcriptional regulation of MT genes is complicated because they are subject to two distinct forms of regulation, a low basal level and a high level induced by various external factors (Hamer, 1986 and introduction). The induced level of transcription is probably mediated by trans-acting factors which are either specific to MT (eg metal regulatory factors, Imbert et al., 1989) or to a number of genes (eg AP1 or the glucocorticoid receptor Angel et al., 1987; Karin et al., 1984c). It is entirely possible that the basal regulatory mechanisms are operative in 229 cells but that the expression of pMK is undetectable. The
sheep MT-1a gene has a very low basal level of expression in skin cells and in transgenic mice (Peterson et al., 1986 & 1988a; Shanahan et al., 1989). It is worth remembering that pMK expression in the transgenic mouse liver after cadmium induction was only 1% that of the endogenous MT-1 gene (Brinster et al., 1981; Palmiter et al., 1982a). Failure of 229 cells to respond to zinc, a typical and potent inducer of MT gene transcription, would, in this case, suggest that the gene was inactive, not merely unresponsive to zinc. It is clearly necessary to establish that 229 cells are able to maintain basal level transcription and achieve induction of the resident family of sheep MT genes (Peterson et al., 1988a & 1988b).

Use of RNA Probes Specific to the Members of the Sheep Metallothionein Gene Family

Gene specific sense transcripts were prepared in vitro using SP6 RNA polymerase and appropriately linearised vectors containing gene specific regions for the four, transcribed sheep MT genes (kind gift, Julian Mercer, Murdoch Institute). To generate anti-sense ribo-probes, the gene specific fragments from the SP6 vectors were gel purified and inserted in appropriate orientation into vectors bearing the T7 phage RNA polymerase promoter (pTZ18r and pTZ19r, Mead et al., 1986). Full length sense and anti-sense transcripts were visualised by denaturing polyacrylamide-urea gel electrophoresis (data not shown). To check that the anti-sense probes were specific for their proposed targets, approximately 1ng of each target (sense RNA) was loaded in quadruplicate onto a denaturing agarose gel and subjected to northern analysis (see materials and methods). Each set of four targets (Ia, Ib, Ic and II) was hybridised to approximately 7ng of one of the anti-sense probes labelled with [α-32P]dUTP to a specific activity of 7 x 10^7 cpm/μg and then autoradiographed. The result of a four hour exposure is shown in Figure 6. Clearly, the probes are specific. However some cross-hybridisation does occur most notably in the case of the Ia probe and the Ic and II
FIGURE 6

SPECIFICITY OF SHEEP METALLOTHIONEIN GENE SPECIFIC RIBO-PROBES.

1 ng of each sense transcript was loaded in quadruplicate onto a denaturing agarose gel and subjected to northern analysis. Each set of four targets was hybridised to approximately 7 ng of one of the four anti-sense transcripts labelled to $7 \times 10^7$ cpm/µg. The final hybridisation wash was 20 mM NaHPO$_4$:1% SDS at 68°C for 1 hour. The filters were autoradiographed for 4 hours at -70°C against intensifying screens using pre-flashed X-ray film.
T7 RNA probe

- lc
- lb
- la

- II
- lb
- la

- II
- lc
- lb
- la

- III
- lc
- lb
- la
transcripts, and the Ib probe and the Ic transcript. The effect of non-specific signals will be minimal provided that Ia is not a minor transcript and Ic and II major transcripts. The same applies for the Ib probe and the Ic transcript.

Expression of the Sheep Metallothionein Gene Family in 229 Skin Fibroblasts: Response to Zinc and Copper

Figure 7 shows that there are marked differences in the level of expression of the four transcribed genes of the sheep MT family in 229 skin cells. Only the II gene shows an appreciable basal level of expression. The basal level expression of Ia is just detectable on this overnight (approximately 12 hour) exposure. Basal expression of Ib is detectable on longer exposure (4 days) whilst Ic is still undetected (data not shown). This data is in rough accord with previous results obtained with sheep skin fibroblasts (Peterson et al., 1988a), where the basal levels of Ia, Ib and Ic expression were below the level of detection in RNA dot hybridisation assays.

The addition of ZnSO_4 to the culture medium 8 hours prior to the RNA preparation induced all four genes. The hierarchy of induced expression is Ia>II>Ib>Ic. It is not immediately obvious which of Ia and II is the most abundant partly because there are two distinct, zinc inducible Ia messages, the smaller one being the more abundant. The larger message could represent the use of alternative initiation and/or termination signals or merely an unprocessed precursor RNA. The sequence of the sMT-Ia gene reveals a single but variant polyadenylation signal (AGTAAA) with no additional site within 400bp of this site (Peterson et al., 1988b). It is clear that Ia undergoes a larger induction than II. The Ic transcript is, by some margin, the least abundant transcript. This disagrees with a previous finding in similar cells that Ib was barely, if at all, expressed and that the abundance of the Ic message in the presence of zinc was 22% of the highest expressor, Ia (Peterson et al, 1988).
FIGURE 7

EXPRESSION OF MEMBERS OF THE SHEEP METALLOTHIONEIN GENE FAMILY IN 229 SKIN FIBROBLASTS.

Total RNA was prepared from 229 skin cells following 8 hours growth in medium supplemented with the indicated concentrations of heavy metal ions. 20 µg of cellular RNA and 50 µg of sense target transcripts were loaded onto four denaturing agarose gels and subjected to northern analysis using one of the gene specific riboprobes (materials and methods). The filters were autoradiographed overnight at -70°C against intensifying screens using pre-flashed film.
Another difference from previously published data was the failure to show a peak in expression at 100\(\mu\)M zinc. In all four cases the peak would appear to lie somewhere beyond 125\(\mu\)M zinc. Induction of mouse MT genes in cultured cells also reaches a peak at or before 100\(\mu\)M (Yeagle & Palmiter, 1985). Whether this reflects specific features of this experiment for example the zinc stock solution or the culture medium used or reflects a characteristic of these cells is unclear.

With reference to the small degree of non-specific hybridisation between probes and targets, since Ia is a major transcript and Ic very minor no problem arises from the observed hybridisation of the Ia probe to the Ic target (Figure 6). Both II and Ia are major expressors so cross-hybridisation here can be ignored. Both Ib and Ic are minor expressors so hybridisation of Ib to Ic will have little impact on either. The only instance where non-specific signals may distort the outcome is the hybridisation of the Ic probe to Ia and II transcripts. Longer exposure of Figure 6 reveals that this can occur and considering the vast amounts of Ia and II specific message in the RNA population this may account for at least some of the Ic signal.

No evidence is presented to explain why Ic is the marginally expressed gene in 229 (Welsh Blackface sheep) cells as opposed to Ib in previous studies on skin fibroblasts derived from Australian sheep (Peterson et al., 1988a). It is possible that during the period of divergence of the breeds subtle differences in expression might have emerged. It was suggested (Peterson et al., 1988a) that the Ib gene was poorly expressed because of its position close to both Ic and Ia genes in the family gene array and was therefore subjected to some form of transcriptional interference. If this is the case then a rearrangement of the gene family in 229 might commit the Ic gene to such an effect instead.

Attempts to induce the sheep MT gene family with Cu\(^{2+}\) failed in all four cases. Any differences apparent in the copper treated samples are slight enough to be explicable by unequal loadings of
RNA. Previously (Peterson et al., 1986 & 1988a), the same copper treatment had induced the sheep la gene approximately 25 fold albeit to only a quarter of the zinc induced maximum. The reasons for the lack of response are not clear. Trivially there may have been an error in the administration, but this seems unlikely. If the lack of response to this particular treatment with copper is a real phenomenon then certain issues are worth considering. Firstly, these cells may have a markedly different dose or time response to copper ions compared to the sheep cells used in previous studies. Certainly different breeds of sheep were used, probably Merino and Welsh Blackface. It is possible that the time course of the induction was incorrect. When heavy metals are administered to mice the time course of induction can vary with respect to both tissue and inducer (Durnham & Palmiter, 1981). However, this may well be due to variation in the rate of accumulation and metabolism of the different metal ions in various tissues and therefore may not be relevant to cultured cells. Alternatively, the concentration of copper used (up to 250µM) may have been inadequate. For instance, had the dose response curve taken an exponential form then even at values approaching the optimum concentration, the expression might still be very low. This type of response is not untypical of MTs (Durnham & Palmiter, 1981; Yagle & Palmiter, 1985).

If the cells are indeed not poisoned by the copper administration and do not increase the synthesis of MT RNA as a consequence of the treatment then three possible explanations emerge. Firstly that copper elicits a post-translational effect. The stability of MTs bound to different heavy metals has been shown to vary (Durnham & Palmiter, 1981 and refs therein). Secondly the basal level of MT expression adequately meets the copper challenge. Thirdly, copper chelators other than MTe a) present in the medium or b) produced by the cells such as Cu-chelatin (Winge et al., 1975). With respect to the last point the inclusion of 10% calf serum in the culture medium effectively removes 80% of the available copper probably because of the copper binding capacity of serum albumin (Peterson et al., 1986). It is possible that the calf serum
used in this experiment was able to bind a somewhat greater proportion of the copper. Alternatively, the medium contained a 10% supplement of tryptose phosphate broth (DIFCO) which might also sequester copper.

To explore these possibilities RNA was prepared from a pool of transformed 229 cells transfected with plasmids pM\text{I}ne\text{o} and pMTLT (the neomycin phosphotransferase gene and SV40 tumour antigen coding sequences respectively) placed under the control of the mMT-I promoter) following a variety of copper induction regimes. Northern analysis was performed using the II and Ia strand-specific probes (panels a and b in Figure 8 respectively). As discussed below the expression of the sheep metallothionein gene family in these cells is typical and similar to that in the normal 229 skin cells. The tryptose phosphate broth (TPB) medium supplement was omitted in all cases. The omission of TPB has little effect upon the induction because 250\text{M} is still ineffective (lanes 4). When the medium is supplemented with 500\text{M} CuSO$_4$ (lanes 3) both II and Ia are induced though to a lesser extent than achieved by 125\text{M} zinc (lanes 5).

Interestingly, when the usual medium 199 is replaced by DMEM effective induction is achieved by 250\text{M} copper (lanes 1). Serum was removed and the concentration of copper reduced to 30\text{M} (lane 2 in panels a and b). In the absence of serum this induction regime has been demonstrated to achieve maximal copper induction in sheep skin cells (Peterson et al 1986 & 1988a). However, Figure 8 shows this regime to be ineffective, on the contrary, the removal of serum from the medium causes a reduction in the expression of the II gene. These results suggest that medium 199 contains a copper chelator which effectively removes most of the available copper. Therefore it appears that the previous failure to detect copper induction was an artefact of the culture medium and not a characteristic of the cells.
FIGURE 8

RESPONSE TO VARIOUS COPPER INDUCTION REGIMES.

Total RNA was prepared from transformed 229 skin cells (pMThneo:pMTLT pool see figure 9) cultured in medium supplemented with CuSO$_4$ and subjected to northern analysis. 10μg of RNA was applied to each track. The filters were probed with either the sMT-II (panel a) or the sMT-Ia (panel b) gene specific probes.

Lane 1: D'MEM medium 10% foetal calf serum, 250μM CuSO$_4$.
Lane 2: 199 medium, no serum, 30μM CuSO$_4$.
Lane 3: 199 medium, 10% foetal calf serum, 500μM CuSO$_4$.
Lane 4: 199 medium, 10% foetal calf serum, 250μM CuSO$_4$.
Lane 5: 199 medium, 10% foetal calf serum, 125μM ZnSO$_4$.
Lane 6: 199 medium, 10% foetal calf serum, no induction.
Lane 7: 50pg of sense sMT RNA target transcripts (Ia and II).
Expression of the Sheep Metallothionein Gene Family in Other Cell-lines.

Expression of the sheep MT gene family has been investigated in other cell lines some of which are transformed derivatives of 229 cells. The purpose of the study was two-fold, firstly to establish that the pattern of expression described for 229 cells was typical and secondly to see whether the pattern of expression was altered by the transformed phenotype produced by expression of SV40 Tag. The results are shown in Figure 9. In general, these data confirm those previously obtained from study of the 229 skin cells. Transformed derivatives of 229 cells show an elevated basal level of expression of Ia, Ib and particularly II but generally the levels of expression in the presence of zinc, are greater in 229 parental cells. In addition the secondary, larger Ia, transcript is more prominent in 229 cells. A comparison of 229 cells (from an adult sheep) with foetal skin cells reveals that levels of expression were generally greater in the adult cells. Induction of the II message was reduced in the foetal cells compared to that in the adult. Transformation of FSF with SV40 Tag results in an elevated expression relative to the non-transformed parent cells (FSF), the opposite is true of 229 cells transformed with SV40 Tag.

In no instance did copper invoke a significant response (see above for explanation). The hierarchy of expression is observed in all cases, and therefore unaffected by the expression of SV40 Tag and not different in cells derived from foetal tissue. The data presented here do suggest that of Ia and II, II is the major expressor certainly in FSF cells and transformed derivatives of 229. It is of interest to note that transformed adult cells resemble normal foetal cells, and transformed foetal cells resemble normal adult cells. Normal foetal and transformed adult cells both divide more rapidly than normal adult cells. Increased proliferation is likely to be accompanied by a raised demand for zinc which is a component of many enzymes. This increased demand for zinc could be met by increased metallothionein synthesis. This may offer some
FIGURE 9

EXPRESSION OF THE SHEEP METALLOTHIONEIN GENE FAMILY IN NORMAL AND TRANSFORMED SHEEP SKIN CELL-LINES.

Total cellular RNA was prepared from various cell-lines following 8 hours growth in normal medium or in medium supplemented with either 100μM ZnSO₄ or 200μM CuSO₄. 15 μg of each cellular RNA sample was loaded in quadruplicate onto de-naturing agarose gels and subjected to northern analysis using the sheep metallothionein gene specific ribo-probes as before (figure 7). Filters were autoradiographed overnight.

CELL-LINES.

229: primary skin fibroblasts from sheep 229.

229: pMK: pSV3neo: pooled clones of 229 cells expressing pMK (HSV-tk) and pSV3neo (neomycin phosphotransferase and SV40 early region).

229: pMTLT: pMlneo: pooled clones of 229 cells expressing pMTLT (SV40 large T antigen) and pMlneo (neomycin phosphotransferase).

FSF: Foetal skin fibroblasts.

FSF pSV3neo: pooled clones of FSF expressing pSV3neo.
explanation as to the elevated basal level expression, particularly of the II gene, in transformed adult and normal foetal cells. In support of this both the hMT-IIa gene and the sMT-II gene are regulated by serum and activators of protein kinase C both of which stimulate cell proliferation (Imbra & Karin, 1987 and Paul Smith, data not shown).

The observation that expression of SV40 Tag and therefore a transformed phenotype does not radically alter the pattern of MT gene expression in sheep skin cells is of significance to work presented in this thesis since the transformed cell lines are a more amenable system in which to study the expression and regulation of mouse MT-I fusion gene constructs.

SUMMARY.

The sheep metallothionein gene family appears to be expressed normally in skin cells derived from transgenic sheep 229. All four genes respond to zinc administration and the Ia and II genes at least, respond to copper. There are minor differences between the data presented in this thesis and that already published (Peterson et al., 1988a), perhaps most notably the observation here that Ic and not Ib is the least abundant transcript. In view of the normal expression of the endogenous sheep MT genes in 229 skin cells, defects in the mechanisms responsible for MT gene expression and regulation by heavy metals cannot account for the inactivity of the pMK transgene. The expression and regulation of the gene family is not significantly altered in 229 cells transformed by plasmids encoding the SV40 early region, this is of significance to work presented later in this thesis.
CHAPTER 4

THE MOUSE MT-I PROMOTER IS FUNCTIONAL IN SHEEP SKIN FIBROBLASTS

The apparent inactivity of the single copy of pMK in 229 skin fibroblasts may simply reflect the inability of the resident sheep transcriptional machinery to utilize the mouse MT-I promoter, even though 229 fibroblasts are fully competent in the expression and regulation of the resident sheep MT gene family. Given the conserved patterns of regulation of MT genes, absolute species incompatibility is unlikely. In support of this view the trout MT-B promoter is active in mouse and human cells (Imbert et al., 1989; Zafarrulah et al., 1988), the sheep MT-Ia promoter is active in transgenic mice (Shanahan et al., 1989) and the mouse MT-I promoter: bovine growth hormone and human growth hormone releasing factor fusion genes are expressed in transgenic sheep (Pursel et al., 1989). It is apparent that species does not present a barrier to the appropriate expression of transgenes provided all the relevant cis-acting DNA control elements are present. What can be considered a possibility however, is that the mouse MT-I promoter is not utilized in sheep skin cells particularly when inserted at random sites into the sheep genome. Skin is not mentioned as a prime site of MT synthesis though clearly the sheep genes are transcribed in cells derived from this tissue (Peterson et al., 1986 & 1988a and this thesis) and human MT genes are transcribed in human skin fibroblasts (Richards et al., 1984). If the mouse MT gene is not expressed or only marginally active in mouse skin, this may dictate inactivity in sheep skin. For this type of effect there is a precedent illustrated by several transgenic experiments. Where subtle differences in the expression of genes occur between species, it is common for the site of expression to be dictated by the foreign gene (see introduction and Aronow et al., 1989; Koopman et al., 1989; Kelsey et al., 1987; Gordon et al., 1987).
In order to prove that the mouse MT-I promoter is not inherently inactive in sheep skin cells it was necessary to introduce DNA constructs harbouring this promoter, in particular it was necessary to investigate the activity of the promoter when introduced at random sites into the sheep genome (ie stable transfection). Progress in this respect was hampered by difficulty in obtaining reliable transfection of primary sheep skin fibroblasts (see appendix). Although the standard calcium-phosphate:DNA co-precipitation technique (Graham & Van der Eb., 1973) was eventually effective in stably transfecting the cells, in the initial experiments the procedure was toxic to the cells and did not result in stable transfection. Therefore an early decision was made to establish electroporation as a reliable method of gene transfer. In the case of stable transfection electroporation proved reliable but was less efficient in transient assays. Therefore much of the data presented in this chapter and later in the thesis is confined to stable transformation and analysis of the expression of mMT-I fusion genes in the resulting clones or pools of cells.

In addition to the absence of an effective transient expression system, two additional problems presented themselves. Firstly since the cells retain an endogenous thymidine kinase activity, tk+ cells can be selected for expression of HSV-tk, but the selection is slow, requires titration of drug and supplement levels, the efficiency is very low, and the method is not widely used (Mercola et al., 1980). Therefore, in practical terms stable transformation to the HSV-tk+ phenotype is only indirectly selectable by co-transfection with a selectable marker such as neomycin phosphotransferase or hygromycin B phosphotransferase (Southern & Berg, 1982; Bochlinger & Diggleman, 1984). Secondly since the cells are not immortal they have a finite life-span in culture during which time there is a regressive change in growth characteristics eventually leading to crisis and senescence and as a consequence a relatively short period during which they can be transfected reliably. Furthermore, any clones arising from the transfections are likely to have a restricted potential for
### TABLE 4

**STABLE TRANSFECTION OF 229 SKIN CELLS**

<table>
<thead>
<tr>
<th>DNA</th>
<th>Colonies</th>
<th>Plates</th>
<th>Average</th>
<th>Efficiency&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSV2neo: pSV3gpt</td>
<td>393</td>
<td>6</td>
<td>66</td>
<td>0.029 % *</td>
</tr>
<tr>
<td>pSV2neo: pMTLT</td>
<td>152</td>
<td>6</td>
<td>25</td>
<td>0.0109 % *</td>
</tr>
<tr>
<td>pMTneo: pSV3gpt</td>
<td>182</td>
<td>6</td>
<td>30</td>
<td>0.013 % *</td>
</tr>
<tr>
<td>pMTLT: pMTneo</td>
<td>29</td>
<td>6</td>
<td>5</td>
<td>0.0022 % *</td>
</tr>
<tr>
<td>pSV2neo</td>
<td>190</td>
<td>5</td>
<td>38</td>
<td>0.0115 % §</td>
</tr>
<tr>
<td>pMTneo</td>
<td>54</td>
<td>6</td>
<td>9</td>
<td>0.0032 % §</td>
</tr>
</tbody>
</table>

* Colonies counted 14 days after electroporation.

§ Colonies counted 20 days after electroporation.

10μg of each plasmid (linearised) used, selection in medium supplemented with 400μg/ml G418 commenced 72 hours after electroporation.

<sup>a</sup> number of colonies expressed as a percentage of the number of cells plated after the electroporation.

The table below illustrates the transfection efficiencies obtained with each combination of the 4 DNAs.

<table>
<thead>
<tr>
<th>DNA</th>
<th>pSV3gpt</th>
<th>pMTLT</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSV2neo</td>
<td>0.0115</td>
<td>0.0290</td>
</tr>
<tr>
<td>pMTneo</td>
<td>0.0032</td>
<td>0.0130</td>
</tr>
</tbody>
</table>

- 96 -
propagation since they will have undergone a substantial number of population doublings during clonal expansion. Indeed, since the plating efficiency at low seeding density is very low (see Chapter 7) clonal expansion is probably only possible on a sub-population of the cells.

In order to circumvent these problems two mMT-I promoter fusion genes were constructed. One linking the promoter to the coding region for neomycin phosphotransferase (pMTneo) and the other to the SV40 large tumour antigen gene (pMTLT). The pMTneo construct provided a selectable marker under the control of the mMT-I promoter and pMTLT provides a means of increasing the in vitro life-span of the 229 skin cells by virtue of SV40 Tag expression which increases the number of population doublings prior to crisis and senescence (see later Chapter 7).

Co-transfection Using SV40 Tag to Promote Clonal Expansion

229 skin fibroblasts were electroporated (see materials and methods) in the presence of 10µg of each of the linearised plasmids listed in Table 4. Cells were selected for neo" by growth in medium supplemented with G418 (GIBCO), no zinc was added to the medium. Each pair of plasmids included one bearing a selectable marker, neo", and one the SV40 Tag gene to promote cellular proliferation. The promoters used to drive the two genes were either the SV40 early promoter (pSV2neo and pSV3gpt: Southern & Berg, 1982; Mulligan & Berg, 1980) or the mouse MT-I promoter (pMTneo and pMTLT). All combinations of the four plasmids were used. The results are shown in Table 4.

When both the neo" and Tag were driven by the SV40 early promoter the transfection efficiency was highest. When both were driven by the mMT-I promoter the frequency of stable transfection was lowest. Intermediate transfection efficiencies were obtained when an mMT-I and an SV40 driven construct were used together.
Plasmid pMTLT contains the SV40 early coding region (Tag and tag) fused to the mouse metallothionein I promoter.

Plasmid pMTneo contains the neomycin phosphotransferase gene and HSV-tk 3' sequences fused to the mouse metallothionein I promoter.

Plasmid pSV2neo contains the neomycin phosphotransferase gene fused to the SV40 early region promoter/enhancer. (Southern & Berg, 1982).

Plasmid pSV3gpt contains the bacterial xanthine-guanine phosphoribosyltransferase gene and the SV40 early coding region both under the control of the SV40 early region promoter/enhancer. (Mulligan & Berg, 1980).
At a minimum, these results indicate that the mMT-I promoter can function when inserted into the chromosomes of sheep skin fibroblasts and that silence of the transgenic pMK insert does not arise because of some deficiency on the part of the cells in providing the correct repertoire of trans-acting factors required to establish basal level expression. Moreover, the function of the mMT-I promoter is not dependent on the presence of an adjacent SV40 early region promoter (it is likely that co-transfection results from physical linkage of the two genes prior to integration) since pMTneo and pMTLT used in conjunction give rise to colonies.

To assess these results in terms of the relative promoter strengths of mMT-I and the SV40 early region promoter in sheep cells one must make some assumptions. Firstly, one must assume that the efficiency with which neo-R or Tag α expression is only or largely determined by 5' promoter sequences and not by sequences 3' to the gene which may influence the stability and processing of the mRNA. Plasmid pMTneo has 3' regions from the HSV-tk gene. All the others use 3' sequences from the SV40 early region.

Secondly, if, as we believe, co-transfection is usually achieved because the two input plasmids become physically linked prior to integration then one has to assume that the frequency at which this occurs is similar for each combination of plasmids. If this assumption is wrong then the results may reflect the frequency of co-transfection rather than the relative abilities of the promoter regions to promote transcription. Our knowledge of the mechanisms by which foreign DNA molecules interact and recombine prior to integration (Bishop & Smith, 1989) predicts that the predominant process by which intra-cellular DNA molecules bearing regions of homology become physically linked is by extra-chromosomal homologous recombination. End to end joining plays a lesser role. Were end to end joining to be the primary route towards physical linkage of foreign DNA molecules then co-transfection of pSV2neo and pSV3gpt would probably be favoured since they have complementary Eco RI sticky ends whereas pMTneo and pMTLT have Sal I and Bam HI ends.
respectively. In the case of electric field mediated gene transfer it is not known whether the protruding termini of the DNA molecules are required to be complementary in order to undergo intra-cellular ligation. This is a requirement for microinjected DNA but not for DNA transferred by the calcium phosphate or DEAE-dextran technique (Bishop & Smith, 1989 for review and refs therein). One might predict that intra-cellular ligation of electroporated DNA, like microinjected DNA, might require complementary ends. Extra-chromosomal homologous recombination between DNA molecules proceeds via a non-conservative mechanism involving the free ends of DNA molecules (Lin et al. 1986, 1990a & 1990b; Wake et al., 1985 and Bishop & Smith, 1989 for review). Aligned linear molecules are poorer substrates for this type of recombination than non-aligned DNA molecules. With respect to their homologous regions, when linearised for the purposes of this transfection, pSV2neo and pSV3gpt constitute a pair of aligned DNA molecules whereas pMTneo:pMTLT, pMTneo:pSV3gpt and pMTLT:pSV2neo represent non-aligned DNA molecules. Although it is advisable to bear these considerations in mind, predictions as to their likely effect on the physical linkage of DNA molecules cannot be made with any certainty. It will suffice to conclude that, overall there would appear to be no obvious preference for the physical linkage of any of the pairs of plasmids.

