INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

A GENETIC SCREEN FOR DEVELOPMENTALLY REGULATED GENES IN MICE

by

RICARDO A. BATTAGLINO

A dissertation submitted to the Graduate Faculty in Biomedical Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The City University of New York.

1998

This Manuscript has been read and accepted for the Graduate Faculty of Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

1.8.1998
Date

Francesco Ramirez, PhD
Chair of Examinig Committee

1/8/98
Date

Terry Krulwich, PhD
Executive Officer

David Colman, Ph.D.

Achim Gossler, Ph.D.

David Sassoon, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

A GENETIC SCREEN FOR DEVELOPMENTALLY REGULATED GENES IN MICE

by

RICARDO A. BATTAGLINO

Advisor: Heidi Stuhlmann, Ph.D.

This project was aimed at the identification of developmentally regulated genes in mice. The early mid-gestation stage of development (E6.0 to E9.0) is particularly important because a number of complex events take place at that time (gastrulation, onset of organogenesis, establishment of the A-P axis) which will set the basic body plan and affect the subsequent development of the embryo. Identification of genes with a possible role in the aforementioned processes will increase our understanding of mouse development. I designed a promoter-trap strategy to screen for genes that are regulated during early-mid gestation. Mouse ES cells were infected with promoter-trap retroviral vectors that carry a promoter-less human alkaline phosphatase (*PLAP*) reporter gene. Infected ES cells were induced to differentiate *in vitro* into embryoid bodies and clones were selected on the basis of their differential expression of the *PLAP* reporter gene. The rationale to this approach assumes that a number of cellular and molecular events that take place during the differentiation of ES cells into embryoid bodies recapitulate events that occur *in vivo* during early post-implantation

development. One ES clone, promoter trap clone IVE38, was isolated and tested for expression of the *PLAP* reporter gene during embryogenesis. To this end, chimeric embryos were generated by aggregating E38 ES cells with tetraploid embryos. *PLAP* reporter gene expression was found to localize, mainly, to the paraxial (somitic) mesoderm in E9.0 embryos. Expression was also observed in the neural tube. *PLAP* expression data suggests a role for the trapped endogenous locus in somitic mesoderm differentiation.

To Leslie Rae, for all she means to me...

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Although extensive morphological description of the early stages of mouse development has been documented, the molecular mechanisms regulating these processes are not fully understood. The analysis of early mid-gestation development at the molecular level has been hampered by the size of the embryo, the heterogeneity of cell types arising at this stage and the difficulty to access and manipulate the post-implantation embryo *in utero*. Traditional genetic analysis, based on the generation and characterization of large numbers of mutants, has been successfully used to dissect developmental pathways in organisms like *C. elegans* or *D. melanogaster*. This type of genetic screening, however, has been precluded in mice due to the long generation time and genome complexity characteristic of mammals.

1.1 Early Mouse Embryogenesis: A Brief Overview (46, 135)

The gestation period in the mouse is 19-20 days, depending on the strain. One remarkable feature of mouse early embryonic development is that it is much slower than many lower vertebrates or invertebrates such as *Xenopus*, sea urchin or *Drosophila*. In these species, complex embryos are formed by 24 hours after fertilization. In contrast, the mouse embryo has gone through only one cell division by that time and the rate of cell division will remain slow up to the stage of implantation, which occurs at day 4.5 post coitum (p.c.).

Around day 3.0 p.c., the first differentiation event takes place. Cleavage number 4, leading to the 16-cell stage embryo, ultimately results in the segregation of two cell lineages. The group of cells on the surface of the 16-cell stage embryo will give rise to an epithelial layer (the trophoectoderm, TE) whereas the inner cells give rise to a group of undifferentiated cells (the inner cell mass or ICM). This process of differentiation actually starts with the compaction of the 8-cell morula, when blastomeres undergo an increase in the surface in contact with each other, and ends with the formation of the blastocyst, which consists of a layer of TE cells surrounding a cavity (the blastocoel) and the ICM (approximately 30 cells).

Shortly before implantation, a second round of differentiation takes place which involves the formation of another epithelial layer, the primitive or embryonic endoderm, originating from cells of the ICM. The remaining cells of the ICM give rise to the embryonic ectoderm. After implantation, there is a dramatic increase in the cell division rate, especially in the embryonic ectoderm, from which the ectodermal, mesodermal and endodermal tissues of the fetus will be derived, as well as the mesodermal components of the extraembryonic membranes.

Between day 5 and day 9 p.c., a complex series of developmental processes occurs that will establish the basic body plan. The main event is gastrulation, which leads to the establishment of the definitive germ layers of the embryo. Gastrulation starts at about day 6.5 p.c. with the formation of the primitive streak and the delamination of a group of cells of the

primitive ectoderm and its accumulation as a distinct layer of cells between the primitive ectoderm and the visceral endoderm. The mesoderm delamination starts at the posterior end of the embryo and progresses anteriorly. The signals that trigger the initiation of gastrulation are currently unknown. The start of expression of several genes, such as *Brachyury* and *goosecoid*, however, strongly correlates both temporally and spatially with the appearance of the primitive streak (7, 12). At the anterior end of the primitive streak there is a discrete bilaminar structure, the node, which is homologous to Hensen's node in the chick and the dorsal blastopore lip of *Xenopus* The ventral layer of the node, known as the notochordal plate, gives rise to the notochord, a rod-like structure which is an axial structure and the defining feature of the phylum. At day 7.5 p.c., the neural plate, a thickened, medial band of ectoderm anterior to the primitive streak, is induced by signals emanating from the underlying mesoderm to form the neuroectoderm. Later, the neural plate folds up, giving rise to the neural folds. This "V" shaped structure will fuse along the dorsal midline forming the neural tube. Along with neural tube formation, mesodermal cells not involved in notochord formation will give rise to the paraxial or somitic mesoderm, which runs parallel along each side of the notochord and the neural tube. Starting at day 8.0 p.c., the neural tube closure proceeds from the level of the 5th somite both rostrally and caudally to be completed at day 9.0 and 10.5 p.c., respectively.

Some of the cells originated in the early primitive streak will contribute to the formation of the extraembryonic membranes. By

day 8.5 p.c. a variety of mesodermal derivatives can be observed in the extraembryonic region. Mesoderm cells, together with embryonic ectoderm, will form the amnion (day 7.0 p.c.). Mesodermal cells together with extraembryonic ectodermal cells will give rise to the chorion (about day 7.5 p.c.). The allantois appears at day 7.5 p.c. as a protrusion of mesodermal cells arising from the posterior border of the embryonic ectoderm. Mesodermal cells from the posterior end of the primitive streak will give rise to blood islands in the visceral yolk sac.

Beginning at about day 8.0 p.c., the paraxial mesoderm is divided into paired blocks of cells, or somites, along the anterior-posterior axis of the embryo. The somites are a manifestation of the metameric nature of chordates. Although transient, they play a critical role in organizing the anterior-posterior segmental pattern of the embryo. Somites will give rise to the vertebrae and ribs, connective layers of the skin, and skeletal muscle components of the back, body wall and limbs. Lateral to the somitic mesoderm, the intermediate mesoderm develops. This mesoderm compartment will produce the pronephric tubule, precursor of the kidney and genital ducts. Further laterally, the lateral mesoderm extends. This divides in two portions which surround the celomic cavity. The dorsal portion, the somatopleura, associates with overlying ectoderm and is continuous with the mesoderm of the amnion. The ventral portion, the splachnopleura, associates with endoderm and is continuous with the mesodermal component of the visceral yolk sac. Lateral mesoderm provides mesenchyme that will participate in the

formation of viscera, connective tissues of the limbs, lining of the body cavity and mesenteries

1.2 Embryonic Stem (ES) Cells

In recent years, the dominant emerging technology used for studying mouse developmental genetics has been the manipulation of murine embryonic stem (ES) cells *in vitro*. Mouse ES cells are continuously growing cells derived from the ICM of day 3.5 blastocysts (32, 67, 88, 89). Several properties are unique to ES cells and thus make them appropriate tools to study embryonic development. They can be grown and maintained in culture as undifferentiated cells by cocultivating them with embryonic fibroblast cells or by supplementing the medium with a differentiation inhibitory factor, LIF (leukemia inhibitory factor). When kept under these conditions, ES cells retain their pluripotency and can be reintroduced into blastocysts where they will contribute to the formation of somatic as well as germ line tissue in the resulting chimeric animal (13, 14, 31, 140). In addition, when grown in suspension culture, ES cells can be induced to undergo differentiation *in vitro*. Under these conditions, attachment to the substrate is prevented and complex structures form, with endoderm, ectoderm and mesoderm. These are called cystic embryoid bodies (EBs) because they resemble in many respects the mouse post-implantation embryo at the day 6-8 egg-cylinder stage (30, 89).

Although EBs do not display normal morphogenesis, they can sustain ES cell differentiation into a variety of cell lineages. A

number of molecular markers involved in regulatory events during mouse embryonic development were found to be expressed in EBs (130, A. Leahy, unpublished results). Moreover, expression of *Brachyury*, *Nkx-1.1* and *Pax* genes in EBs was found to be regulated upon addition of factors like activin or bFGF, a phenomenon also observed *in vivo* (144). The formation of EBs represents a powerful *in vitro* model system to study the processes of cell differentiation and cell lineage commitment. This system has several advantages over equivalent *in vivo* studies, the main one being that many mature cell types as well as their precursors, whose access *in vivo* is virtually impossible, can be readily accessed and studied.

To date, the ES *in vitro* differentiation system has been successfully used to study the development of the hematopoietic, endothelial, muscle and neuronal lineages. Several studies have shown that cells of the erythroid, myeloid and lymphoid arise *in vitro* in differentiating EBs (23, 30, 55, 64, 73, 101, 134). The process is very efficient, reproducible and can proceed without exogenously added factors, suggesting that the EBs offer a favorable environment for hematopoiesis. One very interesting observation is that the kinetics of early hematopoietic development recapitulates the approximate temporal sequence of events observed in hematopoiesis *in vivo*. (55, 69, 101). In addition, EBs can be generated from genetically altered cells, allowing a characterization of the role of a gene on the emergence of certain cell types, at times more precise than that achieved *in vivo*. Using the *in vitro* differentiation system with GATA-1 defficient cells, the GATA-1 deffect was characterized as a block late in erythroid differentiation.. In another similar study,

EBs generated from GATA-2 -/- ES cells showed significant decrease in the number of primitive ertythoid precursors and absense of definitive precursors (56). Endothelial cells were observed in cystic embryoid bodies, surrounding primitive erythroid cells that develop in the blood islands (30). More recent studies have shown that EBs can also support vasculogenesis as well as angiogenic response (86, 131). Foci of contractile cells, a sign of cardiomyocyte development, can be regularly observed in differentiating EBs in culture (30). In addition, when EBs are allowed to attach to tissue culture dishes, myocytes arise within two weeks, which later fuse to form myotubes (91). These myotubes express the Ca^{++} channel proteins and nicotinic receptors characteristic of muscle cells. Studies at the molecular level have shown that the myogenic program occurs in EBs, and that the kinetics of activation of the genes involved in the process (*myf-5*, *myogenin*, *MyoD*) parallels that observed in normal embryogenesis (91). Cultured EBs can also be induced to generate neuron-like cells by adding retinoic acid. These cells express neural specific markers such as the class II β-tubulin and neurofilament M subunit. Physiological studies show that these cells can generate action potentials and express the appropriate ion channels (3). Taken together, these results show that the mechanisms governing cell lineage establishment and commitment within EBs are very similar to those observed in the mouse, and the ES *in vitro* differentiation is a very powerful model system.

The ES cell technology has been mainly oriented towards reverse genetic approaches, that is, the selective disruption of already known genes via homologous recombination and subsequent

analysis of the genetic consequences of the mutation in transgenic mice derived from the transformed cell line (120). This approach has by no means been exhausted. However, an alternative approach aimed at the generation of new mutations would be very useful since it would expand the number of genes to be studied.

Over the past few years, ES cells have been shown to be good targets for insertional mutagenesis by several groups (25, 37, 61, 63, 87, 106, 108, 128). Since they can form highly organized cystic embryoid bodies which are, to a certain extent, analogous to post-implantation embryos, their use should allow one to study *in vitro* the involvement of specific genes in cell lineage specification or differentiation events, that may be relevant to developoment *in vivo*.

1.3 Identification of Murine Genes Involved in Development

Since developmental defects caused by mutations can not easily be detected and studied as the embryo develops, a number of approaches have been designed to isolate and characterize genes in mice, which are not based on the detection of defects *in vivo* caused by mutations. Cloning of mouse genes on the basis of sequence homology to *D. melanogaster* and *X. laevis* genes that are known to play a role in development led to the identification of, among others, the Hox gene clusters, *sonic hedgehog* and the mouse *goosecoid* gene (12, 47, 58). Genes which were identified on the basis of their role in normal and malignant cell proliferation were shown to also participate in normal development (for review (1, 77). The screening of stage-and tissue-specific cDNA libraries also proved successful for

identifying novel genes and studying their activation patterns (11, 17, 80). In every case, however, the method imposed a bias towards the isolation of specific gene families or stage- or tissue-specific genes, depending on the source of the cDNA library used to perform the screening.

Genes isolated by the aforementioned methods most likely represent only a fraction of those which are developmentally regulated. For many of these genes, a role in mouse development was determined *a posteriori* by functional disruption of the gene via homologous recombination in ES cells. This method proved very useful in elucidating the developmental relevance of many candidate genes. However, since this approach requires previous knowledge of the gene, its use is limited.

1.4 Insertional Mutagenesis in Mice

The production and analysis of insertion mutants allows a different approach to study the molecular basis of embryonic development. These mutants can be generated by introducing foreign DNA into the germ line by either DNA microinjection into zygotes, electroporation of DNA into ES cells or retroviral infection of early embryos or ES cells (42, 70, 92, 93, 113).

Retroviruses replicate by a complex life cycle which involves reverse transcription of the RNA genome into double stranded DNA by a virus-encoded RNA-dependent DNA polymerase, integration of double stranded DNA into the genome of the infected cell, transcription and translation of the integrated provirus as a cellular gene and lastly, packaging of genome-length viral transcripts into

virions. As a result of reverse transcription, untranslated sequences near the 5' end (U5) and the 3' end (U3) of the RNA genome are duplicated such that the integrated provirus is flanked by long terminal repeats (LTR) containing a copy of both the U3 and the U5 element separated by a direct repeat sequence (R) (123, 125). Stably integrated proviruses are vertically transmitted from one cell to its progeny. This particular life cycle enables retroviruses to move or deliver genes into mammalian genomes (5, 25, 49, 59, 102, 107, 109-111, 124, 132, 138).

Moloney murine leukemia virus (Mo-MLV), a type C exogenous retrovirus (133), or Mo-MLV-based vectors have been successfully employed to infect mouse embryos *in vitro* before implantation, or *in utero* after implantation, as well as ES cells (48, 50, 79, 114-116). Following DNA microinjection or retroviral infection, the foreign DNA becomes integrated into the host genome and can act both as a mutagen, disrupting the structure and function of the endogenous gene, and as a "tag", allowing its subsequent isolation and characterization. Mutants produced this way were first detected by visible abnormalities associated with the DNA insertion, or because they resulted in prenatal death. However, subtle or weak mutations that are likely to alter rather than arrest development may not be detected.

The frequency with which an integration results in a mutant phenotype is quite high, approximately 5% for retrovirus- and 10% for transgene-induced mutations (70, 93), demonstrating the potential of exogenously introduced DNA to generate mutations. Several factors, however, undermine the use of retroviruses and

microinjected transgenes as genetic tools to study mammalian development. First, the large size of the mammalian genome, on the order of 10^9 bp, makes it necessary to screen a large number of integration events in order to detect mutations in any specific gene. In addition, since mammalian genomes are diploid, most of the mutations resulting from proviral and plasmid integrations are recessive. Mouse strains carrying insertions have to be bred to homozygosity making this approach both time and labor consuming. Secondly, DNA microinjection frequently leads to deletions and major DNA rearrangements at the site of integration. Deletions may remove part of the inserted locus. Rearrangements, including inversions and translocations, may complicate the identification of the mutated gene. In the case of retroviral infection, the presence of certain regulatory elements within the LTRs, (in particular the enhancer and the poly(A) signal) can interfere with the detection of cellular promoters and limits the ability of retroviruses to promote functional gene fusions, in other words, to be mutagenic. Finally, this strategy does not offer a way to select for developmentally regulated genes versus housekeeping genes, whose disruption could also produce a developmentally abnormal or lethal phenotype. As a result of these limitations, only a few genes associated with retrovirus- or transgene-induced mutations have been molecularly cloned and fully characterized (43, 102, 105, 132, 141, 145, 146).

1.5 Entrapment Vectors

In order to increase the efficiency of insertion mutagenesis, several types of vectors have been developed which allow the

identification of integrations into expressed genes from a background of random integration events. In addition, since entrapment depends on random integration into a genome locus, the entrapment event will not be, in principle, influenced by the abundance of the transcript, allowing the detection of low copy mRNAs.

1.5.1 *Promoter Trap Vectors*

Promoter trap vectors contain coding sequences of either a selectable marker or a reporter gene. They are designed to detect DNA elements with transcription regulatory activity upon integration into an exon. As a result of the insertion, a read-through transcript is generated, which contains both host and reporter gene sequences. A retroviral promoter trap vector was devised based on the insertion of a selectable marker (for example *his* or *neo*) into the U3 region of the retroviral genome. Selection for the expression of the selectable marker led to the isolation of cell clones with host/reporter fusion transcripts initiated by adjacent cellular sequences (127-129).

The use of this type of vector presents some limitations. First, the vector has to integrate into an exon to give rise to resistant colonies. Second, since expression of the selectable marker depends upon the trapping event, this type of vector was only able to detect promoters that were active at the time of selection. This approach, however, has been successfully used for the identification of a murine homologue of the yeast RNA1 gene, which was shown to be required for post-implantation development (28). More recently, a mutation in the Eck receptor tyrosine kinase gene was generated by a U3-βgeo (a fusion protein with β-galactosidase and neomycin

phosphotransferase activity) gene trap insertion (100). The insertion was introduced into the germ line and *in vivo* expression of the endogenous locus reflected expression *in vitro* (22).

A different version of this vector type contains a reporter gene in the U3 region of the LTR and a selectable gene between the LTRs. Expression of the *neo* gene is driven by an internal promoter, allowing isolation of provirus-containing cell clones regardless of the trapping event (83, 84). This type of vector has recently been used in an *in vitro* screen aimed at identifying and isolating genes involved in myeloid differentiation (53). Using this approach, genes were identified whose expression was down-regulated upon differentiation of the progenitor cells into neutrophils. In a different study, genes induced by IL-3 withdrawal in hematopoieitic precursor cells were identified (94).

1.5.2 *Gene Trap Vectors*

Gene trap vectors contain a selectable marker or reporter gene downstream from a splice acceptor (SA) consensus sequence. SA sequences derived from different genes have been used in gene trap vectors. The mouse *engrailed-2* gene (41), the Mo-MLV *env* gene (6) and the adenovirus major late transcript intron 1/exon 2 boundary (37), are some examples. Following integration of the vector in the correct orientation into an intron, and splicing from a cellular Splice Donor onto the vector Splice Acceptor, a read-through transcript initiated in cellular flanking sequences, will be generated (37, 41). The reporter gene may or may not contain an ATG initiation codon. The inclusion of an ATG allows one to detect integrations into an

unstranslated first exon.When an ATG is not included, the readthrough message will be translated into a fusion protein with β-gal activity if the fusion after the splicing event is in frame with β-gal. Since intronic sequences are more abundant than exonic, the number of target sites and the probability of trapping a cellular gene should be greater. To date, gene traps have been successfully employed to identify novel genes (36-38, 41, 118) or to generate mutations in already known genes in order to study their effects *in vivo* (39, 122). Gene-trap mutagenesis has been performed almost exclusively in ES cells. However, differentiated cell lines are also amenable to this approach (39, 40, 53, 94).

1.5.3 *Enhancer Trap Vectors*

In addition to gene trap and promoter trap vectors, a third kind of vector has been used to mutate and identify new genes. These vectors are referred to as enhancer traps and contain a reporter gene or a selectable marker driven by a basal promoter, ideally only a TATA element. Enhancer traps can detect an increase in the transcriptional level that occurs as a consequence of integration near a genomic enhancer and, in addition, any tissue- or stage-specificity conferred by such an element (2, 60). This approach led to the identification of several genes in *Drosophila* (9, 10, 44, 78, 137). However, enhancer traps are less mutagenic than gene or promoter traps because integrations leading to the activation of the marker do not necessarily disrupt the endogenous gene. In addition, in more complex organisms like mice, cloning of genes mutated by

this type of vector may be very difficult since enhancer effects can be exerted over long physical distances.

Entrapment vectors can be introduced into cells as plasmid DNA by electroporation or by retroviral-vector infection. The introduction of foreign DNA into cells by infection with retroviral vectors rather than by electroporation of plasmid DNA has several potential advantages. First, retroviruses cleanly integrate into cellular chromosomes with little host DNA rearrangement. Only a two to four base pair duplication at the site of integration occurs as opposed to major DNA rearrangements and head-to-tail integrations frequently observed after DNA microinjection or DNA electroporation, respectively. Secondly, at a low multiplicity of infection (m.o.i.) retroviruses integrate as single copy elements per genome. Finally, proviral integrations are often found at the 5' end of genes, in DNaseI-hypersensitive sites and in open chromatin regions (90, 99, 104, 126).

The use of reporter gene-containing constructs to study the expression of endogenous tagged genes is based upon the assumption that the pattern of reporter gene expression, both temporal and spatial, will reflect that of the endogenous gene. The requirement for gene/promoter trap insertions to occur within a transcription unit in order to be detected makes such an assumption reasonable. In fact, most of the enhancer trap insertions analyzed to date in *Drosophila* showed a β-galactosidase expression pattern similar to that of flanking cellular sequences (8-10, 60, 130, 137). In chimeric mouse embryos derived from ES cells transfected with *lacZ* entrapment vectors, a good correlation was observed between the distribution of

the endogenous transcript and the sites of β-galactosidase expression (108).

1.6 Reporter Genes

Reporter genes code for enzymes whose activities can be easily visualized by staining with a chromogenic substrate. Visualization can also be achieved *in vivo*, as in the case of the jellyfish green fluorescent protein (GFP). The use of reporter genes to monitor transcriptionally active regions of the genome was first reported in bacteria (19) and involved the introduction of a reporter construct into the bacterial genome, whose expression was dependent upon the acquisition of cis-acting DNA elements. This approach has also been recently applied to higher eukaryotic organisms (for review 7, 94). As a reverse genetic approach, genes are identified on the basis of their pattern of expression before the phenotype associated with the mutation is known.

1.6.1 *E.coli lacZ*

The bacterial β-galactosidase encoded by the *Escherichia coli lacZ* gene has been extensively used as a reporter gene. Many features justify its use. The gene is developmentally neutral, as shown by the normal developmental potential of *lacZ* expressing mouse embryos. The method of detection, based on β-galactosidase-mediated cleavage of the chromogen 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-GAL) to yield a blue precipitate, is very sensitive. This allows the detection of activity in samples ranging

from single cells to whole embryos (81, 96). Finally, under the appropriate conditions, the background activity is very low. In studies of mammalian development, the *lacZ* gene has been employed to monitor the pattern of expression of an endogenous gene tagged with a *lacZ*-containing replacement vector by homologous recombination (54), to determine the activity of cis-acting elements in *lacZ* transgenic mice (15, 57), to mutate genes in ES cells and examine their pattern of expression upon *in vivo* differentiation in chimeric embryos (15, 37, 41, 57) and to mark cells with replication-defective *lacZ* retroviruses in cell lineage analysis of cerebral cortex, retina, skin and muscle (81, 96).

1.6.2 *PLAP*

More recently, the human placental alkaline phosphatase (*PLAP*), has been successfully used as a reporter gene. *PLAP* was used in clonal boundary analysis of the mouse retina (35). This reporter gene has proven to be at least as useful as the *lacZ* gene in the context of retroviral vectors with regards to high viral titer, stability of expression and identification of expressing cells *in vivo* (33-35). The histochemical detection assay, based on cleavage of the chromogen 5-bromo-4-chloro-3-indolyl phosphate (X-P) and reduction of nitroblue tetrazolium (NBT) to give a purple precipitate, is also very sensitive. The *PLAP* cDNA has been cloned, sequenced and expressed into mammalian cells. In addition, since the enzyme is expressed endogenously almost exclusively in human placenta, no major background expression is observed in mouse cells. Although some rodent alkaline phosphatase (AP) activities have been

described, these can be inhibited specifically without affecting the human enzyme activity. In addition, mouse AP activities can be inhibited by heat treatment under conditions that do not affect the human *PLAP* (35, 74). To date, *PLAP* has not been employed as a reporter gene in gene trap or promoter trap vectors.