Transfection Using Single Plasmids

When 229 skin cells are transfected with a single plasmid bearing only a selectable marker and no SV40 Tag gene, colonies do arise though their number is reduced and subsequent cloning and expansion into cell-lines is much less successful when compared to transfectants expressing SV40 Tag. In addition, the success of the transfection is much more sensitive to the health and age of the cells used when transformation to a drug resistant phenotype is not accompanied by an increase in proliferative capacity such as occurs when SV40 Tag is co-expressed with the selectable marker.
Using the same batch of cells at the same time as those cells transfected to produce the results presented in Table 4, pSV2neo, pMTneo, were introduced into 229 skin cells. The results are shown in Table 4. Taken together (see comparison of transfection efficiencies obtained with each combination of the four DNAs below Table 4) the results of both the single and co-transfections indicate several points worthy of attention.

a) The SV40 early region promoter is more efficient in promoting transcription in a stable transformation assay than is the mMT-I promoter. The single transfection data show that the difference is 3.6 fold. In a co-transfection assay replacing either of the SV40 constructs with the equivalent mMT-I construct results in a more than two fold reduction in transfection efficiency. When both neo and Tag are provided by mMT plasmids the reduction in transformation efficiency is 13 fold.

b) Whereas pSV3gpt (Tag) is able to increase colony formation by both pMTneo and pSV2neo (4 and 2.5 fold respectively), pMTLT does not increase colony formation by either. This probably reflects a deficiency on the part of pMTLT in establishing and maintaining sufficient Tag expression required to transform the cells.

c) pSV2neo is 3.6 fold more efficient than is pMTneo. If pSVneo has a probability of being expressed of 1 then the probability of pMTneo being expressed is 0.277. The probability of two SV40 promoter based plasmids being expressed is also 1 but the probability of two pMT plasmids being expressed would be 0.277 x 0.277 = 0.0077, if this is multiplied by 0.029, the transformation efficiency of pSV2neo:pSV3gpt, then the answer, 0.00223, is the transformation efficiency of pMTneo:pMTLT. This is to say that the deficiencies of pMTneo and pMTLT relative to their pSV counterparts is approximately equivalent and that in conjunction the shortfall in colony formation is the product of the individual deficiencies.
The correlation between promoter strength and frequency of stable transformation is well established (Luciw et al., 1983; Spandidos & Wlikie, 1983). There is no reason to suppose that the SV40 based DNA integrates at a higher frequency than either pMTneo or pMTLT. Therefore it seems that mMT based plasmids integrate into the chromosomes and fail to establish transcription more frequently than do SV40 based plasmids. When transfected singly, pSV2neo is nearly four times as efficient as pMTneo. This might be interpreted as meaning that four times as many pSV2neo integrants are active as are pMTneo integrants. The reason for such "silent" integrations is obscure but probably relates to a position effect phenomena whereby the SV40 early region is less sensitive to the cis-acting effects of the flanking DNA whether it be plasmid or chromosomal. The essentially random nature of chromosomal integration is likely to expose the transgene or transgenic array to a selection of chromosomal environments some of which may exclude the establishment of transcription.
All metallothionein genes studied so far are characterised by their response to heavy metal ions (Hamer, 1986). Exposure to heavy metal ions results in an increase in mRNA and protein levels.

To verify that the mMT-I fusion genes which stably expressed when introduced into primary skin fibroblasts (Table 4) were correctly regulated by heavy metals several clones and pools of clones were analysed. Several clonal cell lines were established from pMTLT:pMTneo transfectants, RNA was prepared from zinc treated and non-treated cells and subjected to northern analysis using a neomycin strand-specific RNA probe. The results are shown in Figure 10, panel A. All the cell-lines produce a message of around 1200 bases in length which is strongly induced by zinc (8 hour exposure, 100μM). In addition, 4/5 cell-lines produce a larger, less abundant, zinc inducible transcript. The level of the larger message does not correlate with that of the smaller one. For instance, in clone 2 this transcript is more abundant than suggested by the level of the major one. The band at around 1800 bases is probably an artefact of the RNA probe. Non-specific hybridisation is a common difficulty experienced when using RNA probes. This particular band may serve some purpose in confirming roughly equivalent loadings in the tracks. Both the basal and the induced expression varies between clones, though when induced, 9, 7 and 3 are quite similar. In all cases the basal level of expression is readily detected, even after a six hour exposure. This may be contrasted with endogenous sheep MT genes of which only the basal expression of sMT-II is detectable after a similar exposure (see Chapter 3). The hybridisation conditions including the concentration and specific activity of probe were the same in both cases and the results are therefore comparable. The mMT-I promoter is therefore able to establish
FIGURE 10

ANALYSIS OF FOREIGN AND ENDOGENOUS METALLOTHIONEIN GENE EXPRESSION IN CLONAL CELL-LINES

Total RNA was prepared from clonal cell-lines derived following transfection of 229 skin cells with plasmids pMTneo and pMTLT. RNA was prepared from control (-) and zinc treated (8 hours, 100μM ZnSO₄) cells (+). 10μg of each RNA sample was subjected to northern analysis using various probes.

Lane C: 5μg RNA from zinc treated cos-7 cells transiently expressing pMTneo
Lane M: neomycin sense strand RNA markers.

Panel A: neomycin strand specific RNA probe. 
           Autoradiographed for 6 hours.

Panel B: sheep MT-II strand specific RNA probe. 
           Autoradiographed for 12 hours.

Panel C: SV40 early region DNA probe. 
           Autoradiographed for 3 days.
correct expression in a sheep cell environment. The absolute level of expression is less than that observed for the sheep Ia and II genes, but probably greater than the Ib gene and certainly more than the Ic gene.

When the filter was re-hybridised with the sMT-II gene specific probe the relative strength of signal between the tracks was similar to that observed with the neomycin specific probe (Figure 10 panel B). When an identical northern blot was hybridised to an SV40 earlyregion specific DNA probe (Figure 10, panel C) the only deviation from the neo and sMT-II results was that clone 3 and not 6 shows the highest induced and basal level expression. This deviation might be explained if pMTLT and pMTneo lie either at separate loci or are not equivalently represented in clone 3. The similar results obtained with the three probes suggest that variation in expression between clones may be determined not by stochastic phenomena such as the arrangement of foreign DNA and its position of integration but by variation in the ability of individual clones to express MT genes and respond to zinc induction. It is easier to explain this result for the level of induced expression because, unlike the basal expression, it is determined by the ability of the cell to respond to heavy metal challenge and induce the expression of genes which are already active, albeit at a basal level. It is plausible that the response to zinc may override the constraints placed upon expression by the structure and location of the foreign DNA. In contrast, the basal expression of the transfected sequences is presumed to relate to the ability of the mMT-I promoter to establish transcription at a particular chromosomal locus. However, Figure 10 indicates that a clone of cells expressing high levels (relative to other clones) of the sMT-II gene is likely to express high levels of a foreign mMT-I fusion gene. This contradicts the theory that the expression of foreign genes is determined by the structure and chromosomal location of the foreign DNA integrant. This theory predicts that the levels of foreign and endogenous gene expression should vary independently between clones of cells or transgenic mice. The expression of an hβ-
globin genes (without DCRs) in transgenic mice provides a good example of this, though there is little variation in the endogenous gene expression (Ryan et al., 1989b). A better example is provided by the expression of an mMT-I human growth hormone fusion gene in several tissues of four transgenic mice (Palmiter et al., 1983). In this case the ratio of foreign to endogenous gene expression is quite consistent in the liver, lung and intestines of four transgenic mice. In other tissues, notably the heart and kidney, the ratios do vary, sometimes dramatically, between mice, i.e. the expression of the foreign and endogenous genes are not strongly related. Therefore, this particular transgenic example supports both the data presented in Figure 10 and the more common position effect phenomena observed in transgenic animals.

Expression of pMTneo in non-transformed 229 skin cells and mouse L-cells

To establish that correct regulation of the mMT-I promoter occurs in sheep cells in the absence of SV40 Tag expression, clones of pMTneo and pSV2neo transfectants shown in Table 4 were pooled and analysed. For comparison pools of mouse L-cell clones transfected with the same DNAs were also analysed. The result of northern analysis using a neomycin strand specific probe is shown in Figure 11 (panel A). Note that the predominant neo specific transcript in cells expressing pSV2neo is larger than in cells expressing pMTneo (lanes 1, 2 & 4, 5, 6, 8, 9, 10 respectively). It is apparent that pMTneo is regulated by zinc but not by copper in sheep cells (lanes 3, 4 & 5). The level of expression is lower than that observed in L-cells (lanes 8, 9 & 10) possibly reflecting a species preference of the mMT-I promoter. The gene is also regulated by copper in L-cells (lane 8). The reason for the lack of response to copper by the same gene in sheep cells is probably due to the medium used to culture the sheep cells (see Chapter 3). It is of interest to note that the same difference in expression between L-cells and 229 cells is not seen with pSV2neo. This is possibly a feature of the lack of cell
Mouse L-cells and 229 skin cells were transfected with plasmids pSV2neo, pMTneo and B1,2,polyIII neo (tk-neo). Pools of stably transfected clones were expanded and RNA prepared. 20μg of total cellular RNA was loaded in each track and subjected to northern analysis. The resulting filter was hybridised to a neomycin RNA probe (panel A) and later to an actin DNA probe (panel B) to confirm equivalent loadings. In both cases the filters were autoradiographed for 6 hours at -70°C.

Lane M: RNA markers, neomycin gene sense strand specific.
Lane 1: 229 pSV2neo clones.
Lane 2: 229 tk-neo clones.
Lane 3: 229 pMTneo clones, induction 200μM CuSO₄ for 8 hours.
Lane 4: 229 pMTneo clones, induction 100μM ZnSO₄ for 8 hours.
Lane 5: 229 pMTneo clones, no induction.
Lane 6: L-cell pSV2neo clones.
Lane 7: L-cell tk-neo.
Lane 8: L-cell pMTneo clones, induction 200μM CuSO₄ for 8 hours.
Lane 9: L-cell pMTneo clones, induction 100μM ZnSO₄ for 8 hours.
Lane 10: L-cell pMTneo clones, no induction.
specificity of the SV40 early region promoter (Spandidos & Wilkie, 1983) and supports the stable transfection data in which SV40 based constructs produced more stable colonies than did mMT-I fusion genes in sheep cells (Table 4). The tk promoter-neomycin coding region construct is expressed at a similarly low level in both cell-types. Panel B of the same figure shows the same filter re-hybridised to an human skeletal muscle α-actin probe (Hanauer et al., 1983) and demonstrates that variation in the level of neomycin specific message is not due to unequal loading in the tracks (note, the actin probe only allows comparison between RNA samples from the same cell-type).

Sheep cells can transcribe pMK and produce functional HSV-tk

It is formally possible that sheep cells are simply unable to transcribe pMK or produce HSV-tk. With this in mind pMK and pSV3neo (5 to 1 molar ratio) were co-transfected by electroporation into 229 skin cells. The pSV3neo served two functions, firstly to provide a selectable marker and secondly to produce SV40 Tag which stimulates cell proliferation and facilitates subsequent analysis.

The colonies of one plate (approximately 40 clones) were pooled and expanded to produce large numbers of cells for RNA and HSV-tk analysis. Figure 12 shows the RNA. The behaviour of pMK is similar to that previously observed for sheep MT genes (see earlier). Induction is observed up to and beyond 125μM ZnSO₄. No effect is noticed with copper (not shown). The basal level of expression is high and more reminiscent of the sheep MT-II gene than of the sMT-Ia gene. Using a double-stranded HSV-tk specific DNA probe four bands are visible, three of which are induced by zinc. The expected size of a correctly initiated transcript running to the HSV-tk polyadenylation site is around 1300 bases. This is the major transcript. There is no evidence of transcription initiating downstream of the cognate CAP site as observed in transgenic mice and cell-lines bearing HSV-tk genes (Al Shawi et al., 1988; Roberts & Axel, 1982). The other zinc-inducible transcripts might represent precursor RNAs.
Total cellular RNA was prepared from a pool of 229 clones transfected with plasmids pMK and pSV3neo following an 8 hour treatment with various zinc concentrations and subjected to northern analysis using an HSV-tk specific DNA probe. The X-ray film was exposed for 10 days. The position of RNA size markers are indicated.

Lane 1: 20μg RNA, no induction.
Lane 2: 20μg RNA, 50μM ZnSO₄.
Lane 3: 20μg RNA, 75μM ZnSO₄.
Lane 4: 20μg RNA, 100μM ZnSO₄.
Lane 5: 20μg RNA, 125μM ZnSO₄.
TABLE 5
PRODUCTION OF FUNCTIONAL HSV-tk IN SHEEP CELLS.

<table>
<thead>
<tr>
<th>CELL-TYPE</th>
<th>METAL TREATMENT</th>
<th>HSV-TK ACTIVITY pmol/min/mg ptn</th>
<th>INDUCTION FACTOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>229: pMK: pSV3neo</td>
<td>no treatment</td>
<td>14.2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25µM ZnSO₄</td>
<td>20.6</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>50µM</td>
<td>22.8</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>75µM</td>
<td>40.7</td>
<td>2.87</td>
</tr>
<tr>
<td></td>
<td>100µM</td>
<td>64.3</td>
<td>4.53</td>
</tr>
<tr>
<td></td>
<td>125µM</td>
<td>90.4</td>
<td>6.37</td>
</tr>
<tr>
<td></td>
<td>50µM CuSO₄</td>
<td>13.8</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>100µM</td>
<td>14</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>150µM</td>
<td>12.3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>200µM</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>250µM</td>
<td>17</td>
<td>1.2</td>
</tr>
</tbody>
</table>

229 cells               -                                0                               -
229 cells               +100µM Zn²⁺                   0                               -

All assays were performed in the presence of TTP to inhibit cellular thymidine kinase activity (see materials and methods). HSV-tk activities are expressed as pmol thymidine converted to thymidine monophosphate per minute per milligram protein in the reaction.

The reactions were assayed after 60 minutes by spotting aliquots of each onto two DEAE discs which were then washed and counted.
or aberrantly terminated transcripts. Since they are induced by zinc it is likely that they initiate from the mMT-I promoter. The band at around 1800 bases is not induced by zinc and probably represents an anti-sense transcript because it is not detected using an anti-sense RNA probe (data not shown). Largely speaking, the results of the tk assay (Table 5) mirror those of the northern analysis though the extent of induction of the RNA is not reflected in the HSV-tk levels which are maximally induced to a value only six times that of the basal level. The reasons for this are not clear but may be an artefact of the assay or reflect a limit on the intra-cellular HSV-tk protein.
Transient expression visualized in individual cells

The use of antibodies to detect transient expression in individual cells can provide a rapid assay of transfection efficiency (see appendix in this thesis). The results of such an assay can also provide an insight into the relative strengths of promoters. The assay is extremely sensitive but less easy to quantitate than conventional transient assays. Quantitation may take two forms, firstly the number or percentage of expressing cells and secondly the level of expression within individual cells.

Validity of the assay.

This type of assay has been used to measure the efficiency of gene expression of various recombinant plasmids bearing altered SV40 early region promoters (Moreau et al., 1981; Weintraub et al., 1988) and also to ascertain the influence of steroid hormones on the expression of the chicken ovalbumin gene in primary cultured oviduct tubular gland cells (Gaub et al., 1986). The assay has quantitative limitations and cannot measure the expression of the cell population as a whole as conventional transient assays do (eg RNA levels or enzyme assays eg CAT). However, the results obtained with the two types of assay are in agreement (Moreau et al., 1981; Benoist & Chambon, 1981; Weintraub et al., 1986). Discrepancies between the two types of assay are likely to arise in three ways. First, low expression in large numbers of cells might go undetected by the staining method but cumulatively might be positive in a conventional assay. Secondly, if the level of expression within a cell is less with one construct but the number of expressing cells is the same as observed with others, then the methods will give different answers. The relationship between the number of expressing nuclei and expression at the population level has been shown to be non-linear.
beyond a certain point (Dierich et al., 1987). Thirdly, the assay is liable to saturation beyond which no increase in promoter efficiency is detectable (Dierich et al., 1987).

Quantitation of the level of immunofluorescence within individual cells is possible by measuring the time needed to collect a specified number of photons, this being inversely related to the number of fluorescent antibody molecules (Weintraub, 1988). However, the real issue is whether the amount of fluorescence or colour staining is directly related to the expression of the target antigen. Though there is a correlation, the fidelity is not known.

Promoter strength correlates with the frequency of stable transformation (Luciw et al., 1983; Spandidos & Wilkie, 1983). Therefore, it is reasonable to assume that the features of a promoter which determine this frequency also bear upon the percentage of cells expressing from the same promoter in a transient assay (Moreau et al., 1981; Luciw et al., 1983). As discussed earlier in this thesis, the SV40 promoter is more efficient at achieving stable transformation than is the mMT-I promoter in sheep cells. It is therefore of value to see whether these data are confirmed by assay of the number of transiently expressing cells.

Using pSV3neo and pMTLT, SV40 Tag is transiently expressed in the nucleus of 229 skin cells

Using SV40 Tag as the target protein, an anti-Tag monoclonal antibody and an anti-IgG horse radish peroxidase conjugated antibody as the first and second antibodies respectively, it is possible to stain the nuclei of expressing cells. Since SV40 Tag is located in the nucleus, non-expressing and expressing cells are readily distinguishable. The assay provides an estimate of the number of expressing cells after transfection and allows discrimination between various levels of expression.
To further confirm that the mMT-I promoter is active when introduced into sheep cells, pMTLT was introduced into 229 skin cells by electroporation. Following electroporation around 10⁶ cells were plated per well of a 24-well dish (15 mm diameter). After 48 hours the cells were fixed and stained for SV40 Tag. Prior to staining some of the cells were treated overnight (16 hours) with 100µM ZnSO₄. As a comparison pSV3neo, in which the Tag expression is driven by its natural promoter/enhancer sequences from the SV40 early region, was also transfected. The number of stained cells per well was counted and typical expressing cells photographed.

The results are represented in Figure 13. Panel A shows the expression of pSV3neo without zinc induction. Between 500 and 600 expressing cells were counted per well. Contrast this with panel B which represents the expression of pMTLT in the absence of zinc. The only obvious, stained nuclei are those of two cells which appear to have just divided. Around 10 expressing cells were counted in each of two wells. The expression in these cells is high and it is notable that low or medium expressing cells are not readily apparent, the same also applies to panels A and C. Panel C shows expression of pMTLT following zinc treatment of cells from the same transfection as represented in panel B. Two points are worthy of note i) the total number of cells is reduced possibly due to the toxicity of the zinc induction and more importantly ii) the number of expressing cells is higher than that observed in the absence of zinc, approximately 100 per well.

These results confirm that the SV40 early region promoter/enhancer sequences are more potent than the mMT-I 5' region in stimulating transcription in sheep skin fibroblasts. This is apparently because more cells express pSV3neo than pMTLT, thus correlating with the respective performances of these promoters in a stable transformation assay (see Chapter 4). The major difference between pSV3neo and pMTLT appears to be one of the number of expressing cells rather than the level of expression within cells.
229 skin cells were electroporated in the presence of either plasmid pSV3neo or pMTLT (20µg/ml supercoiled plasmid). 10⁶ cells were plated in each well of a 24-well dish. Some of the pMTLT transfecants were grown in medium supplemented with 100µM ZnSO₄ for 16 hours prior to fixing and indirect immuno-staining at 48 hours post-electroporation. Following staining the cells were photographed.

PANEL A: pSV3neo.

PANEL B: pMTLT, no zinc treatment.

PANEL C: pMTLT, zinc treated.
The effect of zinc upon the transient expression of pMTLT

Since the concentration of zinc which elicits the maximum response from mMT-I genes (Yagle & Palmiter, 1985) is mildly toxic to cells, its use in a stable transformation assay is not really appropriate. Increased expression may be offset by the toxicity of zinc. When monitoring transient expression over a short time period toxicity is less of an issue. The results presented here (Figure 13, panels B & C) suggest that in a transient assay zinc stimulates expression by increasing both the number and percentage of cells expressing SV40 Tag. At a population level, when transient expression of plasmid pMK in BHK cells is measured by HSV-tk assays the stimulation by zinc is around ten fold (data not shown). It is plausible that the increase in the number of expressing cells following zinc stimulation could account for this.

It might be argued that the basal level of pMTLT expression is below the limit of detection and that only exceptional cells show densely staining nuclei in the absence of a heavy metal challenge. However, indirect evidence suggests that basal level expression should be detectable. Firstly, in transient assays following calcium phosphate mediated transfection (either pMTneo RNA in cos-7 cells or pMK HSV-tk activity in BHK cells) the basal level expression is readily detectable (data not shown). Secondly, when expression of stably integrated pMTLT is monitored by northern analysis (Figure 10, Chapter 5) or by indirect immuno-staining (Figure 14) in sheep cells the basal expression (Figure 14, panel A) is readily detectable. It is a prediction rather than an established fact that transient expression in individual cells is likely to be higher than that within stably transfected cells. In support of this, when equal amounts of RNA from cos-7 cells transiently and stably transfected by the CaPO₄ method with pMTneo are analysed, the basal level of mRNA is similar in both cases (Figure 15). In the case of the transient assay it is likely that an unknown majority of cells will not have received or failed to express the DNA. Indirectly, this suggests that the level of expression in a transiently expressing
Figure 14

REGULATION BY ZINC OF SV40 Tag EXPRESSION IN 229 SKIN CELLS STABLY TRANSFECTED WITH pMTLT

Clone MT3 cells (see Chapter 7) which stably express pMTLT and pMTneo were analysed for SV40 Tag expression by indirect immuno-staining. Panel A shows control cells which did not receive a zinc treatment and Panel B show cells which had been treated for 16 hours with growth medium supplemented with 100μM ZnSO₄ prior to the immuno-staining.
A: no Zn

B: +Zn
FIGURE 15.

TRANSIENT AND STABLE EXPRESSION OF pMTneo IN COS-7 CELLS

Cos-7 cells were transfected with plasmid pMTneo by the calcium phosphate-DNA co-precipitation technique (see materials and methods). For transient expression total cellular RNA was prepared 48 hours post-transfection. For stable expression cells were selected in growth medium supplemented with 400μg/ml G418 for 14 days. Total cellular RNA was prepared from pools of G418 resistant clones. For zinc induction 100μg/ml ZnSO₄ was added to the culture medium 8 hours prior to RNA preparation. 20μg of each RNA was subjected to Northern analysis using a random primed neomycin specific DNA probe. The filter was exposed to X-ray film for 36 hours at -70°C.

Lane 1: Stable expression, no zinc treatment.
Lane 2: Stable expression, zinc treated.
Lane 3: Transient expression, no zinc treatment.
Lane 4: Transient expression, zinc treated.
cell will be more than that in a cell expressing stably integrated copies. Therefore it is likely that basal level, transient expression of pMTLT would be detected by indirect immuno-staining and that those cells visibly expressing pMTLT in the absence of zinc represent true basal level expression.

The intensity of staining appears similar whether the cells are expressing pMTLT (in the presence or absence of zinc) or pSV3neo. This may be due to the quantitative limitations of the assay. Densely staining nuclei may represent a saturation point beyond which any increase in the amount of target antigen (in this case Tag) will not be detected or alternatively indicate that expression within the cell has reached a limit. Both must occur at some point (Dierich et al., 1987). Since cells expressing pSV3neo (or pSV3gpt see appendix:1) exhibit varying staining intensity it is likely that neither the expression nor the assay has reached saturation levels in all cases.

In summary, in the absence of zinc the number of cells detectably expressing pMTLT is around 2% of that expressing pSV3neo. Following zinc induction the number of cells expressing pMTLT increases to 20% of that expressing pSV3neo. The major effect of zinc administration or substitution of the mMT-I promoter with that from the SV40 early region is to increase the number of expressing cells and not the level of expression within individual cells.

The expression of pMTLT in stably transfected sheep cells has been analysed, zinc can increase the Tag expression in individual cells (see Figure 14). This is in contrast to the result obtained following transient expression. The effect of zinc upon chromosomal genes must be to increase the total number of RNA polymerase molecules active on the template ie to increase the rate of transcription initiation. This is the effect of classical enhancer sequences (Treisman & Maniatis, 1985; Weber & Schaffner, 1985). In a transient assay, if the only effect of zinc was to increase the expression in individual cells then the expected outcome of a transient immuno-staining assay would be a low level of

-118-
expression in some cells in the absence of zinc and a higher level of expression in a similar number of cells following zinc treatment. This is not the result obtained. Instead, the transient expression data suggest a role for zinc in the establishment of active transcriptional complexes.

Transient expression is mostly derived from expression of non-integrated copies of transfected DNA. Therefore the level of transient expression within a cell can be determined by both the number of active templates and the rate of transcription at active templates. The increase in the number of expressing cells following zinc treatment suggests that the metal increases the probability of transcription being established (in a given cell). If this is true then one might argue that only a few templates are active following transfection. This may arise because the amount of DNA is limiting. If the DNA were in vast excess then zinc might be expected to produce cells expressing at levels considerably above those seen in the absence of zinc simply because the metal establishes transcription at more templates. However, when the amount of DNA is limiting and possibly non-randomly distributed between cells it is logical to suppose that the chances of a cell expressing pMTLT is greater following zinc induction. This can explain why the number and percentage of expressing cells increases following zinc administration. It is possible that the amount of DNA entering the cells by electroporation is limiting or that only a few templates are active because when BHK cells are transfected by both the CaPO₄:DNA co-precipitation technique and by electroporation, the level of expression in some of the cells transfected by the former method is greater than that detected following electroporation (data not shown).

In conclusion, the mechanism by which zinc is able to increase transient gene expression is likely to be two-fold, firstly by increasing the probability of transcription occurring at a given template and secondly by stimulating higher expression from an active template. When the template is limiting an increase in the
probability of a template being active will have a larger effect upon the number of expressing cells than upon the expression within individual cells. Analysis of the SV40 early region enhancer using anti-Tag immuno-staining provides evidence to support this theory.

Enhancer mediated activation of transcription: As assayed by indirect immuno-staining.

Enhancers can dramatically activate the expression of linked genes. The effect is independent of orientation and distance from the promoter (Serfling et al., 1985 for review). In population level transient assays an enhancer can stimulate expression several hundred fold (Banerji et al., 1981; Spandidos & Wilkie, 1983; Benoist & Chambon, 1981) and in stable transfection assays enhancers increase the colony formation (Luciw et al., 1983; Spandidos & Wilkie, 1983). Stable transformation of tk-mouse L cells can be enhanced 20-40 fold when the HSV-tk gene is linked to either the SV40 or rous sarcoma virus enhancers (Luciw et al., 1983). Enhancers function, at least in part, by establishing stable transcription complexes (Mattaj et al., 1985; Wang & Calame, 1986). Additionally, the presence of an enhancer increases the number of polymerase molecules engaged in the transcription of the linked DNA (Treisman & Maniatis, 1985; Weber & Schaffner, 1985). Although this implies that the polymerase density per gene copy is greater in the presence of an enhancer, it allows no conclusion as to how this is achieved. In a transient assay an enhancer could increase the polymerase density on active templates or increase the number of active templates.

When the transient expression in individual cells is used to assay SV40 enhancer sequences, the number of cells expressing the SV40 Tag marker is increased when the Tag gene is linked to an enhancer sequence (Moreau et al., 1981). In this assay wild-type SV40 sequences gave the full range of fluorescent intensity whereas poorly expressed recombinant molecules gave a low number of weakly fluorescent cells. This does not necessarily favour either mechanism
of transcriptional enhancement, enhancer mediated increase of the number of active templates or enhancer mediated increase in the polymerase density per active template could both explain the effect. In this study, the use of micro-injection instead of DEAE-Dextran transfection improved the performance of poorly expressed constructs. In some instances the improvement was very large from <1% to >80% of the number of cells expressing the wild-type SV40 control DNA. This suggests that the outcome may depend upon the amount of DNA successfully transferred to the nucleus. Micro-injection guarantees the transfer of 150-300, mostly undamaged, DNA molecules directly to the nucleus, this is probably more than the average for DEAE-dextran transfection (McCutchen & Payano, 1968) where DNA is vulnerable to damage (Wake et al., 1984). When the DNA concentration is raised, constructs bearing a disabled enhancer can promote transcription in a larger percentage of cells, however even under such circumstances a large number of cells still do not express. One interpretation of this is that in the absence of an enhancer the probability of transcription occurring at a given template is low and that expression under such circumstances is not due to low transcription of available molecules but relatively high transcription of a very small number of them. The effect of increasing the concentration of the transferred DNA when the enhancer is absent or disabled would be to increase the chance of stable transcription being established not to allow more templates to be poorly transcribed.

Recently these observations have been extended and refined (Weintraub et al., 1988). The critical feature of this work being that conditions were established whereby co-expression of two unlinked markers was very rare, and by implication expression of either marker in individual cells was probably derived from a single active, template (Weintraub, 1988). The study compares the expression at single cell resolution of genes linked to an SV40 wild type enhancer and those linked to a mutated or deleted enhancer. At the population level the enhancer-less construct was expressed at 1% of the level of that with a functional enhancer. When expression of
the same constructs were analysed for transient expression of SV40 Tag by indirect immuno-staining the enhancer-less construct was expressed in 0.1-1% of the number of cells which expressed Tag when linked to a functional SV40 enhancer. This is in accord with previous data (Moreau et al., 1981 and see above). Quantitative fluorescence microscopy revealed that the level of expression within individual cells was similar regardless of the presence of an active enhancer. 100 cells expressing both constructs were assayed and the average fluorescence of cells expressing the enhancer gene was only 2-3 fold above those expressing the gene in the absence of the enhancer. Clearly this is not sufficient to account for the 100 fold difference in transient expression observed at the population level. The most obvious conclusion to be drawn from these results is that there is a low, but finite, probability of full transcription being established at a given template in the absence of an enhancer. The effect of an enhancer is to dramatically increase the chances of this occurring. Sub-cloning experiments suggested that the transcription complexes are stable for both constructs. Since only one or a few cells of a sub-clone expressed the marker the assertion that expression is derived from a small number of active templates is strengthened. These results cannot discount the possibility that without an enhancer Tag is expressed at very low levels in the large number of cells which presumably receive DNA but do not significantly express the marker, and that the enhancer can stimulate transcription from these poorly transcribed templates. However, this does not alter the major conclusion, that enhancers catalyse the formation fully active templates.

This data confirms that when the template is limiting the effect of an enhancer in a transient assay is to increase the percentage of those cells receiving DNA which express the marker. To explain the effect of zinc upon the transient expression of pMTLT it is necessary to suppose that the enhancing properties of the mMT-I promoter are dependent upon or dramatically increased by the presence of zinc.
Metallothionein promoters as inducible enhancer elements.

The promoter regions of metallothionein genes can function as bona fide inducible enhancers. The response to heavy metals is mediated through the MREs which themselves possess the properties of an enhancer type element (Serfling et al., 1985b; Karin et al., 1987; Westin & Schaffner, 1988). The difference between the enhancer elements in the SV40 early region promoter and MREs is that the former is constitutively active whilst the latter is only highly active under certain conditions, for instance during a heavy metal challenge. This is because the activity of the SV40 enhancer is mediated by trans-acting factors which are constitutively active in many cell-types (though SV40 transcription can be induced by TPA via the API trans-activator protein; ref Angel et al., 1987). In contrast, it appears that occupancy of the MREs upstream of metallothionein genes is dependant upon the presence of heavy metals (Mueller et al., 1988), i.e., the activity of such metal regulatory proteins is conditional upon a heavy metal challenge. When the challenge is removed the expression reverts to basal level. This is to say that the enhancing properties of the MREs are only stable in the presence of heavy metals. Presumably this applies to inducible enhancers in general (Maniatis et al., 1987). The activity of a chicken ovalbumin promoter in oviduct tubular gland cells is repressed in the absence of steroid hormones, the repression is mediated in cis and removal of this sequence leads to a high constitutive activity. In a transient immunofluorescence assay the number of expressing cells is low in medium depleted of steroids. Addition of steroids leads to a 2-3 fold increase in the number of expressing cells (Moreau et al., 1981). Cells were scored either as positive or negative. The prevalent level of expression in the presence or absence of the steroid was not mentioned but it is clear that basal level expression is readily detectable (26% of the number of cells expressing the SV40 early region promoter driven SV40 Tag positive control). This result adds credence to the results presented in this thesis regarding the expression of pMTLT in sheep fibroblasts. However there is no evidence that the heavy metal regulation of MT
gene expression arises by reversal of a suppression mechanism which operates in the absence of the induction (Karin et al., 1987; Hamer, 1986; Stuart et al., 1985, 1984; Culotta & Hamer, 1989; Searle et al., 1985; Carter et al., 1985). For instance none of a variety of mutated mMT-I promoter regions assayed showed significantly elevated basal level expression (Stuart et al., 1984; Carter et al., 1985).

Metallothionein promoters contain regions which determine the basal level of activity (Stuart et al., 1985; Carter et al., 1985; Peterson et al., 1988b) and such sequences of varying match to the consensus are found in mouse, human and sheep genes. The sMT-Ia promoter which lacks an identifiable basal level element is expressed at very low levels in the absence of heavy metals (Peterson et al., 1986, 1988a & 1988b and this thesis), the sMT-II gene contains this element with strong identity to the consensus and shows a much higher basal level of expression (Peterson et al., 1988a & 1988b and this thesis). The activity of this type of sequence and its relation to constitutive and induced expression is best characterised in the hMT-IIa gene which shows exceptionally high basal level activity compared to the multiple human type I genes (Heguy et al., 1986; Richards et al., 1984). It is now apparent that the high basal activity of both the hMT-IIa and the sMT-II promoters in cultured cells is due to factors present in serum in particular activators of protein kinase C (Imbra and Karin, 1987; and Paul Smith data not shown). The mMT-I promoter is unresponsive to activators of protein kinase C (Mueller et al., 1988). In the hMT-IIa gene the basal level enhancer element, which is flanked by two MREs, functions independantly of the MREs (Karin et al., 1987) but the converse is not true, ie the MREs do not function efficiently when the basal element is deleted. This is in disagreement with the data regarding the murine MREs, which can function independantly of other MT promoter elements when multiple copies are present (Westin & Schaffner, 1988), suggesting that the promoter of the hMT-IIa and perhaps the sMT-II genes may be atypical.
Therefore it appears that a fundamental difference exists between the regulation of the hMT-IIa gene and the mMT-I gene particularly regarding roles of the MREs. Those belonging to the human gene do not appear to function in the absence of the basal level enhancer which itself is sufficient to establish basal level expression, on the other hand the murine MREs are able to function in the absence of other MT promoter elements (Searle et al., 1985; Culotta & Hamer, 1989; Westin & Schaffner, 1988). These observations may offer some explanation for the increase in the number of sheep cells expressing pMTLT following zinc induction since they imply that, when active, the MREs can promote the establishment of transcription in addition to increasing a basal level of transcription following heavy metal induction.
CONCLUSIONS REGARDING THE INACTIVITY OF THE 229 TRANSGENIC INSERT

Southern analysis of the pMK sequences in 229 DNA indicate that they are structurally intact and therefore do not suggest that re-arrangement of the transgene can account for the failure to express. Skin fibroblast cells derived from sheep 229 are proficient in the expression of the endogenous sheep MT gene family. In each case, when challenged, 229 skin cells express mMT-I fusion genes correctly. The fusion genes, including pMK, are active and zinc inducible following stable transfer into the sheep skin cell genome and pMTLT is expressed transiently in the same cells. Therefore the mMT-I promoter is not inert in sheep skin cells. However, both the transient and stable transformation data do indicate that the mMT-I driven fusion genes are expressed less efficiently than those driven by the SV40 early region promoter. As discussed in Chapter 6, the transient expression data indicates that mMT-I fusions genes appear to be deficient in establishing transcription since the number of cells expressing pMTLT is around 2% of that expressing an equivalent SV40 based construct. The stable transformation data (discussed in Chapter 4) indicate that the mMT-I fusion genes are frequently not expressed when inserted at random sites into the sheep genome because pSV2neo produces more than 3 times the number of stable colonies when introduced into 229 skin cells than does pMTneo. Presumably this relates to the structure of the foreign DNA integrant and the site of chromosomal integration. The data presented in this thesis are consistent with a similar cause for the failure of the 229 insert to express HSV-tk mRNA or protein. Since there is only one copy of the foreign pMK molecule at the transgenic integration site and other mMT-I fusion genes are active when inserted into other sites in the genome the failure of the transgene to be expressed probably reflects an overriding, negative, chromosomal position effect phenomenon such as that described recently in transgenic mice (Al-Shawi et al., 1990 and also see introduction).
The observation that mMT-I promoter is weak, with respect to its ability to establish transcription in sheep cells, compared to the SV40 early region promoter may have important implications for its use in the genetic manipulation of livestock. It appears that most mMT-I fusion genes in transgenic mice are liable to "silent" integration (see Table 1 in this thesis); in some cases the frequency of expression can be extremely low (see Table 1). Growth promoting genes fused to the mMT-I promoter have been used to produce transgenic livestock. Again, these genes appear to be prone to "silent" integration (see Table 2). In transgenic mouse experiments, substitution of the mMT-I promoter with the mAlbumin, mH-2K^a (major histocompatibility class I promoter) or mTransferrin promoters can raise the frequency of hGH expression to 100% as opposed to 70% with the mMT-I fusion gene (Palmiter & Brinster, 1986). All 8 transgenic mice bearing an mAlbumin: Hepatitis B surface antigen fusion expressed whereas only a third of those bearing the equivalent mMT-I construct did so (Palmiter & Brinster, 1986). Since the frequency of production of transgenic pigs and sheep is up to ten-fold lower than that achieved in mice, the likelihood of non-expression is an important consideration. For instance, 1000 transferred eggs can be expected to generate less than 5 expressing transgenic pigs or sheep (Purcel et al., 1989; Rexroad et al., 1989). If large numbers of expressing transgenic animals are a prime requirement, the results presented in this thesis, and in numerous transgenic experiments suggest that the mMT-I promoter is not ideal in this respect.
CHAPTER 7

ATTEMPTS TO EXTEND THE \textit{in vitro} LIFESPAN OF 229 SKIN FIBROBLASTS BY TRANSFECTION OF VIRAL AND CELLULAR ONCOGENES

This line of study was pursued for several reasons. Primarily because the 229 fibroblast cell-line derived from sheep skin has a finite life-span in culture and the period during which it is amenable to study, is limited. Clonal expansion and analysis of clones arising from stable transfection experiments is not always possible with the primary cell-line but would be if it were feasible to extend its life-span. In addition to these practical considerations, the study is of topical interest since the field of \textit{in vitro} transformation of primary mammalian cultured cells is largely polarised between studies on rodent and human cells. Human and rodent cultured cells differ greatly in their susceptibility to \textit{in vitro} transformation and results obtained with rodent cells cannot be extrapolated to include human cells.

Identification and characterisation of transformed cells

Transformation is used to describe the acquisition of tumour cell characteristics ie it refers to the transition from normal to neoplastic growth. Neoplastic growth is characterised by several \textit{in vivo} and \textit{in vitro} assays; these include altered morphology, loss of contact and density dependent growth, anchorage independent growth and immortality. Cell-lines which exhibit at least one of these characteristics may be considered transformed.

Several properties of transformed cells relate to changes to the cell surface which affect cell to cell and cell to surface contacts. In normal, non-transformed cells, cell to cell and cell to surface contacts often, though not always, dictate that they cease to divide when the cell monolayer is confluent. Normal cells also
generally require attachment to a solid support before they can divide and grow, this is termed anchorage dependent growth. In contrast, transformed cells can often continue to divide when the monolayer is confluent and as a consequence acquire a less flattened morphology and eventually form piles of rounded cells or foci. Transformed cells may also exhibit anchorage independent growth when plated in a soft agar suspension (Macpherson & Montagnier, 1964).

Transformed cells have a tendency to grow in a disorganised fashion. This may be particularly obvious if growth of the normal cells is very ordered.

Cells derived from tissue explants may not grow at all (eg mammalian hepatocytes) or undergo a fixed number of population doublings prior to a characteristic senescence (eg human skin fibroblasts). Transformed cells are usually immortal or have an increased proliferative capacity compared with their non-transformed counterparts.

The use of SV40 Tag to transform sheep skin cells

Apart from its essential role in SV40 origin mediated DNA replication, the SV40 large T antigen (Tag) can generate varying degrees of neoplastic transformation on a wide variety of cell lines. Therefore Tag can be considered a pernicious oncogene and an obvious candidate for initial attempts to transform primary skin fibroblasts. The constructs used in this work encode the whole SV40 early coding region and therefore produce both Tag and the small tumour antigen (tag). The role of tag in cellular transformation is not clear though it may be required in some circumstances (see later). Therefore, here I refer to Tag only.

The plasmide used contained the coding section of the SV40 early region linked either to the natural early region promoter such as in pSV3gpt and pSV3neo (Mulligan & Berg, 1980; Southern & Berg,
1982) or linked to the mMT-I promoter, pMTLT. pSV3gpt and pMTneo were co-transfected into 229 skin cells and some of the cells were selected in growth medium supplemented with G418. The remainder were allowed to grow without selection. Around 10 days after plating rapidly growing colonies were evident (Figure 16, Frame a). At 14 days the selection was complete and cloning rings were used to isolate several colonies. Most of these were expanded into cell-lines. All the colonies from one plate were Tag positive as assayed by indirect immuno-staining. This indicates a large selective advantage of co-transfected over pMTneo singly transfected clones. Two weeks after plating some of the cells which had been allowed to grow without selection had developed into foci (Figure 16, frames b & c). Prior to this they had appeared as rapidly growing cells with a morphology distinct from that exhibited by the surrounding cells. Some of these foci were cloned and expanded into cell-lines.

Three of the clonal cell-lines were analysed in more detail using the parental skin cells, which had been maintained during the period of selection and clonal expansion, for comparison.

SV40 Tag producing cell-lines grow more rapidly to a higher saturation density than primary 229 skin fibroblasts

Ten thousand cells were plated in each well of a 24 well plate and fed every day following. At day 3 and every day there after the cells were recovered from the wells and counted. The results are shown in Figure 17. During the same 11 day period the three transformed cell lines reached a saturation density of around 500,000 per cm² as opposed to 30,000 per cm² for 229. To reach this cell-density the transformed cells outgrew the monolayer and at high cell-density the cells clumped and eventually sheets of cells became detached from the dish. Typical of non-transformed fibroblasts the 229 parental cells became quiescent at the stage of the confluent mono-layer.
FIGURE 16

ALTERED MORPHOLOGY AND GROWTH OF 229 SKN CELLS FOLLOWING TRANSFORMATION BY SV40 Tag

PANEL A: shows a rapidly growing colony transformed with plasmids pSV3gpt and pMTneo ten days following electroporation.

PANEL B: shows the border between non-transformed (top right) and transformed (bottom left) cells ten days following electroporation.

PANEL C: shows a focus (centre) of transformed cells flanked by non-transformed cells (left). Picture taken 14 days following electroporation.
FIGURE 17

SV40 Tag PRODUCING CELL-LINES GROW MORE RAPIDLY TO A HIGHER SATURATION DENSITY THAN PRIMARY 229 SKIN FIBROBLASTS

10,000 cells were plated in each well of a 24-well plate and re-fed every day following. At day three and every day thereafter the cells were trypsinised and counted. The figure opposite shows the growth of three transformed cell-lines (NS, CT3 and CT2) and the parental 229 line as a comparison.
[10^5 Cells]

24 Hours  Days in culture
**TABLE 6**

**PLATING EFFICIENCIES OF TRANSFORMED CELL-LINES ON PLASTIC CULTURE DISHES AND IN SOFT AGAR**

<table>
<thead>
<tr>
<th>CELL-LINE</th>
<th>PLASTIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT3</td>
<td>14</td>
</tr>
<tr>
<td>NS</td>
<td>21</td>
</tr>
<tr>
<td>CT2</td>
<td>7</td>
</tr>
<tr>
<td>229 parental</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SOFT AGAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

* 500 and 1000 cells were plated onto 9 cm tissue culture dishes and allowed to grow for 9 days. Colonies were fixed, GIEEMSA stained and counted. The plating efficiency is expressed as the percentage of cells plated developing into colonies.

* 5000 cells were plated in soft agar (Machpherson & Montagnier, 1964) in 60mm culture dishes and allowed to grow for 21 days. The number of colonies in several 1 cm² areas were counted. The plating efficiency is expressed as the percentage of cells plated developing into colonies.
The plating efficiency of the transformed cell-lines is greatly increased

9cm plates were seeded with 500 or 1000 cells. After 9 days the colonies were fixed, stained and counted. The results are shown in Table 6. The parental 229 cells did not form colonies presumably because their proliferative potential did not permit clonal outgrowth from a single viable cell. The transformed cell-lines exhibited plating efficiencies ranging from 7-21%. Presumably, expression of Tag increased the proliferative capacity of the cells and thus enabled them to form colonies.

The transformed cell-lines will form colonies in soft agar

Many transformed cell-lines show anchorage independent growth as assayed by colony formation in semi-solid media eg soft agar (Macpherson & Montaignier, 1964). The results of such an assay are shown in Table 6. The frequencies may be over-estimated since the colonies were generally small and not readily distinguishable from clumps of dead cells. The small size of the colonies probably indicates that growth in semi-solid media was restricted compared to that on plastic culture dishes.

The plating efficiencies in soft agar and on culture dishes are not correlated. For instance NS had the highest plating efficiency on plastic but the lowest in agar.

Sheep skin cells expressing SV40 Tag have a greatly increased proliferative capacity

The transformed sheep skin cell-lines exhibit several of the traits typical of neoplastic cells. Relative to the non-transformed parental cells they have a decreased doubling time, reach confluence at a higher cell-density, have a greatly increased plating
efficiency, are not contact inhibited and show some degree of anchorage independent growth. However the cells eventually undergo a reproducible crisis and senescence after approximately 3 months in culture (data not shown).

Thus, they are not immortalised. Several explanations for this may be proposed. Firstly, SV40 Tag may lack functions necessary for immortalisation. Secondly, it may be impossible to immortalise the cells, perhaps because dominant senescent mechanisms are in operation (Smith & Periera-Smith, 1990). Thirdly, beyond the natural crisis point, proliferation is conditional upon the expression of SV40 Tag and that this may be progressively lost during prolonged culture.

Since SV40 completely transforms rodent fibroblasts (Tooze, 1980; Lanford et al., 1985; Bishop, 1985), it is not usually co-transfected with another oncogene. In this respect SV40 Tag differs from the tumour antigens of another papovavirus, polyoma virus middle and large T antigens are both required to achieve full transformation (Raasoulzadegan et al., 1982). Cooperation between transforming genes extends to cellular oncogenes such as ras and myc (Land et al., 1983). It is an attractive hypothesis that the transforming abilities of SV40 Tag are, at least in part related to its ability to interact with cellular proteins involved in the regulation of the cell-cycle. The best characterised of these interactions is with the cellular phosphoprotein p53 (Jenkins & Sturzbecher, 1988). More recently, an interaction with the protein product of the classic anti-oncogene allele, retinoblastoma (De Caprio et al., 1988) has been detected. Under certain circumstances mutated p53 can cooperate with activated ras oncogenes and itself immortalise rodent fibroblasts (Jenkins et al., 1984).

To determine the extent of the increased proliferative capacity of SV40 transformed sheep cells and to explore the possibility that p53 might immortalise 229 skin cells or cooperate with SV40 Tag in transformation the two genes were co-transfected.
TABLE 7
EFFECT OF MURINE p53 AND SV40 Tag UPON COLONY FORMATION IN 229 SKIN FIBROBLASTS

<table>
<thead>
<tr>
<th>DNA</th>
<th>COLONIES</th>
<th>NO. OF PLATES</th>
<th>AVERAGE</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSV2neo</td>
<td>140</td>
<td>2</td>
<td>70</td>
<td>0.02 % #</td>
</tr>
<tr>
<td>pSV2neo: pSV3gpt</td>
<td>294</td>
<td>2</td>
<td>147</td>
<td>0.039 % #</td>
</tr>
<tr>
<td>p4JJKan</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>0.001 % #</td>
</tr>
<tr>
<td>p4JJKan: pSV3gpt</td>
<td>213</td>
<td>2</td>
<td>107</td>
<td>0.03 % #</td>
</tr>
</tbody>
</table>

NO. OF FOCI

<table>
<thead>
<tr>
<th>DNA</th>
<th>NO. OF FOCI</th>
<th>AVERAGE</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSV2neo</td>
<td>256</td>
<td>256</td>
<td>0.068 % §</td>
</tr>
<tr>
<td>p4JJKan: pSV3gpt</td>
<td>252</td>
<td>252</td>
<td>0.07 % §</td>
</tr>
</tbody>
</table>

229 skin cells were electroporated in the presence of 13μg/ml of each linearised plasmid. Selection for G418 resistance was performed by supplementing growth medium with 400μg/ml geneticin (Gibco, G418) 72 hours following electroporation.

# colonies counted 14 days after electroporation

§ Foci counted 15 days after electroporation, growth in medium supplemented with 5% foetal calf serum (fcs) as opposed to 10% fcs.
The Tag was provided by plasmid pSV3gpt (as before) and p53 by plasmid p4JJkan which also contains the neomycin phosphotransferase gene linked to the SV40 early region promoter/enhancer (Jenkins et al., 1984). This construct alone promotes immortality and allows subsequent transformation by the ras oncogene in adult rat chondrocytes (Jenkins et al., 1984). As controls in the experiment, pSV3gpt and pSV2neo were co-transfected and both pSV2neo and p4JJkan were transfected alone. Two out of three plates were selected in medium supplemented with G418 and the remaining plate was grown in non-selective medium supplemented with only 5% foetal calf serum as opposed to the normal 10%.

It is clear (Table 7), that transfection of p4JJkan and possibly by implication expression of murine p53 reduces colony formation by 229 skin fibroblasts. pSV2neo is twenty fold more effective in this respect. The colonies derived from transfection of pSV2neo were more vigorous than those derived from p4JJkan and could eventually be cloned (see later Table 10). Since the sequences which provoke survival in medium supplemented with G418 are the same in each case, it is highly unlikely that failure to attain G418 resistance can, alone, explain the result. This view is strengthened by the results obtained when pSV3gpt was co-transfected with each of the plasmids. The ratio of colonies formed, pSV2neo/p4JJkan, both in the presence of pSV3gpt, was 1.3, as opposed to 20 in the absence of pSV3gpt, a 15 fold decrease in the ratio (Table 7). Thus, the mutant murine p53 appears to prejudice the survival of skin cells, and Tag appears to counteract this. It is possible that Tag provides a dominant stimulus to proliferation regardless of the presence of murine p53. Alternatively and perhaps more likely, Tag and murine p53 form complexes within the cell and in this way p53 is removed from active circulation. The fate of ovine p53 in cells expressing Tag is not known. The evolutionary conservation of the Tag:p53 interaction (Jenkins & Sturzbecher, 1988) predicts that some Tag will form complexes with ovine p53. Therefore the level of ovine p53 may determine the ability of Tag to override the effects of the murine p53 and transform the cells. The levels of murine and ovine
Plasmid pSV2neo contains the neomycin phosphotransferase gene fused to the SV40 early region promoter/enhancer (Southern & Berg, 1982).

Plasmid pSV3gpt contains the bacterial xanthine-guanine phosphoribosyltransferase gene and the SV40 early coding region both under the control of the SV40 early region promoter/enhancer. (Mulligan & Berg, 1980).

Plasmid p4JJkan (Jenkins et al., 1984) contains the neomycin phosphotransferase gene under the control of the SV40 early region promoter/enhancer, as in pSV2neo. Expression of a transforming murine p53 cDNA clone is directed by the Rous Sarcoma Virus long terminal repeat (RSV-LTR). For both neo and p53 transcripts, splicing and polyadenylation is directed by appropriate sequences from pSV2neo.
The cell-lines were passaged as described in text. 229 cells were re-plated between passages at $5.3 \times 10^{3}$ cells/cm$^2$ and the transformed lines at $2.6 \times 10^{3}$/cm$^2$. During the pre-crisis growth period the cells were passaged every 5-7 days.

**TABLE 8**

<table>
<thead>
<tr>
<th>CELL-LINE</th>
<th>POPULATION DOUBLINGS AT SENESCENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>229 Parental line</td>
<td>32</td>
</tr>
<tr>
<td>pSV3gpt:pSV2neo 2</td>
<td>72</td>
</tr>
<tr>
<td>pSV3gpt:pSV2neo 3</td>
<td>70</td>
</tr>
<tr>
<td>pSV3gpt:pSV2neo 4</td>
<td>72</td>
</tr>
<tr>
<td>pSV3gpt:pSV2neo 5</td>
<td>73</td>
</tr>
<tr>
<td>pSV3gpt:p4JJKan 1</td>
<td>66</td>
</tr>
<tr>
<td>pSV3gpt:p4JJKan 2</td>
<td>66</td>
</tr>
<tr>
<td>pSV3gpt:p4JJKan 3</td>
<td>63</td>
</tr>
<tr>
<td>pSV3gpt:p4JJKan 4</td>
<td>52</td>
</tr>
<tr>
<td>pSV3gpt:p4JJKan 5</td>
<td>73</td>
</tr>
<tr>
<td>pMTLT:pMTneo 1</td>
<td>62</td>
</tr>
<tr>
<td>pMTLT:pMTneo 2</td>
<td>42</td>
</tr>
<tr>
<td>pMTLT:pMTneo 3</td>
<td>70</td>
</tr>
<tr>
<td>pMTLT:pMTneo 3Z</td>
<td>46</td>
</tr>
<tr>
<td>pMTLT:pMTneo 4</td>
<td>63</td>
</tr>
</tbody>
</table>

**CALCULATION OF DOUBLING TIME BETWEEN PASSAGES**

$$T = \frac{d \log 2}{\log (H/P)}$$

$T$ = doubling time

$d$ = time between passages

$H$ = number of cells harvested

$P$ = number of cells plated

Population doublings between passages = $d/T$
p53 protein are not known. However, the primary cells and those expressing SV40 Tag contain similar amounts of ovine p53 mRNA as assayed by northern analysis using a murine cDNA probe, the level of message is quite low compared to the amount of hybridising species present in cells co-transfected with p4JJkan and pSV3gpt (data not shown).