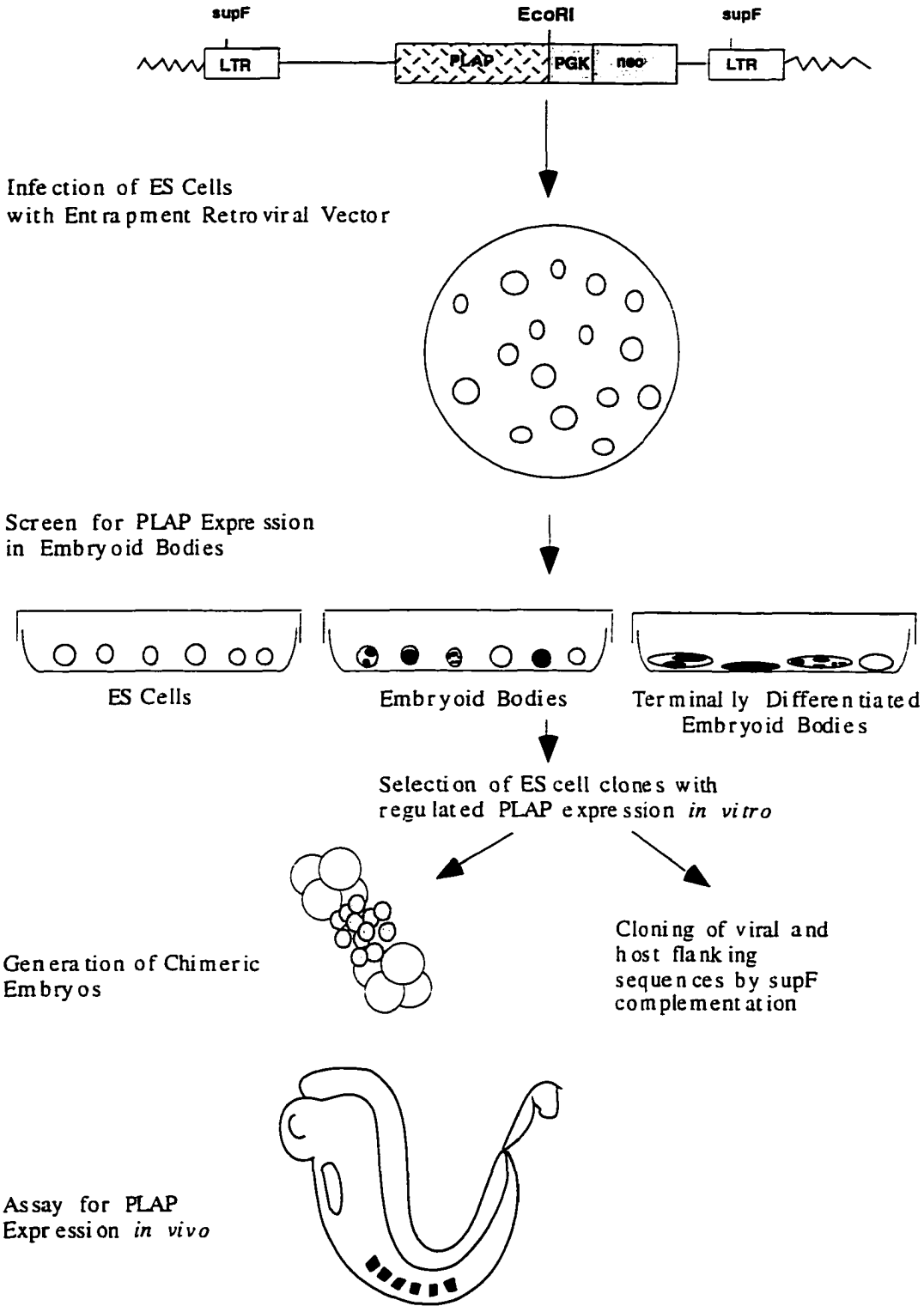
1.7 A Genetic Screen: Rationale and Experimental Design

The genetic control of mammalian development requires both temporal and spatial regulation of gene expression. As in any multicellular differentiation process, it is believed that the development of the early post-implantation mouse embryo is accompanied and driven by a cascade of gene expression events. It has been observed that a large number of genes that are essential for normal embryonic development are subject to a tight spatial and temporal control, suggesting that genes showing regulated patterns of expression during embryogenesis might also be important for mouse development.

Identification of developmentally regulated genes would increase our knowledge of mammalian early development. The early-mid gestation stage, between day 6 and day 9 p.c. in the mouse, is particularly important for development, because a number of developmentally complex events take place at that time (gastrulation, migration of primordial germ cells and neural crest cells, onset of organogenesis) which establish the basic body plan and affect the subsequent pattern of embryogenesis. However, a detailed molecular analysis of the mouse embryo at the early post-implantation stage has been hindered by the size of the embryo, the

heterogeneity of cell types arising at that time and the difficulty to access the early post-implantation embryo *in utero*.

A strategy was devised to isolate and identify genes involved in mammalian post-implantation development, which takes advantage of the differential expression most likely displayed by those genes. I developed a series of retroviral vectors and used them as insertion mutagens in mouse ES cells. The vectors, delivered into ES cells by retroviral infection, contain a promoter-less reporter gene, *E.coli lacZ* or human *Placental Alkaline Phosphatase (PLAP)*, whose expression is dependent upon transcription initiated from a cellular sequence which is adjacent to the integration site. A screen for potential developmentally regulated genes was performed by monitoring changes in the expression of the reporter gene upon induction of differentiation of infected ES cells *in vitro*, in embryoid bodies, and *in vivo* in developing chimeric embryos. Consequently, candidate genes were identified based on their expression pattern. Like some screens performed in the past ,where entrapment lines were isolated in which the reporter gene was not expressed in undifferentiated ES cells(41, 36), this screen was not limited to genes that were already expressed in undifferentiated ES cells.In contrast, in other screening efforts (4, 142), entrapment events were detected only if the trapped loci were expressed in undifferentiated ES cells



GENETIC SCREEN PROTOCOL

CHAPTER 2

MATERIALS AND METHODS

2.1 Construction of Promoter Trap and Gene Trap Vectors

Promoter trap and gene trap retroviral vectors were derived from Moloney-Murine Leukemia Virus (Mo-MLV). The 5' end of the vectors was generated by digesting the Mov9 plasmid (24) with EcoRI-XhoI. A 2.3 Kb fragment was produced, which extends up to the XhoI site at position 1560 of Mo-MLV to include the 5' LTR of the provirus, the packaging site Ψ and approximately 1 Kb of *gag* sequence (133). The 3' end of the vector was derived from a 1 Kb fragment obtained by digesting pLTR $^{sup-1}$ (85) with ClaI-PstI. This fragment contains Mo-MLV sequences from position 7674, including the entire 3' LTR of the virus with the *E.coli* suppressor F (*supF*) gene plus 50 bp of mouse flanking sequence (85). The fragment was ligated to a ClaI-PstI pSP72 vector to create pLTR*supF* . A 294 bp deletion, from a PvuII site at position 7935 to a SacI site at position 8223, was generated by partially digesting pLTR*supF* with PvuII. The resulting linear fragment was isolated and recut with SacI. The deleted fragment contains the viral enhancer and promoter (133).

The ATG$^+$ *lacZ* reporter gene was excised as a 3.5 Kb NotI fragment from pNASSβ (Clontech). The ATG$^-$ *lacZ* gene was generated as follows, a 830 bp *lacZ* fragment was amplified by PCR using the primers LacZ1 (5' **ggt acc** cgt cgt ttt aca acg tgg t 3') and LacZ2 (5' aaa a**at cga t**aa ttt cac cgc cga aa 3'). The amplified fragment does not contain the ATG translation initiation codon and is flanked by an upstream KpnI site and a downstream ClaI site. This fragment was ligated to a 2.5 Kb ClaI-EcoRI fragment containing the remainder of

the coding sequence downstream of the ClaI site, and to a KpnI-EcoRI psP72 vector (Promega). The resulting plasmid was named pATG$^-$*lacZ*. The ATG$^+$*PLAP* gene was constructed as follows, Cla12-AP (provided by D. Goff and C. Cepko) was digested partially with EcoRI, blunted, ligated to XhoI linkers and recut with an excess of EcoRI and XhoI. The resulting 1.94 Kb fragment was subcloned into pKS EcoRI-XhoI. For the gene trap vectors, a splice acceptor (SA) consensus sequence derived from the Mo-MLV *env* gene was generated as follows, the Mov9 plasmid was digested with BglII and BamHI. A 1130 bp fragment was isolated and XhoI linkers were added to the BamHI site. After recutting with KpnI, a 170 bp fragment was produced and subcloned into pKS(+) KpnI-XhoI which contains the SA sequence (62). This SA fragment was ligated to ATG$^-$*lacZ* KpnI-EcoRI to generate SA/ATG$^-$*lacZ*.

Both types of vectors also contain a selectable marker, the bacterial neomycin (*neo*) gene, whose expression is driven from an internal promoter (the *phosphoglycerokinase* gene, PGK, promoter) which has been shown to be constitutively expressed in ES cells (110, 121). The PGK *neo* cassettes were constructed as follows, PGK-*neo1* was excised as a 1.3 Kb EcoRI-XhoI fragment from pMC1 PGK *neo* polyA BglII. This cassette contains a 65 bp fragment including the polyA sequence from the PGK gene. PGK-*neo2* is a 1.8 Kb EcoRI-SalI fragment excised from pGEM7-PGK *neo* polyA. This cassette contains a 500 bp fragment that includes the polyA sequence from the SV40 early region. PGK-*neo3* is a 1.3 Kb EcoRI-BamHI fragment from pGEM7-PGK *neo* polyA. This cassette does not contain a polyA sequence.

2.2 Establishment of Virus Producing Cell Lines

The promoter and gene trap constructs were introduced by electroporation into the helper virus-free packaging cell line Ψ2 (66) or ΨCRE (27). For this, 5 μg $CsCl_2$ purified, supercoiled DNA were transfected into Ψ2 cells using a Gene Pulser (Bio-Rad). Cells were trypsinized, counted, spun down and resuspended in HEPES Buffered Saline (HBS) pH 7.05. 5 x 10^6 cells were subjected to a square electric pulse (300 V, 250 μF) for approximately 4 msec. Salmon sperm DNA (500 μg) was used as a carrier. After electroporation, cells were allowed to recover for 10 min at room temperature and plated in 10 cm tissue culture dishes using DMEM supplemented with 10% neonatal calf serum. 24 h after the transfection, cells were overlaid with fresh medium and 48 h after the transfection, cells were passaged into selection medium containing G418 (GIBCO) at a final concentration of 1 mg dry powder/ml. The medium was changed every other day during selection. G418-resistant clones become macroscopically visible after 5 days. Clones were picked 8 to 10 days after starting the selection, expanded and frozen down for further characterization.

2.3 Isolation of High Molecular Weight DNA

High molecular weight DNA was extracted from Ψ2 producer clones as follows, cells were harvested from a 10 cm dish using a rubber policeman, washed in PBS, transferred to 1.5 ml Eppendorf tubes and spun down. Cell pellets were frozen down, resuspended in 0.5 ml ice cold PBS and incubated on ice for 15 min. The pellets, consisting mostly of cell nuclei, were then resuspended in 0.3 ml cold

nuclear buffer (75 mM NaCl, 25 mM EDTA, pH 8.0). Resuspended nuclei were then lysed by adding 0.4 ml lysis buffer (10 mM NaCl, 10 mM Tris pH 8.0, 10 mM EDTA, 1% SDS and 400 mg/ml proteinase K) followed by a 4 h incubation at 50°C. The lysates were extracted twice with Tris pH 9.5-saturated Phenol and once with chloroform. DNA was precipitated by adding a 1/10 volume of 5.0 M NaCl and 0.7 ml isopropanol followed by 1 min of centrifugation. DNA pellets were air-dried and resuspended in TE by incubation overnight at 37°C.

High molecular weight DNA from ES cell clones was extracted according to the following protocol, ES cell clones were grown to confluence in 24-well dishes, without a feeder layer. The wells were then washed with PBS. Cells were lysed by adding 0.5ml lysis buffer (100 mM Tris pH 8.5, 5 mM EDTA pH 8.0, 0.2% SDS, 200 mM NaCl and 100 μg/ml proteinase K) followed by a 1h incubation at 37°C and agitation on a rocking plate for 30 min. DNA was precipitated by adding 0.5ml isopropanol with continuous agitation. DNA pellets were transferred to a tube, air-dried and resuspended in TE by incubation overnight at 37°C.

2.4 Southern Blot Analysis

Southern blot analysis of high molecular weight DNA was performed according to standard procedures (95). High molecular weight DNA was digested with the appropriate restriction endonucleases and size-fractionated on 1% agarose gels. Gels were rinsed for 10-15 min in 1/200 diluted HCl, denatured in denaturation solution (1.5 M NaCl, 0.5 M NaOH) 1x for 15 min and 1x for 30 min and neutralized in neutralization solution (1.5 M NaCl, 1 M

Tris-HCl pH=8.0) 1x for 15 min and 1x for 30 min. DNA was blotted overnight onto nitrocellulose membranes in 10X SSC buffer. Filters were baked for 1h in a vacuum oven at 80°C. Prehybridization was carried out in sealed plastic bags at 67°C for >2 h in prehybridization solution (10X Denhardt's, 5x SET, 0.1% SDS and 100mg/ml sonicated and denatured salmon sperm DNA). Hybridization was carried out in fresh prehybridization solution for 24 h at 67°C. A double stranded-DNA probe (1 - 2 x 10^7 cpm) labeled with [a-^{32}P]dCTP by the hexamer method (95) was denatured by boiling for 5 min and added to the prehybridization solution.

2.5 Embryonic Stem Cell Culture

The ES cell line R1 (kindly provided by Drs. Andras Nagy and Janet Rossant, Mount Sinai Hospital, Toronto) was maintained and propagated in an undifferentiated state on embryonic fibroblasts (EF) feeder cells on gelatin-coated dishes. EF feeder cells were mitotically inactivated by γ-irradiation. Dishes were coated with 0.7% gelatin and stored at 4°C for 30 min. The growth medium consisted of DMEM (GIBCO or Whittaker) supplemented with 15% Fetal Bovine Serum (FBS, HyClone), 20 mM HEPES pH 7.3, 0.1 mM non-essential amino acids (GIBCO), 0.1 mM β-mercaptoethanol, Penicillin/Streptomycin and 500 units/ml LIF, (ESGRO-GIBCO). The ES cell line D3 (kindly provided by Dr. Thomas Lufkin) was cultured without EF feeder cells, and growth medium was supplemented with 1000 units/ml LIF. The medium was changed every day and the ES cells were passaged every other day and, maintained at low density(~50% confluence) to avoid spontaneous differentiation.

2.6 Virus Titration

The Ψ2 virus producing clones were analyzed for their infectious titer on both NIH 3T3 cells and D3 ES cells. For this, sub-confluent (50-80%) 10 cm dishes with virus producing cells were overlaid with fresh medium. 12 to 16 h later, virus-containing medium was harvested, filtered through sterile 0.2 mm filters and serially diluted in 1:10 steps in medium containing 4 μg/ml polybrene (Sigma). Confluent 3T3 or D3 cells were trypsinized and 1/10 of the cells were seeded one day before the infection and infected using 1 ml of virus dilution each. 4 h later, fresh medium was added to the cells and forty-eight hours later cells were trypsinized, diluted 1:10 and passaged into G418-containing medium (1 mg/ml, dry powder). Seven to ten days after starting the selection, G418-resistant clones were fixed with methanol, stained with Giemsa solution and counted.

2.7 Infection of ES Cells

Retroviral vectors were introduced into ES cells by infection. For this, $1x10^6$ cells were seeded in 100 mm dishes prepared with inactivated feeder cells. The following day, cells were overlaid with supernatant from sub-confluent virus producing ΨCRE clones containing 200-500 $G418^r$-colony forming units ($G418^r$-CFU, as determined on ES cells) supplemented with 4μg/ml polybrene. This represents a m.o.i. significantly lower than 1 and ensures that only one provirus is present per cell. Infected clones were selected in G418-containing medium and picked seven to ten days after starting selection. Selected clones were expanded and stored as frozen

aliquots for further studies. Alternatively, ES cells were infected by cocultivation with mitotically inactivated virus-producing ΨCRE cells as previously described (116). Briefly, one 10 cm dish with producer cells was grown to 70-80% confluence. 5ml of fresh medium containing 10 μg/ml of mitomycin C were added to the cells. Cells were incubated for 3h at 37°C, carefully washed three times with PBS to remove the mitomycin C, trypsinized and plated with embryonic fibroblast feeder cells in ES cell medium supplemented with 4 μg/ml polybrene together. 0.5-1.0 x 10^6 ES cells were seeded for cocultivation into the same dish. Cells were cocultivated for 48 h, and fresh medium plus polybrene was added daily. At that time, cells were trypsinized and plated onto new dishes with feeder cells. Individual ES cell clones were obtained by plating appropriate dilutions of the ES cells, usually 1:50 to 1:100, onto plates containing selection medium. G418 resistant clones were picked seven to ten days after beginning the selection.

2.8 *In vitro* Differentiation of ES Cells

ES cells were induced to differentiate *in vitro* using a modification of previously described protocols (89, 134). For this, individual ES cell clones or pools of clones were plated for two days onto gelatinized 6 cm dishes without feeder cells in embryoid body (EB) medium (ES cell medium supplemented with 10% FBS and no LIF). After two days, small clumps of 10 to 20 cells each were removed from the dishes by careful trypsinization and plated into bacterial Petri dishes to avoid attachment. Half the volume of EB medium was changed daily. ES cell clumps cultured in suspension in

EB medium give rise to simple embryoid bodies with endoderm and ectoderm layers after 4 days and cystic embryoid bodies after 6 days. EBs were allowed to develop for 12 days. Samples were collected at 4, 7 and 12 days, fixed and stained for *PLAP* activity. To induce terminal differentiation, 4-day EBs were plated back onto gelatinized tissue culture dishes, allowed to reattach to the surface and further cultured for 7 days in EB medium.

2.9 *PLAP* and β-galactosidase Staining

PLAP staining was performed on cells, EBs and embryos using a modification of established procedures (35). Samples were washed with PBS, and fixed with 2% paraformaldehyde solution for 5 to 10 min (cells and EBs) and >2 h (embryos). In addition, ES cells, EBs and mouse embryos had to be heat-treated for 30 to 40 min at 70°C in PBS plus 10 mM $MgCl_2$ to inactivate mouse embryonic AP activities. Samples were equilibrated for 10 min in *PLAP* buffer (100 mM Tris-Cl pH 9.5, 100 mM NaCl, 10 mM $MgCl_2$) and stained overnight at room temperature without light using the chromogenic substrates 5-bromo-4-chloro-3-indolyl phosphate (BCIP, 0.1 mg/ml) and nitroblue tetrazolium (NBT, 1 mg/ml) in the presence of levamisole (Sigma, 0.24 mg/ml), an inhibitor specific for rodent AP activity.

For b-galactosidase staining, cells were washed with PBS and fixed in 2% paraformaldehyde solution at room temperature for 5 to 10 min. Fixed cells were equilibrated with ice-cold PBS plus 2 mM $MgCl_2$ for 10 min at 4°C, and with PBS plus 2 mM $MgCl_2$, 0.01% Sodium Deoxycholate and 0.02% NP40 for 10 min at 4°C. Cells were stained overnight at 37°C in X-gal PBS solution (2 mM $MgCl_2$, 0.01%

Sodium Deoxycholate, 0.02% NP40, 35 mM $K_3Fe(CN)_6$, 35 mM $K_4Fe(CN)_6$ x 3H20 and 1 mg/ml 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside, X-gal)

2.10 Cloning of Proviral and Flanking Genomic DNA

2.10.1 *Sup F Cloning*

Proviral and flanking mouse genomic sequences from selected ES cell clones were isolated by complementation cloning on a *supF* deficient bacterial host. To this end, DNA fragments from infected cells were ligated to a *sup* F^- λ phage vector and recombinant phages were selectively grown on *sup* F^- MC1061 cells (85). The genomic insert was prepared as follows, 5 μg of genomic DNA from the selected clone was digested with EcoRI or EcoRI and SacI (2 units enzyme/μg DNA) for 2 h at 37°C. The DNA was then ethanol precipitated and resuspended in water to a final concentration of 0.1-0.2 mg/ml. λ zap (Stratagene, a *sup* F^- insertion vector) was used to clone the *supF* containing genomic fragments. l zap arms were prepared as follows, λ zap (10 μg) was incubated overnight at 12°C with 100 units of T4 ligase (New England Biolabs) to ligate the cos ends. Cos end-ligated phage DNA was digested with either EcoRI or EcoRI plus SacI. λ zap EcoRI arms were further dephosphorylated with Calf Intestinal Phosphatase (CIP, New England Biolabs) as described (95) to prevent vector self-ligation.

Genomic fragments (0.2 μg) were ligated to the arms of the λ zap vector (1 μg) at 12°C for 24 h in a 5 μl total ligation volume. The ligation reaction was subsequently packaged into phage particles using the Gigapack Gold Packaging Extract (Stratagene) according to

the manufacturer's instructions. The packaging reaction was stopped by addition of 0.5 ml SM buffer (NaCl 5.8 g/l, $MgSO_4$. 7H2O 2 g/l, Tris-HCl (pH 7.5) 50 mM, gelatin 0.01%) and 20 ml chloroform, followed by vortexing and centrifugation at 12,000 rpm for 30 sec. Resuspended viruses were used to infect the *supF* deficient MC1061 P3 bacteria (Invitrogen). 200 µl of virus were added to 600µl of MC1061 bacteria (OD_{600}=1, resuspended in 10 mM $MgSO_4$) and incubated at 37°C for 15 min. Following adsorption of the phage particles, 6ml LB melted top agar (0.7% Agarose) were added and cells were plated on 150 mm Petri dishes with LB agar. Dishes were incubated overnight at 37°C.

2.10.2 *DNA Screening*

Plates were chilled for 1 h at 4°C after incubation. Plaques were transferred onto nitrocellulose membranes for 1-2 min. The membranes were denatured by submerging them in denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 2 min, and neutralized in neutralization solution (1.5 M NaCl, 1 M Tris-HCl pH=8.0) for 5 min. Filters were baked for 1h in a vacuum oven at 80°C, prehybridized in prehybridization solution (10X Denhardt's, 5X SET, 0.1% SDS and 100µg/ml sonicated and denatured salmon sperm DNA) for >2 h. Hybridization to radiolabeled probes was carried out in prehybridization solution as previously described. Petri dishes were stored at 4°C to be used after screening. Phage plaques were picked and transferred to an Eppendorf tube containing 0.5 ml SM Buffer and 20 ml chloroform, vortexed to release the phage particles and incubated on a shaker for 1-2 h at room temperature.

2.10.3 *In vivo Excision of pBluescript SK(-) from λ zap*

The λ zap vector allows *in vivo* excision. Cloned inserts are recovered as a phagemid (pBluescript SK-) containing the cloned inserts. Phagemids containing *sup F* inserts were excised as follows, 250 µl of resuspended phage stock (>1 x 10^5 phage particles) were combined in a 50 ml conical tube with 200 µl freshly grown XL1-Blue MRF' cells (OD_{600}=1 in 10 mM $MgSO_4$) and 1 µl f1 helper phage (ExAssist, Stratagene). The mixture was incubated at 37°C for 15 min to allow phagemid adsorption. 3 ml LB broth were added followed by further incubation at 37°C between 3 h to overnight. During this incubation step, a single stranded DNA molecule was synthesized, circularized and packaged into fiiamentous phage particles by helper phage encoded-proteins which were then secreted into the medium. The mixture was heated at 70°C for 15 min to kill the bacterial cells and spun at 4000 x g for 15 min. The phagemid-containing supernatant (1 µl) was added to SOLR cells (200 µl, OD_{600}=1 in 10 mM $MgSO_4$). The mixture was incubated at 37°C for 15 min, and plated on LB-ampicillin (50 mg/ml) plates containing X-gal (1 mg/ml). Plates were incubated overnight at 37°C. Ampicillin resistant SOLR colonies were picked and grown overnight in 3 ml LB broth at 37°C for plasmid DNA restriction analysis (95).

2.10.4 *Sequencing*

Flanking genomic DNA was sequenced with a commercial sequencing kit (Sequenase Version 2.0, USB), using T3 or T7 primers, and synthetic oligonucleotide primers specific for the U3 region (5'

GAG TGA TTG ACT ACC C 3') or U5 region (5' GCT AGC TTG CCA AAC C 3') of the Mo-MLV LTRs, respectively.

2.11 RNA Preparation and Northern Analysis

Total RNA from virus producing cells, ES cells and EBs was isolated using the TriPure Isolation Reagent (Boehringer Mannheim) following the manufacturer's instructions. The RNA was recovered from the solution by isopropanol precipitation, resuspended in DEPC-treated water, and stored at -80°C. Northern Blot analysis was performed according to established protocols (95). Briefly, 10mg RNA samples were size-fractionated on 1% agarose-formaldehyde gels. RNA was blotted overnight onto nylon membranes in 20X SSC buffer. Nylon membranes were UV-crosslinked for 2 min. Prehybridization was carried out in sealed plastic bags at 42°C for >2 h in prehybridization solution (5X SSC, 5mM Na_2PO_4, 0.1% SDS, 1X Denhardt's solution, 50% Formamide and 100µg/ml sonicated and denatured salmon sperm DNA). Hybridization was carried out in fresh prehybridization solution for 24 h at 42°C. A double stranded-DNA probe (1 x 10^7 cpm) labeled with [α-^{32}P]dCTP by the hexamer method (95) was denatured by boiling for 5 min and added to the prehybridization solution.

2.12 RT-PCR Analysis

Total RNA was reverse transcribed to cDNA essentially as described (95). 5 µg total RNA were resuspended in 10 µl H_20, heated at 65°C for 10 min and placed immediately on ice. 85 µl Reverse transcription mix (5 mM dNTPs 10 µl , 5X Buffer (Gibco-BRL) 20 µl,

0.1 M DTT 10 µl, 1 µg/µl reverse/antisense primer 5µl, H20 40 µl) were added to each sample. 5 µl Mo-MLV Reverse transcriptase (Gibco-BRL, 200 U/µl) were added followed by incubation at 42°C for 1.5 h. cDNA was sequentially extracted with phenol, phenol chloroform and chloroform, and ethanol precipitated. cDNA was then resuspended in 25 µl H_2O, and 5 µl aliquots were used as a template for the PCR reactions, using the following 5' primers: AP3 (forward/sense) and AP4 (reverse/antisense) and 3' primers AP1(forward/sense) and AP2 (reverse/antisense). The sequences are, AP4: 5' gtg gct cca ctg tct ggc ac 3', AP3: 5' ctg ctg ctg ggc ctg agg 3', AP2: 5' act ggg ctg tag ctg ctg gc 3', AP1: 5' gtt gct tcc tct gct ggc c 3'. Primers hgapdh1 (5' ctc atg acc aca gtc cat gcc atc 3') and hgapdh2 (5' tcg ttg tca tac cag gaa atg agc 3'), that amplify a 431 bp fragment of the glyceraldehyde 3-phosphate dehydrogenase (gapdh) cDNA were used in positive control reactions. PCR reactions were carried out in a Perkin-Elmer-Cetus DNA thermocycler. The cDNA template was denatured at 94°C for 1 min and annealed to the primers at 58°C for 1 min. The temperature was then ramped to 72°C and elongation by Taq DNA polymerase (Perkin-Elmer-Cetus) proceeded for 1 min in buffer supplied by the manufacturer supplemented with 3mM $MgCl_2$. The cycle was repeated 30 times. PCR products were visualized on ethidium bromide-stained 1.4% agarose gels and transferred to nylon membranes for Southern analysis with radiolabeled *PLAP* probe.

2.13 Generation of Chimeric Embryos

Totally ES cell derived mouse embryos were generated by aggregating tetraploid embryos and ES cells, essentially as described

(72). Tetraploid embryos were generated by electrofusion of two-cell-stage embryos. Late two-cell-stage embryos were obtained from oviducts of superovulated and mated CD1 females 44 to 46 h after treatment with human chorionic gonadotrophin (hCG). Electrofusion was carried out by placing the embryos between the electrodes of the electrofusion chamber connected to a CF-100 Pulser (BLS, Budapest, Hungary). The embryos were exposed to a short square electric pulse (100 V for 20-30 μsec) applied perpendicular to the plane of contact between the two blastomeres. The fusion was performed in non-electrolyte solution (0.3 M mannitol). After applying the pulse, the embryos were transferred into a drop of M2 medium (46) to remove the mannitol, and then transferred back into a drop of M16 medium under oil, for incubation at 37°C in 95% air/5% CO_2. 24 h later, 80-90% of the electrofused embryos developed to the four cell-stage and were selected for aggregation. The zonae pellucidae were removed by acid Tyrode's treatment (46).

ES cell-clones from the R1 line were trypsinized to single cell suspension and plated at low density two days before the aggregation to produce uniformly sized clumps of 10-20 cells. Clumps of loosely connected cells were placed between two four cell-stage tetraploid embryos in aggregation wells covered with M16 medium overlaid with mineral oil (Sigma). These aggregation "sandwiches" were incubated overnight at 37°C in 95% air/5% CO_2. The following day, 50 to 80% of the aggregates had developed to the compacted morula or blastocyst stage and were transferred into the uteri of day 2.5 pseudopregnant females. Embryos were dissected at day 8.5 and 9.5 p.c.

2.14 RNA *In situ* Hybridization

Embryoid bodies were harvested on days 4, 6, and 12. Procedures for embedding, sectioning, and *in situ* hybridization were exactly as described (98). Briefly, embryoid bodies fixed in 4% paraformaldehyde were embedded in paraffin at 60°C, following a series of dehydratring washes in ethanol and a final 30 min wash in Xylene. Material sectioned at 6 mm was deparaffinized and digested with proteinase K prior to overnight hybridization with ^{35}S-radiolabeled antisense mRNA probes (20 µl, $4x10^4$ cpm/ml). Post-hybridization washes of increasing stringency were included to prevent background signals associated with non-specific hybridization. Slides were dipped in Kodak NBT-2 emulsion, dried overnight and developed for 5 days to 2 weeks at 4°C. A radiolabeled sense transcript was included as a negative control.

CHAPTER 3

CONSTRUCTION AND CHARACTERIZATION OF ENTRAPMENT VECTORS

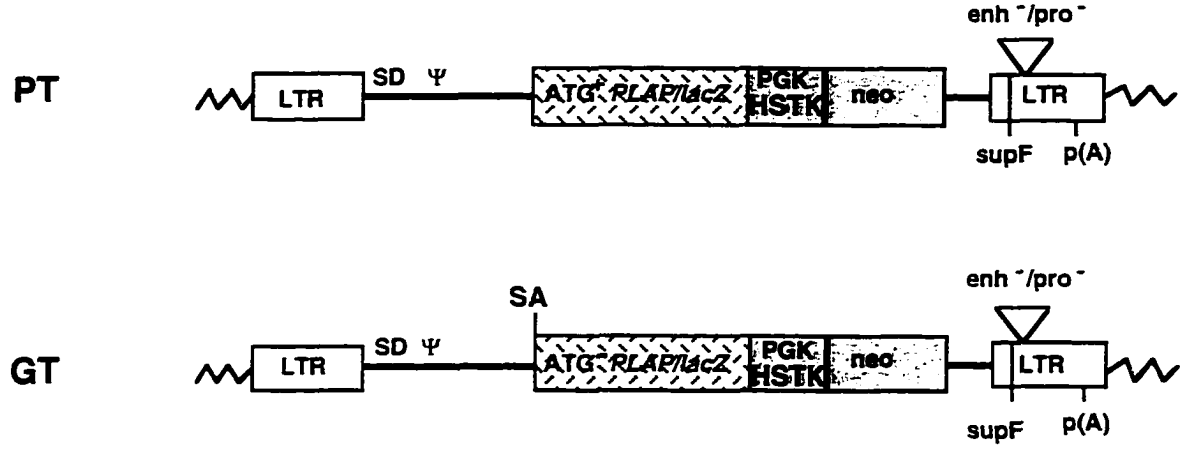
3.1 Generation of Promoter Trap and Gene Trap Retroviral Vectors with *lacZ* as a Reporter Gene

I constructed two types of *lacZ* Mo-MLV-based vectors to infect and mutagenize ES cells, promoter trap and gene trap vectors. The prototypes for both vectors and a list of constructs are depicted in Figure 1. They contain Mo-MLV-derived sequences, including the 5' and 3' LTRs, and sequences needed for reverse transcription and packaging of viral RNA. The promoter and enhancer elements contained in the 3' LTR were deleted. In addition, all vectors contain a promoterless reporter gene, the bacterial *lacZ* gene, whose expression is dependent on integration of the vector in the proximity of an active cellular promoter.

In the promoter trap vectors, expression of the reporter gene depends upon the generation of a read-through transcript originating in upstream cellular sequences. The *lacZ* reporter gene was derived from pNASSβ (Clontech) and has a sequence derived from the *D. melanogaster* alcohol dehydrogenase (Adh) gene which provides the ATG initiation codon.

In the gene trap vectors, a splice acceptor consensus sequence derived from the Mo-MLV *env* gene (62) was inserted upstream of an ATG$^-$ *lacZ* gene. Integration of this vector into an intron of an endogenous gene in the correct orientation was expected to generate a spliced read-through transcript. Translation of this spliced message

Vector Structure



Vector Constructs

1. **WT/P1-GT/S**
2. **WT/P1-GT/AS**
3. **E/P1-GT/S**
4. **E/P1-GT/AS**
5. **EP/P1-GT/S**
6. **EP/P1-GT/AS**
7. **EP/S1-GT/AS**
8. **EP/P2-GT/AS**
9. **EP/P1-PT/S**
10. **EP/P1-PT/AS**
11. **EP/S1-PT/AS**
12. **EP/P2-PT/AS**
13. **EP/P3-PT/S**
14. **EP/P3-GT/S**

Figure 1. *Structure of the Promoter Trap and Gene Trap Constructs* **WT**: wild type 3' LTR, **E**: enhancer⁻ 3' LTR, **EP**: enhancer⁻/promoter⁻ 3' LTR, **SD:** splice donor site of Mo-MLV; **Ψ:** packaging signal; **P1**: PGK promoter-*neo*-PGK polyA cassette, **P2**: PGK promoter-*neo*-SV40 polyA, **P3**: PGK promoter-*neo* cassette, **S1**: TK promoter-*neo*-PGK polyA cassette, **PT**: promoter trap, **GT**: gene trap, **S**: *lacZ -neo* cassette in sense orientation, **AS**: *lacZ -neo* cassette in antisense orientation.

will result in expression of a fusion protein with β-galactosidase activity.

Both types of vectors also contain a selectable marker, the bacterial neomycin phosphotransferase (*neo*) gene. Its expression is driven by an internal promoter (either the mouse *phosphoglycerokinase*, PGK, or the Herpes simplex virus *thymidine kinase*, HSTK, gene promoter) which has been shown to be constitutively expressed in ES cells (110, 121). Therefore, ES cell clones carrying an integrated provirus can be selected by using the neomycin analog G418, regardless of activation of the inserted gene in undifferentiated ES cells. The *lacZ -neo* cassette was inserted in both orientations relative to the transcriptional orientation of the virus. In the sense orientation constructs, after transfection into viral packaging cell lines, transcripts initiated from both the viral LTR and the internal promoter terminate at the viral polyadenylation signal present in the 3' LTR. In the antisense constructs, the polyadenylation signal is provided by sequences derived from either the PGK gene or the SV40 early region. This design allows transcription of the genome-length RNA in the packaging cell line regardless of the orientation of the *lacZ-neo* cassette. In addition, the 3' LTR contains a copy of the bacterial *sup F* suppressor gene, which can be used for cloning of proviral and flanking cellular sequences using a *sup* F^- bacterial strain.

The 5' LTR in the plasmid constructs is wildtype, whereas the 3' LTR is deleted for the viral enhancer and promoter. Following virus production and reverse transcription, this deletion is duplicated into the LTR. Therefore, in infected ES cells, the integrated provirus

contains deletions at enhancer and promoter sequences in both LTRs. The boundaries of this deletion were confirmed by sequencing of the plasmid vector and integrated DNA from infected ES cells.

3.2 Establishment of Virus Producing Cell Lines

The promoter and gene trap constructs were introduced by electroporation into the packaging cell line Ψ2. This cell line contains a defective Mo-MLV provirus which provides all functions needed *in trans* for the generation of a viral particle (66). The integrated provirus has a deletion in the Ψ (packaging) signal, which prevents the generation of helper-virus particles. Salmon sperm DNA was used as a carrier during electroporation. Salmon sperm DNA has also been shown to prevent multiple integrations of plasmid DNA. Infection of ES cells with retroviruses containing a promoter-less reporter gene was essential for our screening because expression of the reporter gene must depend upon the activity of the trapped cellular locus. Therefore, the generation of Ψ2 clones with single or low copy integrants was important to lower the probability of recombination between integrated retroviral genomes. Such a recombination event could lead to the reconstitution of a wildtype 5' LTR, which in turn would drive constitutive expression of the reporter gene in infected differentiating cells.

Transfected cells were plated for selection on G418 medium. G418-resistant clones became visible after 5 days. 8 to 10 days after starting the selection, clones were big enough to be picked and 12 Ψ2 stable transfectants per construct were picked and expanded.

3.3 Characterization of Virus Producer Cell Lines

3.3.1 *DNA Analysis*

These clones were analyzed for the integrity and number of integrated proviruses. For that, high molecular DNA from Ψ2 transfected clones was digested with either XbaI or KpnI and EcoRI endonucleases, and analyzed by Southern blot. Filters were hybridized to a 1.3 Kb ^{32}P labeled *neo* fragment. Both XbaI and KpnI cut the proviral DNA only once on each LTR. Therefore, the length of the *neo*-hybridizing fragment is diagnostic for the integrity of the integrated provirus. EcoRI, on the other hand, cuts the proviral vector once between the LTRs at the *neo-lacZ* junction. Therefore, the number of *neo*-hybridizing fragments is diagnostic of the number of integrated proviruses of a particular clone. KpnI was used in the enhancer$^-$ and enhancer$^-$/promoter$^-$ clones, because the XbaI site in position 8113 was removed in these clones as a result of the enhancer and promoter deletion.

In Figures 2 and 3, Southern blot analysis of DNA from Ψ2 clones transfected with GT constructs #4, #5, #7 and #8 is shown. Several Ψ2 clones from constructs #7 and #8 contained a 7.0 Kb and 7.5 Kb KpnI *neo* fragment, respectively, indicative of an intact provirus (Figure 2A *top*: lanes a, c, e and h and 2B *top*: lanes l and m). Each clone produced a single EcoRI *neo* fragment, indicating that these clones contained a single integrated construct (Figure 2A *bottom*: lanes a, c, h; and 2B *bottom*: lanes l and m). Several clones from constructs #4 and #5 had a 6.5 Kb and 5.6 Kb *neo*-hybridizing band, respectively (Figure 3A *top*: lanes c, d, e, i and k; and 3B *top*: lanes o and r). These clones showed either one (3A *bottom* lanes c, d,

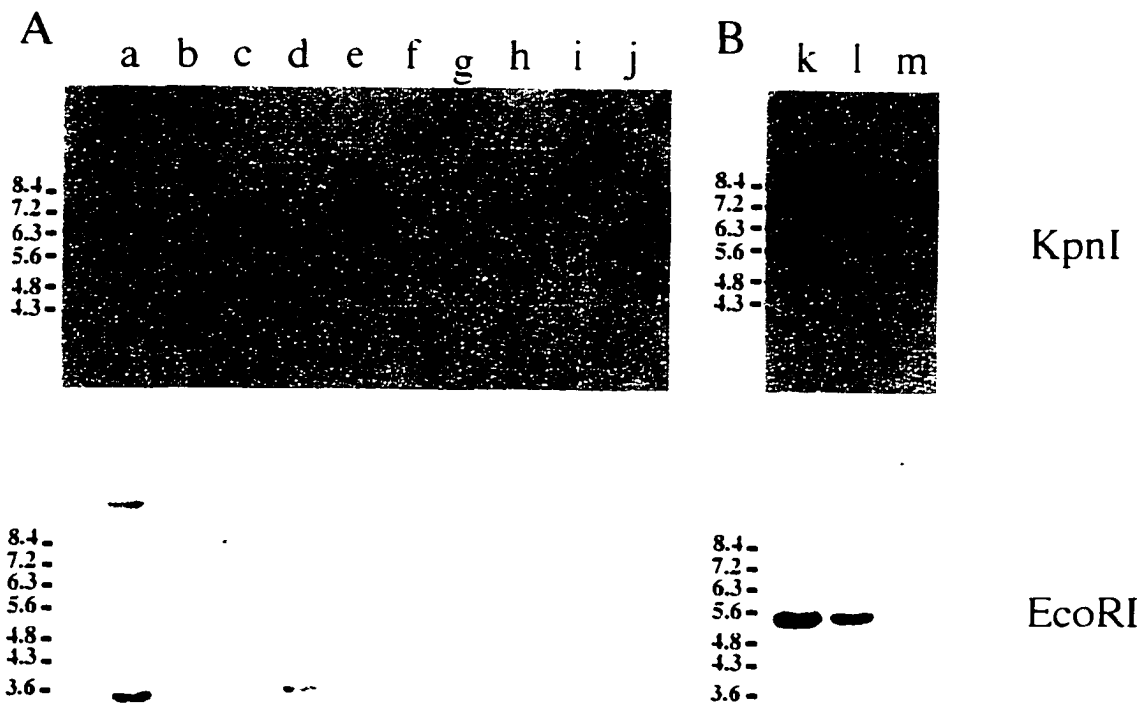


Figure 2. *Southern Blot Analysis of Gene Trap virus Ψ2 Producer Lines*

A) Genomic DNA (10 μg) from 10 independent Ψ2 clones transfected with construct #7 (EP/S1-GT/AS, lanes a-j) was digested with KpnI (*top*) or EcoRI (*bottom*). B) Genomic DNA from 3 independent Ψ2 clones transfected with construct #8 (EP/P2-GT/AS, lanes k-m) was digested with KpnI (*top*) or EcoRI (*bottom*). Digested DNA was subjected to Southern analysis using a radiolabeled *neo* probe. Molecular weight marker: BstEII digested λ DNA.

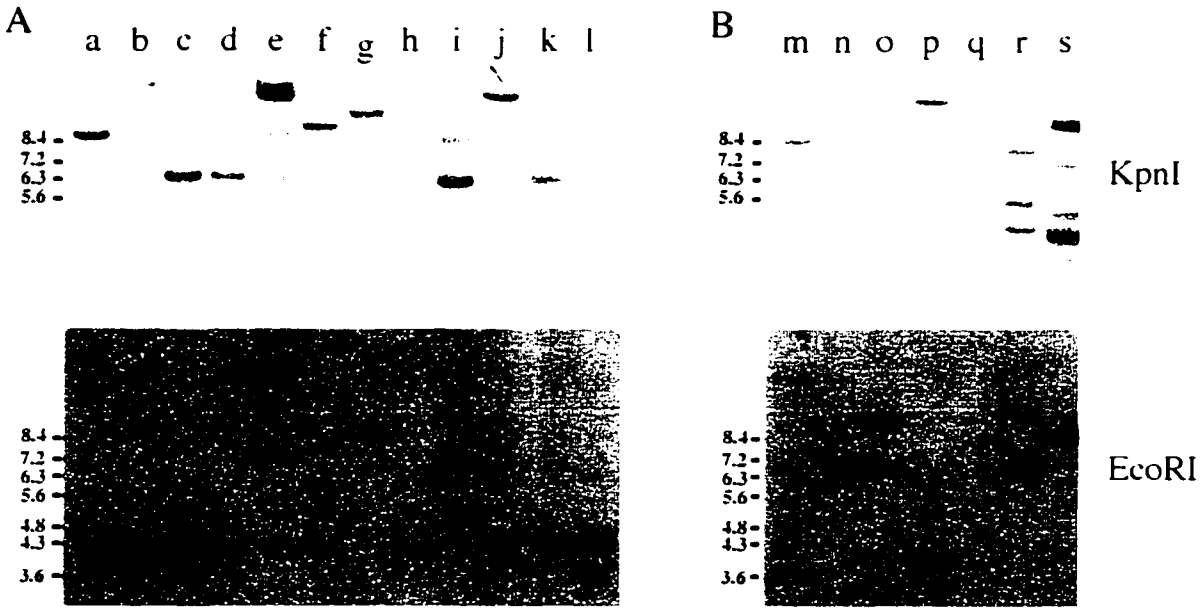


Figure 3. *Southern Blot Analysis of Gene Trap virus Ψ2 Producer Lines*

A) Genomic DNA (10 μg) from 12 independent Ψ2 clones transfected with construct #4 (E/P1-GT/AS, lanes a-l) was digested with KpnI (*top*) or EcoRI (*bottom*). B) Genomic DNA from 7 independent Ψ2 clones transfected with construct #5 (EP/P1-GT/S, lanes m-s) was digested with KpnI (*top*) or EcoRI (*bottom*). Digested DNA was subjected to Southern analysis using a radiolabeled *neo* probe. Molecular weight marker: BstEII digested λ DNA.

e and k), two (3B *bottom* lanes o and r) or three (3A *bottom* lane i) EcoRI *neo* fragments.

3.3.2 *Viral Titer*

Single or low copy Ψ2 transfectants were selected and subsequently analyzed for their infectious titer on both NIH 3T3 cells and D3 ES cells (see Materials and Methods) These results are shown in Table 1 (Gene Trap Vectors) and Table 2 (Promoter Trap Vectors). No significant difference was observed in the infectious titer obtained with the different constructs. Both internal promoters used, *HSV-tk* and *PGK*, proved equally efficient to drive *neo* expression in D3 ES cells. Similarly, no difference in the infectious titer was observed when a polyadenylation signal was placed downstream of the *neo* gene (in the antisense constructs) or when no polyadenylation signal was used at all (in the sense constructs). With all the retroviral constructs, producer clones were generated which produced viruses. Viral titers were as high as 10^5 G418-resistant CFU/ml. Infectious G418-resistant titers, when determined in ES cells, were consistently 10 to 100 fold lower than those in 3T3 cells, in agreement with previously published data (110). These titers were, however sufficiently high to efficiently infect ES cells.

3.3.3 *Trapping Efficiency*

In order to determine the fraction of G418 resistant clones that was able to express *lacZ*, parallel dishes were infected and G418 resistant clones were grown, fixed, and stained with X-gal.16 Gene

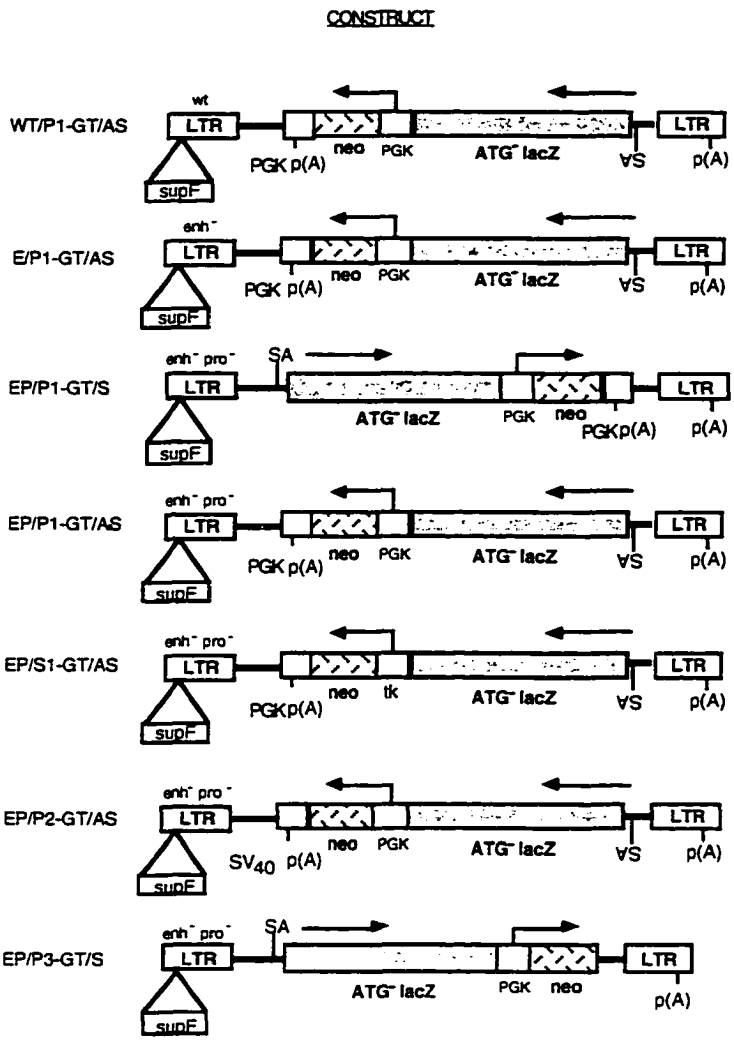


Table 1. *Infectious Titer of Ψ2 Clones Stably Transfected with Gene Trap Vectors*

LTR: long terminal repeats, enh⁻pro⁻: deletion of Mo-MLV enhancer and promoter sequences, S: sense orientation. AS: antisense orientation, SA: splice acceptor site from Mo-MLV env gene, *lacZ*: E. coli *lacZ* gene, *neo*: Tn5 neomycin resistance gene, tk: Herpes simplex virus thymidine kinase promoter, PGK: phosphoglycerokinase promoter, *supF*: bacterial suppressor F gene, p(A): polyadenylation signal.

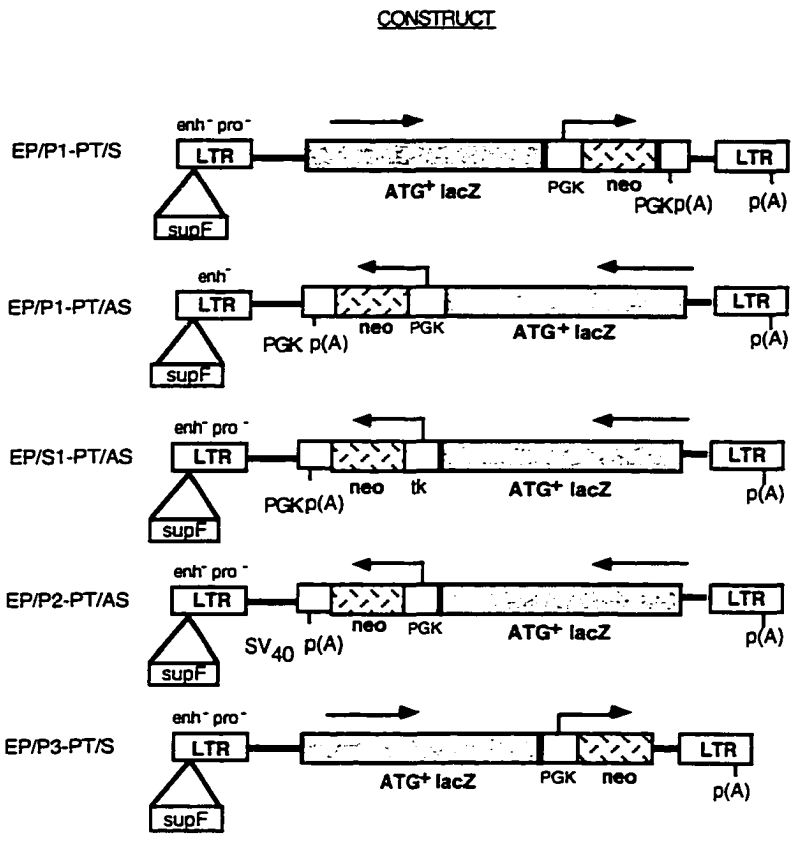


Table 2. *Infectious Titer of Ψ2 Clones Stably Transfected with Promoter Trap Vectors*

LTR: long terminal repeats, enh⁻pro⁻ : deletion of Mo-MLV enhancer and promoter sequences, S: sense orientation. AS: antisense orientation, *lacZ*: E. coli *lacZ* gene, *neo*: Tn5 neomycin resistance gene, tk: Herpes simplex virus thymidine kinase promoter, PGK: phosphoglycerokinase promoter, *supF*: bacterial suppressor F gene, p(A): polyadenylation signal

trap vectors and 9 promoter trap vectors were analyzed. In a typical experiment, ES cells were overlaid with serial decimal dilutions of virus-containing supernatants, so as to generate from 10^2 to 10^4 G418 resistant colonies/10cm dish. None of these infected cell clones was able to express β-galactosidase. This was an unexpected result because between 1 and 10% of the G418 resistant clones were shown to express β-galactosidase using similar double expression vectors (31, 35). Lack of *β*-galactosidase expression with our promoter and enhancer trap vectors could have resulted from recombination of vector plasmid DNA during transfection, or transcriptional inhibition of the *lacZ* gene by retroviral sequences.