Transformed cells characteristically outgrow the monolayer (see earlier) particularly in medium containing reduced serum. Under conditions of low serum only cells producing SV40 Tag will form foci. To determine the efficiency of transfection the number of foci obtained following co-transfection of pSV3gpt with each of pSV2neo and p4JJKan was scored. Transfection of p4JJKan:pSV3gpt and pSV2neo:pSV3gpt result in the same number of foci, this indicates that the efficiency of the transfection is the same in both cases and cannot account for differences in the number of G418 resistant colonies obtained. The number of foci is greater than the frequency of G418⁻ colony formation, this presumably reflects the frequency of co-transfection.

In order to monitor the extent of the increase in proliferative capacity of sheep cell-lines expressing SV40 Tag, four colonies derived from pSV2neo:pSV3gpt co-transfection and five from p4JJKan:pSV3gpt were isolated with cloning rings and expanded into cell-lines. These lines and the parental cell-line used in the transfection were propagated and passaged concurrently until crisis.

All the cell-lines were passaged every 5-7 days depending upon the time taken to reach confluence. To achieve this the parental cells were re-plated after each passage at twice the density of the transformed cells. At each passage the cells were counted to allow an estimate of the number of population doublings achieved since plating. Table 8 shows the number of population doublings achieved, and Figure 18 illustrates the proliferative histories of the transformed and non-transformed 229 skin cells. The parental cell-line exhibited a steady rate of growth during the
FIGURE 18

PROLIFERATIVE HISTORIES OF NON-TRANSFORMED 229 SKIN CELLS AND SV40 Tag TRANSFORMED DAUGHTER CLONES

Transformed clonal cell-lines were established and passaged as described in the text and Table 8. The number of population doublings achieved between passages was calculated as described in Table 8 and plotted against the number of days in culture.

229: parental cell-line.
○: pSV3gpt:pSV2neo transformant.
●: pSV3gpt:p4JJKan transformant.
■: pMTLT:pMTneo transformant.
first 35-40 days of the study. In contrast, the transformed cell-lines maintained their growth characteristics for 60-70 days. Therefore the useful lifespan in terms of days of consistent growth in culture is nearly doubled by transformation with SV40 Tag. The onset of crisis was taken to be the point at which the growth curve becomes shallower i.e. when the time to reach confluence increases. The end-point or senescence was taken to be the passage after which the cells failed to reach confluence. Towards the end of their life-span in culture, the cell-lines showed a decreased growth rate. It is likely that this was, in part, due to a decreased plating efficiency; however both features are characteristic of senescing cells. Figure 18 shows the proliferative histories of the parental cell-line and the shortest- and longest-lived transformed lines. The end-points of the other lines are also indicated. The life-span of the transformed lines is only slightly longer than that of the parental line. However, at senescence all the transformed lines, with the exception of p4JJkan:pSV3gpt:4 had undergone more than twice the number of doublings achieved by the 229 cells. This represents a vast increase in the proliferative capacity of the cells following transformation. At the end points the difference is \(10^{12}\) fold. Without this, detailed analysis of clones of sheep cells expressing stably integrated mMT-I fusion genes (see chapter 5) would not be possible.

The growth of 8 of the transformed cell-lines is remarkably similar, crisis and senescence occurring at approximately the same time. The number of population doublings was calculated using the apparent doubling time of the cells between passages. It should be noted that this may be an underestimate since it assumes 100% plating efficiencies. At the plating densities used it is unlikely that the parental or the transformed lines exhibit 100% plating efficiency. Since it is unlikely that the cells grow at a constant rate between passages, the doubling time is a mean value.

The culture conditions used in this study might be regarded as stringent since the re-plating density of the cells at each
passage was \(2.6 \times 10^3/\text{cm}^2\) for the transformed lines and \(5.3 \times 10^3/\text{cm}^2\) for the parental cell-line. These conditions may affect the time taken to reach crisis but not the end result i.e., senescence of the culture. Whilst the cells maintained their growth characteristics this translated into 4-6 days between passages. With prolonged culture all the cell-lines exhibited a change in phenotype. These changes included an increased resistance to dispersal with trypsin, increased doubling time, larger size and flatter morphology. In the case of the transformed lines the number of cells adhering initially and then rounding up and detaching increased at higher passage numbers. In this respect they differ from the parental cell-line which tends to settle into an attached but quiescent state rather than die when they approach terminal senescence.

Since the proliferative histories of the 9 transformed cell-lines are similar (Table 8, Figure 18). It can be concluded that the substitution of pSV2neo by p4KJJKan in the co-transfection has little effect upon the end result. One could highlight that the cell-lines bearing p4JJKan had slightly lower and more variable proliferative capacities than those expressing pSV2neo. However, in view of the dramatic reduction in colony formation observed when p4JJKan is substituted for pSV2neo in single transfections the slightly reduced longevity of transformed lines bearing p4JJKan may not be significant.

It appears that senescence of sheep skin fibroblasts is a reproducible phenomenon which is enormously delayed by the expression of SV40 Tag. One of the effects of SV40 Tag on cell proliferation is to stimulate cellular DNA synthesis (Wright et al., 1989; Chou & Martin, 1975 and Lumpkin et al., 1986). If, beyond the number of population doublings achieved by normal cells, the transformed lines are exclusively dependant upon Tag to stimulate DNA synthesis then a loss or reduction of Tag expression might be expected to precipitate the onset of crisis and senescence. Immunostaining shows that during the onset of crisis the cell-lines continue to synthesise SV40 Tag.
Cells transformed with plasmid pMTLT also have an extended but finite lifespan in culture.

The SV40 Tag represses transcription from its natural promoter, the SV40 early region promoter/enhancer (Levin & Manley, 1989; Stahl & Knippers, 1987). In all the transformed lines studied above Tag synthesis was under the control of the SV40 early region promoter/enhancer. Therefore, cell-lines derived from co-transfection of pMTLT and pMTneo were generated and passaged to senescence as before. The mMT-I promoter is inducible by zinc and therefore allows control of Tag expression, in particular the levels of Tag might be raised in an attempt to override cell crisis and the onset of senescence should Tag be the limiting factor which prevents immortalisation.

Five colonies were isolated, expanded into cell-lines and then passaged. The transfections were performed on the same batch of cells within a couple of days of the pSV3gpt transfections illustrated in Table 7. The results were therefore directly comparable. During selection of the transformants some of the cells were grown in medium supplemented with 50μM ZnSO₄ to induce expression of the mMT-I fusion genes. Only 2-3 colonies appeared on the plates grown in the absence of zinc and in general the colonies grew more slowly than their pSV3gpt transformed counterparts. The results of growth in zinc were more difficult to interpret. Although many cells survived the selection, few distinct colonies were evident. This was presumed to be due to secondary colonies arising from cells "spilling over" from a small number of primary colonies.

Clones were expanded and passaged in the absence of zinc. The proliferative histories of the lines are shown in figure 18 and Table 8. All five lines showed increased proliferative capacity relative to the parental cell-line and in general this capacity is similar to that of the pSV3gpt transformed lines. Two clones, MT2 and MT3Z, were, however, relatively short lived. Clone MT3 was lost
due to infection of the cultures. The expression and induction of Tag in these lines was confirmed by indirect immuno-staining.

Failure to synthesise SV40 Tag at high levels does not account for the senescence of transformed cell-lines

During crisis, the pMTLT transformed cell-lines still synthesize Tag and despite this continue to senesce. The photographs in Figure 19 illustrate this. Frame (A) shows clone MT3Z during the latter stages of crisis and approaching senescence. The cells are large, flat and irregularly shaped, some have obviously failed to divide. Despite this, most cells express Tag. Here, heterogeneity of expression between cells is marked. There is no obvious correlation between normal cell morphology and Tag expression. Ongoing Tag synthesis is demonstrated by the response to zinc induction (Frame B). During the early and in the latter stages of crisis the culture medium of some of the pMTLT cell-lines was supplemented with zinc to determine the effect of increasing Tag expression. No conclusive effect of zinc was detected on the three cell-lines studied (data not shown). During crisis two of the cell-lines were cultured in medium containing various concentrations of zinc. In neither case was a beneficial effect of zinc noticed, the progress of senescence was not obviously delayed. It would appear that Tag expression may be necessary but not sufficient for continued proliferation of the cells. Possibly dominant factors provoking senescence are in operation and these can override the growth promoting influence of SV40 Tag.

Clearly, MT-I promoter dependent Tag expression can extend the in vitro lifespan of 229 skin fibroblasts in a similar way to SV40 early region promoter dependent Tag expression. This indicates that there is not some obscure interaction between the SV40 promoter and Tag which might cause premature crisis. Reasons for the failure of SV40 Tag antigen to immortalise sheep skin cells are discussed later.
FIGURE 19

pMTLT TRANSFORMED CELL-LINES SYNTHESISE SV40 Tag DURING CRISIS

The photographs opposite represent the results of an anti-Tag indirect immuno-staining assay performed upon cells from clone MT3Z during the crisis period when the cell division is progressively balanced by increased cell death and failure to complete mitosis.


Panel B: MT3Z cells following 16 hours growth in medium supplemented with 100μM ZnSO₄.
Utility of SV40 transformed sheep cells

The proliferative capacity of fourteen SV40 early region transformed sheep cell clones have been analysed. With three exceptions (p4JJKan:pSV3gpt 4 and pMTLT:pMTneo 2 and 3Z) the growth and eventual senescence is quite consistent regardless of the promoter driving Tag expression and the nature of the foreign DNA integrant. Possibly this suggests that the level of Tag expression may not correlate with the degree of transformation. The ability to reproducibly extend the useful life-span of non-immortal cultured cells is valuable aside from its relevance to the study of senescent mechanisms and progression to neoplasia. The transformed cell-lines can be expanded rapidly to allow large stocks to be stored in liquid nitrogen. These frozen stocks remain viable and can be recovered with high efficiency. The increased proliferative capacity of the transformed cells is such that after 10 days in culture there is a 10 fold increase in the number of cells available for experimentation. The data presented in chapter 3 and 5 establish that endogenous and transfected metallothionein gene expression is not altered by SV40 transformation and therefore these cells offer an improved system in which to study the regulation at least of metallothionein genes. Since the cells have a far higher plating efficiency than their normal counterparts they will probably be more useful in stable transformation assays.

DISCUSSION: SV40 VIRUS MEDIATED TRANSFORMATION OF CULTURED CELLS

Simian virus 40 is a DNA tumour virus which can transform a variety of primary and established cultured cells and cause tumours in newborn hamsters (Tooze, 1980). The transforming sequences encoded by the early region are both necessary and sufficient for transformation (Tooze, 1980 for review). The transforming region encodes two proteins the large (Tag) and the small (tag) tumour antigens. The tumour antigens are encoded by overlapping genes. Around 95% of Tag is found in the nucleus the rest is located in the plasma membrane, tag is found in the cytoplasm (Lanford et al., 1985...
and refs therein). The transforming activity of SV40 has largely been assigned to the nuclear Tag though a cooperative role has been proposed for the plasma membrane fraction (Lanford et al., 1985). The tag is not essential for transformation but may assist in the transformation of quiescent cells (Seif & Martin, 1979).

Among the transforming oncogenes SV40 Tag is exceptional because it can achieve complete transformation of cells single-handedly (Lanford et al., 1985; Kriegler et al., 1984). In contrast, the related polyoma virus middle and large T antigens are both required and effect distinct steps in the transformation process. Polyoma large T can only immortalise the cells, middle T is required for full transformation (Rassoulzadegan et al., 1982). Two-step neoplastic conversion is a theme observed by other viral and cellular oncogenes (Land et al., 1983). It is not clear how the SV40 Tag is able to transform cells in one step but it is presumed to reflect the multi-functional nature of the protein and may require the cytoplasmic as well as the nuclear fraction of Tag (Lanford et al., 1985). During the SV40 virus life-cycle Tag is an absolute requirement for viral DNA replication and regulates the transcription of viral and cellular genes (Stahl & Knippers, 1987). Tag can also promote cellular DNA synthesis even in senescent cells (Lumpkin et al., 1986). It has recently been suggested that the unique transforming potency of the SV40 Tag depends upon its ability to form complexes with two cellular proteins p105-Retinoblastoma and p53 both of which are implicated in cell-cycle control (Jenkins & Sturzbecher, 1988; De Caprio et al., 1988). In contrast to SV40, the transforming ability of an adenovirus requires two viral proteins Ela and Elb which complex with p105-Rb and p53 respectively (Whyte et al., 1988; Sarnow et al., 1982).

The major factor determining the extent of transformation achieved by SV40 is the cell-type. In particular, cells of rodent origin are relatively easy to transform whereas transformation of human cells is very rare. This is true for any in vitro transforming treatments including exposure to ionising radiation and chemical
cancer, and the response to carcinogen challenge (DiPaolo, 1983). The frequency with
which primary cells escape cellular senescence shows the same species
dependence. Untreated, rodent cells become immortal with high
probability (Kraemer et al., 1986). Spontaneous immortalisation of
human diploid skin fibroblasts has not been reported (Wright et al.,
1989). Rodent cells are readily immortalised by chemical
carcinogens, a similar effect on human fibroblasts is rare
(McCormick & Maher, 1988). Several viral and cellular oncogenes can
immortalise rodent cells (Bishop, 1985) but fail to do so when
introduced into human diploid fibroblasts (Sager et al., 1983; Fry
et al., 1988; O'Brien et al., 1986).

Immortality, the escape from Senescence.

One of the steps towards cell-transformation is
immortalisation. Normal animal cells show a characteristic loss of
proliferative ability beyond a defined number of population
doublings. This is referred to as senescence. Since escape from
senescence is not observed in untreated normal human fibroblasts,
they provide a good system in which to study cellular senescence and
immortalisation.

The major loss of function associated with senescence is
the ability to synthesise DNA. Studies involving fusion of
proliferating and senescent cells have demonstrated that senescent
cells are able to inhibit DNA synthesis in the nucleus of
proliferation competent cells (for review see Smith, 1990). Similar
studies demonstrated that similar inhibition could be achieved in
cells immortalised by chemical carcinogen treatment but not in
immortal SV40 or polyoma transformed cells (Smith, 1990). This
suggested the existence of at least two classes of immortal cells
which differed in their ability to respond to the inhibitory
properties of senescent cells.
Early studies assaying the tumorigenicity of immortal: normal human cell-hybrids indicated that the immortal phenotype was dominant (Smith, 1990) however it was later shown that 70% of SV40 immortal: normal cell hybrids demonstrated severely limited lifespans and that the previous, contradictory data was probably due to selection for rapidly growing cell-hybrids (Periera-Smith & Smith, 1981). The generality of this phenomenon was confirmed with a number of hybrids involving the fusion of a variety of independently derived SV40 established cell-lines (Periera-Smith & Smith, 1983). By fusing a variety of immortal cell-lines with each other a series of complementation groups were identified which suggested that the immortalisation process related to a limited number of genes or processes. Interestingly, 9 out of 10 lines established by SV40 fell into the same complementation group, suggesting a common route to immortalisation following SV40 transformation (Periera-Smith & Smith, 1988). Using DNA synthesis to report senescent and immortalisation characteristics, RNA isolated from senescent human fibroblasts strongly inhibits DNA synthesis when micro-injected into pre-crisis human fibroblasts. On the other hand cellular DNA synthesis can be stimulated by the injection of SV40 Tag but not by expression of either the protooncogenic or oncogenic forms of c-Ha-ras despite the ability of the same DNA to stimulate cellular DNA synthesis in quiescent fibroblasts (Lumpkin et al., 1986). These results confirm the strict control of cell proliferation exercised by normal human cells.

Transformation of human skin fibroblasts by SV40 Tag: Parallels with sheep skin fibroblasts.

When either SV40 virus or plasmids encoding the small and large tumour antigens are introduced into human fibroblasts transformed clones or foci may be picked and sub-cultured. The frequency with which human diploid fibroblasts are transformed by SV40 sequences is around 10^-5 (0.001% efficiency, Sager et al., 1983). The transformed cells can be isolated as foci outgrowing the
monolayer, as colonies in soft agar or as drug resistant clones (Sager et al., 1983; O'Brien et al., 1986; Smith & Periera-Smith, 1990; Wright et al., 1989). Sheep cells transformed by the SV40 early region can be similarly isolated though at a considerably higher frequency (upto 0.1%). This may, in part, relate to the efficiency of electroporation over the techniques commonly used to transfect human cells (Donahue & Stein, 1988).

Typically, SV40 can extend the lifespan of normal human fibroblasts by around 20 population doublings (Wright et al., 1989). The life-span of clones of SV40 transformed sheep cells is 30-50 population doublings greater than that of non-transformed clones (see later). Both human and sheep transformed cells eventually succumb to crisis during which time cell division is progressively balanced by cell death. This is distinct from senescence observed in normal cells which do not die but simply fail to divide. This observation carries with it the strong implication that SV40 cannot immortalise normal human or sheep skin fibroblasts. However, occasionally foci of immortalised cells will emerge from a culture of SV40 transformed human skin cells in crisis. The frequency of this has been estimated, around half of SV40 transformed human cell clones can produce immortal clones at a frequency of approximately $3 \times 10^{-7}$ (Wright et al., 1989). During the course of the present work 17 SV40 transformed sheep skin cell-lines have been maintained until crisis. At the time they were discarded, none showed foci of immortal cells. However, given the frequency of immortalisation of similarly transformed human cells it is plausible that the number of cells retained in crisis was insufficient to guarantee detection of an immortalised clone. This seems a likely explanation for the failure to isolate immortal sheep cell clones if the event(s) that lead to the immortalisation occurs relatively late during the crisis. Therefore it seems probable that sheep and human cells behave in a similar way in response to the transforming properties of SV40. Low frequency immortalisation in SV40 transformed human cells does not necessarily alter the earlier conclusion that expression of the SV40 tumour antigens is insufficient but necessary
to immortalise the cells. The tumours antigens are necessary for in vitro immortalisation because immortal foci are not observed in senescent cultures of normal cells (Wright et al., 1989). Tag synthesis is insufficient to immortalise sheep skin cells because, in crisis they still synthesise Tag (see earlier) yet no immortalised clones were derived.

It has been demonstrated in human fibroblasts that Tag is required for proliferation of cells beyond both the senescent point of normal cells and subsequently of those immortal cells emerging during crisis of the transformed cells. Removal of Tag by switching to a non-permissive temperature in the case of a temperature sensitive Tag (Michalovitz et al., 1989) or removal of dexamethasone in the other case of Tag synthesis driven from a glucocorticoid responsive promoter (Wright et al., 1989), resulted in a reversible cellular senescence. A two-stage model has been proposed to explain the behaviour of SV40 transformed human fibroblasts and is also likely to apply to sheep fibroblasts. The model proposes two distinct mortality stages. The first causes senescence of normal cells and is overridden by SV40 Tag. The second mortality stage marks the onset of crisis of the transformed cells from which a very small number emerge as immortalised lines. The growth of these immortalised clones can be interrupted by the re-emergence of mortality stage 1 such as occurs when Tag synthesis is blocked.

The role of Tag in the immortalisation process is two-fold. Firstly, proliferation beyond mortality stage 1 is conditional upon its expression and secondly an indirect and unproven role in the generation of mutants able to survive mortality stage 2. The nature of the second role is unclear but is probably dependant upon the first role. The extended proliferative capacity of the cells following by-pass of mortality stage 1 would increase the probability of a fortuitous mutation or perhaps induce chromosomal aberrations thereby inactivating or counter-acting the gene(s) involved in mortality stage 2. This hypothesis can offer an explanation as to why human cells are relatively resistant to
chemical carcinogens since it appears to demand a statistically unlikely "two hit" mechanism. An extended prediction of the model being that rodent cells either lack or possess a less rigorous mortality stage 2 since SV40 Tag, myc, polyoma virus Tag transfected rodent cells are immortal upon isolation (Land et al., 1983; Bishop, 1985). The question arises whether any transfected oncogenes can override either of the proposed immortality stages and cooperate with SV40 Tag to induce immortality with high probability. The tumour associated cellular protein p53 which, in some cases can immortalise rodent cells, does not appear to fulfill any of these roles in sheep cells though clearly the effect of this gene is not neutral given the dramatic reduction in colony formation when introduced without Tag (see earlier). This would appear to contrast with experiments upon primary rodent cells in which p53 can enhance transformation by SV40 Tag (Michalovitz et al., 1986). The next section includes a preliminary investigation into the possibility that other viral or cellular oncogenes can also transform sheep skin fibroblasts.

ATTEMPTS TO TRANSFORM SHEEP SKIN FIBROBLASTS WITH OTHER VIRAL AND CELLULAR ONCOGENES

Numerous studies have confirmed that rodent cells are susceptible to transformation by a wide variety of oncogenes. The same is not true of normal human cells which are generally resistant to transformation by the same oncogenes. Human and sheep skin fibroblasts appear to be similarly affected by SV40 mediated transformation. They show some characteristics of transformed cells and their in vitro life-span is extended, but SV40 transformed clones are not immortal upon isolation. Immortalisation of SV40 transformed human cells occurs subsequently and with low frequency possibly as a result of mutation. It is likely, though not proven, that sheep cells would behave in the same way. In the light of this data it was of interest to see if the similarity between sheep and
human cells extends to their reaction to other viral and cellular oncogenes.

Five combinations of oncogenes were used, c-myc (plasmid pSV-c-myc-1, in which the myc gene is driven by the SV40 early region promoter), activated c-Ha-ras, bovine papilloma virus whole genomic fragment (plasmid pBVI), ras and myc and finally ras and p53. All these oncogenes either alone or in conjunction have a transforming effect upon primary rodent fibroblasts (Land et al., 1983; Jenkins et al., 1984; Bishop, 1985; Lambert et al., 1988; Land et al., 1986; Cerni et al., 1989). Plasmid pY3 which directs resistance to hygromycin was co-transfected when single oncogenes were used. The transfections were performed at the same time and on the same cells as used in the previous section to transform with SV40 sequences. Therefore the data were directly comparable and negative results are not due to failure of transfection or inability of the cells to respond to any proliferative stimuli. The results are shown in Table 9.

Previous data has shown that colony formation in cells transfected with plasmid p4JJkan, encoding neo r and murine p53, is very poor (see Table 7). This is clearly not improved by co-transfection with c-Ha-ras. The same combination can fully transform primary rodent cells (Jenkins et al., 1984). Those colonies surviving the selection were small, not vigorous and could not be cloned.

Colony formation with c-myc/pY3 and pBVI/pY3 was good and the cells appeared to possess a morphology distinct from normal 229 skin cells which indicated vigorous growth. Six colonies of each were cloned, four pBVI/pY3 and one myc/pY3 lines were established and passaged until senescence. Plates transfected with c-Ha-ras/pY3 were initially similar to those transfected with pPV1 or c-myc. However around 12 days after plating the colonies started to regress. The cells became rounded in appearance and eventually detached. Consequently, no colonies were counted or cloned. It is
**TABLE 9**

TRANSFECTION OF ADDITIONAL VIRAL AND CELLULAR ONCOGENES

<table>
<thead>
<tr>
<th>DNA</th>
<th>COLONIES</th>
<th>PLATES</th>
<th>AVERAGE</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-myc:pY3</td>
<td>63</td>
<td>3</td>
<td>21</td>
<td>0.0105 %*</td>
</tr>
<tr>
<td>Ha-ras:pY3</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pBV1:pY3</td>
<td>75</td>
<td>2</td>
<td>37.5</td>
<td>0.0189 %*</td>
</tr>
<tr>
<td>p4JJKan:Ha-ras 16</td>
<td>16</td>
<td>2</td>
<td>37.5</td>
<td>0.004 %*</td>
</tr>
<tr>
<td>Ha-ras:myc</td>
<td>N/A*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* 229 cells were electroporated in the presence of 20µg/ml linearised DNA (test plasmid). Plasmid pY3, conferring resistance to hygromycin, was included at 6µg/ml when required. Colonies selected in medium supplemented with 100µg/ml hygromycin for 14 days.

* 229 cells were electroporated in the presence of 13µg/ml of each linearised plasmid DNA. Colonies selected in medium supplemented with 400µg/ml G418 for 18 days.

* 229 cells were electroporated in the presence of 20µg/ml of each linearised plasmid. No selection. 3/6 plates maintained in medium containing reduced serum (5%) as opposed to 10%.
Plasmid p4JJKan (Jenkins et al., 1984) contains the neomycin phosphotransferase gene under the control of the SV40 early region promoter/enhancer, as in pSV2neo. Expression of a transforming murine p53 cDNA clone is directed by the Rous Sarcoma Virus long terminal repeat (RSV-LTR). For both neo and p53 transcripts, splicing and polyadenylation is directed by appropriate sequences from pSV2neo.

In plasmid pSV c-myc-1 (Land et al., 1983), expression of exons 2 and 3 of the cellular mouse myc gene is driven by the SV40 early region promoter/enhancer.

Plasmid pBPV-1 (Campo & Coggins, 1982) contains the whole BPV-1 genome cloned into the Hind III site of pAT153.

Plasmid pUCEJ 6.6 (Chiaho & Weinberg, 1982) contains the Ha-ras oncogene from the EJ/T24 human bladder carcinoma cell-line cloned into the Barn HI site of pUC13.

Plasmid pY3 (Blochinger & Diggelmann, 1984) contains the hygromycin B phosphotransferase gene 3' to Moloney Sarcoma virus enhancer elements. TATA box cap site and ribosome binding site and 5' to viral the polyadenylation signal.
<table>
<thead>
<tr>
<th>CELL-LINE</th>
<th>POPULATION DOUBLINGS AT SENESCENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSV2neo 1</td>
<td>19</td>
</tr>
<tr>
<td>pSV2neo 2</td>
<td>9</td>
</tr>
<tr>
<td>pSV2neo 3</td>
<td>14</td>
</tr>
<tr>
<td>Myc 1</td>
<td>31</td>
</tr>
<tr>
<td>RM 1</td>
<td>20</td>
</tr>
<tr>
<td>RM 2</td>
<td>15</td>
</tr>
<tr>
<td>RM 3</td>
<td>19</td>
</tr>
<tr>
<td>RM 5</td>
<td>19</td>
</tr>
<tr>
<td>RM 6</td>
<td>20</td>
</tr>
<tr>
<td>BPV a</td>
<td>19</td>
</tr>
<tr>
<td>BPV b</td>
<td>14</td>
</tr>
<tr>
<td>BPV c</td>
<td>14</td>
</tr>
<tr>
<td>BPV d</td>
<td>13</td>
</tr>
</tbody>
</table>

The cell-lines were passaged as described in the text. Between passages the cells were re-plated at $5.6 \times 10^3$ /cm$^2$.