3.4 Recombination within the Viral Genome

Recombination of plasmid DNA could have occurred during transfection of the constructs into Ψ2 cells. Such events could lead to the generation of an integrated provirus containing either 5' and 3' enhancer- and promoterless LTRs or 5' and 3' wild type LTRs. I ruled out the first possibility because Ψ2 clones containing proviruses with both LTRs deleted for viral enhancer and promoter would result in lack of transcription of viral RNA, and therefore, no viral particles would be produced. However, as shown in Tables 1 and 2, the Ψ2 clones do produce viruses. The second type of recombination event, on the other hand, would result in the generation of viruses with 5' and 3' wildtype LTRs. Such viruses would still be able to transduce *neo* resistance to infected ES cells since the PGK promoter is active in those cells. However, *lacZ* expression would be inhibited due to methylation and inactivation of the wildtype LTR, as has been shown

previously (51). To test for this possibility, I analyzed DNA from Ψ2 producer clones for the presence of wildtype LTRs by Southern blot. XbaI cuts the proviral DNA in the U3 region of the LTR at a site (position 8113) that is deleted in both enhancer⁻ and enhancer⁻/promoter⁻ constructs. Therefore, the size of the XbaI restriction fragment is diagnostic of the recombination event. Clone 2D, a WT/P1-GT/AS clone which contains 5' and 3' wild type LTRs, was included as a control. This clone produced a genome-length 7.4 Kb XbaI fragment (Figure 4, lane b). A XbaI fragment with a provirus-length size (7.2 Kb) was observed in only one (10K, an E⁻P⁻/P1-PT/AS clone) out of twelve tested Ψ2 producer clones, suggesting that a conversion of the enhancer⁻/promoter⁻ 3' LTR into a wildtype LTR took place in that clone (Figure 4, lane o). Therefore, this could not account for the inability to transduce *lacZ* into ES cells observed in every retroviral vector. Besides, I could not detect β-galactosidase activity in transfected cells as was expected in transfected cells containing a construct where the wildtype LTR drives *lacZ*.

3.5 Transcriptional Inhibition Mediated by Retroviral Sequences

All our vectors contain approximately 1.4 Kb of viral sequences 3' of the 5' LTR, including 1 Kb of *gag* sequences. I tested whether viral sequences contained in our vectors were inhibiting transcription of the *lacZ* gene. This possibility was suggested by the fact that a number of our promoter trap constructs were unable to express b-galactosidase when transiently transfected into NIH 3T3 cells, where the viral LTR is known to have strong promoter activity.

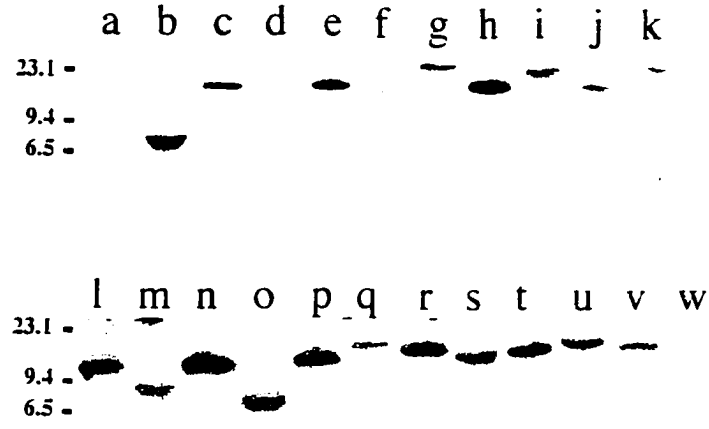


Figure 4. *Southern Blot Analysis of Gene and Promoter Trap Virus Ψ2 Producer Lines*

Genomic DNA (10 μg) from Ψ2 producer clones was digested with XbaI and subjected to Southern analysis using a radiolabeled *neo* probe. Control lanes contained plasmid DNA from the different constructs. Lanes **a**: construct #2 WT/P1-GT/AS; **b**: clone 2D; **c**: construct #4 E/P1-GT/AS; **d**: clone 4C; **e**: construct #6 EP/P1-GT/AS; **f** and g: clones 6E and 6J; **h**: construct #7 EP/S1-GT/AS; **i**: clone 7A; **j**: construct #8 EP/P2-GT/AS; **k**: clone 8E; **l**: construct #9 EP/P1-PT/S; **m**: clone 9B; **n**: construct #10 EP/P1-PT/AS; **o**: clone 10K; **p**: construct #11 EP/S1-PT/AS; **q**: clone 11D; **r**: construct #12 EP/P2-PT/AS; **s**: clone 12K; **t**: construct #13 EP/P3-PT/S; **u**: clone 13H; **v**: construct #14 EP/P3-GT/S; **w**: clone 14F. Molecular weight marker: HindIII digested λ DNA.

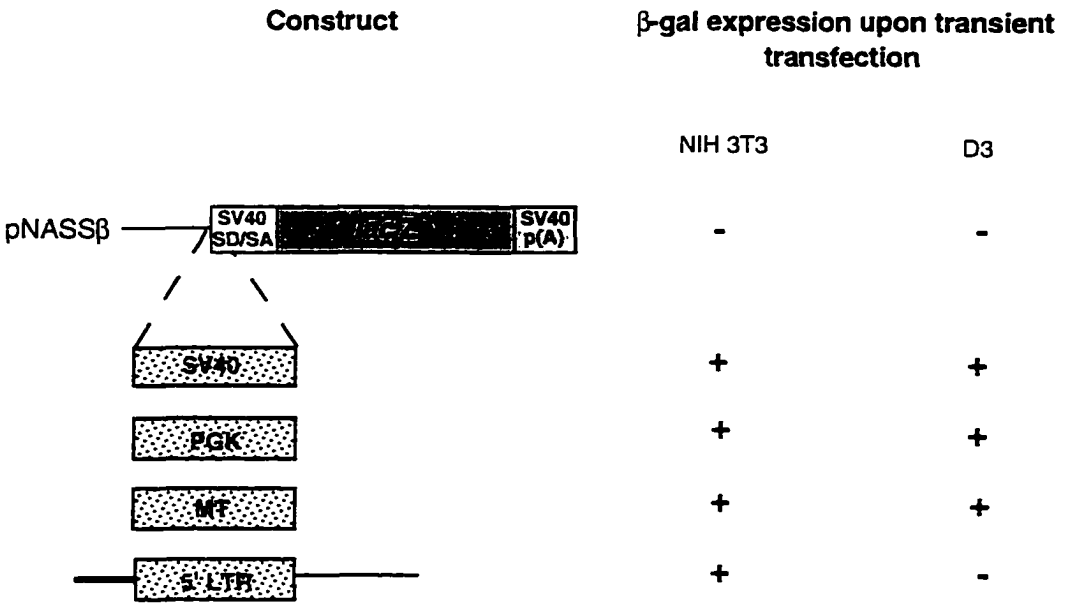


Figure 5. *β-gal Expression in 3T3 and D3 ES Cells Transiently Transfected with pNASSβ-Derived Constructs*

SV40 SD/SA: a 180 bp fragment containing the 16s/19s late gene boundary, including the splice donor and splice acceptor signals
SV40 p(A): polyadenylation signal from SV40 early region.

To examine this possibility, a series of pNASSβ-derived plasmids was constructed which contained the *lacZ* gene under the transcriptional control of the following promoters (Figure 5), mouse PGK, simian virus 40 (SV40) early promoter, mouse methallothionine (MT) gene promoter and a Mo-MLV fragment including the 5' LTR plus adjacent mouse and viral sequences. The Mo-MLV fragment was identical to that used as a 5' LTR to construct the entrapment vectors. These constructs were introduced into both NIH 3T3 and D3 ES cells by the Calcium phosphate co-precipitation method. 36 h after transfection, the cells were stained with X-gal for transient b-galactosidase expression. All four tested promoters were able to drive *lacZ* expression under these conditions in 3T3 cells, and all of them, except for the 5' LTR-pNASSβ, expressed β-galactosidase in D3 ES cells. This result showed that the viral 5' LTR was an active promoter in 3T3 cells and that viral sequences downstream of the 5' LTR did not have any inhibitory effect on reporter gene transcription. Therefore, the viral sequences could not be responsible for the lack of β-galactosidase expression observed with the retroviral constructs. The 5' LTR-pNASSβ construct includes a 180 bp SV40 fragment containing the late protein gene 16s/19s splice donor and splice acceptor signal upstream of the β-galactosidase gene. This SV40 intron is required for *lacZ* expression in this context and its presence may stabilize the message, and thus promote translation. The promoter and gene trap constructs lacked this 180 bp fragment. Therefore, it was possible that inclusion of this signal would be necessary for *lacZ* expression in our retroviral constructs. This was later confirmed, when a new PT-LZ vector (PT-LZ/SV40) including

the SV40 intron was costructed and transfected into producer cells. While producer cells transfected with PT-LZ/SV40 were able to express β-gal (not shown), PT-LZ/SV40 viruses, which lacked the intron as a result of splicing occurred in the producer cells, were unable to tranduce *lacZ* into infected cells.

3.6 Construction of a Promoter Trap Vector with *PLAP* as a Reporter Gene

A Promoter trap vector was constructed, PT-AP, which contained the human *Placental Alkaline Phosphatase (PLAP)* gene as a reporter gene (Figure 6). I chose *PLAP* because it has proven to be at least as useful as the *lacZ* gene in the context of retroviral vectors with regards to high viral titer, stability of expression and identification of expressing cells *in vivo* (33-35). In addition, it does not seem to show the variability in expression previously reported for *lacZ* (37, 65). PGK was the internal promoter used, and the *PLAP-PGKneo* cassette was inserted in the sense orientation relative to the retroviral LTRs, a vector configuration that had been shown to be suitable for the production of infectious viruses. The construct was tested for *PLAP* expression upon transient transfection, in NIH 3T3 cells and D3 ES cells. For this, 10μg supercoiled DNA were transfected by calcium phosphate coprecipitation into cells. Transfected cells were stained with X-P/NBT for *PLAP* expression. Transfected cells showed expression of the reporter gene. The PT-AP construct was then tested for reporter gene expression upon stable transfection in 3T3 cells. Stable transfectants were generated by electroporating 5μg

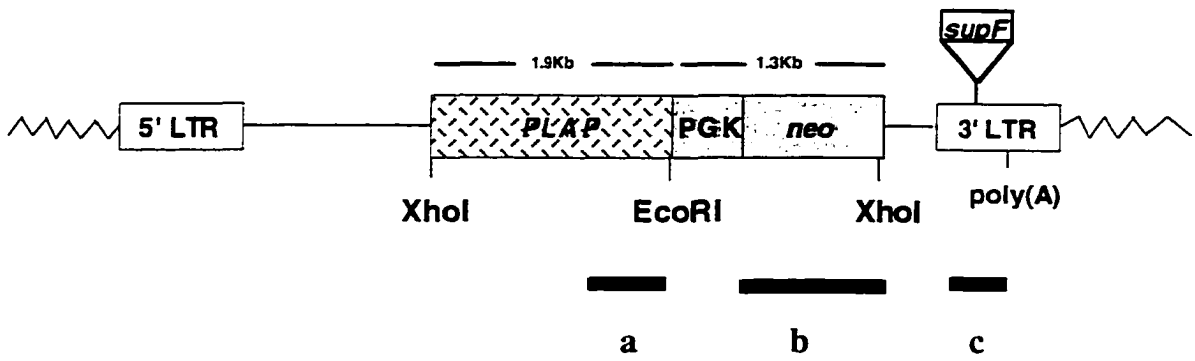


Figure 6. *Structure of the PT-AP Vector*
Thick bars depict the relative positions of the *PLAP* (a), *neo* (b) and *supF* (c) probes. Thin bars indicate the sizes of the *PLAP* (1.9Kb) and PGK*neo* (1.3Kb) cassettes

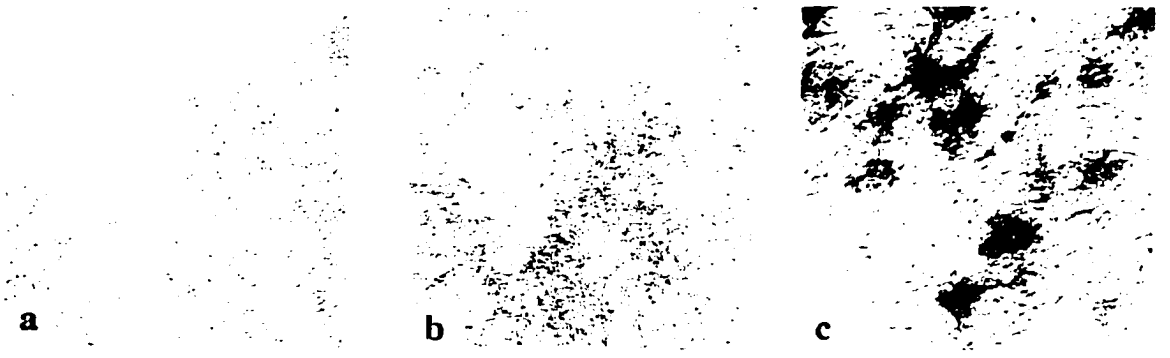


Figure 7. *Reporter Gene Expression in NIH 3T3 Cells Stably Transfected with the PT-AP Vector*
A) Mock transfected clone. B) clone transfected with *PT-AP* DNA. C) Detail of clone in B) showing perinuclear localization of *PLAP* staining. The picture in 7C correspnds to the border of a *PLAP* positive clone, in a confluent (aproximately 10^4 G418 resistant clones). *PLAP* positive cells are mixed with neighboring (negative) ones.

BglII linearized plasmid DNA. The conditions used in this experiment were previously described (see Establishment of Virus-Producing Cell Lines). G418-resistant clones were stained for *PLAP*. Reporter gene expression was observed in stably-transfected cells as well. (Figure 7). *PLAP* expression was found to be perinuclear (Figure 7C), in agreement with the apparent membrane association of *PLAP* previously reported (35). These results showed that the constructs were able to confer *neo* resistance and to express the reporter gene upon both transient and stable transfection. Stable ΨCRE packaging lines were generated by electroporation of BglII linearized PT-AP DNA. ΨCRE cells (27) were generated by sequential transfection with two Y defective Moloney Leukemia Viruses. These cells were shown to be safer than other packaging cell lines with regards to helper virus production. ΨCRE cells were electroporated and passaged into selection medium as previously described.

3.7 Characterization of PT-AP Vector

3.7.1 *Virus Titer*

Transfected ΨCRE clones were pre-screened for their ability to produce infectious viruses in small-scale infections. 3T3 cells were infected on 6-well dishes with undiluted samples of virus-producing cell supernatant. This method allowed me to pre-screen larger numbers of transfected clones, before testing them for the actual virus titer. 24 ΨCRE clones were picked after growing the transfected cells in selection medium for 10 days. These clones were expanded for DNA preparation and virus-supernatant production. Aliquots

were frozen down at -80°C. Six PT-AP clones were found to produce infectious viruses at high titer (not shown).

3.7.2 *DNA Analysis*

DNA from the PT-AP ΨCRE producer clones was analyzed by Southern blot for the number and integrity of the integrated retroviral constructs. EcoRI or KpnI digested DNA was size fractionated on agarose gels and blotted onto nitrocellulose membranes. Membranes were hybridized to a *neo* probe. In all six clones, a 5.6 Kb KpnI fragment was detected, corresponding to the intact retroviral construct. (Figure 8A). When DNA was digested with EcoRI (Figure 8B), one *neo*-hybridizing fragment was detected in one clone (clone #4), two fragments were detected in four clones (#3, #6, #11, #18), and 6 fragments were detected in clone (#19), indicating that these clones had one, two and six integrated plasmids respectively.

3.7.3 *Trapping Efficiency*

To further characterize these clones, supernatants were titrated on both 3T3 and R1 ES cells. Three of these clones produced viruses with titers in the order of 10^4 G418 resistant CFU/ml. When parallel plates of infected cells were stained for *PLAP*, positive clones were found at a frequency of ~1%, as shown in Table 3.

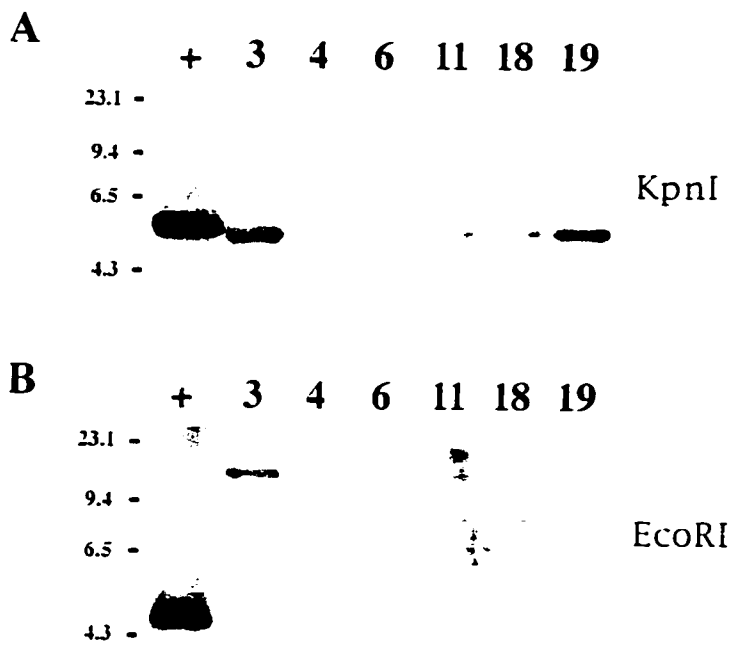


Figure 8. *Southern Blot Analysis of PT-AP Virus ΨCRE Producer Lines*

Genomic DNA (10 μg) from six ΨCRE independent clones (3, 4, 6, 11, 18 and 19) was digested with KpnI (A) or EcoRI (B). Digested DNA was subjected to Southern analysis using a *neo* probe. 10 pg of PT-AP plasmid DNA were used as a control (lane +). Molecular weight marker: HindIII digested λ DNA.

Ψcre CLONE	3T3		ES	
	AP-positive / total G418^r	%	AP-positive / total G418^r	%
3	20 / 2,000	1	2 / 200	1
11	9 / 300	3	6 / 500	1.2
19	18 / 1,500	1.2	7 / 800	0.9

Table 3. *Trapping Efficiency of PT-AP Viruses*
Supernatants from clones 3, 11 and 19 were used to infect NIH 3T3 and R1 ES cells. Appropriate dilutions of infected cells were plated in selection medium and allowed to grow for a week. Resistant clones were stained with Giemsa solution and counted. Duplicate dishes were prepared and stained for *PLAP* activity.

CHAPTER 4

IN VITRO SCREENING

4.1 Generation of PT-AP ES Cell Lines

PT-AP clone #3 was selected as the PT-AP virus producing cell line. The virus titer was $2x10^2$ G418 resistant CFU/ml as measured on R1 ES cells (Table 3). This titer was 10 times lower than that measured on 3T3 cells ($2x10^3$ G418 resistant CFU/ml), consistent with previous reports (110). To improve the efficiency of the infection, R1 ES cells were infected by cocultivation with virus producing cells. In two independent experiments, 10^5 (A) and 10^6 (B) R1 ES cells were cocultivated with $5x10^6$ mitomycin C-treated producer cells. The target cell (R1 cells)/virus producer cell ratio in these experiments was 1/50 and 1/5, respectively. After 72 (A) and 48 h (B) of cocultivation, cells were trypsinized and plated into selection medium at decreasing dilutions (1/5, 1/10, 1/20, 1/50), One week later, 512 clones were picked and transferred into 96-well plates, expanded and frozen down at -80°C.

4.2 *PLAP* Expression in Differentiating ES Cell Pools

The first phase of the screen consisted of identifying promoter trap ES cell lines with differential *PLAP* expression upon differentiation of ES Cells *in vitro*. Cells were induced to differentiate *in vitro* by withdrawing LIF from the culture medium and preventing cell attachment to the culture dish surface. After 4 days of culture in suspension, ES cells formed simple embryoid bodies (EBs) consisting of an outer layer of primitive endoderm-like cells, an inner layer of primitive ectoderm-like cells and a basal lamina in

between them. Simple EB formation recapitulates *in vitro* the primitive ectoderm-primitive endoderm lineage separation that occurs *in vivo* shortly after implantation. After 6 days in suspension, EBs further differentiate into cystic embryoid bodies, which resemble to a certain extent the egg-cylinder stage embryo. They develop a visible cavity, as well as blood islands and contracting cardiomyocytes. In addition, simple EBs can be transferred back to gelatinized tissue culture dishes where reattachment to the surface triggers outgrowth and terminal differentiation of cells. Differentiated cells were observed following 4 to 7 days in culture, including hematopoietic cells, contracting cardiomyocytes, fibroblasts and neuronal like-cells.

4.3 Identification of PT-AP ES Cell Clones with Differential *PLAP* Expression

Infected ES cells were thawed out into 96-well plates and expanded into 24-well plates. When cells reached 50% confluence, they were trypsinized, pooled in groups of 12 individual clones, and plated onto 10cm gelatinized dishes without feeder cells in EB medium. Two days later, small clumps of differentiating ES cells were observed in the dishes. Clumps were briefly trypsinized and transferred to bacterial 10 cm dishes. Simple EBs could be seen after 2 days in suspension. Aliquots were taken after 4 days and 7 days. EBs were fixed in paraformaldehyde and stained for *PLAP*. 40 pools each of 12 ES cell clones were generated and analyzed. EBs were produced from 3 pools (ID, IE and IVG) which showed expression of the reporter gene after differentiation. In any given pool, EB staining

patterns varied. Positive pools had 10 to 20% positive EBs, as was expected from a pool composed of 12 ES clones. Additionally, a distinctive pattern of *PLAP* staining was observed within each individual EB. Since ES cells were plated on gelatinized dishes as single cells prior to the initiation of the suspension culture, each EB should be derived from individual ES cells. Therefore, restricted *PLAP* staining within EBs is expected to be a reflection of the expression of the trapped gene in that particular ES clone.

Individual clones belonging to positive pools of ES cells were analyzed for *PLAP* expression after *in vitro* differentiation. 5 promoter trap ES clones were originally identified which showed *PLAP* expression at different stages of differentiation (Table 4 and Figure 9). 2 of the pools (pool IIIC and IVE) contained 2 positive clones each. One of the clones, IIIC4, identified as a ES cell clone with *PLAP* expression in 4-day EBs, failed to display a reproducible pattern in repeated experiments, and was not used in further experiments. 3 clones (IIIC5, IIID7 and IVE35) showed no *PLAP* activity in the undifferentiated stage. One clone, IVE38, showed weak, uniform staining in undifferentiated cells. In all of them, *PLAP* staining was clearly detected after inducing cells to differentiate. In every case, staining was reproducibly observed in more than 50% of the stained EBs.

Expression of the *PLAP* reporter gene was clearly observed in 4-day and 7-day EBs (IIIC5 and IIID7) and was restricted to a few small groups of cells inside the EBs. Clone IVE35 displayed a much different pattern of expression. Undifferentiated cells did not express *PLAP* at all. 4 day EBs showed "patchy" expression, restricted to

small groups of cells. In 12-day EBs, groups of *PLAP* expressing cells extended across the entire EB. Terminally differentiated cells, on the other hand, were almost completely negative for *PLAP* expression. Clone IVE38 represented the most remarkable example of differential expression upon EB formation. Undifferentiated cells stained uniformly but weakly, displaying a pale purple tint clearly distinguishable from background staining. 4-day EBs exhibited a dramatic increase of *PLAP* expression. Positive cells were observed throughout the EBs. Sections of *PLAP*-stained EBs, however, revealed that staining was confined to the internal mass of cells, which is the EB equivalent of the primitive ectoderm in the egg-cylinder embryo (not shown). In 12-day EBs, the domain of reporter gene expression was further restricted to internal groups of cells, presumably derived from primitive ectoderm-like progenitors. Finally, no staining was observed in terminally differentiated cells. Cystic EBs had an elaborate and thicker outer layer of cells, many of them presenting complex ball-like expanded structures. Whole-mount (Figure 9) as well as sectioned stained EBs (not shown) clearly showed that *PLAP* expression was completely excluded from this domain. This outer layer is equivalent to the primitive endoderm in the early post-implantation embryo which will give rise to extraembryonic tissues in the developing embryo. The embryo proper, on the other hand, is entirely derived from the primitive ectoderm layer.

PLAP staining upon in vitro differentiation

CLONE	Undifferentiated cells	4 days	7 days		12 days	Terminally differentiated
III C4	+	ND	ND			
III C5	-	-/+	-/+	-/+	-	
III D7	-	-/+	+	-/+	-/+	
IV E35	-/+	+	++	++	-/+	
IV E38	+	+++	++	+	-	

Total =480

Table 4. *PT-AP ES Cell Clones with Differential PLAP Expression* Summary of the screening. A total of 500 clones were analyzed. Five clones were isolated which showed differential *PLAP* expression

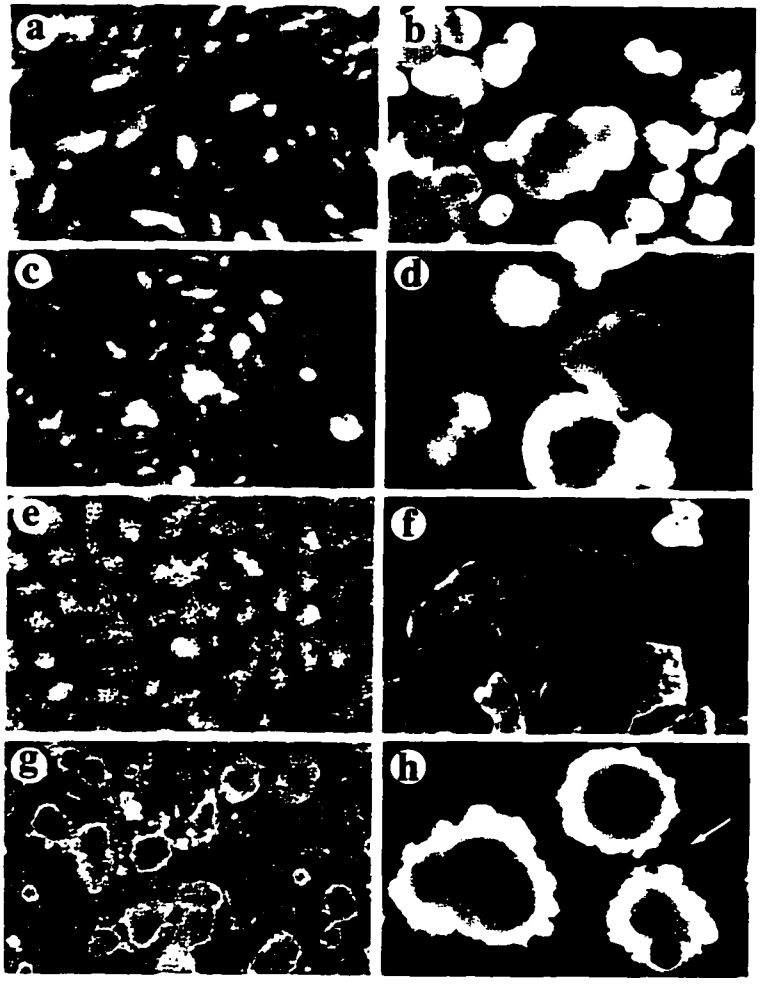


Figure 9. *Differential PLAP Expression in ES cell Clones with PT-AP Insertions*

ES cells from individual clones were induced to differentiate into embryoid bodies (EBs). Aliquots from EBs were taken at different times of differentiation, fixed and stained for AP activity and photographed using a Nikon SMZ-U strereomiscroscope. Undifferentiated ES cells and day-7 EBs from PT-AP clone IIIC5 (a, b); PT-AP clone IIID7 (c, d); PT-AP clone IVE35 (e, f) and PT-AP clone IVE38 (g, h). The outer layer of primitive endoderm-like cells in day-7 EBs from IVE38, shows no *PLAP* activity (h, white arrow). *PLAP* expression is restricted to the inner layer of primitive ectoderm-like cells (h, black arrow)

4.4 Characterization of PT-AP ES Cell Clones

4.4.1 *PT-AP ES Cell Clones Contain Single or Low Copy Number of Proviruses*

I performed Southern blot analysis of genomic DNA isolated from the promoter trap ES cell lines to characterize them for the number of integrated proviruses and the site of integration into the genome. I took advantage of a unique EcoRI site within the retroviral vector located at the *PLAP*-PGK *neo* junction (Figure 6). Therefore, DNA from infected cells digested with EcoRI produces one 5'-specific and one 3'-specific fragment per integrated provirus, each containing viral and flanking cellular sequences. These fragments can be detected by probing the Southern filters with a 5' probe (*sup F*), or a 3' probe (*neo*).

Clones IIIC4, IIIC5 and IIID7 contain one integrated provirus. The sizes of the EcoRI 5' integration fragments were 4.0 Kb, 4.2 Kb and 4.2 Kb, respectively (Figure 10A, *lanes a, b and c*; for clone IIID7 see also Figure 11). The sizes of the 3' fragments were 8.5 Kb (IIIC4), 3.0 Kb (IIIC5) and 14.0 Kb (IIID7) (Figure 10B, *lanes a, b and c*). ES cell clones IVE35 and IVE38, obtained by cocultivation with PT-AP virus producing ΨCRE cells, contained more than one integrated retroviral vector per genome. DNA analysis revealed the presence of two proviruses in clone IVE35 (Figure10A, *lane d* and Figure 10B, *lane d*) and three proviruses in clone IVE38 (Figure 10A, *lane e* and Figure 10B, *lane e*). Furthermore, an additional band of sub molar intensity is visible in DNA from clones IIIC5 and IIID7. These additional bands probably reflect infection with multiple virus particles during cocultivation.

The *sup F* probe could actually detect both 5' and 3' integration sites, because infected cells contain a copy of the *sup F* gene in both proviral LTRs. The *sup F* probe was found to produce a weak hybridization signal under the condition used for the Southern analysis, probably due to its small size of 200 bp (Figure 10B). I decided to utilize a full length (1.9 Kb) *PLAP* probe as a 5' specific probe. A potential problem with this probe, however, was that several mouse AP genes exist which have extensive sequence homology with the human *PLAP*. At least two of them are expressed during mouse embryogenesis (45). Sequence comparison analysis of the human *PLAP* with mouse AP genes, performed using the BLAST algorithm, revealed a highly human-specific fragment located at the 3' end of the cDNA. A 380 bp SmaI-SmaI fragment containing 60 bp of coding sequence and ~320 bp of 3' UTR was excised from the plasmid AP-12 (H. Stuhlmann, unpublished), subcloned into a pKS-SmaI vector and used as a source of 5' the specific probe.

To test the full length and the human-specific *PLAP* probes, DNA from uninfected RI cells and from PT-AP clones IIID7 and IVE38 was analyzed by Southern blot. Triplicate filters were probed in parallel with *neo*, the full length *PLAP* and the 3' human-specific *PLAP* probes. The results are shown in Figure 11. At least four bands were detected in the uninfected RI cell lane when the Southern filter was probed with the full length *PLAP* probe (Figure 11B, lane R1). These bands probably correspond to the endogenous mouse AP genes. No bands were observed when the 380 bp *PLAP* probe was used (Figure 11C, lane R1). One 4.0 Kb band was observed in the IIID7 lane and three bands (4.4, 7.0, and 16 Kb) in the IVE38 lane

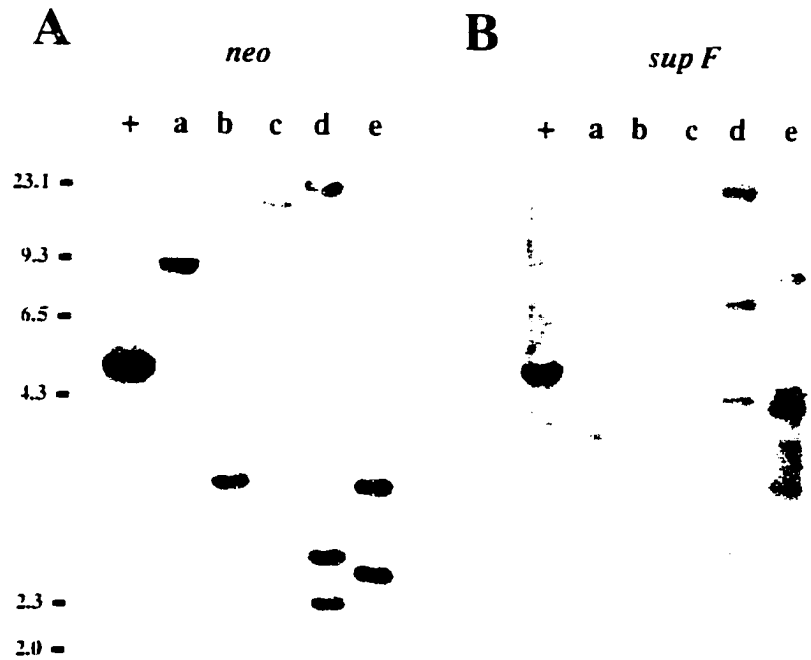


Figure 10. *Southern Blot Analysis of DNA from ES Cell Clones with PT-AP Vector Insertions*
High molecular weight DNA (10 μg) from ES cell clones infected with PT-AP virus was digested with EcoRI and subjected to Southern Analysis, using a ^{32}P-labeled *neo* probe (A) or a ^{32}P-labeled *supF* probe (B). Lanes: **+**. PT-AP plasmid DNA (10 pg); **a**. IIIC4; **b**. IIIC5; **c**. IIID7; **d**. IVE38; **e**. IVE35. Molecular weight marker: HindIII digested λ DNA.

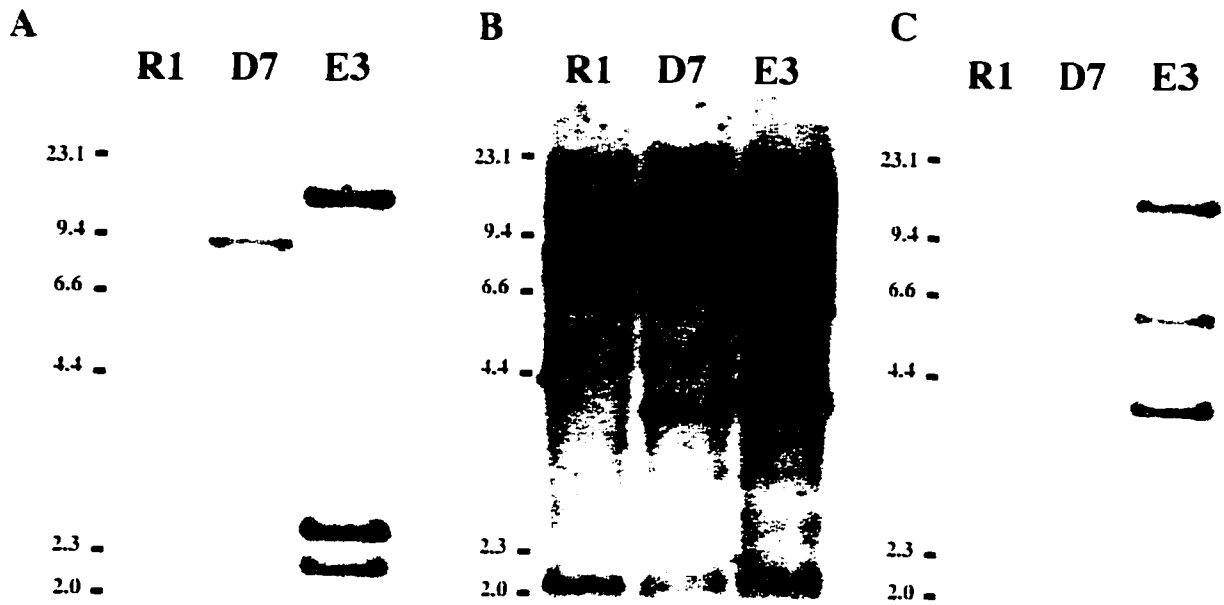


Figure 11. *Southern Blot Analysis Using the Human-Specific PLAP Probe*

Genomic DNA (10 μg) from uninfected R1 cells (R1), PT-AP line IIID7 (D7) and PT-AP line IVE38 (E3) was digested with EcoRI and subjected to Southern Analysis, using a *neo* (A) full length *PLAP* (B) or a human specific *PLAP* (C) probe. Molecular weight marker: HindIII digested λ DNA.

(Figure 11C lanes D7 and E3), confirming that this probe was specific for the human *PLAP* gene. One 14 Kb band was observed for clone IIID7 when the *neo* probe was used (Figure 11A, lane D7) and three bands of 2.3, 2.8 and 16 Kb were observed for clone IVE38 (Figure 11A, lane E3).

4.4.2 *SupF Cloning of IVE35 and IVE38 Host Flanking Sequences*

Three bands were detected with the *neo* probe in the IVE38 lane, corresponding to the 3' integration sites, of 2.3, 2.8 and 16 Kb. Three bands were also detected with the *PLAP* specific probe, of 4.4, 7.0 and 16 Kb (Figure 11A, lane E3 and 11C, lane E3). Two bands were detected with the *neo* probe in the IVE35 lane of 2.5 and 3.0 Kb (Figure 10A lane e) and two additional bands were detected with the *sup F* probe (the 5' integration sites) of 4.0 and 8.2 Kb (Figure 10B, lane e). These two clones had a very interesting pattern of *PLAP* expression *in vitro* (Table 4 and Figure 9). In addition, they showed the most consistent PLAP staining. Over 90% of the EBs generated with both IVE35 and IVE38, displayed the same staining pattern, and the same result was obtained in several independent experiments. Besides, the level of PLAP expression was the highest (at the protein-activity level) among all five PT clones. On the assumption that such high expression could result from translation of an abundant readthrough message (particularly in the case of clone IVE38) I decided to study those clones, in spite of the potential difficulties that, to identify the trapped locus, could arise from the presence of multiple integration events. To further characterize the endogenous loci trapped by the vectors, viral and flanking mouse

genomic DNA was isolated by selective cloning in *sup F* deficient bacteria. A library of EcoRI-digested genomic DNA was constructed for each ES cell clone in the vector λ zap. The library was plated on the non-permissive bacterial host strain MC1061. In a typical experiment, ~10^6 plaque forming units (PFU) were used to infect 600ml (OD600=1) of MC1061 bacteria and plated on one or two 150mm Petri dishes. *sup F* containing recombinant phages gave rise to plaques. The number of plaques obtained varied from 10 to 100 in different experiments. In order to distinguish positive from false positive signals, Southern analysis was performed. Plaques were transferred to nitrocellulose membranes and filter-hybridized to a [α-^{32}P] dCTP-labeled *PLAP* or *neo* fragment. Two positive plaques were detected with the *PLAP* probe and four with the *neo* probe (Figure 12A and 12B) on the IVE38 clone filter. 7 positive plaques were detected on the IVE35 filter when hybridized to both probes (Figure 13). The *sup F+ recombinant* phage number to total phage number ratio was then 2/10^6 and 4/10^6 for clone IVE38 and 3/10^6 for clone IVE35. These values, in the order of 10^{-6}, are consistent with the number of integrated entrapment vectors per genome (two for IVE35 and three for IVE38). A number of plaques were observed on some of the plates which did not hybridized to either probe (70-80 plaques, clearly seen in Figure 13). These false *sup F* positive plaques, usually smaller than their *sup* F^+ counterparts, probably are *sup* F^- revertants, generated during phage replication. Since the *sup* F^- phenotype of λ zap is the result of a single point mutation, a number of revertants is expected. The ratio of false positive to positive plaques ranged from 1 to 10 in different experiments,

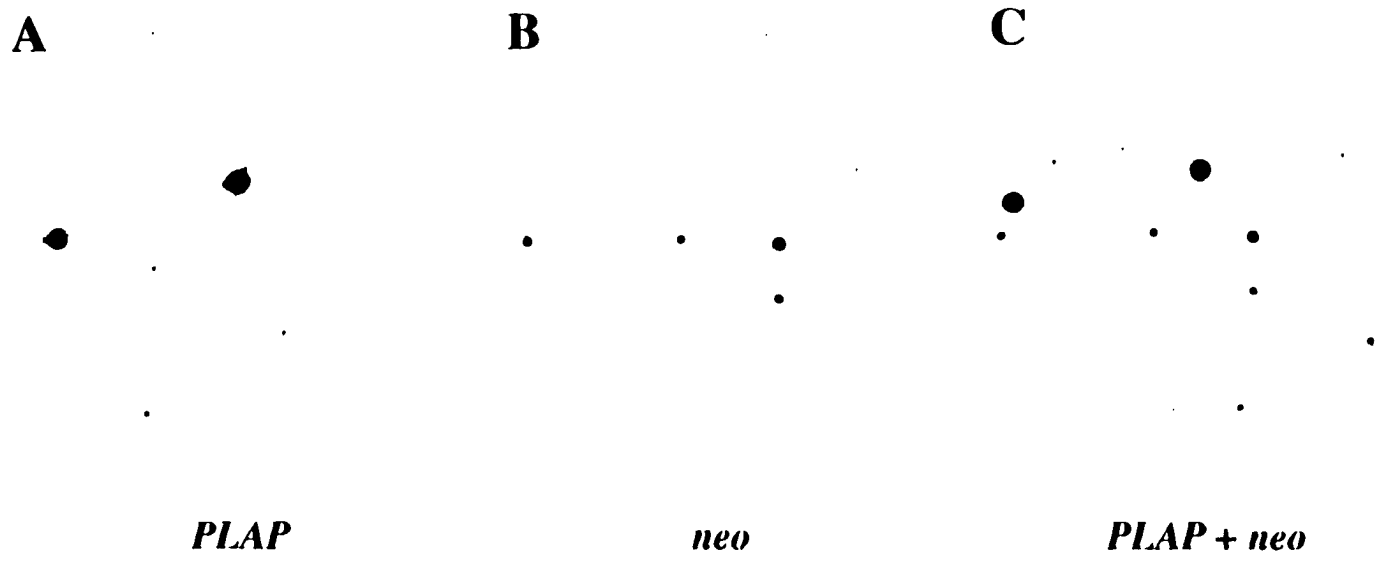
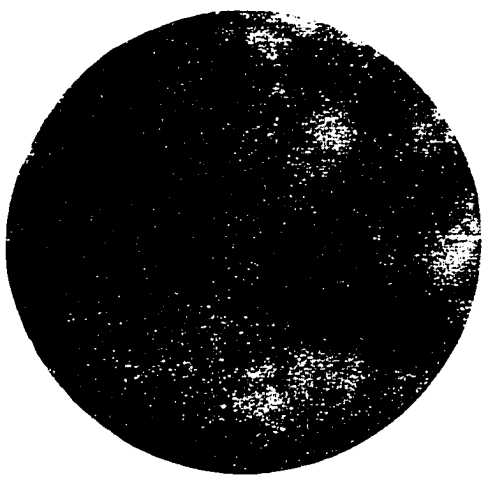


Figure 12. *Sup F Cloning of iVE38 Clone Integrated Proviruses*
A recombinant IVE38-1 zap (EcoRI) library was used to infect MC1061 bacteria. The infected cells were plated on LB plates. Plaques were transferred to nitrocellulose membranes and hybridized to ^{32}P labeled *PLAP* (A), *neo* (B) and *PLAP+neo* (C) probes.



PLAP + neo

Figure 13. *Sup F Cloning of IVE35 Clone Integrated Proviruses*
A recombinant IVE35-λzap (EcoRI) library was used to infect MC1061 bacteria. The infected cells were plated on LB plates. Plaques were transferred to nitrocellulose membranes and hybridized to a mix of ^{32}P labeled *PLAP+neo* probes (10^6 cpm each).

depending upon the batch of bacteria and phage preparation used for the infection. Five colonies obtained from each phagemid rescue experiment were isolated.

Restriction analysis showed that two different inserts had been cloned out from the IVE38 clone (termed C1 and C4), and three from IVE35 (B3, B14, and B2)(Table 5). The identity of these clones was confirmed by Southern blot analysis, using radiolabeled *PLAP*, *neo*, B2 and C4 host flanking probes. B2 host flanking probe was derived from pSK-B2 plasmid by PCR amplification using T7 and U3 primers. C4 host flanking probe was derived from pSK-C4 plasmid by PCR amplification using T3 and U3 primers. DNA from two 5' insertion clones, B2 and C4, and three 3' insertion clones, B3, B14 and C1 was digested with EcoRI and XhoI, size fractionated on agarose gels and blotted onto nitrocellulose filters. In B2 and C4 plasmids, no *neo* hybridizing band was detected (Figure 14A, *left* and 14B, *left*). On the other hand, the *PLAP* probe detected a 1.9 Kb band, corresponding to the 1.9 Kb *PLAP* cDNA (Figure 14A, *center* and 14B, *center*). Cellular flanking probes detected a 2.4 Kb (B2) and 2.5 Kb band (C4) (Figure 14A, *right* and 14B, *right*). Analysis of the 3' insertion clones B3, B14 and C1 produced analogous results. While all three clones generated a 1.3 Kb *neo* band corresponding to the EcoRI-XhoI PGK-*neo* cassette (Figure 15A, 15B and 15C, *left*), no *PLAP* hybridizing band was detected (Figure 15A, 15B and 15C, *right*), confirming that B3, B14 and C1 were 3' integration clones.

Four EcoRI integration fragments from clone IVE38 (two 3', 16 Kb and 2.0 Kb; and two 5', 16 and 7 Kb) could not be cloned out using this approach due to their size. λ zap is an insertion vector that most

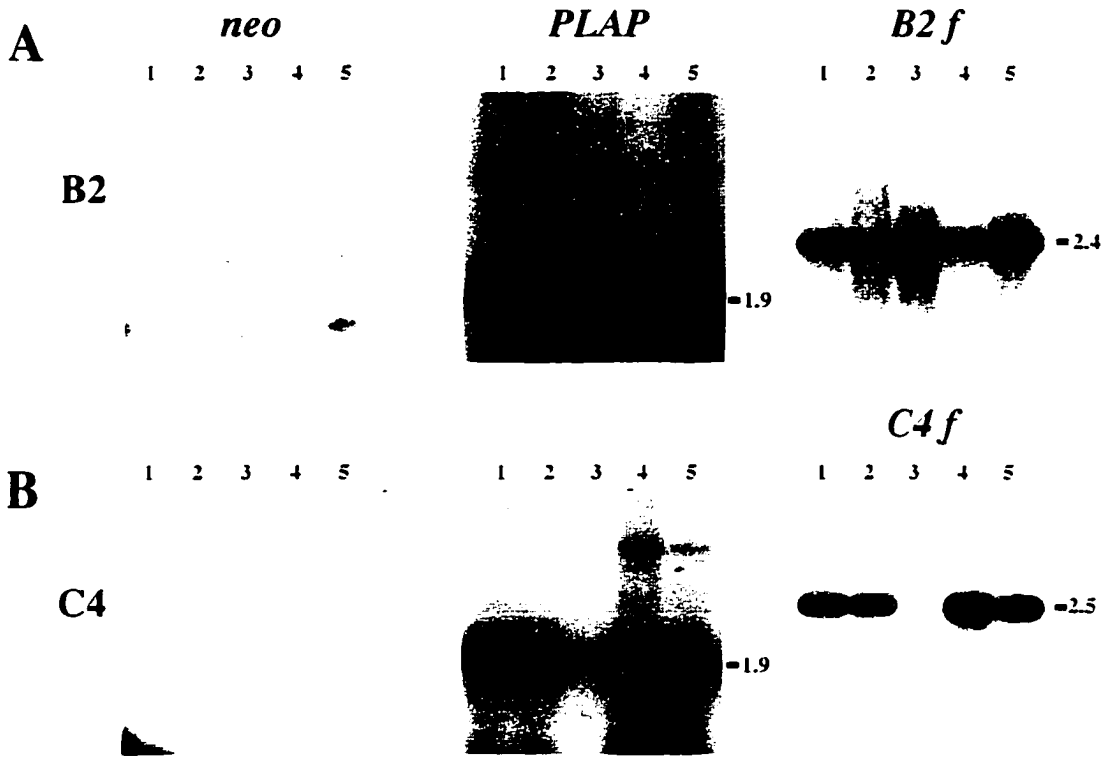


Figure 14. *Restriction Analysis of 5' Integration Clones*
Plaques from B2 (14 A) and C4 (14 B) 5' integration clones were isolated and grown for *in vivo* excision of the *sup F* containing inserts as phagemid particles. SOLR bacteria were "transformed" with phagemids and plated on LB+Amp plates. 5 colonies per clone were picked and grown for DNA preparation. DNA from these 5 colonies (lanes 1 through 5) was digested with EcoRI and XhoI and subjected to Southern analysis using radiolabeled *neo* (14A and 14B, *left panel*), *PLAP* (14A and 14B, *center panel*) and flanking probes (14A and 14B, *right panel*). A 1.9 Kb *PLAP* hybridizing fragment, corresponding to the XhoI-EcoRI *PLAP* cDNA, can be observed (14 A and 14 B, *center*). A 2.4 Kb (B2) and 2.5 Kb (C4) fragment can be detected with the B2 and C4 flanking probes, respectively.