**Calculation of Doubling Time Between Passages**

\[
T = \frac{d \cdot \log_2}{\log(H/P)}
\]

$T$ = doubling time

$d$ = time between passages

$H$ = number of cells harvested

$P$ = number of cells plated

Population doublings between passages $= d/T$. 
not clear why this should occur, possibly the ras protein is toxic to the cells. This behaviour does not compare to that of human skin fibroblasts which, though not always sensitive to ras mediated transformation, nonetheless form colonies (Sager et al., 1983).

The cells transfected with ras/myc were assayed for focus formation in medium containing either 10% or 5% foetal calf serum. A small number of patches of morphologically altered cells appeared on the plates maintained in 10% serum. These patches did not develop into foci as those transfected with SV40 were able to do (see earlier). Approximately one month after the transfection no true foci were visible though patches of morphologically altered cells were evident. Cells from seven of these patches were isolated using cloning rings, five survived and were expanded into cell-lines and propagated.

Three pSV2neo transfectant colonies (see Table 7) were cloned and propagated as controls. All the putatively transformed cell lines shared one characteristic, immediately after cloning they reverted to a normal morphology indistinguishable from the parental 229 skin fibroblasts and the pSV2neo controls. All the lines senesced within 20 population doublings with the exception of myc which persisted for around 30 generations (see Table 10).

The plasmid used to carry c-myc sequences pSV-c-myc 1 has been successfully used to immortalise rodent fibroblasts and cooperate with ras in full tumorigenic conversion of the same cells (Land et al., 1983 & 1986). The promoter in the plasmid is from the SV40 early region which is known to function in these sheep cells (see earlier). The proliferative capacity of the myc line is similar to that of the parental 229 fibroblasts plus the generations required during clonal expansion. Had several myc lines been established and shown similar proliferative capacities one might propose an effect of myc upon the proliferation of 229 skin cells. However, in this case it is not possible to reach such a conclusion. When DNA from the myc line was digested and subjected to southern
analysis using a c-myc probe the banding pattern was identical to that obtained with control 229 DNA. Therefore it is possible either that the line did not ever contain the transfected myc or had lost it at some stage. In this case one must assume that individual 229 cells do not possess equivalent proliferative capacities and that the myc line was the result of selection for one such cell.

The cell-lines derived from pBV1/pY3 co-transfection all senesced between 13 and 18 population doublings after being established (Table 10). This is within the range of the pSV2neo control lines and therefore suggests that the cells either have no BPV DNA or receive no proliferative advantage from BPV transforming sequences. DNA from all the cell-lines possessed sequences which hybridised to a BPV specific probe whereas 229 control DNA did not. Southern analysis of uncut genomic DNA revealed no evidence of episomal BPV DNA. The banding pattern in digested tracks was complex and consistent with there being multiple, single copy integrants (data not shown). It therefore seems that either the BPV DNA was not expressed or that expression did not lead to a recognisable phenotypic change. Bovine papilloma virus typically induces fibroepitheliomas in vivo, can readily transform established rodent fibroblast cell-lines (for review see Lambert et al., 1988) and leads to the delayed formation of tumours in transgenic mice (Lacey et al., 1986). Immortalisation of primary rat embryo fibroblasts by BPV has been demonstrated (Cerni et al., 1989). The frequency of immortalisation is high (50%), somewhat more efficient than c-myc in the same study. Three viral open reading frames are required for and appear to play distinct roles in the immortalisation though given the complex molecular biology of BPV it is not clear how many proteins are involved in the immortalisation process. Failure to replicate and transcribe BPV has been established to correlate with lack of transformation in a variety of cultured cells (Antmann & Sauer, 1982). Only a minority (2/11) of bovine skin cultures can be transformed by the virus (Antmann & Sauer, 1982). In some lines (notably not those of primate origin) brief treatment of infected cells with the tumour promoter (TPA) resulted in the initiation of
expression and DNA synthesis and in the case of mouse embryo fibroblasts caused tumorigenic conversion (Antmanr & Sauer, 1982). It is not known whether sheep skin cells carrying integrated BPV sequences can be similarly effected by TPA administration.

The proliferative capacities of the ras/myc derived cell-lines is also within the range set by the pSV2neo control lines. Southern analysis indicated that all the lines had ras sequences not present in the control track but none appeared to have any additional myc sequence. Northern analysis of these lines indicated that all the lines expressed very low levels of ras message above that in control tracks; one line, rm5, expressed rather more than the other lines (data not shown).

Different laboratories report varied findings when human skin fibroblasts are transfected with activated ras genes. In some cases no transformation is observed and the cells appear normal (Sager et al., 1983). In other instances transfected clones were initially transformed at high frequency but eventually lost the transformed phenotype and succumbed to senescence at a population doubling level similar to that of normal cells (Fry et al., 1988). The reversion of phenotype might explain the results presented here since clearly some kind of morphological alteration was apparent following transfection of ras/myc. Failure to observe clear foci may be related to the number of cell divisions achieved before the cells became confluent. If the time to reach confluence is short then expression of the transformed phenotype may not be possible or may even be suppressed by the presence of normal cells which at confluence are contact inhibited and therefore quiescent. It has been proposed that the ability of cells to express a transformed phenotype is influenced by nearby cells (Land et al., 1986). As highlighted (Land et al., 1986), removal of non-transfected cells by co-selection for drug resistance may allow transformed clones to emerge but this has little relevance to in vivo neoplastic conversion. One cause of the discrepancies between reports appears to be the plasmid construct used to direct ras expression, in
particular the level of expression seems to correlate with the transformed phenotype. By these criteria the ras encoding plasmid used in these studies is poorly expressed and does not transform human cells (Sager et al., 1983; Fry et al., 1988).

The results presented in this section, although preliminary do not indicate that any of the oncogenes used have a detectable influence on the proliferative capacity of 229 skin cells in the way that SV40 Tag is clearly able to. The study was not pursued thoroughly with respect to establishing integration and expression of the foreign genes, however there is no reason to suppose that the genes were not integrated in at least some of the transfected cells.

Summary: General conclusions regarding the transformation of sheep skin fibroblasts.

Only SV40 Tag expression caused indisputable transformation of sheep skin cells. This is in agreement with an earlier report in which sheep fibroblasts of unknown tissue of origin were transfected with the SV40 early region and pools of foci were passaged until senescence. The cells in this report survived for 42 passages though the corresponding number of population doublings was not mentioned (Stromberg et al., 1980). The reproducible senescence of the SV40 transformed sheep cell-lines described in this thesis and the failure of other oncogenes to elicit a detectable phenotypic change indicates that senescence is rigorously enforced in sheep skin cells. Therefore sheep skin fibroblasts display similar characteristics to fibroblasts derived from human skin and may be a useful alternative system in which to study multi-step in vitro tumorigenesis and cellular senescence.
USE OF ELECTROPORATION TO INTRODUCE DNA INTO SHEEP SKIN FIBROBLASTS AND OTHER CELL-LINES.

As mentioned earlier, the initial attempts to transflect sheep skin fibroblasts using the standard calcium phosphate technique (Graham & Van der Eb, 1973) were unsuccessful. Exposure to the calcium phosphate was toxic to the cells which then took some time to recover normal growth. No stable transfectants were obtained using plasmid DNAs which are effective when introduced by electroporation. At the same time the technique was effective with BHK cells excluding the possibility of some trivial problem with the technique. Two variants of the technique were also attempted. Strontium phosphate-DNA co-precipitate has been effective in transfecting cells to which calcium phosphate is toxic (Brash et al., 1987) or causes differentiation. This was effective in stably transfecting BHK cells but ineffective on 229 skin cells. The second modification replaces the buffer Hepes with BES at pH 6.95 and allows the precipitate to form slowly by incubation of the DNA with the cells in a 3% CO₂ atmosphere (Chen & Okayama, 1987). This is reported to offer a marked increase in the efficiency of stable transformation compared to the standard protocol but was also ineffective on 229 skin cells.

Transfection of cultured cells by electroporation.

Electroporation is held to be the most widely applicable method of transflecting DNA into cultured cells, a very large variety of cell-types have now been successfully electroporated (Andreason & Evans, 1988; Chu et al., 1987; Knutson & Yee, 1987). The most commonly used transfection technique, calcium phosphate-DNA co-precipitation is technically simple and often very efficient (Chen & Okayama, 1987). However several cells types are notoriously
refractory to transfection by this method, these cell-types include those grown in suspension eg lymphocytes, and some primary cell-lines (Kaspa et al., 1986; Potter et al., 1984; Toneguzzo., 1986; Boggs et al., 1986). The efficiency of other transfection techniques such as DEAE dextran (McCutchen & Payeno, 1968), which also depend upon chemical treatment of the cells, depends very much upon the cell-type and may be intolerably low in some cases. Electroporation is applicable to a wide range of cell types because it involves physical and not chemical manipulation of the cell. When placed in an electric-field the membrane integrity of the cells is disturbed. At an appropriate electric-field strength this disturbance takes the form of membrane pores through which molecules may pass into and out of the cell. These pores represent reversible disruptions of the membrane structure. When the electric-field is too low no or very few pores will be generated, when it is too high the membrane will be irreversibly damaged and the cell is killed. This is a simplistic picture since all three types of process are likely to occur at all voltages. Successful electroporation depends upon finding conditions where the largest number of cells are reversibly permeated. Several papers deal with adaptation of several different electroporation devices to suit many cell-types. The major, consistent theme to emerge is that optimal conditions for a given apparatus and cell-type must be determined empirically (Chu et al., 1987; Andreason & Evans, 1988; Knutson & Yee, 1987). The transfection efficiency by electroporation has variously been shown to be electric-field dependant, pulse shape and decay time dependant, temperature dependant. Biological parameters such as cell-type, electroporation media and topology and concentration of DNA may also influence the efficiency. Cheaper apparatus depending upon electrophoresis rated power supplies may show stochastic variation in efficiency owing to the non-reproducibility of the pulses generated (Andreason & Evans, 1988).

The electroporation used in this study, a replica of that described by Chu et al (1987), was kindly constructed by the Edinburgh University Genetics Department Workshop. The device is
FIGURE 20

CIRCUIT DIAGRAM OF THE ELECTROPORATION DEVICE

The device is a replica of that used by Chu et al (1987).

P: Power supply.
S1: Switch 1.
V: Voltmeter.
C: Capacitor (each 540μF rated to 450V).
S2: Switch 2.
E: Platinum electrodes.
### TABLE 11

**EXPRESSION OF HSV-tk ACTIVITY IN BHK tk- CELLS FOLLOWING ELECTROPORATION AT VARIOUS VOLTAGES**

<table>
<thead>
<tr>
<th>Cell extract</th>
<th>tk activity: pmol/min/mg protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30 minutes</td>
</tr>
<tr>
<td>BHK tk+</td>
<td></td>
</tr>
<tr>
<td>250 V</td>
<td>0.64</td>
</tr>
<tr>
<td>300 V</td>
<td>1.48</td>
</tr>
<tr>
<td>350 V</td>
<td>2.58</td>
</tr>
<tr>
<td>BHK tk-</td>
<td>0</td>
</tr>
</tbody>
</table>

BHK tk- were electroporated at various voltages in the presence of 20 μg/ml supercoiled plasmid pSVEtk (tk gene under the control of the SV40 early region enhancer). Cells were assayed for HSV-tk activity in the presence of TTP (see materials and methods) 52 hours following electroporation.

Viability was determined by plating a known number of cells following electroporation and counting them 72 hours later. Viability is expressed as a percentage of the number of control cells which had been plated just prior to the voltage shock.
<table>
<thead>
<tr>
<th>Voltage</th>
<th>Colonies</th>
<th>Plates</th>
<th>Average</th>
<th>Efficiency^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 V</td>
<td>1047</td>
<td>3</td>
<td>349</td>
<td>0.116 %</td>
</tr>
<tr>
<td>250 V</td>
<td>1424</td>
<td>3</td>
<td>475</td>
<td>0.16 %</td>
</tr>
<tr>
<td>300 V</td>
<td>553</td>
<td>3</td>
<td>184</td>
<td>0.06 %</td>
</tr>
<tr>
<td>350 V</td>
<td>122</td>
<td>3</td>
<td>41</td>
<td>0.0136 %</td>
</tr>
</tbody>
</table>

BHK tk- cells were electroporated in the presence of 20\(\mu\)g/ml linearised plasmid pSV3neo, counted and re-plated. 72 hours following electroporation the growth medium was supplemented with 400\(\mu\)g/ml G418. Selection was complete 10 days later at which time the number of colonies was counted.

^ Efficiency is expressed as the percentage of the cells plated which develop into colonies under selection.
simple and inexpensive, the major financial outlay being upon re-
useable platinum electrodes. The circuitry is illustrated in Figure
20. Operation (discussed in more detail in Materials and Methods)
involves charging the two capacitors (total capacitance 1080µF, each
capacitor is rated to 450V) using a standard electrophoresis rated
power-pack (Pharmacia 500V). The capacitors are discharged through
the cell:DNA suspension via platinum electrodes. Unlike many
commercially available devices the only parameter which can be
varied is the voltage. Essentially the voltage determines the
strength of the electric field (volts/cm). The voltage pulse
undergoes an exponential decay with a time constant equal to the
product of the capacitance and the resistance of the electroporation
buffer. The decay time is around 7 mssec (Chu et al., 1987) with
isotonic buffers. Since the capacitance is fixed and most studies
recommend the use of isotonic buffers (Chu et al., 1987; Andreason &
Evans, 1988), the decay time and the voltage cannot be altered
independently. When the potential to alter the pulse time exists it
can have a major influence upon the transfection efficiency
(Andreason & Evans, 1988) but must be determined empirically. The
major task then in standardising this device is to find an optimum
voltage for each cell-line to suit transient or stable transfection
assays.

Successful Electroporation of BHK tk- Cells.

Initial experiments following published guide-lines for the
apparatus (Chu et al., 1987) were performed on BHK tk- cells which
are easier to culture and more amenable to transfection than sheep
skin cells. Transient HSV-tk assays were used to monitor
transfection with plasmid pSVEtk as the reporter DNA. The results
are shown in Table 11. The HSV-tk activity per mg protein increases
with the voltage. This voltage dependant increase correlates with a
decrease in the cell viability. This probably means that the total
HSV-tk activity in all surviving cells is similar at all three
voltages but that the activity per viable cell increases with voltage.

A similar experiment was performed to optimise the voltage for stable transformation using linearised plasmid pSV3neo (encoding SV40 Tag and neo'). The results are shown in Table 12. The optimum is clearly 250 V and therefore not predicted by the transient tk assay. Since a transient assay optimum depends upon the percentage of expressing cells and the stable assay optimum upon the total number of expressing cells it is plausible that the voltage optimum will be different for the two assays.

These data and hypothesis are confirmed by a transient SV40 Tag indirect immuno-staining assay. Cells were electroporated with plasmid pSV3gpt which encodes the Tag under the control of the SV40 early region promoter, and processed for immuno-staining 48 hours later. The results are shown in Figure 21. The percentage of expressing cells is directly related to the voltage and hence inversely related to the cell viability. This explains the transient HSV-tk expression data. The total number of transfected cells is greater at 200 V and 250 V thus accounting for the stable transfection optimum.

This result indicates that a similar relation between transient HSV-tk expression and voltage arises because at an optimum voltage the percentage of expressing cells is maximal rather than the level of expression within individual cells. Although the level of expression within individual cells does vary there is no evidence that this varies between different voltages. This could be interpreted in two ways, firstly that the number of templates does not limit the expression and secondly that the number of DNA molecules entering a particular cell does not depend upon the voltage of the shock. As mentioned in Chapter 6 it is possible that the template is limiting the expression firstly because of the variation in expression observed between transfected cells on the same plate and secondly because calcium-phosphate transfection of
FIGURE 21

TRANSIENT EXPRESSION OF SV40 Tag IN BHK CELLS FOLLOWING ELECTROPORATION AT VARIOUS VOLTAGES

BHK cells were electroporated at various voltages in the presence of 20µg/ml supercoiled plasmid pSV3gpt. 48 hours following electroporation the cells were fixed and expression of SV40 Tag was detected by indirect immuno-staining.
BHK cells with the same plasmid leads to higher levels of staining in some of the cells. Notionally it is possible to suppose that the amount of template taken up by a cell is dependant upon the voltage. A relationship between the strength of the electric field (volts/cm) and the number and size of membrane pores is suggested by the inverse relation between cell viability and the electric field strength. Since the frequency and size of membrane pores will depend upon the strength of the electric field. Below a certain voltage an insignificant number of pores are generated because the electric field is too weak to disrupt the molecular interactions in the membrane. Extending this one might assume that as the electric field increases so will the number and size of the membrane pores thereby increasing the opportunity for DNA in the medium to enter the cell. When the electric field is too high the membranes of an unacceptable number of cells are irreversibly damaged. However, the data presented here suggest that if there is any increase in the number of templates molecules entering cells at higher voltages it is not sufficient to allow a large increase in expression. Perhaps this suggests that a very fine line is drawn between transfection and cell death. This is to say that a cell may tolerate a very limited number of membrane pores, below this no transfer of DNA occurs and above this the cell is rendered non-viable.

The variable level of expression between cells may also relate to the cell-cycle position of the cells either at the time of the assay or at the time of DNA entry into the nucleus. Since evidence suggests that once established, an active transcriptional complex is stable (Weintraub, 1988) it seems plausible that the state of the cell during the period of DNA entry may be critical to the establishment of transcription. There may be a period of the cell-cycle which is pre-disposed towards the assembly of active transcriptional complexes and that this is a factor which determines the level of expression within a given cell. In view of this it is interesting that Chu et al (1987) report a two to ten fold decrease in transient expression when fully confluent cells were used for the electroporation. Confluent cells are presumably quiescent and
blocked in the G₀ phase of the cell-cycle. It is not clear how much of the decrease in signal can be accounted by any loss of cell viability incurred during quiescence.

Transfection of Sheep Skin Cells by Electroporation

Attempts to transiently express HSV-tk encoding plasmids (including pMK) in sheep skin cells were unsuccessful, presumably because the percentage of expressing cells was low enough to render the HSV-tk activity below the level of detection amongst the high background of cellular tk.

Various assays were performed to pursue the issue. Initially nuclear DNA was prepared from electroporated cells during the period when expression would be transient and analysed for the presence of transfected sequences. This indicated that foreign DNA did enter the cells following electroporation (see later). Subsequently the transient immuno-staining and stable transformation assays were used to optimise the technique for the cells.

The results of the Tag immuno-staining assays are presented in Table 13 and illustrated in Figure 22. As with BHK cells the percentage of expressing cells is directly related to the voltage and inversely related to the cell viability. At 300V the cell survival is very low but positive cells are apparent. The cell survival at 250V and 200V is similar but the number of expressing cells at the higher voltage is perhaps 2-3 fold greater. The voltage optimum by these criteria is clearly 250V. Again the level of expression between positive cells is not consistent.

It was of topical and practical interest to ascertain the utility of this assay in predicting the stable transfection frequency. To this end, 229 skin cells were electroporated in the presence of linearised pSV3neo (Tag and neo⁺) and pY3 (hygro⁻), 15µg/ml and 3µg/ml respectively. Linearised DNAs give higher stable
FIGURE 22

TRANSIENT EXPRESSION OF SV40 Tag IN 229 SKIN CELLS.

229 skin cells were electroporated at 200 V, 250 V and 300 V in the presence of 20μg/ml supercoiled plasmid pSV3gpt. 48 hours following electroporation indirect immuno-staining was used to detect those cells expressing SV40 Tag.
TABLE 13

TRANSFECTION OF 229 SKIN CELLS BY ELECTROPORATION

<table>
<thead>
<tr>
<th>VOLTAGE</th>
<th>COLONIES</th>
<th>PLATES</th>
<th>AVERAGE</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 V</td>
<td>82</td>
<td>2</td>
<td>41</td>
<td>0.02%</td>
</tr>
<tr>
<td>225 V</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>250 V</td>
<td>130</td>
<td>2</td>
<td>65</td>
<td>0.0325</td>
</tr>
<tr>
<td>275 V</td>
<td>67</td>
<td>2</td>
<td>33.5</td>
<td>0.017</td>
</tr>
</tbody>
</table>

229 skin cells were electroporated in the presence of 15μg/ml linearised (EcoRI) pSV3neo (encoding neo<sup>r</sup> and SV40 Tag). 2 x 10<sup>6</sup> cells were seeded onto 6 cm plates and grown for 72 hours prior to selection in growth medium supplemented with 400 μg/ml G418. Colonies were counted eleven days after electroporation. Efficiency is expressed as the percentage of cells plated which developed into colonies.

<table>
<thead>
<tr>
<th>VOLTAGE</th>
<th>TOTAL NUMBER OF Tag EXPRESSING CELLS#</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 V</td>
<td>560</td>
</tr>
<tr>
<td>225 V</td>
<td>56</td>
</tr>
<tr>
<td>250 V</td>
<td>900</td>
</tr>
<tr>
<td>275 V</td>
<td>280</td>
</tr>
</tbody>
</table>

* 10<sup>6</sup> cells from the above electroporations were plated in each well of a 24-well culture dish. Cells were analysed for expression of SV40 Tag by indirect immuno-staining 48 hours later. The total number of positively staining cells in 15 fields of view was counted and the average multiplied by the number of fields of view per well of the culture dish.
transformation frequencies than circular DNAs, this is a common theme in gene transfer (Bishop & Smith, 1989). Some of the cells were set aside for immuno-staining and the rest for selection in medium supplemented with either G418 or hygromycin. The results of the immuno-staining and stable transfection assays are shown in Table 13. The correlation between the number of positively staining cells and stable colonies is fair, 250V is clearly the optimum. It is not clear why the 225V data interrupts the general trend but it is re-assuring to see that the poor stable transformation frequency at this voltage is predicted by the transient assay. This suggests that the immuno-staining assay is a rapid and informative assay of transfection efficiency. The optimum voltage for a transient enzyme or RNA assay may not be the same as the optimum for stable transformation since the former is dependant upon the percentage of expressing cells and the latter upon the total number of expressing cells. The immuno-staining assay permits a rapid evaluation of electroporation conditions to suit both assays.

Electroporation therefore meets the requirements of a stable transformation assay in sheep cells ie it allows the generation of several hundred stable colonies per voltage shock (Table 13) depending upon the strength of the promoter (Chapter 4, Table 4) and whether SV40 Tag is co-expressed with the selectable marker (Chapter 7, Table 7). Three primary sheep skin fibroblast cell-lines have been successfully transfected by electroporation. However several unsuccessful attempts to electroporate such cell-lines in the early stages of the work suggest that certain criteria are met particularly when working with non-established cell-lines. Although possibly all cell-lines can be electroporated successfully, the efficiency still varies with the cell-type and also the passage number of the cells if they are not immortal. In such cases it is advisable to use early passage cells of which stocks should be kept in liquid nitrogen. When required the cells should be recovered, amplified and assayed immediately by the immuno-staining method. The remainder of the cells should be used within a few days. Over a short period the cells may be transfected with consistent results.
this is particularly important when comparisons are to be made between transfection experiments (see Chapter 7 on *in vitro* transformation of sheep skin cells).

Recovery of DNA from Cells after Electroporation.

Initial attempts to monitor the transient expression of HSV-tk encoding plasmids in sheep cells were unsuccessful. Therefore to determine whether any DNA was entering the cell following electroporation, nuclear DNA was extracted from the cells 72 hours after electroporation (see materials and methods) and the foreign DNA detected by southern analysis. The results of such an assay are shown in Figure 23a. Sheep skin cells were compared to BHK cells and an immortal mouse liver derived cell-line, H3. The nuclear DNA was digested and equal amounts (20μg) loaded in each track. All cell-lines show a voltage dependant appearance of the foreign DNA, plasmid pUCEJ6.6 (Chiaho & Weinberg, 1982) which contains the activated Ha-ras gene. The probe was Ha ras specific and cross-hybridisation to the resident genes explains the presence of faint bands other than the expected 6.6 kb band and also indicates that tracks 6 and 11 are overloaded. Since equal amounts of DNA were loaded in each track, the strength of the signal is dependant upon the amount of foreign DNA per cell and therefore the result is consistent with the expression data discussed above ie that the percentage of transfected cells is directly related to the voltage.

In order to extend the relation between DNA recovery and expression, an SV40 origin containing plasmid (pSV2neo; Southern & Berg, 1982) was introduced into cos-7 cells which will support replication of any plasmid bearing an SV40 origin of replication. The use of such a system also allows confirmation that the recovered foreign DNA is intra-cellular and has not simply become trapped on or in the cell membrane during the electric shock and then carried through the nuclear isolation and DNA extraction procedures without actually entering the cell. Because the foreign DNA has been
propagated in a dam- bacterial host strain the Adenosine residues in the sequence GATC will have been methylated. When such DNA replicates in a mammalian cell the unmethylated A's are lost because mammalian cells are not able to methylate A residues. The restriction enzyme MboI only cleaves the GATC sequence when the A is unmethylated and will therefore distinguish replicated and non-replicated DNA. Cos-7 cells were electroporated with supercoiled plasmid pSV2neo at three voltages. Some of the cells were eventually selected in medium supplemented with G418 and the rest set aside for nuclear DNA preparation 72 hours following the electroporation. Each sample of nuclear DNA was divided in two, one half was analysed uncut the other digested with MboI. The results of the southern analysis of the recovered DNA are illustrated in Figure 23b. No colonies were obtained at either 250V or 300V. The reason is apparent from the southern analysis. Lanes 7 and 8 show the uncut and MboI cut 200V samples. Several times more DNA is recovered from cells electroporated at 200V than either 250V or 300V (lanes 9, 10, 11 and 12). The signals in the tracks represent the total amount of foreign DNA recovered and not the amount recovered per cell, this agrees with the stable transformation result. Therefore at 250V and 300V the number of viable, transfected cells is very much reduced compared to 200V. In all cases most of the recovered DNA can be cleaved to completion with MboI indicating that it has been replicated and was therefore intra-cellular. The uncut recovered samples show an increase in the proportion of the open-circular and linear forms compared to the input plasmid (lane 5). Lanes 1 and 2 represent nuclear DNA extracted from cells which were mixed with DNA but not voltage shocked. Lane 1 is spoilt but on longer exposure lane 2 (MboI digested) shows bands corresponding to the open circular, linear and MboI digested pSV2neo and resembles lanes 8, 10 and 12. This suggests that DNA can enter mammalian cells without any external encouragement and that electroporation may artificially enhance the natural passage of molecules across the plasma membrane. Control DNA from cells not exposed to the plasmid DNA is clearly negative (lanes 3 & 4).

-173-contd.
FIGURE 23

RECOVERY OF FOREIGN DNA FROM CULTURED CELLS FOLLOWING ELECTROPORATION

Panel A.

229 skin cells, BHK cells and a cell-line spontaneously established from primary mouse hepatocytes (H3) were electroporated in the presence of 20µg/ml supercoiled plasmid pUCEJ6.6 (Chiaho & Weinberg, 1982) and seeded into 9 cm culture dishes. 72 hours following electroporation nuclear DNA was prepared from the cells, digested with Bam HI and subjected to southern analysis using a probe specific for the Ha-ras insert from pUCEJ6.6. Control DNA was prepared from cells which had been mixed with DNA but not voltage shocked and grown for 72 hours. 20µg of nuclear DNA was loaded in each track.

Lane 1: 21pg Bam HI fragment from pUCEJ6.6.
Lane 2: 105pg as lane 1.
Lane 3: 210pg as lane 1.
Lane 4: 229 control DNA.
Lane 5: 229 DNA 200 V electroporation.
Lane 6: 229 DNA 250 V electroporation.
Lane 7: H3 control DNA.
Lane 8: H3 DNA 200 V electroporation.
Lane 9: H3 DNA 250 V electroporation.
Lane 10: BHK control DNA
Lane 11: BHK DNA 250 V electroporation.
Lane 12: BHK DNA 300 V electroporation.
FIGURE 23 continued

Panel B.

Cos-7 cells were electroporated in the presence of 20μg/ml supercoiled plasmid pSV2neo (Southern & Berg, 1982) plated and nuclear DNA prepared 72 hours later. Half (by volume) of each DNA sample was digested with MboI, the digested and undigested samples were subjected to southern analysis using a neomycin specific DNA probe.

Lane 1: uncut control DNA (from cells mixed with DNA but not voltage shocked)
Lane 2: MboI digest control DNA.
Lane 3: uncut cos-7 DNA.
Lane 4: MboI digest cos-7 DNA.
Lane 5: 100 pg, input form pSV2neo.
Lane 6: 100pg Sau 3A digest pSV2neo.
Lane 7: uncut DNA, 200 V electroporation.
Lane 8: MboI digest, 200 V electroporation.
Lane 9: uncut DNA, 250 V electroporation.
Lane 10: Mbo digest, 250 V electroporation.
Lane 11: Uncut DNA, 300 V electroporation.
Lane 12: MboI DNA, 300 V electroporation.
Lane 13: open circular (upper band), linear (middle), supercoiled (lower) pSV2neo.
Additional DNA recovery experiments were performed to establish the intra-cellular topological alterations occurring to linear and super-coiled DNAs; the results and their implications regarding the assembly of arrays of foreign DNA molecules are discussed (Bishop & Smith, 1989).
As outlined in the introduction and elsewhere in this thesis there are several reasons to explain the inactivity of the 229 insert. The results discussed in earlier chapters primarily deal with the elimination of certain possibilities namely the non-functional rearrangement of the foreign DNA, incompatibility of the mMT-I promoter and sheep cells and the inability of the 229 cells to synthesise the resident sMT genes. These studies suggest that the insert is not grossly re-arranged and that both the sheep MT genes and mMT-I fusion genes are active in 229 skin cells. It is suggested that unless the gene is mutated, the likely cause of the non-expression is a negative position effect such as demonstrated recently (Al-Shawi et al., 1990). This is only a prediction, formal proof requires recovery of the transgene by molecular cloning techniques and re-introduction into another region of the genome. The same procedure would allow the question of mutational inactivation to be addressed. Using probes derived from the regions flanking the foreign DNA in the clone, the nature of the chromosomal DNA which had received the foreign DNA might be investigated firstly by southern analysis followed by an option to clone the normal locus from non-transgenic sheep DNA. This type of analysis has proven useful in the dissection of the events leading to and occurring at the integration of foreign DNA into the chromosomes (Wilkie & Palmiter 1987).

Construction of a Random Genomic Library from 229 Skin Cells.

This was constructed according to standard procedures (see materials and methods) using a size fractionated Mbo I partial digest of 229 DNA as the donor and EMBL 301 as the recipient bacteriophage λ vector (Lathe et al., 1987). The ligation was
TABLE 14:

RANDOM GENOMIC LIBRARY CONSTRUCTED FROM 229 DNA

<table>
<thead>
<tr>
<th>STRAIN</th>
<th>GENOTYPE</th>
<th>PFU/LIGATION</th>
<th>PFU/µg DONOR DNA</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM259</td>
<td></td>
<td>2.8 x 10^6</td>
<td>4.3 x 10^6</td>
<td>12%</td>
</tr>
<tr>
<td>NM621</td>
<td>mcrA mcrB</td>
<td>2.35 x 10^6</td>
<td>3.6 x 10^6</td>
<td>100%</td>
</tr>
<tr>
<td>DL538</td>
<td>mcrA mcrB</td>
<td>1.86 x 10^6</td>
<td>2.84 x 10^6</td>
<td>79%</td>
</tr>
</tbody>
</table>

LIBRARY OF SIZE FRACTIONATED HINDIII FRAGMENTS ENRICHED FOR THE 229 INSERT

<table>
<thead>
<tr>
<th>STRAIN</th>
<th>GENOTYPE</th>
<th>PFU/LIGATION</th>
<th>PFU/µg DONOR DNA</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL358</td>
<td>mcrA mcrB</td>
<td>1.95 x 10^6</td>
<td>4.5 x 10^6</td>
<td>100%</td>
</tr>
<tr>
<td>NM259</td>
<td></td>
<td>3.5 x 10^5</td>
<td>8.16 x 10^5</td>
<td>18%</td>
</tr>
<tr>
<td>Q358</td>
<td></td>
<td>5.2 x 10^5</td>
<td>1.2 x 10^6</td>
<td>27%</td>
</tr>
</tbody>
</table>

Experimental details are given in Materials and Methods.
packaged and titred on three different host bacterial strains. The results are shown in Table 14. Strains NM621 and DL538 are both McrA⁻:McrB⁻. The Mcr (modified cytosine restriction) system of _E. coli_ restricts incoming DNA containing methyl-cytosine (Rayleigh et al., 1988) and has been shown to present a barrier to the representation of specific clones in genomic libraries constructed from DNA of mouse, human and _neurospora crassa_ sources (Rayleigh, 1988). Between 10 and 20 fold under representation has been reported when Mcr proficient hosts are used (Rayleigh, 1988 and refs therein). Whittacker et al (1988) report that the total yield of recombinant phage can be enhanced nearly 10 fold by using Mcr⁻ hosts. The Mcr system does not affect the recovery of vector genomes, therefore an increase in the recovery of plaques when using an Mcr⁻ host indicates that the clones are recombinant. This is confirmed by the results presented in Table 14. Host DL538 was used for large scale plating and screening since it is preferred over NM621 because of its _sbcC_ mutation which is relevant to the stabilisation of palindromic sequences (Whittaker et al., 1988). At least 1.6 x 10⁶ were screened using an HSV-tk specific DNA probe. No positive plaques were obtained.

**Attempt to clone a HindIII fragment spanning the 229 insert**

Figure 2, panel A shows a genomic southern blot hybridised to a probe specific for the pBR322 and _mMT-I_ sequences in pMK. A HindIII digest reveals a single band of around 16kb. The position of the 5' HindIII site is mapped using a BglII/HindIII digest. The BglII site lies at the splice between the _mMT-I_ promoter and the HSV-tk gene. A BglII digest gives a band of around 13·3kb, digestion of this band with HindIII reduces it to 10·4kb implying that the 5' HindIII site lies between the two BglII sites and allows the 3' HindIII site to be mapped. Therefore the foreign DNA insert in 229 is contained within a 16kb HindIII fragment with 4·7kb and 3kb of 5' and 3' sheep flanking sequences respectively. This fragment is suitable for cloning in the vector XL47 (Loenen & Brammer, 1980).
229 genomic DNA was digested to completion with HindIII and size fractionated on sucrose gradients. The fractions were blotted onto nylon membranes and hybridised to an HSV-tk specific probe. Those fractions containing the most HSK-tk positive sequence were pooled and ligated to purified HindIII λL47 arms (see materials and methods for details). The ligations were packaged and titred on three hosts, Q358 (Mcr⁺, natural host), NM621 and DL538 as before. The results are shown in Table 14. Around 7 x 10⁶ plaques were screened using an HSV-tk specific probe. No positive plaques were obtained.
MATERIALS AND METHODS

CELL-CULTURE.

Cells and media: Sheep skin fibroblasts and transformed daughter clones were maintained in medium 199 with Earle's salts (GIBCO or SIGMA) supplemented with 10% foetal bovine serum (GIBCO or nbl) and 10% tryptose phosphate broth (Difco) (unless otherwise stated). BHK cells, mouse L-cells and Cos-7 cells were grown in Dulbecco's modified eagle medium (GIBCO) supplemented with 10% (BHK cells and L-cells) or 5% (Cos-7 cells) foetal bovine serum.

Maintenance of cell-lines: Sheep skin fibroblast cell-lines were grown to confluence in 75 cm² falcon flasks (GIBCO or Costar), detached with trypsin-EDTA, counted and 4 x 10⁵ cells re-plated. SV40 Tag transformed daughter clones were passaged in a similar fashion except that 2 x 10⁵ cells were re-plated at each passage. Pre-crisis, under these culture conditions both the parent cell-line and the transformed daughter clones reached confluence in 5-7 days. The permanent cell-lines were routinely sub-cultured at 1:10 dilutions.

Cloning of cell-lines: Clones were isolated using stainless steel cloning rings and initially established in 24 well plates.

Assay for growth in soft agar: as described (Macpherson & Montagnier, 1964). Cells were plated in an 0·4% agar suspension.

Assay for plating efficiency: 500 or 1000 cells were seeded onto 90 mm culture dishes. After 7-10 days growth colonies were fixed in a 1:1 methanol:acetone mixture, GIEMSA stained and counted.
ELECTROPORATION.

The method is as described by Chu et al (1987). Subconfluent cells were harvested with trysin-EDTA, pelleted and resuspended in 1 x HeBS ( 20 mM Hepes pH 7·05, 137 mM NaCl, 5 mM KCl, 0·7 mM Na₂HPO₄, 6 mM dextrose) at a cell concentration of at least 3 x 10⁶ cells/ml. DNA contained in a small volume of water or T.E. (10 mM Tris-HCl pH 7·5, 1 mM EDTA ) was added to the desired concentration, typically 20 µg/ml. The cell:DNA suspension was then exposed to a single voltage pulse at room temperature, allowed to stand for 5-10 minutes, with 5 ml of warm culture medium, counted and plated as required. Counting after the shock was necessary because it was not possible to recover the cells without appreciable loss from the electroporation chamber. For stable expression 2 x 10⁹ cells were plated per 60 mm dish, selection was applied on the third day after electroporation. For transient SV40 Tag expression assays 1 x 10⁹ 229 skin cells were plated in each well of a 24-well plate (wells 15 mm diameter). For transient HSV thymidine kinase assays the cells were recovered from the electroporation chamber and divided between two 90 mm plates. Transient assays were performed after 48 hours.

The electroporation apparatus used was a replica of that used by Chu et al (1987) and is illustrated in Figure 20. With switch S1 closed the capacitors (each 540 µF and rated to 450 V giving a total capacitance of 1080 µF) were charged from a standard electrophoresis rated power supply ( rated to 500V). When the capacitors were fully charged switch S1 was opened and the capacitors discharged, by closing switch 2, across platinum electrodes E, which were immersed in the cell suspension. The electrodes were approximately 4 mm apart. Electroporation chambers were gas sterilized (City Hospital, Edinburgh) plastic cuvettes (Kartel). With the electrodes inserted the cuvettes hold approximately 0·75 ml of cell-suspension.
CALCIUM PHOSPHATE MEDIATED TRANSFECTION OF CULTURED CELLS (Graham & Van der Eb, 1973).

On the day before transfection exponentially growing cells were re-plated at \(10^4/cm^2\) on 60 mm or 90 mm culture dishes. The next day the medium was replaced (5 ml and 10 ml for 60 mm and 90 mm plates respectively) and allowed to equilibrate in the incubator. The DNA-calcium phosphate precipitate was prepared as follows; for a 60 mm plate DNA (upto 5 \(\mu\)g) and water were mixed to a volume of 225 \(\mu\)l, to this 25 \(\mu\)l of 2·5 M \(\text{CaCl}_2\) was added and mixed thoroughly. Then 250 \(\mu\)l of 2 x Hepes was added dropwise with mixing. For 90 mm plates the amount of DNA and precipitate added per dish was doubled. The precipitate was allowed to form for upto half an hour and then added slowly, with mixing, to the cells. The plates were returned to the incubator overnight (\(\approx\)16 hours). The precipitate was washed off with PBS (phosphate buffered saline) and fresh medium added. Selection for stable transfection was started 48 hours after initial exposure to the precipitate. Samples for transient expression assays were taken after 36 hours unless otherwise stated. Variations on this method using either strontium chloride instead of calcium chloride or substituting Hepes buffered saline for Hepes buffered saline were performed as described (Brash et al., 1987 and Chen & Okayama, 1987 respectively).

STABLE SELECTION REGIMES.

Cells expressing a functional neomycin phosphotransferase gene were selected by supplementing normal growth medium with 400 \(\mu\)g/ml Geneticin (G418 sulphate, GIBCO) (Southern & Berg, 1982). Control plates cleared in 10-21 days, during which period colonies were counted or cloned as required. Cells expressing hygromycin B phosphotransferase were selected in medium containing 100 \(\mu\)g/ml hygromycin B (Cablechem) (Blochlinger & Diggelman, 1984). Under these conditions control plates were clear after 7-10 days.
Thymidine kinase deficient cells transformed by HSV-tk genes were selected in HAT medium (GIBCO HAT supplement).

TRANSIENT ASSAY: ANTI-SV40 Tag INDIRECT IMMUNO-STAINING.

Indirect antibody staining of transfected cells was carried out as follows. Following electroporation cells were plated in 24-well plates and grown for 48 hours. After two PBS washes the cells were fixed for 5 minutes in a 1:1 acetone: methanol solution and then washed twice with PBS. Then 300 µl of a 1:10 dilution, in PBS, of anti-Tag monoclonal antibody pAB 419 (kind gift, John Jenkins, Marie Curie Research Institute) was added to the cells and incubated at 37°C for 1 hour. The first antibody was then removed and the cells were washed four times with PBS, allowing 2 minutes contact per wash. Then 300 µl of peroxidase conjugated anti-mouse IgG (SAPU, Scottish antibody production unit) diluted 1:50 in PBS plus 5% calf serum was added to the cells followed by incubation for further 1 hour at 37°C. The second antibody was then removed and the cells washed four times for 2 minutes each with PBS. Then 500 µl of the developing solution (2 µl of 30 vol. hydrogen peroxidase and 100 µl 0-dianisidine saturated ethanol in 10 ml PBS) was added. The stain was allowed to develop at room temperature for up to an hour. The plates were then washed with distilled water and photographed as required.

PREPARATION OF NUCLEAR DNA FROM TRANSIENTLY TRANSFECTED CULTURED CELLS.

After electroporation, cells were plated in 90 mm culture dishes and allowed to grow for 72 hours. Nuclear DNA was then prepared essentially as described (Burch & Weintraub, 1983). Cells were harvested by scraping into PBS and recovered by centrifugation. Nuclei were prepared by re-suspending the cell pellet in RSB (10 mM Tris-HCl pH 7.4; 10 mM NaCl; and 3 mM MgCl₂) containing 0.5%
Nonidet-P40. The nuclei were then washed twice in the same buffer, once in RSB alone and finally re-suspended in RSB. DNA was extracted from the nuclei by adding an equal volume of STOP solution (0.6 M NaCl; 20 mM Tris-HCl pH 7.4; 10 mM EDTA; 1% SDS), RNAase was then added to 50 μg/ml and the resulting lysate incubated for 30 minutes at 37 °C. At this point proteinase K was added to a final concentration of 200 μg/ml and the samples were then incubated overnight at 37 °C. The samples were then extracted with an equal volume of phenol:chloroform (1:1) and twice with an equal volume of chloroform. The DNA was then recovered by ethanol precipitation.

THYMIDINE KINASE ASSAYS.

These were performed as described (Al-Shawi et al., 1988). Growing cells were harvested by scraping into ice-cold PBS and washed in the same buffer. The cell pellet was re-suspended in 100-200 μl sonication mix (10 mM KCl; 2 mM MgCl₂; 10 mM Tris-HCl pH 7.5; 1 mM ATP; 10 mM NaF; 1 M e-amino-caproic acid) and placed in a sonicating water bath for 2 hours at 4 °C. Cell debris was pelleted at maximum speed in an eppendorf centrifuge for 15 minutes. The supernatant was transferred to a fresh tube and either used immediately in an assay or stored at -70°C until required. To 12.1 μl of the sonicate supernatant 72.8 μl of Assay mix (150 mM Tris-HCl pH 7.5; 10 mM ATP; 10 mM MgCl₂; 25 mM NaF; 10 mM β-mercaptoethanol) containing 12.1 μCi tritiated thymidine was added. The reaction was then monitored, usually at 30, 60 and 120 minute intervals, by spotting 25 μl of the reaction onto DE81 discs. The discs were then washed four times for 10 minutes each in 10 mM Tris-HCl pH 7.5 and then fixed in 95% ethanol. The discs were transferred to scintillation vials and dried in a vacuum oven. When cool, 4 ml of scintillation fluid was added and the samples were counted. When required to inhibit endogenous, cellular thymidine kinase activity 0.4 mM TTP was included in the assay mix (Jamieson et al., 1974). Thymidine kinase activities were expressed as pmol thymidine converted into thymidine monophosphate per minute per milligram of
protein. Protein was determined by the Bradford method (Bradford, 1976).

PREPARATION OF TOTAL CELLULAR RNA.

Total cellular RNA was prepared by one of two methods, firstly by the guanidine thiocyanate method described by Chirgwin et al (1979) utilizing a cesium chloride cushion and latterly by the RNAzol method (Chomczynski & Sacchi, 1987).

GEL ELECTROPHORESIS OF RNA.

RNA samples were denatured in 24% formamide: 5% formaldehyde: 10 mM NaHPO₄ pH 7.0 at 60°C for 5 minutes and loaded onto 1.8% agarose gels containing 6% formaldehyde: 10 mM NaHPO₄ pH 7.0 and electrophoresed at 0.17 V/cm/hour. The RNA was transferred to nylon membranes (Amersham Hybond-N) by capillary blotting from 20 x SSC and covalently cross-linked by u.v. (Church & Gilbert, 1984).

PREPARATION OF GENOMIC DNA FROM CULTURED CELLS.

Confluent cells were washed twice with PBS and then lysed by the addition of cell-lysis-buffer (0.1 M Tris-HCl pH 7.6; 0.1 M EDTA pH 8.0; 0.6% SDS) containing 50 µg/ml RNAase. The cell-lysate was then scraped into a sterile universal container and incubated for 30 minutes at 37°C. After this period proteinase K was added to a final concentration of 200 µg/ml, and the sample incubated overnight at 37°C. The sample was then gently extracted once with phenol:chloroform (1:1) and then twice with chloroform. The DNA was then dialysed extensively against TNE (10 mM Tris-HCl pH 7.5, 1 mM EDTA; 100 mM NaCl) and then TE (10 mM Tri-HCl pH 7.5; 1 mM EDTA). At this point the concentration was measured by spectrophotometry.
RESTRICTION ENZYMES.

These were purchased from Amersham, Boehringer-Mannheim or Pharmacia and used according to the manufacturers' specifications.

SOUTHERN BLOTTING.

Restriction digests were electrophoresed in an agarose gel of appropriate density (usually 0.8 % unless otherwise stated), transferred to nylon membranes (Amersham Hybond-N) by the technique of Southern (Southern, 1975) and covalently cross-linked to the membrane by u.v. (Church & Gilbert, 1984).

RANDOM PRIMED DNA PROBE SYNTHESIS.

This was performed using the method of Fienberg and Vogelstein (Feinberg & Vogelstein, 1983) which reliably yielded probes of $> 10^9$ cpm/µg specific activity.

SYNTHESIS OF RNA \textit{in vitro} AND PREPARATION OF STRAND-SPECIFIC RNA PROBES.

RNA probes of high specific activity were synthesised using T7 or SP6 RNA polymerase as described (Assubel et al., 1987). The following were mixed in a sterile microcentrifuge tube; 4 µl 5 x transcription buffer (200 mM Tris-HCl pH 8.0; 40 mM MgCl$_2$ (30 mM for SP6); 10 mM spermidine; 250 mM NaCl (no NaCl for SP6)), 1 µl 200 mM DTT, 2 µl 4 mM ATP, CTP, GTP, 5-10 µl ($\alpha$-$^{32}$P)UTP (10 mCi/ml, 800 Ci/mmol), 1 µl placental ribonuclease inhibitor (20-40 units), 1 µl 1 mg/ml template DNA, 1 µl RNA polymerase (5-10 units). To generate the target sense transcripts described in chapter 3 the labelled UTP was replaced by unlabelled UTP at the same concentration as the other ribonucleotides. To calculate the amount of target RNA
synthesised, the reaction was split following the addition of the polymerase with one half being added to a known, small amount of labelled UTP. The reaction was allowed to proceed for 30 minutes at 40 °C for SP6 RNA polymerase or 37 °C for T7 RNA polymerase. At this time the incorporation of [α-32P]UTP into TCA insoluble material was determined and the amount of RNA synthesised calculated. Although the efficiency of incorporation varied with different templates the specific activity of the RNA synthesised is around 10⁵/μg regardless. The template DNA was removed by adding 10 U RNAase free DNAase (Boehringer) and incubating for 15 minutes at 37 °C. 2 μl of 10 mg/ml tRNA as a carrier and water were added to bring the volume to 50 μl. This was extracted once with phenol:chloroform. The aqueous phase was added to 200 μl 2·5 M ammonium acetate and 750 μl ethanol. The RNA was precipitated for 15 minutes in dry ice/ethanol and recovered by spinning for 15 minutes in a microcentrifuge at 4 °C. The RNA pellet was resuspended in 50 μl water and precipitated as before. After a third ammonium acetate/ethanol precipitation the RNA pellet was washed with 75% ethanol/25% 0·1 M sodium acetate, pH 5·2. The probe was then ready to use. Generation of full-length transcripts was confirmed by denaturing polyacrylamide gel electrophoresis using end labelled DNA fragments as molecular weight markers.

5' DNA TERMINUS LABELLING (Phosphate exchange reaction). The method was taken directly from that accompanying the BRL 12 Kb DNA ladder. To a microcentrifuge tube the following were added, 14·μl H₂O, 2 μl DNA, 5 μl 5x exchange buffer (250 mM Imidazole pH 6·3; 60 mM MgCl₂; 75 mM β-mercaptoethanol; 1·5 mM ADP; 2·5 μM ATP; 0·5 mg/ml BSA), 10 μCi 3000 Ci/mmol [γ-32P]-ATP, 5-10 units T4 polynucleotide kinase, The reaction was incubated for 30 minutes at 37 °C, and the incorporation into TCA insoluble material determined.
HYBRIDISATION OF SOUTHERN AND NORTHERN BLOTS.

The method is that described by Church and Gilbert (1984) except that bovine serum albumin is omitted from both the pre-hybridisation and hybridisation steps. The washing protocol is also modified; firstly two, ten minute washes (500 ml each) at 68 °C in wash solution A (1mM Na$_2$EDTA, 40 mM NaHPO$_4$ pH 7.2, 5% SDS) followed by three, 20 minute washes (1 litre each) at 68 °C in wash solution B (1 mM Na$_2$EDTA, 40 mM NaHPO$_4$, 1% SDS). The stated volumes are adequate for 1-2 nylon filters, wash A is increased to 1 litre when two or more filters are being washed. Dry filters were autoradiographed on pre-flashed X-ray film at -70°C using intensifying screens.

PLASMID TRANSFECTIONS AND PREPARATION.

These were carried out as described (Bishop, 1979; Bishop & Davies, 1980). However for pUC based plasmids large scale preparations from unamplified cultures were performed by the alkaline lysis method (Maniatis et al., 1982). In both cases, plasmid DNA was purified by centrifugation to equilibrium in caesium chloride-ethidium bromide density gradients and then passed over a sepharose 2B (Pharmacia) column.

PURIFICATION OF DNA FRAGMENTS FROM AGAROSE GELS.

Fragments larger than 2 Kb were eluted onto dialysis membranes from agarose gels run in 1 x TBE pH 8.3 and purified using Elutips (Schleicher and Schuell). For smaller fragments separated by gel electrophoresis in 1 x Tris-acetate pH 7.9 a gel slice containing the DNA of interest was cut from the gel and the DNA purified by the geneclene method (BIO 101 inc.).
CONSTRUCTION OF pMTneo

Digest to completion with Bgl II and partially with Eco RI and gel purify the fragment containing pBR322, mMT-I and HSV-tk 3' sequences.

Insert the 1.7 kb Bgl II-Eco RI fragment from plasmid B1.2-polylneo containing the neomycin phosphotransferase gene coding sequence.

CONSTRUCTION OF pMTLT

Digest to completion with Bgl II and fill in the ends with E. Coli DNA polymerase klenow fragment.

Digest to completion with Bam HI and purify the 6 kb fragment containing the pBR322 and mMT-I promoter sequences. Ligate to the 2.6 kb Bam HI-Stu I fragment containing the SV40 early region encoding the large (Tag) and small (tag) prepared from pSV3gpt.
CONSTRUCTION OF HYBRID GENES.

Plasmid pMTneo was constructed as follows. Plasmid pMK was digested partially with Eco RI and to completion with Bgl II and the 6.5 Kb fragment containing pBR322, mouse metallothionien I promoter and HSV-tk 3' sequences gel purified. This fragment was ligated with a three-fold molar excess of the 1.7 Kb BglIII-EcoRI fragment from B1,2-polyIIIneo (John Bishop, unpublished) containing the neomycin phosphotransferase gene coding sequence.

Plasmid pMTLT was constructed as follows. Plasmid pMK was digested with Bgl II, the ends filled in with E. Coli DNA polymerase Klenow fragment and then digested with Bam H1. The 6 Kb fragment containing pBR322 and mouse metallothionein I promoter sequences was gel purified and ligated to the 2.6 Kb Bam H1-Stu I fragment from plasmid pSV3gpt (Mulligan & Berg, 1980) containing the SV40 early region coding sequences.