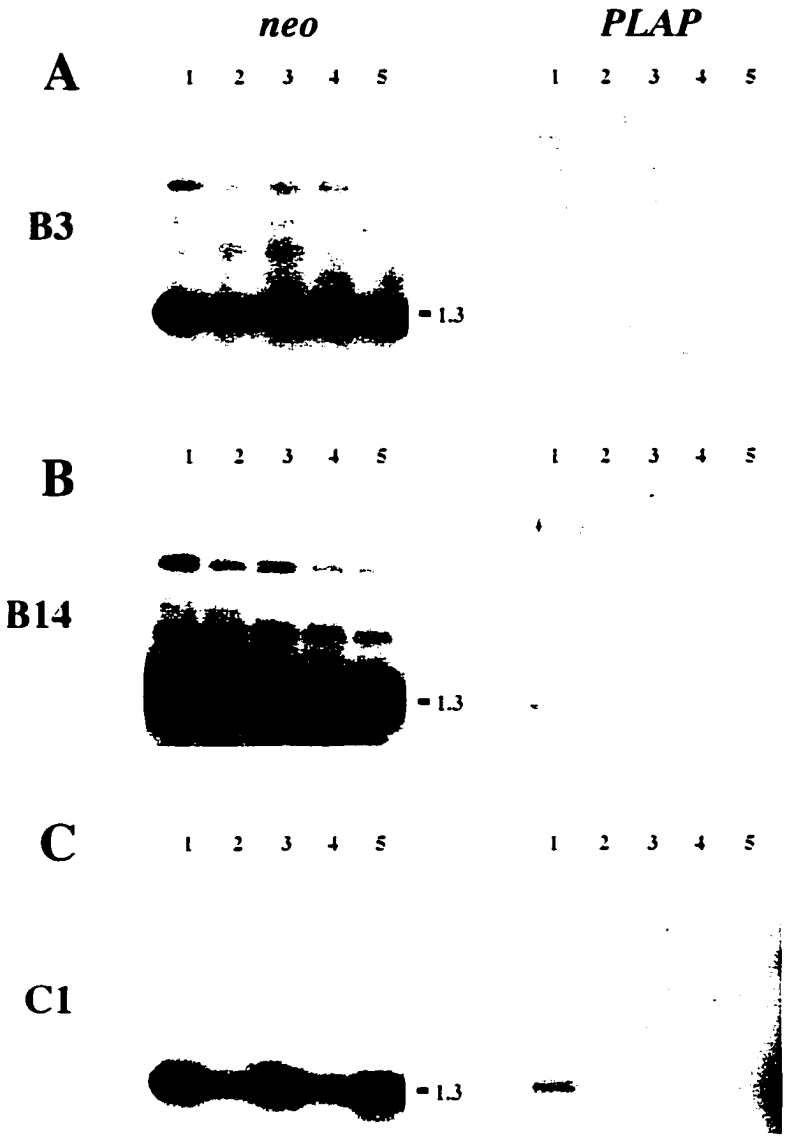


Figure 15. *Restriction Analysis of 3' Integration Clones*
neo positive plaques from B3 (15 A), B14 (15 B) and C1 (15 C) 3' integration clones were isolated and grown for *in vivo* excision of the *sup F* containing inserts as phagemid particles. SOLR bacteria were "transformed" with phagemids and plated on LB+Amp plates. 5 colonies per clone were picked and grown for DNA preparation. DNA from these 5 colonies (lanes 1 through 5) was digested with EcoRI and XhoI and subjected to Southern analysis using a radiolabeled *neo* (14A and 14B, *left panel*) or *PLAP* (14A and 14B, *right panel*) probe. A 1.3 Kb *neo*-hybridizing band, corresponding to the EcoRI-XhoI PGK-neo cassette, could be detected in all 3 clones.

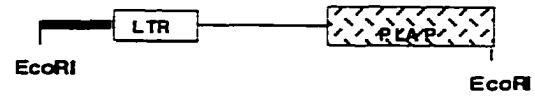
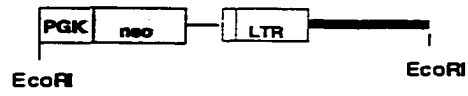
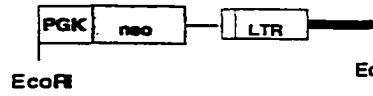
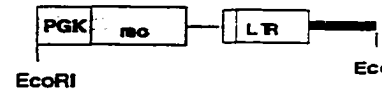
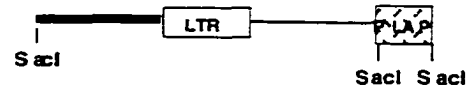
clone	ES-cell clone	flanking sequence size	
B2	IVE3$_5$	0.5 Kb	
B3	IVE3$_5$	1.0 Kb	
B14	IVE3$_5$	0.5 Kb	
C4	IVE3$_8$	0.6 Kb	
C1	IVE3$_8$	0.5 Kb	
C18	IVE3$_8$	0.7 Kb	

Table 5. *Genomic Clones Obtained by sup F Cloning*
Three different *sup F* clones were obtained from PT-AP line IVE35 (B2, B3, B14) and three from IVE38 (C1, C4, C18). These clones contain both viral and cellular flanking (thick lines) sequences.

efficiently takes inserts in the 4-5 Kb size range. Larger inserts generate recombinant genomes that can not be packaged efficiently into virions. To generate smaller *sup F* containing genomic fragments, IVE38 DNA was digested with different combinations of restriction enzymes. EcoRI/SacI digestion produced three 5' integration fragments of 3.3, 2.6 and 2.5 Kb and three 3' fragments of 4.2, 2.3 and 2.1 Kb (not shown). These fragments were in a size range appropriate for cloning in λ zap. A third IVE38 fragment, C18, was cloned into λ zap SacI arms (Figure 16). Upon restriction analysis characterization of clone C18, a 5' integration clone which contains 700 bp of flanking sequence, unexpected results were obtained that revealed the actual configuration of the promoter trap vector. Clone C18 was generated by ligating SacI digested E38 genomic DNA to λ zap SacI arms and was isolated as a positive plaque hybridizing to a *gag* -5' specific- probe (Figure 16 A), which indicated that C18 was a 5' integration clone. SacI digest of the pSK-C18 rescued plasmid produced two inserts of 2.6 and 0.7 Kb, which suggested that the genomic insert was product of partial SacI digestion and that an additional SacI site was present in the *PLAP* fragment. In addition, this 700 bp *PLAP* fragment hybridized to the 3' *PLAP* probe, demonstrating that the *PLAP* cDNA was cloned in reverse orientation relative to the PGK-neo cassette (see Figure 17). The *PLAP* cassette was derived from plasmid Cla12-AP (see Materials and Methods) by partial EcoRI digestion and XhoI linker ligation. The linker was added to the EcoRI site that lies at the very 5' end of the cDNA immediately upstream of a SacI site. This finding, confirmed by further restriction analysis of the PT-AP plasmid DNA (not shown), has consequences on

the examination of the trapping events and will be extensively discussed later.

4.5 Sequence Analysis of the Trapped Loci

Flanking genomic DNA was sequenced using T3 or T7 specific primers and synthetic oligonucleotide primers specific for the U3 region or U5 region of the Mo-MLV LTRs. B14 and B3 are 3' integration sites from clone IVE35 which contain 300 bp and 1000 bp of flanking mouse sequences, respectively (table 5 and Figure 18). B2 is a 5' integration site from clone IVE35 with 400 bp of flanking sequence (Table 5 and Figure 18). Sequence analysis of B14 and B2 revealed no overt homology to any sequence available in the GENEBANK database. The B3 flanking fragment is identical to a mouse 5s RNA gene. C1 is a 3' integration site from clone IVE38 with a 400 bp mouse sequence (Table 5 and Figure 19). C4 is a 5' integration site from clone IVE38 containing 600 bp of flanking genomic sequence (Table 5 and Figure 19). Analysis of integration clones C1 and C4 revealed no sequence homology to any known sequence. Partial sequence analysis of C18 (not shown) revealed that it was different from C4 and had no homology to any sequence available in GENEBANK. Sequence analysis of all integration clones showed that 2 bp from the viral LTRs were lost during integration, as previously reported (119).

4.6 C1 and C4 Integration Clones

To verify that the cloned integration fragments are the host sequences flanking the PT-AP insertions, I performed Southern

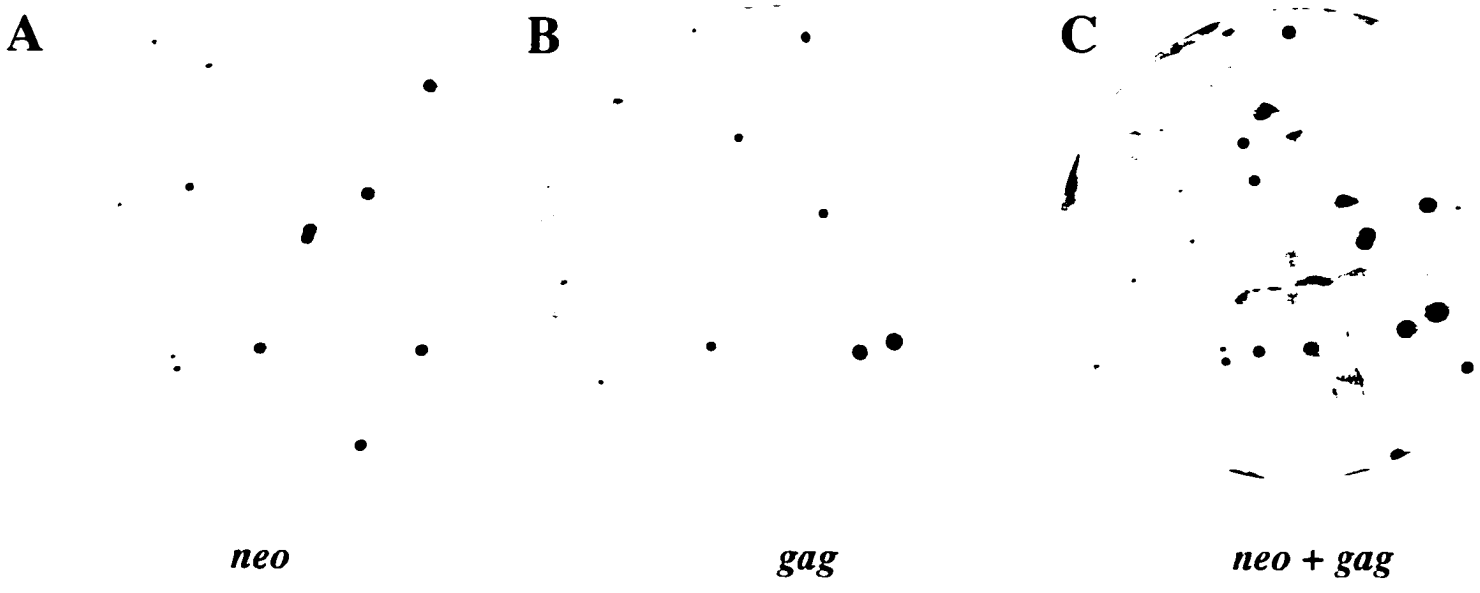


Figure 16. *Sup F Cloning of IVE38 Clone Integrated Proviruses*
A recombinant IVE38-λzap (SacI) library was used to infect MC1061 bacteria. The infected cells were plated on LB plates. Plaques were transferred to nitrocellulose membranes and hybridized to ^{32}P labeled *gag* (A) *neo* (B) and *gag+neo* (C) probes

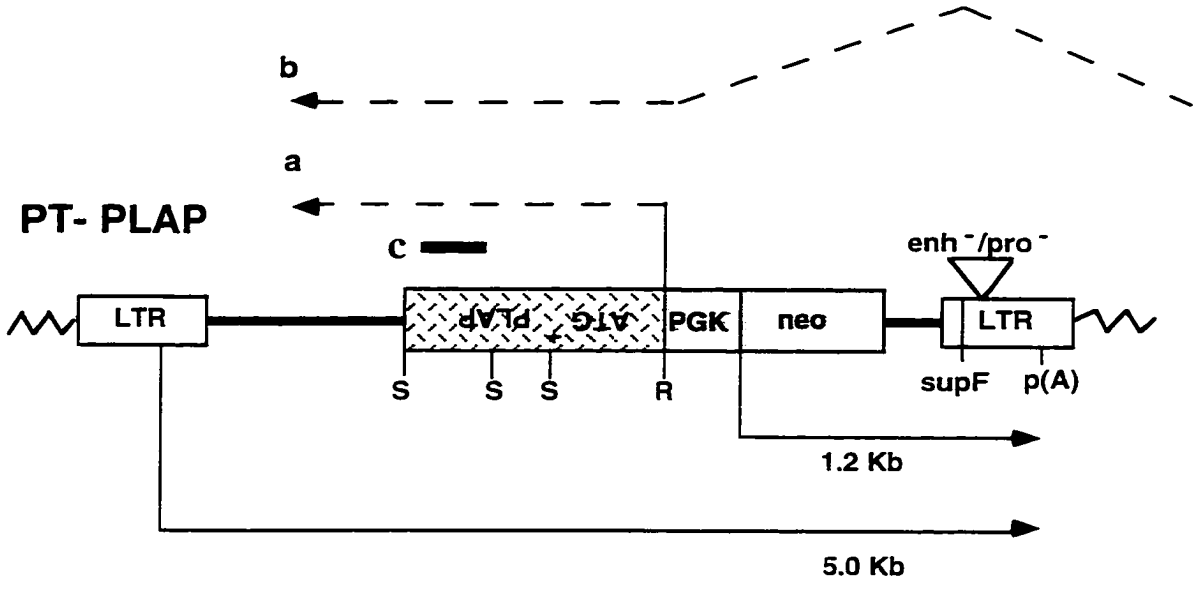


Figure 17. *Structure of the PT-AP Vector*
The actual configuration of the PT-AP vector is depicted. The *PLAP* cDNA was cloned in reverse orientation relative to the PGK-*neo* cassette. The solid line arrows indicate the sizes of the retrovirally-derived transcripts found in producer (5.0 Kb and 1.2 Kb) and infected cells (1.2 Kb). The dotted line arrows, suggest two of the possible mechanisms for PLAP activation in infected cells. The PLAP cDNA could be transcribed from a minimal promoter (the PGK in antisense orientation) under the control of a cellular enhancer (a). Alternatively, the vector could be inserted in an intron, with a cryptic splice acceptor located in the PGK *neo* cassette, accepting splicing from an upstream donor site (b). As a result, the PGK-*neo* cassette would be removed from the message. The thick line (c) depicts the relative position of the human specific *PLAP* probe. R: EcoRI site, S: SacI site.

B14

```
*TCGCGGGGTC TTTCATTAGA AACAGCAGTC TGGAATGAAT GTCACAACCA GCTACCTTTG      60
 CTGGTTAGAT GCTAACAAAG GAACCGTTTG GACTGAATTA CAACTTGAGA AATGAAGCAT
 TTACGATGGT TGTTTTGGGT TTAGATTTTT ATCCCCTGCC AATTATAGTA GTAAAAGCTT     180
 TAGTATTCCG ACCTAATGTA GCCTTTGTTC AGAAATCAGC TTCTTTGTGC AGAGATTTTT
 CTTATGAGTG CTTTTGAATG CATTTGTGTC AGCGCCTTAG TAGTAAAACT GAGCAGAATT C   300
```

B3

```
*TCTTTACAGG GTTGCAGATC TCTTTAGCTC CTTGGATACT TTCTCTAGCT CCTCCACTGG      60
 GGGCCCTGTG CTCCATCCAA TAGCTGACTG TGAGCATCTA CTTATGTGTT TCCTAGGCCC
 CGGCCTAGTC TCACAAGAGG CAGCTATTTC AGGTCCTTTC AGCAAAATCT TCCTAGTGTA     180
 TGCAATCGTG TCATTGTTTG GAGGCTAATT ATGGGATGGA TCTCTGGATA TGGCAGTCTC
 TAAATGGTCC ATCCTTTTGT CTCAGCTCTA AACTTTGTCT CTGTAACTCC TTTCATGGGT     300
 GATTGTTTCC AATTCTAAGA AGGGGAAAAA GTGTCCACAC TTTGGTCTTC ATTCTTCTTC
 AGTTTCATGT GTTTTGCAAA TTGTATCTTA TATCTTGGGT ATACAAAGTT TCTGGGCTAA     420
 TATCCACTTA TCAGTGAGTA CATATCATTT GAGTTCTTTT GTGATTGTGT TACCTCACTC
 AGGATGATGC CCTCCAGGCC ATCCATTTGC CTAGGAATTT CATAAATTCA TTCTTTTTAA     540
 TAGCTGAGTA GTACTCCATT GTGTAAATGT ACCACATTTT TTGTATCCAT TCTTCTGTTG
 GGGGGCATCT GGGTTCTTTC CAGCTTCTGG CTATTATAAA TAAGGCTGCT ATGAACATAG     660
 TGGAGCATGT GTCCTTACTG GTTGGGACAT CTTCTGGATA CATGCCGAGG AGAGGTATTG
 CGGGATCCTC CTGTAGTACT ATCTCCAATT TTCTGAGGAA CCGCCAGACT GATTTCCAGA     780
 GTGGTTGTAC AAGCTTGCAA TCCCACCAAC AATGGAGGAG TGTTCCTCTT TCTCCACATC
 CTTGCCAGCA TCTGCTGTCA CCTGAACTTT TGTGCCATTC TGACTGGTGA GATGGAATCT     900
 TTCTCAGCCA TTCAGTATTC CTCAGGTGCG AATTC
                                                                      1020
```

B2

```
 GAATTCCATG GTATTAATGA TCCTTAGTGC CTAGGATTTC TATCTGTGAT CTGGGGACTG      60
 GGAGCAAGGC TATGTGTGTC CAGTACATCC AGACTGAGTT CTCTGAAGAA GTGCTCTATT
 ATACAGGAAC CCTGATTGCT ATGATAAAAA GGCTGTGTGT ACAACCTCTT CTTAGGTTAT     180
 AAGGAGCTAA AATATAAGGG AGACAAATAA CTAAAAAGAA AATAATAGCT GGACGAGGGT
 AGTACACACC TTTAGTCCCA GCACTTGAGA GGCAGAGGCA GAGGCAGGCA GGTCTCTGTG     300
 AGTTTGAGGC TAGCCTGGTC TACAAAGAGA GTTCCAGGAT ACCGAAGGCT GTATGAAGAA
 AGCTATCTTG AGAAACCAAT AATGATGATG AATATTATAG TACTTAGCAG CTCTCTGAAG     420
 ACTCATATCC TGCAAGCACC TGCCCATGAG TTTGGGGTAT GGTAGGGAGG ACAAGACTGA
 AGGAAAGCAC GCAGAGGGCA TGAGGTAGCA GATGACTAAG TGAACAACCA GACCCTTTCC*     540
```

Figure 18. *Host Fanking Sequences from IVE35 Clone*
3' integration clones B14 and B3, 5' integration clone B2.
*: site of provirus insertion. The flanking EcoRI sites are underlined

C1 + C4

GAATTCTGCA AAATCAGGAA ATCAGAACGT ACCAGTTCTC AGAGCAGTAC ATTTAACTTG 60
ATTGGGAAGG GCACACCCAT GGTCACCCTT TAAAAGGGTT TGTCTTCCTC CGCAACATTT
TGTTTTACCC AGCCTTTTGG AGCTCAGCAC TTTTGGCTTC TTCTCACAGT GTGGGAAAAC 180
AAATTGGAAC ATGACAGTTG CCTCTTAAGA GTGCTTTGTC TGGCCCCAGT GCCTTCTCCA
GGGTGGGGCT CTGTAGAGGT TGTCTGATTG ACTTACAGAT GGAAGGCTAC AGGAAAGGAG 300
TGGAAGGAAC AGGGGTGGGG GTGGGGGCTT CCTTCCACTT CTTGGCCCAG AGAACTGGAT
GTGTTTTATG CAGTCCTTGT TCATGGA**AGA T*AGAT**AATTT AGGCTCTAGG GTATATTTTT 420
AGTTTATTGG TTTAAACTTA AAAATAATAA TAATTCGAAT GTAGCCACCA TTTATTTTTA
GTACTGTAAC TAAACTCTGT TAAGATGCTG TTATTTGTGC TCTCTGAGGG ACTGGATAAC 540
TCCATGTGTG CAGGAATAGG CTGTTCTTGA TGCTGAGTGA ATATTCCTTG CAGAACCACC
CTGATCCTTT CCCTTTCAAA ATCTTAGGTC TAAGAGACGT CCCAGAAGAC ACAATGGGAA 660
TTTCTTCTTC TACCTGGGCC TTGAGTGAAA ACTGTACTGT TTTTCACTGA AGGCCCAGAG
GAAGGAGTTG TAAGTGTTAC TCTACAGGTC TGTCTCAAAC TCCACACTTC CCTCATTTAT 780
GTGCACGGCG GCATTGTGCC ACGAGGCTGG AGACCTTCTA GCCTCCAGCT TCTCCTAGAT
ATTCCTGTTC ATCTGGGGCT AGCAACAGGG ACTCCACATT GGATGAAGTC TGGGTCTCCA 900
CTTTTCTACT TGTAGACACG TTGTGCTCAA GAATCCCCGC CCCGGGGTGG GGTTTCTGGG
TTCCAATTGG GATGTAAAGC GAATAAATAC ATTAATTAAT CCACCCAAAA AAAAAGAATTC 1020

Figure 19. *Host Flanking Sequences from IVE38 Clone*
3' integration clone C1 and 5' integration clone C4.
*: site of provirus insertion. A 4 bp duplication (**AGAT**) is highlighted.
The flanking EcoRI sites are underlined

analysis of genomic DNA from both uninfected R1 and IVE38 cells using radiolabeled genomic fragments C1 and C4 as probes. The results (Figure 20) revealed that C1 and C4 fragments belonged in the same locus. Digestion of genomic DNA from R1 cells with EcoRI produces a fragment of the same size that can be detected using both probes (Figure 20 A *left* and *right*, lanes R). The size of the fragment is 1.0 Kb, which is the expected size of a C1 (400 bp) plus C4 (600 bp) EcoRI fragment. This EcoRI 1.0 Kb genomic fragment is the endogenous locus undisrupted by the proviral integration. The IVE38 clone is heterozygous for the proviral insertion. Therefore, hybridization of EcoRI-digested IVE38 genomic DNA with C1 and C4 probes revealed the presence of a 2.4 and 4.2 Kb band, respectively, in addition to the endogenous 1.0K bp band (Figure 20 A, lanes R IVE38 clone). These results confirm those obtained with *neo* (3' probe) and AP (5' probe), and rigorously prove that the cloned fragments are the host flanking sequences at the site of proviral integration. Similar results were obtained when genomic DNA was digested with both EcoRI and SacI. An endogenous 800 bp band was detected in the R1 lane (Figure 20 A, lanes R/S). Additionally, a 2.3 Kb and 2.5 Kb band can be detected with the C1 and C4 probe, respectively. Southern analysis of genomic DNA using a C18 probe failed to detect specific bands (not shown). Instead, a smear of hybridizing sequences could be observed, suggesting that the fragment contained repetitive sequences.

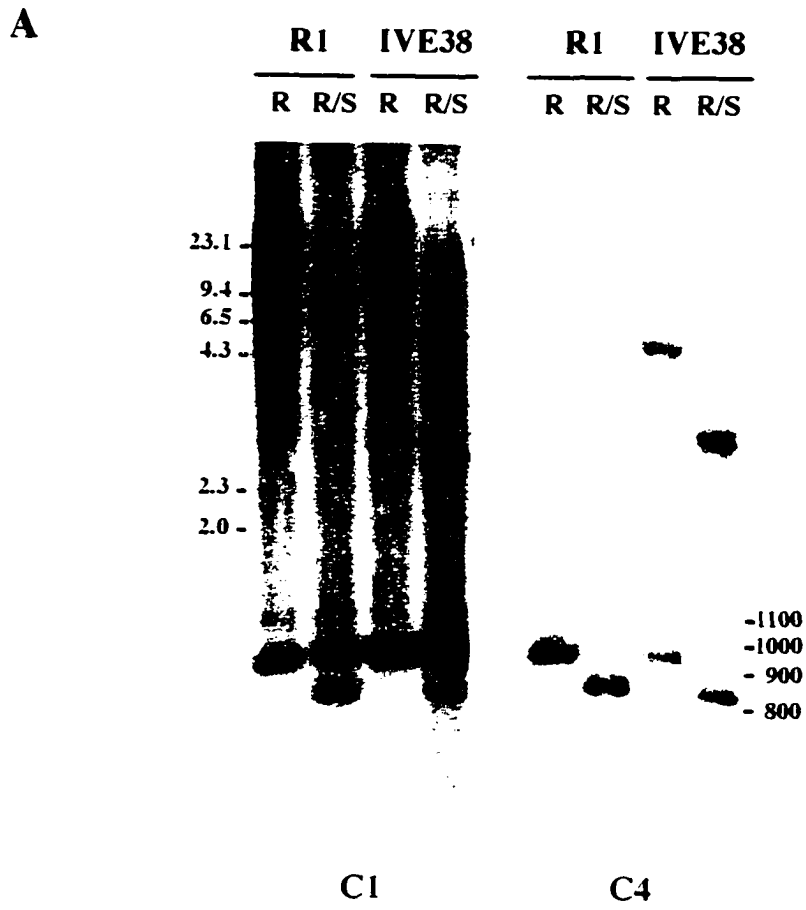


Figure 20. *Southern Analysis of Clone IVE38 DNA Using C1 and C4 Host Flanking Probes*

Genomic DNA from uninfected ES cells (RI) or clone IVE38, was digested with EcoRI (lanes R) or EcoRI plus SacI (lanes R/S). DNA was subjected to Southern analysis using a radiolabeled C1 (left) or C4 (right) flanking probe. Molecular Weight Marker: HindIII digested l DNA (left), 100 bp ladder (right).

4.7 RNA Expression in IVE38 Cells and EBs

In order to further characterize clone IVE38, I performed Northern blot analysis of RNA obtained from IVE38 3-day EBs, using *PLAP* as a probe. This time point was chosen because *PLAP* expression was maximal at this point. I could not detect a *PLAP*-specific message. Several cross-hybridizing bands were detected when full-length *PLAP* was used as a probe, and no specific signal was detected when probing the filters with the 3' human specific *PLAP* fragment (not shown). Failure to detect a IVE38-*PLAP* specific message could be due to a transcript present at levels below the resolution of the Northern analysis.

4.7.1 *A PLAP-Specific Message is Detected in IVE38 RNA*

In order to increase the sensitivity of the analysis, RNA from IVE38 EBs was examined by RT-PCR. Since the human *PLAP* gene is highly homologous to several mouse APs, primers were designed to specifically amplify human but not mouse transcripts. Two pairs of primers were designed: AP4 (reverse/antisense) and AP3 (forward/sense) were used to amplify a 320 bp 5' human-specific fragment. AP2 (reverse/antisense) and AP1 (forward/sense) were used to amplify a 360 bp 3' human-specific fragment. Total RNA was reverse transcribed using the reverse/antisense oligonucleotides as primers, following standard procedures (95). cDNA was later used as a template for PCR reactions. Hgapdh1 and hgapdh2 primers were used as positive controls in parallel reactions. PCR products were then electrophoresed in 1.4% agarose gels stained with ethidium bromide. A 431 bp fragment was specifically amplified with the

primers hgapdh1 and hgapdh2 (not shown), indicating that the reagents and conditions selected for the experiment were adequate. Several faint bands were generated with AP1/AP2 and AP3/AP4 primers (not shown). DNA was transferred to nylon membranes and hybridized to a [α-^{32}P] dCTP-labeled full length *PLAP* probe. RNA from uninfected R1 cells was included as a negative control. Non-reverse transcribed RNA from ES clone IVE38 was used as a template for PCR as a control for possible genomic DNA contamination of the sample. The results in Figure 21 show that only cDNA from clone IVE38 gives rise to a *PLAP*-hybridizing fragment of the expected size (320 bp). No *PLAP*-hybridizing band was detected when either R1 cDNA or IVE38 RNA was used as a PCR template. No amplification product was detected when the 3' primers (AP1/AP2) were used, suggesting that the conditions used for the PCR reaction were suboptimal for this pair of primers.

Amplifiaction of a human PLAP-fragment from PT-AP clone IVE38 cDNA showed that a human *PLAP*-hybridizing message was expressed only in this clone, and further indicated that *PLAP* expression found in EBs is the result of activation of the vector reporter gene by the trapped endogenous locus.

4.7.2 *Host Flanking Sequences do not Detect a IVE38 Specific Message*

Clone E38 has three integrated entrapment vectors (Figure 10). In order to determine which of the trapped loci was responsible for the reporter gene expression observed in IVE38-derived EBs and embryos, RNA expression was studied by Northern analysis using C1,

C4 and C18 host flanking fragments as probes. Total RNA extracted from uninfected R1 ES cells and IVE38-derived EBs was size fractionated in a 1% denaturing agarose gel and blotted onto a nylon membrane. RNA from PT-AP Ψ2 producer clone #3 and mouse brain RNA were included as controls for terminally differentiated cells or tissues, where insertion IVE38 is not expected to be expressed. The filters were hybridized to radiolabeled C1, C4 (not shown) and C18 flanking probes. The results are shown in Figure 22. Neither C1 nor C18 probes can detect a specific message in ES cells or EBs from clone IVE38. C1 and C18 probes detect several non-specific RNA species. Both probes also give rise to background when used in genomic Southern analysis (see below), suggesting the presence of repetitive sequences in these fragments. Conversely, no signal was observed when C4 was used as a probe, suggesting that C4 is not of the transcribed IVE38 locus. Failure to detect a specific message with these probes shows that either they are not part of the trapped expressed locus or that they contain intron sequences and are not part of the message. The same filter was re-hybridized to a *neo* probe (Figure 23A). Two bands were detected in ΨCRE producer clone #3 (Figure 23A, lane PT-AP #3) of 1.2 and 5.0 Kb in size. The 1.2 Kb band corresponds to the *neo* message initiated at the PGK promoter. The larger band is the retroviral genome-length message, initiated at the 5' LTR. No *neo* hybridizing bands are observed in either brain or R1 cell derived RNA (Figure 23A). A 1.2 Kb band is detected in RNA derived from IVE38 cells and EBs, corresponding to the internal *neo* message also detected in the producer cell line. No

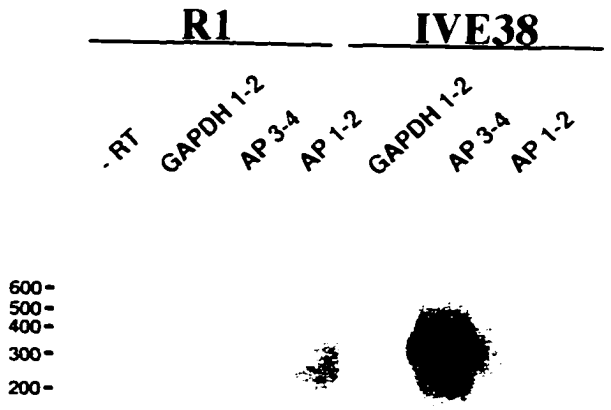


Figure 21. *RT-PCR Analysis of Total RNA from Clone IVE38*
Total RNA from R1 uninfected cells (R1, lanes AP 3-4 and AP 1-2) and IVE38-derived 3 day-EBs (E38, lanes AP 3-4 and AP 1-2) was reverse transcribed to cDNA and PCR amplified. The resulting DNA species were separated on a 1.4% agarose gel, and subjected to Southern analysis using a ^{32}P labeled full length *PLAP* probe. Two sets of primers were used: AP-1 and AP-2, designed to amplify a 360 bp 3' fragment, and AP-3 and AP-4, designed to amplify a 320 bp 5' specific fragment. GAPDH primers 1 and 2, designed to amplify a 431 bp fragment from the glyceraldehyde 3-phosphate dehydrogenase cDNA, were used as positive controls (lane GAPDH 1-2). Non-reverse transcribed IVE38 RNA was included as a negative control (lane -RT). Molecular weight marker: 100 bp ladder (BRL).

other *neo* hybridizing band was detected, even after longer exposure of the filter. The 1.2 Kb band detected in IVE38 cells is stronger in undifferentiated cells and becomes weaker in differentiated 2 day and 6 day EBs (Figure 23A). These results are in agreement with previous reports, where the murine PGK-1 promoter was found to drive broad *lacZ* expression in transgenic mice even though its activity appeared to be regulated so as to be stronger in metabolically active cells, such as undifferentiated ES cells (68). These results suggest that either the *neo* cassette is spliced out and therefore is not a part of the IVE38 message, or that the IVE38 message is below the level of detection. Alternatively, PT-AP insertion IVE38 may be functioning as an enhancer trap rather than as a promoter trap. This hypothesis was suggested by the fact that the PGK promoter was found to have limited bi-directional activity *in vitro* (52). The same blot was reprobed with a murine *oct3* probe, as a control for both the integrity of the RNA and state of differentiation of ES cells in culture. *oct3* is a POU domain transcription factor, expressed at high levels in undifferentiated ES cells and downregulated upon ES cell differentiation (103). While a strong 1.6 Kb band, corresponding to the endogenous *oct3* message, is detected in both R1 and IVE38 undifferentiated cells, the signal gets weaker as *in vitro* differentiation progresses (Figure 23B).

4.8 RNA *In situ* Hybridization Studies

In situ analysis performed on sections of R1-derived 4 day and 12 day EBs (experiments conducted by Amy Leahy), using C1 and C4 probes, confirmed the Northern Blot results. No specific signal was

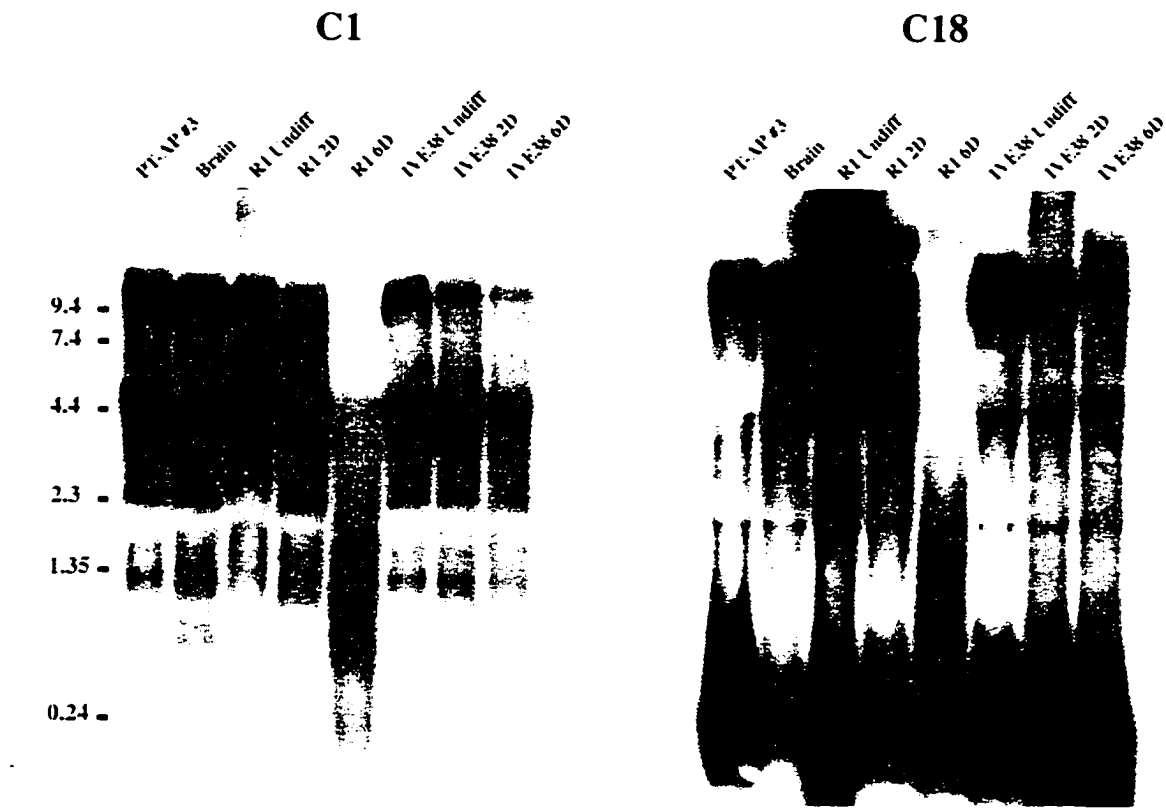


Figure 22. *Northern Blot Analysis of IVE38 Clone RNA*
RNA from uninfected R1 and IVE38 ES cells and EBs was size fractionated on a 1% agarose denaturing gel. The RNA was blotted onto a nylon membrane and sequentially probed with radiolabeled C1 (*left panel*) and C18 (*right panel*) flanking fragment. Molecular weight marker: 0.24-9.5 Kb RNA (BRL).

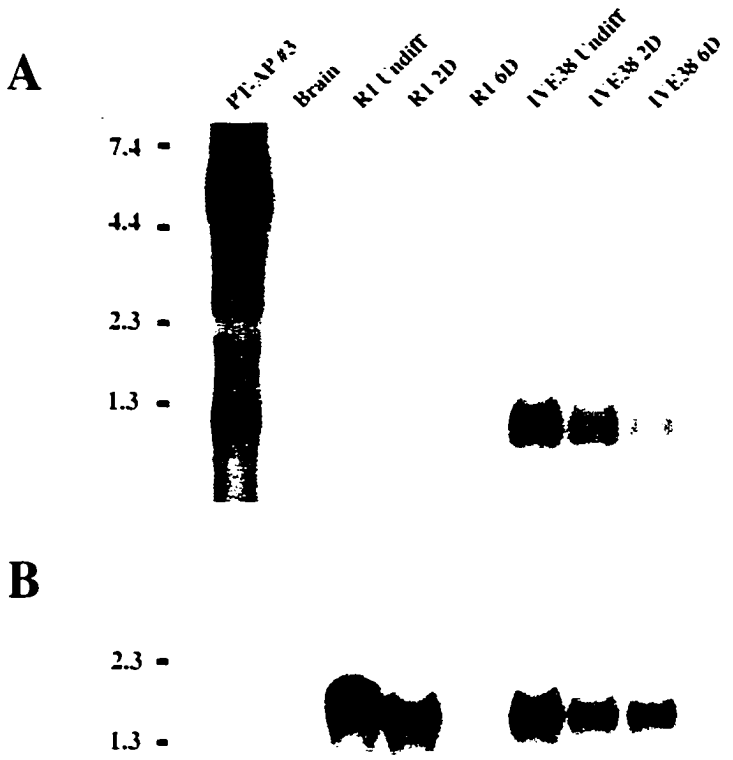


Figure 23. *Northern Blot Analysis of IVE38 Clone RNA*
RNA from uninfected R1 and IVE38 cells and EBs was size fractionated on a 1% agarose denaturing gel. The RNA was blotted onto a nylon membrane (same blot used in Figure 19) and sequentially probed with a radiolabeled *neo* (20 A) and *oct3* (20 B) fragment. Molecular weight marker: 0.24-9.5 Kb RNA (BRL).

detected with either probe (not shown), suggesting that they are not part of the IVE38 message. Alternatively, since clone IVE38 has three proviral insertions, it is possible that the cloned flanking fragments C1, C4 and C18 do not correspond to the trapped active locus. In both situations, these sequences would not be part of the IVE38 message.

4.9 IVE38 Derived Transgenic Embryos Display Restricted *PLAP* Expression

In order to study reporter gene expression in clone IVE38 *in vivo*, ES cells from this clone were used to generate chimeric embryos. For this, IVE38 ES cells were aggregated with tetraploid mouse embryos. As a result of this procedure, embryos are generated in which the embryo proper is almost entirely ES cell-derived and the extraembryonic tissues are derived from the tetraploid embryonic cells (72). Aggregation chimeras were transferred to the uteri of pseudopregnant females. The day of the transfer was considered day 2.5 of gestation. Embryos were isolated at day 8.5 and day 9.5 of gestation and stained as whole mounts for *PLAP* activity. IVE38 ES cells gave rise to apparently normal chimeric embryos, indicating that neither the PT-AP insertion nor the reporter gene had any effect on normal development. However, in a given litter, not all the embryos had reached the same developmental stage, with some of the embryos retarded in their development by 0.5 to 1 days. In day 8.0 p.c. embryos (head fold stage to early somite stage), widespread *PLAP* expression in the embryo proper was found, with stronger expression in the primitive

streak (Figure 24A). At day 9.0-9.5 of gestation, expression was restricted to the somite compartment of the mesoderm, the neural tube and the ventral aspect of the first branchial arch (Figure 24C and 24D). Expression was not observed in somites rostral to the neural tube closure. There was an additional patch of expression in the caudal tip of the embryo, presumably in presomitic mesoderm (Figure 24B and 24C). In day 10.5 embryos, expression was restricted to mature, epithelial somites (not shown). Only one day 10.5 p.c. embryo was generated with clone IVE38 ES cells, and since that result could not be repeated, expression at that stage was not studied in further detail.

These results correlate with those obtained *in vitro* . Undifferentiated IVE38 ES cells showed uniform but weak *PLAP* (Figure 9). Day-4 EBs exhibited a dramatic increase of *PLAP* expression (table 4). Positive cells were observed throughout the EBs. *PLAP*-expressing cells, however, were confined to the internal mass of cells (not shown), which is the EB equivalent of the primitive ectoderm in the egg-cylinder embryo, from which the embryo proper is derived. In day-8.0 IVE38 embryos, *PLAP* expression was widespread, but confined to the embryo proper. In day-12 EBs, the domain of reporter gene expression was further restricted to internal groups of cells. In day 9.5 p.c. IVE38 embryos, *PLAP* expression gets further restricted to a more or less defined compartment, mainly the somites and, to a lesser extent, the neural tube and the tail tip. Day-12 EBs had a thick outer layer of cells. Whole-mount (Figure 9) as well as sectioned stained EBs (not shown)showed that *PLAP* expression was totally excluded from this domain. This outer layer is

equivalent to the primitive endoderm in the early post-implantation embryo which will give rise to extraembryonic tissues in the developing embryo. Similarly, in IVE38 embryos, *PLAP* expression was not detected in extraembryonic tissues at day 8.0, 9.5 (Figure 24) or 10.5 p.c.(not shown)

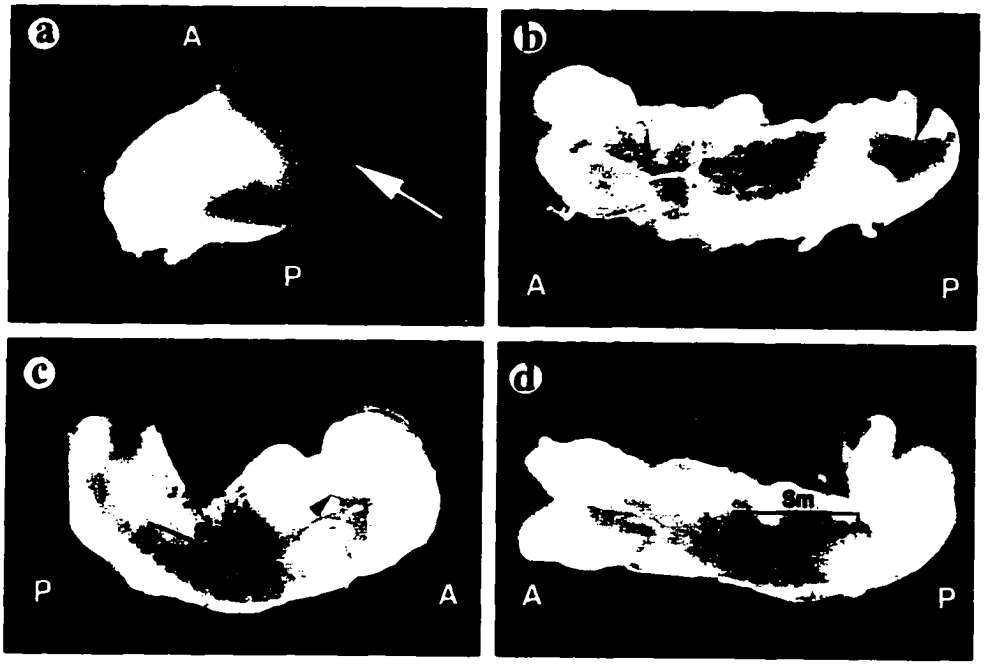


Figure 24. *PLAP Expression in IVE38 Aggregation Chimeras*
Embryos were generated by tetraploid embryo aggregation, transferred to foster recipients, and dissected at different stages. Dissected embryos were fixed with 2% paraformaldehyde, stained for *PLAP* activity and photographed using a Nikon SMZ-U strereomiscroscope. Anterior-Posterior orientation is indicated (A-P). a) Head fold stage embryo, showing the primitive streak (white arrow), b) day 9.0 p.c. embryo (dorsal view) showing *PLAP* staining in the tail tip (grey arrow) and the neural tube (black arrow head), c) day 9.5 p.c. embryo (lateral view) showing *PLAP* staining in mature somites (Sm), tail tip (grey arrow) and ventral aspect of the first branchial arch (grey arrowhead), d) day 9.5 p.c embryo (dorsal view) showing *PLAP* staining in mature somites (Sm) and neural tube (black arrowhead).

CHAPTER 5

DISCUSSION

5.1 Retroviral Vector-Mediated Genetic Screen

The objective of this study was to design an *in vitro* screen aimed at identifying novel genes involved in embryonic development. As a result of this effort, a series of retroviral entrapment vectors was developed. These vectors were characterized for their ability to infect ES cells and for their entrapment efficiency. A screen was performed on 480 infected ES clones. Five ES cell lines were generated and isolated, carrying proviral insertions in a locus whose expression during EB formation and *in vivo* were temporally and spatially restricted. One clone, PT-AP IVE38 was analyzed in more detail.

In any screening protocol, two main factors are to be considered: the method used to generate a population of random independent events and the method used to select specific events from the population according to a certain criterion. In organisms like *Drosophila melanogaster,* the dissection of many developmental genetic pathways was successfully achieved because extensive genetic screening, carried out to near saturation, can be easily performed using chemical or physical mutagenic agents. In addition, large number of mutants can be easily isolated on the basis of morphology, pigmentation, etc. Practical limitations disfavor the application of this kind of random mutagenesis in mice. First, the large size of the genome, with an estimated 10^5 genes, makes this an undertaking of monumental proportions. Second, embryonic lethality, the most obvious criterion used to select for mutants may be not

very useful to identify mutations in development control genes. In an organism like the mouse where embryonic transcription starts as early as the two-cell stage, embryonic lethality will result, in many cases, from mutations in housekeeping genes. Mutations that result in alteration rather than arrest of normal development may be more informative. However, the observation and identification of such mutants is complicated by the fact that the embryo develops *in utero*.

An alternative mutational approach is the use of entrapment vectors in ES cells. This method combines the power of random mutagenesis with the possibility of selecting insertions in expressed genes. Entrapment vectors usually carry a reporter gene whose activity can be easily visualized and whose expression is dependent on integration within an active host gene. The reporter gene pattern of expression mimics that of the trapped gene. Since a number of genes involved in patterning and other developmental processes show restricted expression patterns during embryogenesis, it is speculated that genes whose expression is restricted both temporally and/or spatially may code for developmentally relevant molecules. The use of gene trap and promoter trap vectors has proven to be a powerful means to identify novel murine genes and to study their expression as it changes during development.

To date, two main types of gene trap screens have been conducted. In the first type, ES cells were infected with a gene trap vector carrying a fused β-galactosidase-*neo* reporter gene, ensuring the isolation of cell lines with insertions in genes expressed in ES cells. Approximately 30 such cell lines were introduced into the germ

line of mice and a lethal or visible phenotype was observed in ~25% of them (37). In another screen, (142), ES cells were transfected with a gene trap vector containing a reporter gene and a *neo* gene under different promoters. 30,000 neo^R ES cells were generated, out of which 300 (1%) expressed *lacZ* . Chimeric mice were generated from these ES cells, and ~30 % of them showed restricted expression of the *lacZ* gene at day 8.5 p.c. Both types of vectors rely on integration events occurring in genes already expressed in ES cells, and therefore may miss genes that are not expressed at that stage. The other type of screen (4) includes an *in vitro* pre-selection of infected ES cells, but since a β-galactosidase-*neo* vector is used, a bias is again imposed towards genes already expressed in undifferentiated ES cells. Undifferentiated ES cells are developmentally equivalent to the inner cell mass of day 3.5 blastocysts. Therefore, genes that are switched on after this stage can not be detected.

In our laboratory, we have constructed and characterized three kinds of retroviral vectors, the promoter trap *PLAP* (PT-AP), the gene trap *PLAP* (GT-AP) and the promoter trap IRES-AP (6). The PT-AP carries a promoterless ATG^+ *PLAP* cDNA. The GT-AP carries an ATG^- *PLAP* in addition to a splice acceptor (SA) sequence from the Mo-MLV *env* gene. Promoter trap and gene trap viruses displayed an entrapment efficiency of 1 to 3% in both mouse fibroblasts and embryonic stem cells (6). The PT-IRES-AP vector contains an internal ribosomal entry site (IRES), derived from the encephalomyocarditis virus, upstream of the AP gene. This configuration allows independent translation of the *PLAP* reporter gene from a bicistronic message, thus eliminating the requirement for PLAP fusion in-frame

with the trapped gene. An increased entrapment efficiency was expected as a result of this modification. In fact, the IRES-AP virus had an increased entrapment efficiency (5 to 13% in fibroblasts, 2 to 5% in ES cells) (6).

These vectors were used to perform a large genetic screen for developmental control genes in ES cells. To date, a total of 2,400 ES cell clones which carry random insertions were generated and analyzed for their *PLAP* expression upon *in vitro* differentiation. This includes the 480 PT-AP clones that I generated and analyzed, and 1900 GT-AP and PT-IRES-AP clones generated by Jing-Wei Xiong, (143). 41 clones (1.7%) exhibited differential reporter gene expression during embryoid body formation. 31 trapping events (76%) occurred in genes that were not expressed in undifferentiated ES cells and would have been missed in previously described screens. In 44% of the trapping events, expression of the reporter gene was temporally restricted to differentiating embryoid bodies (143).. 10 selected entrapment ES cell lines were tested in our laboratory for their *in vivo* expression pattern in developing embryos (143). *PLAP* expression was assayed at different stages from day 8 to day 12 p.c. and was found to be restricted both temporally and spatially.

Entrapment vectors have to provide a means to tag the integration site in order to facilitate its molecular cloning. Our vectors contain two copies of the bacterial *supF* gene, which allow selection of viral and host flanking sequences by complementation genomic cloning. This should be the method of choice to use in promoter rather than gene trap insertions. The reason being that *supF* cloning is a "structural" approach, and isolation of a coding piece of DNA

relies on colinearity between DNA and RNA sequences at the site of integration. This situation occurs with vectors inserted in exons, as in the case of promoter traps. For gene traps, on the other hand, active insertions will most likely occur into introns (which are present at a 10:1 ratio to exons in mammalian genomes). In this case, a more "functional" cloning approach should be taken. RACE and other RT-PCR derived protocols, which take advantage of known sequences within the reporter gene, should be the method of choice.

The results obtained by our laboratory and others show that entrapment vectors can be used to efficiently generate and analyze a large number of insertional mutants in ES cells. Improvements in both the vector design and the *in vitro* screening strategy could reduce the number of trapping events that would have to be analyzed. For instance, a reporter gene whose activity could be assayed *in vivo,* like the jellyfish Green Fluorescent Protein (20), would eliminate screening steps. The *in vitro* pre-screening of trapping events has enormous potential. In one study ES cells were screened for their ability to activate and deactivate reporter gene expression in response to Retinoic Acid (36). In another study, a novel class of expressed repetitive sequences in the mouse was identified, whose expression was induced *in vitro* by retinoic acid (94a). In addition, by controlling and directing the formation of EBs, it is possible to screen for genes involved in the establishment or differentiation of specific cell lineages.

Our entrapment vectors do not select for (although they do not exclude) genes that are expressed in ES cells. Clone IVE38 expression, for instance, was weak but widespread in undifferentiated ES cells.