Ligations were performed as recommended (Maniatis et al., 1982) and used to transform competent E. Coli strain HB101. Colonies were selected for ampicillin resistance.

CLONING GENE-SPECIFIC SHEEP METALLOTHIONEIN SEQUENCES INTO T7 VECTORS.

Gene specific Hind III-EcoRI fragments were gel purified from SP64 or SP65 (Boehringer Mannheim) based plasmids (kind gift Julian Mercer, Murdoch Institute) and cloned into HindIII-EcoRI cleaved pTZ18r or pTZ19r (Mead et al., 1986). Ligations were used to transform competent XL1-Blue using X-gal and IPTG in the top agar for blue/white colour selection. White colonies were selected for ampicillin resistance.
CONSTRUCTION OF SHEEP GENOMIC LIBRARIES IN BACTERIOPHAGE λ VECTORS.

Random Library

Preparation of donor DNA.

Total, high molecular weight DNA was prepared from cultured, 229 fibroblasts as described above. The DNA was analysed on 0.3% agarose gels using intact phage λ and T4 DNAs as size markers. The genomic DNA migrated behind the λ band (≈ 50 Kb) but slightly ahead of T4 (≈165 Kb) and was therefore judged to be in excess of 100 Kb. Conditions were established for partial digestion with MboI (Pharmacia) as described (Maniatis et al., 1982) and the digest scaled up using the optimal conditions to digest 100μg of DNA. The scaled up digests were checked on a 0.3% gel against appropriate markers. The digests were extracted with phenol/chloroform, ethanol precipitated, washed with 70% ethanol and resuspended at 100μg/ml in TE. The partial digests were then fractionated on 10-40% sucrose gradients as described (Frischauf, 1987); 50μg was loaded onto a sucrose gradient in Beckman SW40 tubes and centrifuged at 25,000 rpm for 17 hours. Between 40 and 45, twelve drop fractions were collected from each tube and 20μl of every third fraction analysed on a 0.3% agarose gel. A second aliquot of all the fractions spanning the desired 15-22 Kb range were similarly analysed. Those fractions containing most of the fragments between 15 and 22 Kb were pooled, ethanol precipitated directly, washed with 70% ethanol and resuspended in TE at ≈200μg/ml. The size range of the pooled fractions was finally checked on a 0.3% agarose gel.

Preparation of Vector DNA.

The vector used was EMBL 301 (Lathe et al., 1987) which contains sites for the rare-cutting enzymes Not I and Sfi I in the poly-linkers. The "stuffer" fragment was enzymatically removed by
digestion of EMBL 301 with Bam HI and Eco RI and the larger DNA fragments selectively precipitated (Frischauf, 1987), as follows. Complete digestion was checked and then extracted with phenol/chloroform. To the aqueous phase 0.12 vol 3M sodium acetate (pH 6) and 0.6 vol. isopropanol were added. This was incubated on ice for 5 min. and the DNA recovered, (5 min. spin, microcentrifuge). The pellet was washed once with a 1:0.6 mixture of 0.45 M sodium acetate and isopropanol and once with a 1:2.2 mixture of 0.3 M sodium acetate and ethanol. This procedure removes the small polylinker fragments lying between the Bam HI and Eco RI sites. Efficient removal of functional stuffer fragment was assayed by ligating and packaging vector DNA in the presence and absence of added stuffer which had been gel purified from Bam HI cleaved EMBL 301.

Ligations were performed as described (Maniatis et al., 1982) using the recommended ratios and absolute amounts of donor and vector DNA. The ligations were packaged, titred and plated at ≈200 pfu/cm² on library plates. Three host strains were used NM259, NM612 and DL 538 (kind gift of Noreen Murray, Molecular Biology, Edinburgh). Commercial packaging extracts were obtained from Amersham, home-made packaging extracts were kindly provided by Dr. Jane Kinneird.

Non-Random Library

Preparation of donor DNA.

By southern analysis the whole transgenic sequence in 229 DNA was located on a 17 Kb Hind III fragment. Genomic DNA was digested to completion with Hind III and size fractionated on sucrose gradients as described above. A fifth of each fraction containing DNA between 30 and 10 Kb fragment was loaded onto a gel, blotted and the resulting filter probed with an HSV-tk coding sequence DNA probe. The two most strongly hybridising fractions were

-192-
pooled, ethanol precipitated, washed with 70% ethanol and resuspended at around 100μg/ml in T.E.

Preparation of vector.

The vector used was λ L47 (Loanen & Bramner, 1980). Stuffer fragment was physically removed using 10-40% sucrose gradients as described above (and Frischauf, 1987) except that centrifugation was at 26,000 rpm for 24 hours. To aid removal of stuffer fragment λL47 was digested with both Hind III and Sal I, Sal I reduces the stuffer fragment to two smaller fragments. Pooled fractions containing annealed arms were ethanol precipitated and resuspended at around 500μg/ml in TE.

Ligations and packagings were performed as for the random library. The packaging reactions were titred on hosts DL 538, NM 259 and Q 358.

Screening bacteriophage λ genomic libraries.

Duplicate lifts onto nylon membranes (Hybond-N Amersham) were taken and hybridized to an HSV-tk specific random primed probe. Hybridizations were performed in 1 x SSC (0.15 M NaCl, 0.015 M sodium citrate) containing 0.3% low-fat-dried milk and 50 μg of denatured salmon sperm DNA per ml. Washes were in 1 x SSC: 1% SDS. Controls for the lifts and hybridisations were either intact λL47 or EMBL 301 plaques (-Ve control) or HSV-tk positive phage mixed with negative phage at ≈1:20,000 dilution.

COMPOSITION OF SOLUTIONS NOT SPECIFIED IN THE TEXT

1 x SSC 0.15 M NaCl, 0.015 M sodium citrate.
1 x TA pH 7.9 0.05 M Tris-base, 0.02 M sodium acetate, 0.002 M EDTA, 0.01 M NaCl adjusted to pH 7.9 with glacial acetic acid.
1 x TBE pH 8.3 0.089 M Tris-base, 0.089 M boric acid, 0.002 M EDTA.
pTK1. The 3.5 kb HSV thymidine kinase gene (promoter, coding and RNA processing sequences) inserted at the Bam HI site of plasmid pAT153

Plasmid B1.2, polyIII neo contains the HSV-tk gene promoter and RNA processing sequences fused to the neomycin phosphotransferase gene. the Bgl II-Eco RI fragment containing the neomycin phosphotransferase gene and part of the HSV-tk RNA processing sequences was purified and used as described on page 190.
NaHPO₄ pH 7.0 83g/l Na₂HPO₄, 62g/l NaH₂PO₄·2H₂O
NaHPO₄ pH 7.2 70.6 g/l Na₂HPO₄, 1.2% v/v H₂PO₄.

BACTERIAL STRAINS.

HB101: hsdS20 (rB⁻, mB⁻), supE44, ara14, λ⁻, galk2, lacY1, proA2, resL20, xyl-5, mtl-1, recA13. Restriction: (rₚ⁻, mₚ⁻), mcrA (+), mcrB (⁻).

XL1-Blue: recA1, andA1, gyrA96, this, hsdR17 (rₚ⁻, mₚ⁺), supE44, relA1, λ⁻, (lac), (F', proAB, lacI₉ ZΔM15, Tn10 (tetR)).

NM621: recD1009, thy⁻, hsdR, mcrA⁻, mcrB⁻, supE44, thyA.


NM259: hsdR514, mcrA⁻, supE44, sufF.

RECOMBINANT PLASMIDS USED.

pSV2neo (Southern & Berg, 1982)
pSV3neo (Southern & Berg, 1982)
pSV3gpt (Mulligan & Berg, 1980)
PY3 (Blochlinger & Diggleman, 1984)
pUCEJ6-6 (Chiho & Weinberg, 1982)
pSV-c-myc 1 (Land et al., 1983)
p4JJKan (Jenkins et al., 1984)
pBV1 (BPV-1 complete genome in pAT153, Campo & Coggins, 1982)
B1.2. polyIIIneo (John Bishop, unpublished.)
pMTLT this thesis.
pMTneo this thesis.
pTK1 (HSV-tk gene in pAT153, Spandidos & Paul, 1982)
REFERENCES


Ivarie, R. D., B. S. Schacter, and P. H. O'Farrell. 1983. The level of expression of the rat growth hormone gene in liver tumor cells is -202-
at least 8 orders of magnitude less than that in anterior pituitary cells. Mol. Cell. Biol. 3:1460-1467.


-203-


-209-


-211-


Selig, H., and R.G. Martin. 1979. Simian virus 40 small t antigen is not required for the maintenance of tranformation but may act as a promoter (co-carcinogen) during establishment of transformation in resting rat cells. J. Virol. 32:979-988.


-213-


Additional references omitted in error.


### Abbreviations used in Text

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADA</td>
<td>adenosine deaminase</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine 5' triphosphate</td>
</tr>
<tr>
<td>b</td>
<td>bovine</td>
</tr>
<tr>
<td>BHK cells</td>
<td>baby hamster kidney cells</td>
</tr>
<tr>
<td>BPV</td>
<td>bovine papilloma virus</td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyl transferase</td>
</tr>
<tr>
<td>CHO cells</td>
<td>chinese hamster ovary cells</td>
</tr>
<tr>
<td>cpm</td>
<td>counts per minute</td>
</tr>
<tr>
<td>DCR</td>
<td>dominant control region</td>
</tr>
<tr>
<td>DNAase I</td>
<td>deoxyribonuclease</td>
</tr>
<tr>
<td>EC cells</td>
<td>embryonal carcinoma cells</td>
</tr>
<tr>
<td>EDTA</td>
<td>diaminooethane tetra-acetic acid</td>
</tr>
<tr>
<td>FSF</td>
<td>foetal skin fibroblasts</td>
</tr>
<tr>
<td>h</td>
<td>human</td>
</tr>
<tr>
<td>HAT</td>
<td>hypoxanthine, aminopterin, thymidine</td>
</tr>
<tr>
<td>HPRT</td>
<td>hypoxanthine phosphoribosyl transferase</td>
</tr>
<tr>
<td>HSV-tk</td>
<td>herpes simplex virus thymidine kinase</td>
</tr>
<tr>
<td>HTF</td>
<td>HpaII tiny fragments</td>
</tr>
<tr>
<td>kb</td>
<td>kilo-base pair</td>
</tr>
<tr>
<td>m</td>
<td>murine</td>
</tr>
<tr>
<td>MEL cells</td>
<td>mouse erythroleukaemia cells</td>
</tr>
<tr>
<td>MFI</td>
<td>methylation free islands</td>
</tr>
<tr>
<td>µF</td>
<td>micro-farads</td>
</tr>
<tr>
<td>MRE</td>
<td>metal regulatory element</td>
</tr>
<tr>
<td>MT</td>
<td>metallothionein</td>
</tr>
<tr>
<td>neo'</td>
<td>neomycin phosphotransferase gene</td>
</tr>
<tr>
<td>p</td>
<td>porcine</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming units</td>
</tr>
<tr>
<td>r</td>
<td>rat</td>
</tr>
<tr>
<td>s</td>
<td>sheep</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SHS</td>
<td>super hypersensitive site</td>
</tr>
<tr>
<td>SV40</td>
<td>simian virus 40</td>
</tr>
</tbody>
</table>
Tag  
  SV40 large tumour antigen

tag  
  SV40 small tumour antigen

TCA  
  trichloroacetic acid

ts  
  temperature sensitive

TPA  
  12-O-Tetradecanoylphorbol-13-acetate

TTP  
  thymidine triphosphate

V  
  volts
Mechanism of Chromosomal Integration of Microinjected DNA

JOHN O. BISHOP AND PAUL SMITH
Mechanism of Chromosomal Integration of Microinjected DNA

JOHN O. BISHOP AND PAUL SMITH

Department of Genetics, University of Edinburgh
Edinburgh EH9 3JN, Scotland

(Received 2 May 1989, and in revised form 28 August 1989)

Summary. Most transgenic animals have been produced by directly injecting DNA into one of the embryo pronuclei in the period immediately following fertilization. Transgenic animals are produced when DNA becomes integrated into the chromosomes; in most cases the transgenic genotype is transmitted to progeny through the germ line. The characteristics of the foreign DNA include tandem arrangement of multiple copies of the input DNA, predominantly in direct rather than inverted orientation, and illegitimate recombination with the chromosome at the site of integration. Foreign DNA integrated into chromosomes of cultured cells has identical characteristics. We argue that these and other similarities between the integration of foreign DNA into the chromosomes of microinjected embryos on the one hand and on the other of transfected and microinjected cells, strongly suggest that the processes have the same basis. By considering mainly the literature relating to cell transfection and microinjection, we conclude that tandemly arranged concatemers of input foreign DNA are built up by a process of homologous but non-conservative recombination. End-joining and illegitimate recombination events characterize the integration of the concatemers into the chromosome, and also contribute to the formation of the concatemers. We also suggest that these superficially different processes are based on the same opportunistic repair-ligation mechanism; the frequencies of the different types of event reflect both the frequencies with which different sorts of association occur between DNA molecules and their relative stability.

Introduction

Most transgenic animals have been made by directly injecting DNA into one of the pronuclei of a fertilized egg. Other techniques, such as the manipulation of embryonic stem cells and methods based on the use of retroviral vectors, are likely to become widely used in the future. However, the relative simplicity and high success rate of the microinjection method will ensure its continuing use.

When DNA is introduced by direct microinjection the transgenic offspring are found to have the foreign DNA integrated at only one, or rarely at a few chromosomal sites. As far as is known there are no preferred sites for insertion. The number of copies of the input DNA present at a given site may be small or large, and in most cases it probably remains essentially unchanged through generations of hereditary transmission. Here we refer to a set of DNA sequences inserted at a single site as an array. When several transgenic mice are produced by introducing the same DNA fragment, the outcome in terms of gene expression is not uniform. Invariably some of the mice show no expression at all, and their offspring are also expressionally inert (see examples in Palmiter & Brinster, 1986). The proportion of G0 mice which do not express the foreign gene varies greatly from experiment to experiment, and it is not clear to what extent this relates to the properties of the DNA introduced or to the procedures followed (method of DNA preparation, post-fertilization time of injection, mouse strains, etc.). Again, if several expressing lines that contain the same construct are compared, the level of expression in a given tissue varies greatly from line to line (Davis & MacDonald, 1988; Al-Shawi et al., 1988). In addition, the relative levels of expression in different tissues may vary from line to line (Al-Shawi et al., 1988). It is clear that the copy-number does not determine the level of expression. Indeed, copy-number and expression level are virtually uncorrelated in most cases (Palmiter et al., 1982; Overbeek et al., 1986; Davis & MacDonald, 1988). The remaining possible causes of differences in expression are (1) the arrangement of the genes in the array, (2) mutation or rearrangement (for example by illegitimate recombination) of the genes and (3) the possible influence of the chromosomal region around the inserted array, the so-called position effect. As yet
no causal connection has been established between any of these and expression.

Review of experimental data

A TRANSGENIC LINE IS THE RESULT OF A UNIQUE SET OF MOLECULAR EVENTS

It is important to recognize that each transgenic line, consisting as it does of a single G0 mouse and its offspring, is the product of a small set of unique events at the molecular level. These are the events by which a short or long concatemer of incoming DNA fragments is first built up, and then inserted into a site on one of the chromosomes. Of course it is quite possible that many concatemers are built within each embryo; but certainly only one (or a very few) of these is inserted into the chromosomes permanently. We know this because a G0 mouse typically contains a single chromosomal site of insertion (Gordon & Ruddle, 1985). We know too that the molecular events mainly occur before or during the period of DNA synthesis that precedes the breakdown of the pronuclei and the first post-fertilization mitosis (that is, within about 12 h) because about 70% of G0 transgenics are not mosaic (Wilkie et al., 1986).

MULTIPLE COPIES ARE ARRANGED IN DIRECT TANDEM ARRAYS AT THE SITE OF INSERTION

In most cases all the DNA molecules injected into a given pronucleus are identical. In principle, concatemers can be formed in two ways, by the joining end-to-end of linear molecules or by homologous recombination between molecules. Random end-to-end joining should produce head-to-head, tail-to-tail and head-to-tail associations in a ratio of 1:1:2. Homologous recombination, on the other hand, would produce only the head-to-tail arrangement. Analysis of arrays in transgenic animals shows that most adjacent sequences are arranged head-to-tail (Gordon & Ruddle, 1985), leading to the conclusion that arrays are largely built up by homologous recombination. However, the very existence of rare tail-to-tail and head-to-head arrangements shows that end-to-end joining also occurs, and that end-joined molecules contribute to some arrays. Putative illegitimate recombination events also occur, although relatively rarely (see below).

CONCATEMER FORMATION AND CHROMOSOMAL INSERTION OCCUR SEPARATELY

The frequency of homologous recombination between foreign DNA and similar or identical chromosomal sequences in the embryo pronucleus is not known, but by extrapolation from the frequencies observed in injected cells (Thomas et al., 1986; Thomas & Capecchi, 1987; Mansour et al., 1988; Joyner et al., 1989; Zimmer & Gruss, 1989; and see below) we presume it to be very low. On the basis of this assumption we discount models that suppose that arrays of foreign DNA molecules are built up after the insertion of one of them into the chromosome, by recombination between inserted and extrachromosomal copies. This effectively leaves us with two distinct processes, which together produce a concatemeric insertion; concatemer formation and chromosomal insertion.

TRANSFECTED CELLS AS A MODEL FOR INJECTED EMBRYOS

DNA can be introduced into cultured cells in a variety of ways. The most widely used methods have been calcium phosphate co-precipitation, treatment with DEAE-dextran, direct microinjection and latterly electroporation. One outcome of such procedures is the integration of the foreign DNA into the chromosome. Single cells that have integrated the foreign DNA generate monoclonal colonies that can be recovered and propagated. The arrangement of the foreign DNA integrated into the chromosomes of these clones of cells is so similar to the arrangements found in transgenic animals as to suggest that the underlying molecular mechanisms are essentially the same in cultured cells and embryos.

The more striking similarities are as follows. (1) Integration occurs at one or a few chromosomal sites per nucleus (Robins et al., 1981; Huttner et al., 1981; Folger et al., 1982). (2) The frequency of non-homologous integration in microinjected cells and microinjected embryos is similar, from 10% to 30% of cells (Folger et al., 1982, 1985; Thomas et al., 1986) and 15% of embryos (Brinster et al., 1985). (3) Integration occurs more frequently when the molecules introduced are linear than when they are circular (Folger et al., 1982; Chu et al., 1987; Brinster et al., 1985; but note that with microinjection into cultured cells the efficiencies are the same at higher DNA input (Folger et al., 1982)). (4) Integration sites frequently contain concatemers of the incoming DNA. In cellular transfection experiments it has
commonly been the practice to mix the DNA of interest, which often carries a selective marker, with a carrier DNA obtained from a higher eukaryote source. In these cases the marker DNA is found interspersed with carrier DNA in the chromosomal arrays (Perucho et al., 1980). More rarely, transfection of cloned DNA has been carried out without carrier. In such cases tandem arrays of the foreign DNA sequence are found at the sites of integration (Huttner et al., 1981). Furthermore, as in transgenic animals (Brinster et al., 1981), the monomers in the arrays are predominantly arranged head-to-tail, showing that they are built up by some sort of homologous recombination event (Folger et al., 1982). Head-to-head and tail-to-tail neighbours are also found infrequently in transfected cells, showing that end-joining occurs in this situation too, but again makes only a small contribution to the assembly of the arrays. When the foreign DNA is microinjected into the nuclei of cultured cells (also without added DNA carrier) the integrated arrays show the same characteristics (Folger et al., 1982). Microinjection permits a measure of control over the amount of input DNA, allowing a demonstration that more end-joining occurs with greater DNA input (Folger et al., 1982). In both cells and embryos integration of arrays into the chromosomes is mainly by illegitimate (non-homologous) recombination (Anderson et al., 1984; Wilkie & Palmiter, 1987).

The conclusion that the same processes occur in the two cell types (cultured fibroblasts and embryos) is not really surprising. The microinjected embryo pronucleus has entered or is about to enter S phase, embarking on a series of divisions with about the same time interval between them as in cultured fibroblasts. The main known differences are that the embryo male-derived pronucleus is still haploid and its chromosomal DNA may be more highly methylated. Thus, the relevant enzymic complement of the embryo and fibroblast nuclei may be effectively the same. Because of the greater availability and manageability of cultured cells, they have been extensively used in experiments designed to elucidate the molecular mechanisms of extrachromosomal recombination. We describe below conclusions drawn from some of these experiments. Our thesis is that these conclusions may be equally applicable to microinjected embryos.

**Extrachromosomal homologous recombination**

**ARRAY FORMATION BY HOMOLOGOUS RECOMBINATION MUST INVOLVE CIRCULAR OR CIRCULARLY PERMUTED DNA MOLECULES**

Consider the introduction into the nucleus of a cloned DNA fragment or a linearized plasmid. All of the molecules introduced are the same at the outset. Homologous recombination between such molecules can reassert their equivalent parts, but cannot generate a direct tandem array, or even a partial dimer (Fig. 1(b)). If the ends of some of the linear molecules become joined to form circles, array formation becomes possible by homologous recombination between a circle and a linear molecule or between two circles (Fig. 1(c)). Whether this occurs in practice depends of course on the mechanism of recombination; at this stage we are concerned only with formal geometrical considerations. If circles, once formed, are broken at random by nuclease action, circularly permuted linear molecules are generated. Depending
on the mechanism of homologous recombination these circularly permuted linear molecules may form arrays by recombination with each other, with circular molecules, or with the input linear molecules (Fig. 1(d)). If circular molecules were more extensively degraded by random cleavage to give fragments of circularly permuted DNA molecules, and if these fragments engaged in successive rounds of homologous recombination, the result would be a tandem array of perfect repeats of the linear monomer; the only perceptible damage within the array would be any damage to the ends of the input linear molecules that occurred before they were circularized.

There is evidence that recombination between extrachromosomal molecules occurs up to about an hour after microinjection into cultured cells (Folger et al., 1985). Depending on the relative rates of circularization and breakage, this time limit might place constraints on the outcome observed with different forms of input DNA. For instance, assume that circularization occurs, that the random breakage of circles is slow, and that at least one of the molecules involved in a recombination event must be linear. Under these circumstances linear input DNA might appear to recombine more efficiently than circular input DNA. Circular input DNA will generate very few linear molecules with which to recombine, whereas with linear input DNA the molecules that have not become circular will be available to recombine with those that have recircularized. To take a second example, if circularization and random breakage both occurred slowly, and if homologous recombination were confined to pairs of linear molecules, then once again linear input DNA might generate more concatemers than circular input DNA.

In this case, when the input DNA is linear a large pool of linear molecules will be available to recombine with the small number of circularly permuted linear molecules that are produced by circularization followed by random breakage. When the input DNA is circular the pool of linear molecules will be small.

METHODS USED TO STUDY EXTRACHROMOSOMAL RECOMBINATION IN TRANSFECTED CELLS

Extrachromosomal recombination has usually been approached in one of two ways. (1) By introducing a mixture of two DNA molecules into cells (co-transfection). Both molecules carry a copy of a gene that confers a property that can be scored or selected for. The two copies are defective in different ways, so that a non-defective copy may be generated by recombination. (2) Alternatively, the two defective copies are part of a single DNA molecule, and non-defective copies may be generated by intramolecular recombination. Several different systems have been used, each of them imposing different limitations on the range of results that may be obtained. Wilson and his collaborators transfected cells with two mutant forms of simian virus 40 (SV40) or with SV40 DNA fragments, and scored plaque formation after plating (Wake et al., 1984, 1985; Roth et al., 1985). In these experiments only circular recombinant molecules that can be packaged to form viable SV40 virions are scored. Other authors used plasmids and DNA fragments that carry defective (mutant or deleted) neomycin phosphotransferase or herpesvirus thymidine kinase genes and scored colonies of G418-resistant cells or cells transformed to growth in HAT medium (Lin et al., 1984, 1987; Folger et al., 1982, 1985; Anderson & Eliason, 1986). These colonies arise from cells that have integrated recombinant molecules into a chromosomal site. Plasmid rescue has in some cases been used subsequently to isolate parts of the integrated arrays.

LINEAR MOLECULES ARE CIRCULARIZED BY INTRACELLULAR LIGATION

Monkey cells transfected with linear SV40 DNA produce plaques. Since both DNA replication and packaging into the virion require circular molecules, it is clear that circularization must have occurred. When the DNA is cleaved within a region essential for virus multiplication, it is likely that the circularization occurred by the ligation together of the ends of the molecule. Thus, Wake et al. (1984) found that the efficiency of plaque formation with input linearized SV40 DNA was 10 to 20% of that of supercoiled circular DNA and 50 to 100% of that of nicked circular DNA. Introducing a duplicated region into the Tag intron allowed linear SV40 DNA molecules that carried a small terminal duplication within an inessential region (the intron) to be derived. Because the termini were in an inessential region, the different structures produced when the molecules were circularized in different ways could be tolerated. Among the circular molecules recovered after transfection with this DNA, 12% were perfectly joined, 74% had suffered the deletion of a duplicated terminal dinucleotide and 12% resulted from intramolecular recombination between the duplicated regions (Roth et al., 1985).

Other studies have also demonstrated that linear DNA is efficiently ligated within mammalian cells.
FIGURE 2. Intracellular end-joining and random breakage of replicative DNA molecules. Exponentially growing COS7 cells were harvested, mixed with (a) linear or (b) supercoiled plasmid pSV2neo (Southern & Berg, 1982), electroporated (Chu et al., 1987), and plated. After 72 h nuclei were prepared from the cells by lysis with NP40 (Sigma). Nuclear DNA was prepared and analysed using a randomly primed probe (shown in (c)). (a) Lane 1, 100 pg of pSV2neo linearized with EcoRI (input DNA); lane 2, 10 μg of nuclear DNA (not digested); lane 3, 10 μg of nuclear DNA digested with BamHI. (b) Lane 1, 100 pg of supercoiled pSV2neo DNA (input); lane 2, 10 μg of nuclear DNA (not digested); lane 3, 10 μg of nuclear DNA digested with EcoRI. (c) A map of pSV2neo linearized at the unique EcoRI site: oc, open circular DNA; lin, linear DNA; sc, supercoiled DNA; dc, DNA doubly cut (with EcoRI and BamHI). Panel (a) shows that when introduced into COS7 cells, linear pSV2neo is converted to both open circular and supercoiled forms (lane 2). The doubly cut molecules (lane 3) show that when they were recovered, some of the linear molecules were still in their original input form. Panel (b) shows that supercoiled molecules introduced into COS7 cells are converted into both open circular and linear forms (lane 2). No doubly cut molecules were detected when the DNA was digested with a single cut enzyme (lane 3), showing that the linear molecules recovered were randomly broken.
When DNA is transfected by calcium phosphate co-precipitation or by the use of DEAE-dextran, the nature of the ends of the molecules has little effect on ligation efficiency (Wilson et al., 1982; Miller & Temin, 1983; Kopchik & Stacey, 1983, 1984). In contrast, efficient ligation of microinjected DNA appears to require that single-stranded protruding DNA termini are complementary (Kopchik & Stacey, 1983, 1984).

Circular molecules can be visualized by Southern blotting DNA extracted from cells into which linearized plasmid DNA has been electroporated. When...
linear DNA bearing an SV40 replication origin is electroporated into COS7 cells, bands with the mobilities of linear, supercoiled and open circular forms of the input DNA are recovered after 72 hours (Fig. 2(a)). The open circular form is predominant, but it is not known whether this form occurs intracellularly or is produced by nuclease action during cell harvesting and DNA isolation. The DNA samples could be completely digested by \textit{MboI}, showing that the DNA had replicated extensively within the cells, and was not carried through the experiment extracellularly (data not shown). Results similar to those illustrated in Figure 2(a) were obtained when the plasmid could not replicate in the recipient cells (Fig. 3(b)). In this experiment, efficient digestion of the DNA recovered from the cells with a second single-cut enzyme generated two bands, one corresponding to linear DNA produced by cleavage of the circular forms and a second (double-cut) corresponding to cleavage of linear input DNA that had not become circular.

In both experiments (Figs 2 and 3), DNA prepared from cells that had been mixed with plasmid but not voltage-shocked did not contain detectable plasmid DNA sequences (data not shown).

**EVIDENCE FOR THE INVOLVEMENT OF LINEAR MOLECULES IN EXTRACHROMOSOMAL HOMOLOGOUS RECOMBINATION**

Many comparisons of recombination between circular and/or linear input plasmids have used as a quantitative measure the number of stably transformed clones that arise. Interpretation of these experiments is complicated by the fact that stable transformation by two complementary defective input genes requires recombination between them as well as the integration of recombinant molecules into the chromosomes. The most revealing experiments, therefore, are again those involving plaque formation after transfection with SV40 DNA.

Wake et al. (1985) co-transfected CV1 cells with DNA from two SV40 ts+ mutants, and scored plaques at the non-permissive temperature. When one input DNA was linearized appropriately (the question of "appropriate" linearization is taken up below) the frequency of plaque formation was greater by a factor of 3 to 4 than when both were circular. When both were linearized appropriately, the frequency was greater by a factor of 8. At a minimum, these results show clearly that linear DNA molecules recombine more frequently than circular. Note that they do not necessarily imply, nor do any co-infection experiments with circular DNA, that either a pair of circles or a circular and a linear molecule can recombine. Since circular molecules become broken intracellularly (see below), positive results obtained with pairs of circular molecules are quite consistent with the view that all extrachromosomal recombination involves pairs of linear molecules.

**LOW FREQUENCY RECOMBINATION BETWEEN ALIGNED LINEAR MOLECULES**

When SV40 genomes carrying different ts mutations were linearized by cleavage at the same point (with the same restriction enzyme) recombination frequencies were in general less than when both input molecules were circular (Wake et al., 1985). The frequency was significantly greater when the cleavage was made in such a way that the mutations were at opposite ends of the two aligned linear molecules. Much greater frequencies were obtained when the two molecules were linearized at different "offset" sites. The choice of offset was critical, as will be discussed below.

Other experimental approaches allow recombinants to be scored only after recombinant molecules have become integrated into the chromosomes. Such experiments are most convincing when the comparison being made is between pairs of plasmids cleaved in different ways, rather than between cleaved and uncleaved molecules; when all the input molecules are linear it is reasonable to assume that the frequency of integration is the same in each case. Observations in agreement with those described above were reported by Lin et al. (1984), who studied recombination between plasmids that carried deleted forms of the herpesvirus thymidine kinase gene, one with a 5'-deletion, two with 3'-deletions. Pairs of aligned linear 5'- and 3'-deleted forms showed low levels of recombination, while offset pairs showed much higher levels.

**POLARITY OF RECOMBINATION BETWEEN LINEAR MOLECULES**

Several crucial observations from different laboratories show that the relationship between the cleavage sites on the recombining molecules, and the
relationship of those sites to the lesions that are to be repaired by recombination, both strongly affect the frequency of recombination between the sites. In co-transfections with linearized SV40 ts mutations, for example, pairs of fragments that overlapped to an identical extent differed in recombination frequency by a factor of 40-fold (Fig. 4(a)). The difference that led to such dramatically different results was that in one case the mutant sites were closer to the nearest molecular end, while in the other case the non-mutant sites were closer to the nearest end. In the experiments of Lin et al. (1984) with terminally deleted thymidine kinase genes, staggered cuts in the two plasmids increased recombination by 70-fold. These cleavages also placed the lesions (deleted sequences) closest to the ends of the molecules (Fig. (b)). In another type of experiment, Lin et al. (1984) employed plasmids carrying tandemly linked, terminally deleted thymidine kinase genes. One plasmid carried the genes as a direct tandem repeat, and the other as an inverted tandem repeat (Fig. 4(c)). These plasmids were introduced into cells separately so as to study the effect of the different arrangements on intramolecular extrachromosomal recombination. Some prior cleavages led to low recombination frequencies while others led to frequencies up to 100 times greater. Southern blotting experiments with DNA recovered from the stably transformed cellular clones showed that the thymidine kinase gene had been reconstituted. However, the DNA adjacent to the cleavage site was partly or completely absent in each case.

These powerful experiments, in which single differences of arrangement or cleavage produce very different results, form the cornerstone of the proposed mechanisms of extrachromosomal homologous recombination (Lin et al., 1984; Wake et al., 1985).

**TERMINAL INHOMOLOGIES DO NOT PREVENT EXTRACHROMOSOMAL HOMOLOGOUS RECOMBINATION**

In several of the experiments described in the previous section high levels of recombination were observed between plasmids with terminal inhomologies. Wake et al. (1985) set up a series of tests in which the plasmid pairs were otherwise identical and found that short terminal inhomologies had little effect. In all cases examined the selective marker was correctly reconstituted while the regions of inhomology had been lost. The implication of all these results is that the homologous sequences become exposed by single-strand exonuclease attack (Lin et al., 1984) or unwinding of the duplex (Wake et al., 1985). The fact that non-homologous terminal sequences do not affect the outcome, or affect it only marginally, may be taken to show that the exposure of one strand of each participating DNA duplex, near to its end, is part of the normal process of extrachromosomal recombination.

**NON-CONSERVATIVE HOMOLOGOUS RECOMBINATION**

The experiments described above, and indeed all experiments on extrachromosomal recombination of which we are aware (see below), can be explained by the mechanisms proposed by Lin et al. (1984) and Wake et al. (1985), which differ only in detail. Both propose that single strands at the ends of linear
duplexes are exposed, and that exposed complementary strands of two molecules form a duplex. Single-stranded overlaps, which may or may not carry non-homologous sequences, are removed and the hybrid duplex is repaired. The application of the model to explain one of the experiments of Wake et al. (1985) is shown in Figure 5. The model supposes that recombination of a wild-type gene occurs in both cases, but that in one configuration essential viral genes are present in the unwound duplex region and are lost, while in the other they are carried on regions that remain base-paired and are not. The mechanism is thus non-conservative. The relative frequencies observed could be explained according to some types of conservative model on the ground that three recombination events are required to obtain a complete genome in the first case, while in the second only one recombination event is required. Thus the crucial question is whether or not recombination is conservative. Several experiments that demonstrate lack of conservation have been described (Lin et al., 1984; Wake et al., 1984; Rubnitz & Subramani, 1986). One of the clearest of these involved intramolecular recombination within a plasmid with two SV40 replication origins (Chakrabarti & Seidman, 1986). The plasmid also contained two defective copies of the SV40 Tag gene in direct orientation. When recombination reconstituted an active Tag gene, enrichment of any plasmid that carried an SV40 replication origin occurred by replication. The experiment was designed so that conservative recombination events would produce one large plasmid (conversion) or two smaller ones (conservative homologous recombination). In the event, only one of the smaller plasmids was found to have been amplified, demonstrating the non-conservative nature of the process. Recent experiments of Lin et al. (1987) support the idea that exposure of single strands at the ends of the recombining duplexes is due to the action of a 5'-exonuclease.

The experiments of Kucherlapati et al. (1984) have been quoted as demonstrating reciprocal extrachromosomal recombination. On a closer inspection, however, they do not appear to do so. The products of possible reciprocal recombination events that were observed could equally well have been generated during the build-up of the extrachromosomal tandem arrays by successive non-conservative homologous recombinations. We pointed out above that successive recombinations even between short fragments of a single input DNA can build perfect tandemly repeated arrays. The same is true when two DNA molecules that differ in minor ways are introduced, except that the ensuing array will carry interspersed direct copies of the two input molecules and also (if it is large enough) of all possible recombinant combinations.

This alternative interpretation of the data of Kucherlapati et al. (1984) is supported by results presented by Folger et al. (1985). The experiments were similar in that two plasmids carrying different defective neomycin phosphotransferase genes were introduced into cells together (by microinjection in this case). The plasmids also differed in a set of restriction site markers designed to allow detailed analysis of recombinants. Single G418-resistant cell lines were isolated and characterized (whereas Kucherlapati et al. (1984) analysed undefined pools of uncharacterized transformants). Several plasmids were recovered from each cell line by plasmid rescue. Each cell line contained a small number of direct tandem arrays, at least one functional neomycin phosphotransferase gene and a random selection of

**Figure 5.** An explanation of the experimental results obtained with linear SV40 DNA cleaved in different ways, by the non-conservative recombination model. Each line represents a single DNA strand. Strand pairing follows exposure of 1 strand of each duplex, for example by exonuclease attack. The unpaired portion of each single strand is removed by exo- or endonuclease attack, and repair of the duplex follows. In one configuration, when the non-mutant parts of the genes come together, a double deletion of terminal sequence is the inevitable result. In the other configuration, a terminal duplication inevitably occurs; from this a monomer can be generated by a 2nd recombination event. The data are from Wake et al. (1985).
the possible recombinant molecules. Plasmids that could have resulted from reciprocal (conservative) recombination were recovered from only one of seven G418-resistant clones. These observations provide the best evidence that the predominant mechanism of homologous recombination is non-conservative, and that tandem arrays are built up by repeated cycles of homologous non-conservative recombination in the manner described above. The frequency of conservative relative to non-conservative extrachromosomal recombination has been estimated to be 0.01 (Seidman, 1987).

EVIDENCE FOR THE INTRACELLULAR RANDOM BREAKAGE OF CIRCULAR MOLECULES

Many of the experiments that have been described can be explained in terms of the non-conservative mechanism only if it is supposed that linear molecules circularize intracellularly and that circular molecules are broken randomly. One of the experiments described by Wake et al. (1985) is a case in point. It will be recalled that appropriate, but not inappropriate, offset cleavage of two SV40 ts mutant genomes led to a high level of recombination upon co-transfection. If either of the molecules was cleaved appropriately, and the other was uncleaved (i.e. circular), a lower but still rather high level of recombination was observed (see the table in Fig. 5). If either was cleaved inappropriately, and the other was uncleaved, the level of recombination was low. This experiment is most easily explained by supposing that some of the circular molecules happen to be broken randomly within the cell. When the other molecule is linearized appropriately, it forms a pool of molecules able to recombine with the randomly broken molecules that are broken appropriately. When it is linearized inappropriately, the appropriately broken molecules have nothing with which to recombine. Linearization of input circular DNA can be detected by Southern blotting DNA recovered after transfection or electroporation. Figures 2(b) and 3(a) show that much of the foreign DNA recovered from the cells was linear. When this DNA was digested with a single-cut enzyme, only linear DNA was observed. Contrast this with the mixture of linear and double-cut DNA recovered when circular and input linear DNA was digested with a second enzyme (Figs 2(a) and 3(b)). Thus the linear DNA recovered after electroporation of circular molecules was randomly broken.

Integration of foreign DNA into chromosomal sites by illegitimate recombination

Foreign DNA introduced into cultured animal cells becomes integrated into the chromosomes by two distinct mechanisms; illegitimate recombination and homologous recombination. The latter route is available, of course, only when homology exists between the input DNA and resident chromosomal DNA sequences. In cultured cells illegitimate recombination, although itself rare, is 100 to 1000 times more frequent than homologous recombination (Thomas et al., 1986; Thomas & Capecchi, 1987). In transgenic animals produced by pronuclear microinjection homologous recombination has not been detected, and chromosomal insertion has, as far as is known, invariably been by an illegitimate route. Although it is of topical interest we do not address the question of homologous recombination, but confine our discussion to the illegitimate route whereby single DNA sequences and DNA arrays produced by extrachromosomal recombination become integrated.

IS RANDOM CHROMOSOME BREAKAGE REQUIRED FOR INTEGRATION?

The fact that foreign DNA becomes integrated in a minority of injected embryos (or cells), and then usually at a single site, suggests that the frequency of integration is determined by rare intranuclear events. The suggestion has been made that the frequency of chromosomal breaks constrains the number of integration events (Brinster et al., 1985). The observation that the frequency of DNA integration is increased by irradiating transfected cells (Perez et al., 1985) supports this view.

STRUCTURAL ANALYSIS OF INTEGRATION SITES

In this discussion we again adopt the position that the processes that take place in embryos are essentially the same as those that occur in cultured cells. In addition to the similarities cited at the beginning of the article, support for this view comes from structural analysis of chromosomal integration sites in stably transfected cells and transgenic animals (see below).
DOES INTEGRATION TAKE PLACE IN THE ABSENCE OF SEQUENCE SIMILARITY BETWEEN THE RECOMBINING MOLECULES?

There is a serious shortage of sequence data relating to the chromosomal integration of transfected or microinjected DNA. In fact in no case are the nucleotide sequences of both participating molecules known with total certainty on both sides of the boundary between them, and in only four cases are they known with a high degree of probability. These cases are two junctions generated after transfection of cultured cells by the insertion of foreign DNA between two copies of a chromosomal sequence that became tandemly repeated at the time of the insertion event (Wilkie & Palmiter, 1987). No sequence similarity between the pairs of recombining molecules was detected in any of these examples. In other examples chromosomal sequences that are essential for rigorous interpretation of the data are not known. It is of course unsound first to define the boundary as the point at which homology with the incoming DNA is lost, and then to draw a conclusion from the lack of homology between the incoming DNA and the recombinant to one side of that boundary.

On the other hand, the nucleotide sequences of input foreign DNA molecules are generally known and several cases of illegitimate recombination between two input foreign DNA molecules have been detected (see Konopka, 1988). In most of these cases a lack of sequence similarity across the boundary is indisputable. There is every reason to believe that most of these junctions were formed simply by the random end-joining of fragments of the incoming molecules (Wilson et al., 1982). It is known that random end-joining occurs extrachromosomally, particularly when high levels of DNA are input (Folger et al., 1982). It seems reasonable to conclude from the data available that end-joining is also the main mechanism of integration of foreign DNA into the chromosome.

In a minority of cases short and/or partial sequence similarities are found at junctions between fragments of input molecules within chromosomal arrays (Wilson et al., 1982; Anderson et al., 1984; Wilkie & Palmiter, 1987). These sequence similarities suggest the occurrence of homologous recombination with a low selectivity, such as might occur during transient partial base-pairing of exposed single-strands of the two duplexes; in effect a halfway house between end-joining and homologous recombination by the route described in the previous section. In view of the other similarities between illegitimate extrachromosomal recombination and chromosomal integration, it seems likely that instances of short or partial sequence similarity will also be found at some junctions between foreign DNA and the chromosome.

SHORT DIRECT REPEATS OF HOST DNA AT CHROMOSOMAL JUNCTIONS

Short direct terminal repeats of chromosomal DNA characterize the insertion sites of retroviruses and retrotransposons. It is remarkable that in the three cases in which the sequences of both junctions of the integrated DNA are known after transfection into cells (Kato et al. (1986), two cases) or pronuclear microinjection (Wilkie & Palmiter (1987), one case) there are candidate direct repeat sequences at the junctions, as follows:

Kato et al. (1986) C4 and C5 ATATAT C6 and C7 GTTC
Wilkie & Palmiter (1987) J2 and J7 ATCTT

None of these cases is unambiguous however. The first may simply be due to insertion towards one end of a continuous stretch of alternating A and T residues. The second and third depend upon an interpretation of the sequences at the junctions that implies some homology between the chromosome and the inserted sequence (pBR322 in both cases) at one of the two junctions. Additionally, the third is further complicated by the fact that the insertion was accompanied by a 5 kb duplication of host DNA. The most obvious explanation of the 5 kb duplication (see below) is incompatible with an explanation of the putative 5 bp repeat based on the retrovirus model. On balance the available evidence does not support the insertion of transfected or microinjected DNA by a retrovirus-type mechanism. Clearly, more data will be required to resolve the issue.

INSERTION OF SV40 AND POLYOMA DNA INTO THE CHROMOSOME

The insertion of retroviral provirus into the chromosome is part of the life-cycle of the virus, and depends upon specialized viral sequences and functions. The same is not true of polyomavirus and SV40. Insertion is not part of the life-cycle (although
there is evidence of an involvement of Tag in insertion; Della Valle et al., 1981), and the ends of insertions are apparently random sites within the virus genome (Birg et al., 1979; Botchan et al., 1976). The insertions are often tandem arrays (Botchan et al., 1976; Basilico et al., 1979; Birg et al., 1979), probably generated by recombination, at least in non-permissiveness cells (Hacker & Fluck, 1989). Taking into account chromosomal insertions of virus DNA increases significantly the number of cases in which not only are the input and recombinant DNA sequences known but also the resident chromosomal sequence. An SV40—chromosome junction with a short common sequence has been identified by Stringer (1982), but more commonly a lack of sequence similarity at the junction is found (Stringer, 1981; Ruley et al., 1982). In a set of three pairs of integration junctions, no short direct repeats of host DNA were observed (Stringer, 1981).

**FILLER SEQUENCES AT JUNCTIONS WITH THE CHROMOSOME**

Filler DNA has been found sandwiched between integrated foreign DNA sequences that had recombined illegitimately and presumably extrachromosomally (Woodworth-Gutai, 1981; Wilson et al., 1982; Ruley & Fried, 1983; Woodworth-Gutai et al., 1983; Anderson et al., 1984; Wilkie & Palmiter, 1987), and at some virus—chromosome junctions (Botchan et al., 1980; Hayday et al., 1982). Typically, filler DNA is a short sequence but Wilkie & Palmiter (1987) found a 532 bp sequence of unknown origin at a single insert—chromosome junction. Williams & Fried (1986) identified a sequence identical with a 37 bp filler sequence 650 bp upstream from the insertion on the same chromosome in inverted orientation. Filler DNA between sequences presumed to have come together extrachromosomally may represent nothing more than an extra illegitimate recombination event; that is, the only distinction between a filler sequence and a segment (of unidentified origin) may be one of length. This is especially likely to be true when the filler is derived from foreign carrier DNA that is known to have been present in large amounts (Anderson et al., 1984). Filler DNA at a chromosome junction that is derived from elsewhere in the genome may require another explanation.

**DUPLICATIONS AND DELETIONS**

Wilkie & Palmiter (1987) identified a 5 kb duplication of chromosomal DNA bracketing a site of insertion. Deletions of chromosomal DNA in the immediate vicinity of the insertion have been found (SV40 DNA, Botchan et al., 1980; Stringer, 1982; transgenic mice, Woychik et al., 1985). Rearrangements of chromosomal DNA in the region of the insertion occur frequently in transfected cells (Robins et al., 1981; Kato et al., 1986) and in cells infected with SV40 (Williams & Fried, 1986) and polyoma (Hacker & Fluck, 1989). In transgenic mice translocations (Overbeek et al., 1986; Gordon & Ruddle, 1985) and other rearrangements (Covarrubias et al., 1986, 1987) of chromosomal DNA have been reported.

**MODEL OF CHROMOSOMAL INSERTION**

To explain the 5 kb duplication surrounding the foreign DNA array in their transgenic line Myk-103, Wilkie & Palminter (1987) proposed a model that, as they point out, could also explain insertion-associated deletions and other rearrangements. The essence of the model is the proposal that insertion may sometimes involve the eye of the replication loop during chromosomal DNA replication. The involvement of sister chromatids before mitosis would serve the same purpose. An advantage of the model is that it explains how an insertion may follow chromosomal breakage without the broken ends of the chromosome becoming dissociated from each other. If the two duplexes are broken in one way and rejoined via the foreign DNA sequence, the outcome is a duplication, whereas if they are broken in a different way and rejoined via the foreign DNA, the outcome is a deletion (Fig. 6). Either short direct tandem repeats across the insertion or a short deletion would result if the breaks on the two chromatids were close to the same point. In much the same way we can arrive at a rather laboured explanation of the inverted short repeat reported by Williams & Fried (1986) by first supposing that the foreign DNA bridges two breaks on the same chromatid (Fig. 6). If the foreign DNA became integrated into two different chromosomes (or chromatids) a translocation would be the outcome; in effect, unless the foreign DNA is catalysing either breakage or rejoicing (see below), insertion is here an accompaniment to translocation.

**IMPLICATIONS OF THE HIGHER FREQUENCIES OF CHROMOSOMAL INSERTIONS OBTAINED WITH LINEAR DNA**

With both embryo pronuclear injection (Brinster et al., 1985) and injection into the nuclei of cultured
Chromosomal Integration of Foreign DNA

**FIGURE 6.** Insertion of foreign DNA during DNA replication can explain duplication and deletion of chromosomal DNA sequences. The diagrams show the eye formed by 2 replication forks travelling away from the replication origin. (a) A direct duplication; (b) deletion; (c) duplication of a segment of DNA from a remote region of the chromosome. (For the sake of simplicity the diagram has been drawn to show a duplication only; in practice, a deletion or a 2nd duplication would be expected to occur since there is no reason a priori why the breaks to the right in the 2 sister chromatids should occur in equivalent positions.)

- Chromosomal DNA
- Foreign DNA
- Duplicated (a), (c) or deleted (b) chromosomal DNA
- Joins points at which segments of chromosomal DNA are joined

ends of DNA molecules are intimately involved in the integration process. However, they may also have another effect. Following exposure to agents that damage DNA, an increase in cellular DNA ligase activity is observed (Creissen & Shal, 1982). This is probably brought about by ADP-ribosylation of DNA ligase by ADP-ribosyl transferase (ADPRTase; EC 2.4.2.30). An increase in ADPRTase activity follows DNA damage, and inhibitors of ADPRTase inhibit DNA repair (Creissen & Shal, 1982). Inhibitors of ADPRTase also strongly inhibit the stable transformation of cells (Farzaneh et al., 1988). Thus the transforming activity of linear foreign DNA may relate partly to stimulation of DNA ligase activity.

**Conclusion**

A COMMON BASIS FOR HOMOLOGOUS AND ILLEGITIMATE RECOMBINATION?

While we found it helpful to distinguish between extrachromosomal recombination and integration, it seems likely that the two processes have the same mechanistic basis. This may be referred to as opportunistic repair-ligation accompanied by strand exposure and followed by complete repair. We propose (1) that the repair-ligation process joins together DNA strands as and when they are juxtaposed, with an efficiency directly proportional to the stability of the base-pairing association between the strands; and (2) that strand exposure is an ongoing process, quite likely opposed by the repair of the duplex and perhaps also by the destruction of exposed strands. Strand exposure of circularly permuted linear molecules would allow pairs of molecules to form long perfectly matched heteroduplexes that are relatively stable, and therefore persistent even before ligation. Stable heteroduplex structures would be involved in integration events only rarely, simply because the amount of any homologous chromosomal sequence is usually very small, both in absolute terms and relative to the other chromosomal DNA sequences.

Strand exposure would also allow short mismatched heteroduplexes to form between non-homologous sequences. These would be less stable and therefore less likely to encounter the ligase before the strands again separated. However, mismatched duplexes would be formed much more frequently than matched heteroduplexes, and this is especially true of integration events, since the enormous amount of host DNA in the nucleus has the capacity...
to generate large numbers of partial matches. Lastly, entirely dissimilar molecular ends may come together transiently and thus come within the sphere of influence of a molecule of DNA ligase only infrequently.

Wong & Capecchi (1986) microinjected cells with a linearized plasmid carrying an intact APRTase gene, and with two linearized plasmids each carrying a defective APRTase gene such that activity could be restored only by (extrachromosomal) recombination. Both the number of cells that expressed APRTase transiently and the frequency of stable transformation were measured in each case. These experiments established that a cell in which extrachromosomal recombination had taken place was more likely to be stably transformed, thus supporting the idea that the two processes have the same basis.

If the processes do have the same basis, the relative frequencies of the different events might be strongly influenced by the form and the amount of input DNA, and possibly also by the state of the cell. Thus Wong & Capecchi (1987) showed that extrachromosomal recombination is 10 to 15 times more frequent in cells microinjected with linear DNA in S-phase than in those microinjected in G-phase. The frequency of stable transformation, on the other hand, was about the same regardless of the stage in the cell-cycle at which the cells were microinjected (Wong & Capecchi, 1985).

Although we have not dealt with the targeting of input genes to their chromosomal homologues, it is relevant to ask whether the mechanisms of recombination proposed are sufficient to account for that phenomenon. As a rough guide we ask whether the ratio of random insertion to homologous recombination (Rrh) is plausibly accounted for by two factors: (1) the relative “concentrations” of non-homologous and homologous chromosomal targets and (2) the possible stability difference between unmatched and poorly matched single-stranded ends on the one hand and well-matched heteroduplexes on the other. When molecules that shared different sequence lengths with the resident HPRT gene were electroporated into ES cells, Rrh varied between 1000 (9 kb shared sequence) and 25,000 (4 kb shared sequence; Thomas & Capecchi, 1987). The concentration ratio of non-homologous to homologous targets may be taken to be the same as the ratio of the (diploid) C-value to the shared sequence length. Then the stability differences between homologous and heterologous heteroduplexes that would be required in order to account for these values in terms of a unitary model would be from 50 times (4 kb shared sequence) to 600 times (9 kb shared sequence). These values seem plausible.

In cell transfection experiments, Rrh ranges between $10^4$ and $10^5$, again not obviously inconsistent with the operation of the same mechanism. However, in a set of microinjection experiments an Rrh of about 100 was obtained (Thomas et al., 1986), suggesting that a separate mechanism of homologous exchange may exist, but be evident only under some circumstances.

References


Chromosomal Integration of Foreign DNA

297


Southern, P. J. & Berg, P. (1982). Transformation of mam-