This fact does not preclude possible tissue-specific expression and function at later stages. In ubiquitously transcribed genes, expression may be temporally and spatially controlled at the translation level. Alternatively, genes that are strongly transcribed in specific tissues at later stages may be ubiquitously transcribed at low levels before the gene products are required. *nodal* , Myf-5 and *jumonji* are expressed in ES cells (developmentally equivalent to the inner cell mass of day 3.5 p.c. blastocysts) despite of having tissue restricted expression and function *in vivo* at later stages of development (26, 71, 118).

I have used the ES cell *in vitro* differentiation system as a way to pre-screen for entrapment insertions with differential reporter gene expression. EB formation recapitulates molecular and cellular events that take place *in vivo* in post-implantation embryos. ES cell differentiation. However, ES cell differentiation *in vitro* occurs in a morphologically abnormal and disorganized way with no indication of structure or pattern formation. The *in vivo* testing of the selected ES cell clones, therefore, has to complement the *in vitro* study. One PT-AP clone. IVE38, showed expression which was restricted spatially and temporally in differentiating EBs. In day 8 to day 9.5 p.c. embryos derived from clone IVE38, reporter gene expression was mainly localized to the paraxial (somitic) mesoderm, although weaker expression was also observed in the tail bud, the neural tube and the first branchial arch.

The *PLAP* expression data obtained both *in vitro* and *in vivo* with clone IVE38, as well as with other promoter trap ES cell clones in the lab (143), validate the *in vitro* preselection step of the screen.

76% of the selected ES clones , as already discussed, did not show *PLAP* expression in undifferentiated stage. Localization of *PLAP* expression in EBs , however, may be a very limited criterion to select which clones assay *in vivo.*, since EBs have only two topologically distinguishable layers of cells: the outer layer, equivalent to the primitive endoderm and the inner layer, equivalent to the primitive ectoderm of the egg cylinder-stage embryo. In EBs derived from clone IVE38, *PLAP* expression is restricted to the inner layer, and in IVE38-derived day 8.0 p.c. embryos, as expected, expression is restricted to the embryo proper. Temporal regulation of *PLAP* expression in EBs, on the other hand, could be used as a standard, since all cell types studied to date arise in EBs displaying a kinetics of gene expression and cell differentiation that faithfully recapitulate equivalent events *in vivo.*

5.2 Generation and specification of the Paraxial Mesoderm (120b)

Metamerism, the repetition of segmental units (monomers) along the anterior-posterior axis, delineates the body plan of some bilateral organisms. Vertebrates (chordates) are metameric animals. The segmental nature of the embryo will later manifest in the adult animal, not only in the axial skeleton (the vertebral column and ribs are typical metameric structures), but also in the organization of the lateral mesoderm and the peripheral nervous system. The earliest morphological manifestation of metameric pattern in the mouse embryo is the formation of somites. Somites are paired blocks of mesodermal tissue located along the rostro-caudal axis, parallel to

the neural tube. They are derived from a pair of mesodermal stripes, the presomitic mesoderm, that lies anterior to the primitive streak first and the tail bud later. Although the paraxial (somitic) mesoderm is not overtly segmented rostral to the first somite or caudal to the last one, careful examination of these compartments revealed a tandem array of spherical structures, the somitomeres (120). Both the number and morphology of somites and somitomeres are remarkably constant within a given species. This number remains unaltered even after experimental manipulation of the embryo, suggesting that the process is developmentally stable and subject to stringent control.

Generation of segmental units may be achieved by simultaneous partitioning of preexisting tissue. In this case, the number and size of the resulting units will be determined by the amount of tissue available at the time of segmentation, which is usually presaged by expression of specific genes whose domains of expression determine segmental identity. This mechanism is used by *Drosophila* to generate segments in the blastoderm-stage embryo. A similar mechanism results in the formation of rhombomeres in the mouse and chick embryos. Somite formation in vertebrates occurs in a much different way. Segmentation begins *before* the entire precursor tissue (the presomitic mesoderm) is formed. Somites are formed continuously while the precursor tissue is generated from a small population of cells. In this continuously expanding system, segmental identity seems to be achieved by progressive activation of an increasing number of position-specific genes in caudal (newly added) somites. Since not all the segmentable material is available

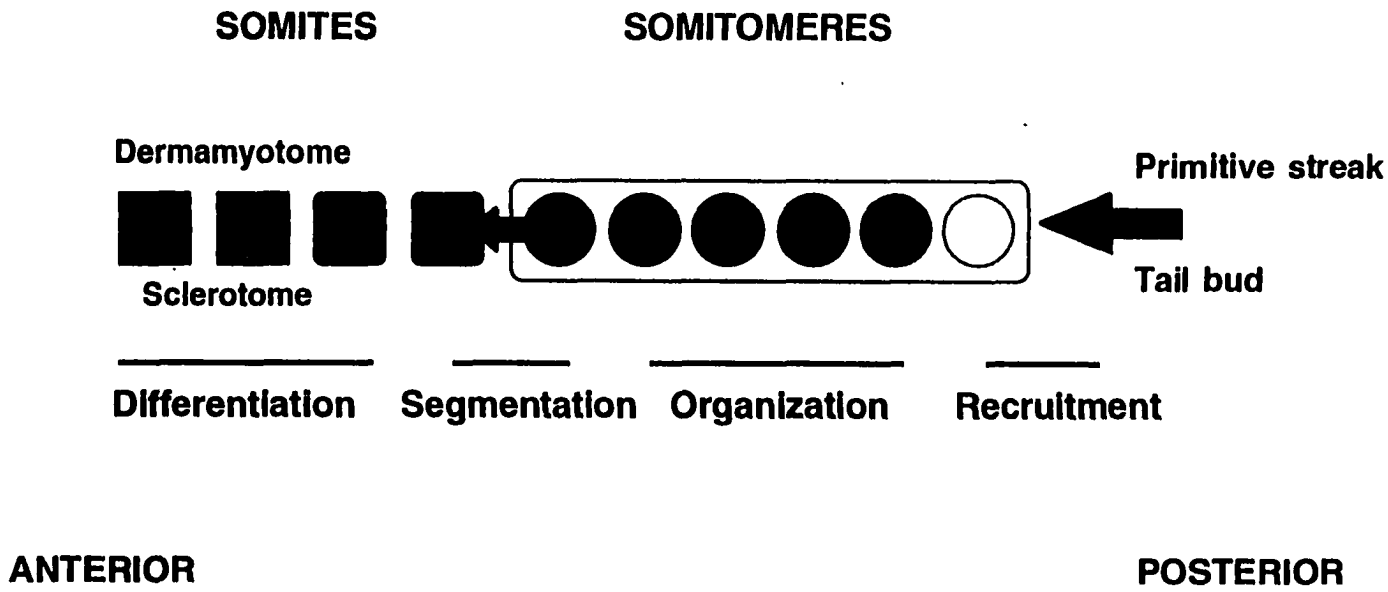


Figure 25. *Specification and Segmentation of the Paraxial Mesoderm* (adapted from 120b) See text for details

from the beginning, somite number and size may be regulated by mechanisms involved in more global processes, like control of cell proliferation and tissue growth. Somite formation involves a series of mesenchyme-epithelial transitions. Cells recruited from the primitive streak undergo a mesenchymal transformation and are incorporated into the presomitic mesoderm, where they form somitomeres. Somitomere maturation progresses as a posterior-anterior wave, culminating in a second epithelial conversion and condensation of the cells into a mature somite. Somites, in turn, go through a third transition, where part of the cell mass loses its epithelial nature, giving rise to a mesenchymal sclerotome (Figure 25).

In addition to the overall anterior-posterior regionalization of the somitic mesoderm, two other polarities are present in each individual segment. First, an anterior-posterior polarity exists, which is established prior to somite segmentation. This is later evidenced by the subdivision of the sclerotome and the exclusive association of motor axons and neural crest cells with the rostral half of the somite. Second, a dorso-ventral polarity exists, which is set after the somite is formed by inductive mechanisms exerted by the neural tube.

With regards to the molecular specification of paraxial mesoderm, it should be noted that cells destined to become somites can first be traced to the primitive ectoderm of the germ layer embryo. It is not clear, however, whether a unique mesodermal precursor tissue ever exists. Fate mapping studies suggest that cells with distinct mesodermal potential emerge from different portions of the primitive streak (119). Gene expression studies showed that different subsets of cells expressed different genes. It is possible

then, that cells from the primitive ectoderm acquire different mesodermal fates as they pass through the primitive streak. It is important to notice that although mesodermal cells may be committed to diverse fates as they leave the primitive streak, transplantation experiments showed that this commitment is not complete and that further reinforcement may be needed for full specification of cell fate even after somites are formed.

Somite segmentation is marked by expression of a number of genes. Possible roles for some of these genes in somitic determination and differentiation may be inferred from their timing as well as their domains of expression. However, further studies, such as analysis of transgenic or null mutant mice, are required in order to determine whether the presence of a gene product is required for determining cell fate or is just a consequence of the committed state of a cell population.

Commitment of a mesodermal cell to a paraxial, medial and lateral mesodermal fate is associated with expression of MF-1, a *Fork Head* domain-containing gene and putative transcriptional regulator (97). This gene is expressed in early fold stage embryos (day 7.5 p.c.) in the entire non-notochord mesoderm, with stronger expression in the paraxial region. MF-1 expression may be regulated by two signals, an anterior one originating within the node, and a more posterior one, originating within the primitive streak. Mesodermal cells, proximal to the node receive both signals, express MF-1 and are committed to form paraxial mesoderm. Distal cells receive only a strong signal coming from the streak and a weak one from the node, and are fated for medial and lateral mesoderm. It is speculated that

FGF family members, especially FGF-4 and *nodal*, a TGF-b like secreted molecule, are some of the signaling molecules involved (76, 146). A receptor protein tyrosine kinase family member, *sek*, is expressed in rhombomeres 3 and 5 and in the presomitic mesoderm. Expression is seen as two stripes in somitomeres #1 and #2, the first and second somitomeres immediately caudal to the most recently formed somite. (75). Expression in somitomere #1 is restricted to the anterior half, in a way that presages the antero-posterior polarity of the future somites. Other transcription factors are expressed later in somite segmentation. *M-twist*, a mouse homologue of the Drosophila *twist* gene that codes for a helix-loop-helix protein, is expressed in the paraxial mesoderm of gastrulation-stage embryos and in the presomitic and somitic compartments of somite stage embryos (139). *Mox-1*, a homeobox containing gene, is expressed between day 7.5 and day 9.5 p.c. in presomitic mesoderm, epithelial and differentiating somites, and in some neural crest cell derived structures (18). *Mox-2*, on the other hand, is not expressed until somites form, but is then expressed during somite differentiation in the entire mature somite. Its expression is restricted to the sclerotome (18). A similar pattern of expression was described for *paraxis* (17), also known as bHLH-EC-2 (82) and *Meso-1* (11). During somite differentiation, however, *paraxis* expression is restricted to the dermatome and sclerotome. This pattern is reciprocal to that of *myf-5*, a member of the *myoD1* family and the first bHLH gene expressed in cells undergoing myogenesis . Interestingly, genetic deletion experiments demonstrated a role for *paraxis* in somite formation, as *paraxis* null mutant mice fail to form epithelial somites

(16). Similarly, in chicks, the paraxial mesoderm fails to undergo segmentation in regions where *paraxis* expression was abolished (112).

Tbx6 (21), a member of the T-box family of putative transcription factors, (whose prototypical member is *T* or *Brachyury*) has also been implicated in paraxial mesodermal determination. This gene has a unique expression pattern. It is first expressed in the primitive streak of day 7.5 p.c. embryos. By day 8.5 p.c. expression is restricted to the presomitic mesoderm and the tail bud, which replaces the primitive streak as the main source of mesoderm around day 10.5 p.c. *wnt-3a,* (117) a member of the WNT family of putative signaling molecules, codes for a secreted protein. Both the expression pattern and the phenotype of *wnt-3a* null mutant mice indicate a role for this gene in paraxial mesoderm formation. At day 7.5 p.c., *wnt-3a* is widely expressed within the streak. At early somite stages, expression becomes restricted to the dorsal and caudal aspects of the streak. As the tail bud forms, expression is found in the caudal tip of the tail bud (the region that continues to generate mesoderm) and in the developing Central Nervous System. *wnt-3a* null mice have a very interesting somite phenotype. At day 9.5 p.c. in homozygous embryos, somites caudal to the forelimb buds are absent. The neural tube is folded and kinked. Rostral to the forelimb buds, organization of the mesodermal derivatives seems to be normal, including the first 7 pairs of somites, which appear to form and differentiate normally. This demonstrates that specification of the rostral most somites follows a wnt-*3a* independent pathway, and adds another level of complexity to the A-P regionalization of the

paraxial mesoderm. Several members of the Paired box (Pax) family of transcription factors are also expressed in the paraxial mesoderm. *Pax-3* and *Pax-7* are expressed in the presomitic plate before segmentation. Expression becomes restricted to the dermomyotome as somites differentiate and to the myogenic progenitors that migrate into the limb buds (136). *Pax-1*, on the other hand, is expressed in the sclerotome of the differentiating somite (29). Additionally, it is important to notice that the *Pax* genes are expressed in the neural tube, where they are required for normal neural development.

5.3 Promoter Trap Insertion IVE38 Expression and Mesoderm Formation

No specific bands were detected with either a full length *PLAP* or a human specific *PLAP* probe on Northern Blots from IVE38 total RNA. If a human-*PLAP* specific message was present, its level was below the resolution of the Northern analysis. That hypothesis was later suggested by the fact that a *PLAP* specific signal could only be detected by RT-PCR. Similar results were obtained when host flanking fragments C1, C4 and C18 were used as probes. These results suggest that these flanking fragments are not part of the IVE38 message. Since IVE38 has three integrated proviruses, it is possible that none of the three cloned fragments correspond to the trapped locus. In addition, in the PT-AP vector, the *PLAP* cDNA is positioned in reverse orientation relative to the retroviral transcriptional orientation and the PGK-*neo* cassette (Figure 17). In a vector with that configuration, activation of the reporter gene may result from: 1) integration into an intron, with splicing occurring

between a cellular donor site and a cryptic acceptor site located in the vector, 2) integration in a locus under the control of a cellular enhancer. In each case, detection of a message using a flanking sequence will be unlikely. In the first scenario, the flanking sequences are likely to be intronic. In the second, the integration site may be physically distant from the trapped enhancer/gene. Further analysis will be required to distinguish between the several possible explanations. However, results shown in Figure 23 shed some light on this problem. When a radiolabeled *neo* fragment was used as a probe for Northern analysis of total RNA from clone IVE38 ES cells and EBs, only a PGK promoter-driven (1.2 Kb) *neo* message was detected, even in longer exposed films (the one in Figure 23 was exposed for 80 hours). This result strongly suggests that there is no colinearity between *PLAP-neo* DNA and RNA. In other words, *neo* is not part of the trapped message. The best way to identify the trapped locus will be to generate a transgenic IVE38 mouse. All three entrapment insertions should segregate in the F1 generation, allowing for the identification of a transgenic line with the proviral insertion responsible for somitic expression of *PLAP* by staining heterozygous embryos. In addition, the generation of a transgenic line carrying the insertion will allow for study of the temporal and spatial pattern of expression. In addition, since the retroviral insertion may have disrupted the normal structure and/or function of the trapped locus, homozygous IVE38 mice would be useful to study the effects of the insertion on somite formation and differentiation *in vivo*.
Alternatively, IVE38 could be used as a cell lineage marker. Since *PLAP* staining is very sensitive and reproducible and IVE38 is

expressed mainly in differentiating somitic mesoderm, transgenic mice generated on a IVE38 background would be useful to study the effects of overexpression of different genes on somite formation and differentiation.

Clone IVE38 *in vivo* expression data allow speculation on the function of the trapped host locus. IVE38 is widely expressed in the embryo proper as early as day 8.0 p.c. (early neural fold stage), with strong expression in the primitive streak. No earlier embryos were stained for *PLAP*, but since IVE38 undifferentiated ES cells clearly showed reporter gene expression (although at low level), it is reasonable to expect the insertion to be expressed earlier in the primitive ectoderm of the egg-cylinder stage embryo. At day 9.0-9.5 p.c., expression is restricted to the somitic mesoderm and the neural tube. *PLAP* staining appears to be stronger in most recently formed somites and weaker in the rostral (more mature) ones. No obvious expression was detected in presomitic mesoderm, apart from the spot located at the caudal tip of the embryo (Figure 21B). This result is not conclusive, and further studies will be required to clarify this point. Altogether, these data show that expression is up-regulated in early somites and probably in the most mature somitomeres. If *PLAP* expression in IVE38 embryos reflects the expression of an endogenous gene, *PLAP* expression data suggests a role for IVE38 in paraxial mesoderm regionalization rather than mesoderm specification.

5.4. Retrospective analysis of the experimental approach.

This section will be dedicated to a critical analysis of the results obtained in this study, in the context of the insertional mutagenesis approach as a whole. Entrapment vectors can be efficiently introduced in ES cells and used as a means to screen the mouse genome for developmentally regulated genes. While one can speculate about the developmental relevance of a developmentally regulated gene, a final conclusion cannot be drawn until the sequence of the gene is known and the mutant phenotype is analyzed.

In this study, the use of retroviral instead of plasmid vectors, was intended to maximize the efficiency of introduction of the vectors into ES cells, while minimizing some of the problems associated with DNA electroporation, namely, rearrangements of the endogenous locus at the site of insertion and integration of multiple copies of vector DNA. As it turned out, two of the PT-AP clones (IVE35 and IVE38) contained several copies of the proviral DNA in different loci. That occurred as a result of cocultivation of producer cells and target ES cells. Cocultivation was performed because then, I considered a viral titer in the order of 10^2 to 10^3 G418 resistant ES colonies to be too low to successfully generate insertional mutant ES cell lines. That assumption was incorrect, and led me to the choice of a method (the cocultivation), that was counter productive, because it resulted in the generation of ES cell clones where the analysis was complicated by the existence of multiple proviral insertions. I chose clones IVE35 and IVE38 for further analysis because, as already discussed, despite the difficulties in the analysis that I envisioned could arise from the fact that they contained several integrated

vectors, these clones displayed the strongest, most consistent and reproducible regulated PLAP expression.

It is difficult for me to judge the prudence of that decision. On one hand, for instance, clone IVE38-derived embryos displayed a PLAP pattern of expression that was regulated temporally and spatially and correlated, to a certain extent, to the expression observed *in vitro*. On the other hand, identification of this potentially interesting locus, could not be achieved during the course of this project, precisely because the clone contained three proviral insertions. In view of this and with the perspective that this experience gave me, I would do things differently if I had to repeat the screen.

I would still use retrovirus-based vectors, since retroviral infection is a very efficient method to generate ES cells clones with vector insertions. I would definitely keep the multiplicity of infection at a low level (10^{-2} to 10^{-3}) to ensure that only one vector is incorporated per cell. The vector would contain an IRES element upstream the reporter gene, since this modification has proven to increase the entrapment efficiency approximately 5 times. By screening the same amount of ES cell clones that I screened in this study (500) 20 to 30 ES cell lines would be identified with regulated expression of the reporter gene. After characterization of those cell lines by Southern Blot Analysis, I would continue the characterization of only those having single proviral insertions.

The *in vitro* prescreening step, which takes advantage of the ability of ES cells to differentiate *in vitro* into Embryoid Bodies, can be successfully used to identify entrapment events that lead to

regulated expression of the reporter gene. 76% of the clones identified in the large screen (as already discussed in section **5.1**) were entrapment events that would have been missed had the in vitro screening not been included. However, limitations inherent to the system (also discussed in section **5.1**) make it suitable to screen for loci subjected to temporal rather than spatial regulation.

Finally, I would analyze the PT clones for *in vivo* expression of the reporter gene, in embryos generated by tetraploid aggregation. Embryos derived from 10 clones could be readily generated and analyzed. Mouse sequences flanking the integration site would be isolated using the *SupF* strategy and used, together with reporter gene probes, to detect the endogenous and fusion message in RNA from the tissues were the reporter gene was expressed. Partial sequence from mouse flanking sequences, would provide additional information, and I would try to clone out those loci that most likely represent novel genes.

BIBLIOGRAPHY

1. **Adamson, E. D.** 1987. Oncogenes in development. Development. **99**:449-471.

2. **Allen, N. D., D.G. Cran, S.C. Barton, S. Hettle, W. Reik, and A. Surani.** 1988. Transgenes as probes for active chromosomal domains in mouse development. Nature:852-855.

3. **Bain, G., D. Kitchens, M. Yao, J. E. Huettner, and D. I. Gottlieb.** 1995. Embryonic Stem Cells Express Neuronal Properties *in Vitro*. Developmental Biology. **168**:342-357.

4. **Baker, R. K., M. A. Haendel, B. J. Swanson, J.C. Shambaugh , B. K. Micales and G. E. Lyons** 1997. *In vitro* Preselection of Gene-Trapped Stem Cell Clones for Characterizing Novel Developmentally Regulated Genes in the Mouse. Developmental Biology. **185**:201-214.

5. **Barklis, E., R.C. Mulligan, and R. Jaenisch.** 1986. Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell. **47**:: 391-399.

6. **Battaglino, R., J. Xiong, J. Jin, and H. Stuhlmann.** 1996. Bifunctional Retroviral Entrapment Vectors with an Alkaline Phosphatase Reporter Gene for Genetic Screening in Mouse Embryonic Stem Cells. Submitted, J. Virol.

7. **Beddington, R., and J. Smith.** 1993. Control of vertebrate gastrulation : inducing signals and responding genes. Curr. Opin. in Gen. and Dev. **3**:655-661.

8. **Bellen, H., C. Wilson, and W.J. Gehring.** 1990. Dissecting the complexity of the nervous system by enhancer detection. BioEssays:199-204.

9. **Bellen, H. J., C.J. O'Kane, C. Wilson, U. Grossniklaus, R. Pearson, and W.J. Gehring.** 1989. P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. Genes & Dev. . **3**:: 1288-1300.

10. **Bier, E., H. Vaessin, S. Sherherd, K. Lee, K. McCall, S. Barbel, L. Acherman, R. Carretto, T. Uemura, E. Grell, L. Y.**

Yan, and Y. N. Jan. 1989. Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. Genes Dev. **3**:1273-1287.

11. **Blanar, M. A., P. H. Crossley, K. G. Peters, E. Steingrimsson, N. Copeland, N. A. Jenkins, G. R. Martin, and W. J. Rutter.** 1995. Meso1, a basic-helix-loop-helix protein involved in mammalian presomitic mesoderm development. Proc. Natl. Acad. Sci. USA. **92**:5870-5874.

12. **Blum, M., S. J. Gaunt, K. W. Y. Cho, H. Steinbeisser, B. Blumberg, D. Bittner, and E. M. de Robertis.** 1992. Gastrulation in the mouse : the role of the homeobox gene *goosecoid*. Cell:1097-1106.

13. **Bradley, A., M. Evans, M.H. Kaufman, and E. Robertson.** 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature: 255-256.

14. **Bradley, A.** 1987. Production and analysis of chimeric mice, p. 113-151. *In* E. J. Robertson (ed.), *Teratocarcinomas and embryonic stem cells: A practical approach*. IRL Press, Oxford.

15. **Brenner, D. G., S. Lin-Chao, and S.N. Cohen.** 1989. Analysis of mammalian cell genetic regulation *in situ* by using retrovirus-derived "portable exons" carrying the *Escherichia coli lacZ* gene. Proc. Natl. Acad. Aci. USA Proc. Natl. Acad. Aci. USA. **86::** 5517-5521.

16. **Burgess, R., A. Rawls, D. Brown , A. Bradley, and E.N. Olson** 1996. Requirement of the *paraxis* gene for somite formation and musculoskeletal patterning. Nature. **384**:570-573.

17. **Burgess, R., P. Cserjesi, K. L. Ligon, and E. N. Olson.** 1995. Paraxis: A Basic Helix-Loop-Helix Protein Expressed in Paraxial Mesoderm and Developing Somites. Developmental Biology. **168**:296-306.

18. **Candia, A. F., J. Hu, J. Crosby, P. A. Lalley, D. Noden, J. H. Nadeau, and C. V. E. Wright.** 1992. Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. Development. **116**:1123-1136.

19. **Casadaban, M. J., and S.N. Cohen.** 1979. Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: in vivo probe for transcriptional control sequences. Proc. Natl. Acad. Sci. USA. **76:**4530-4533.

20. **Chalfie, M., Y. Tu, G. Euskirchen, W. Ward, and D. Prasher.** 1994. Green Fluorescent Protein as a Marker for Gene Expression. . **263:**802-804.

21. **Chapman, D. L., I. Agulnik, S. Hancock, L. M. Silver, and V. E. Papaioannou.** 1996. Tbx6, a Mouse T-Box Gene Implicated in Paraxial Mesoderm Formation at gastrulation. Developmental Biology. **180:**534-542.

22. **Chen, J., A. Nachabah, C. Scherer, P. Ganju, A. Reith, R. Bronson, and H.E Ruley.** 1996. Germ-line inactivation of the murine Eck receptor kinase by gene trap retroviral insertion. Oncogene. **12:**979-988.

23. **Chen, U.** 1992. Differentiation of mouse embryonic stem cells to lympho-hematopoietic *in vitro*. Dev Immunol. **2:**29-50.

24. **Chumakov, I., H. Stuhlmann, K. Harbers, and R. Jaenisch** 1982. Cloning of two genetically transmitted Moloney Leukemia proviral genomes: correlation between biological activity of cloned DNA and viral activation in the animal. Journal of Virology. **42:**1088-1098

25. **Conlon, F., K. Barth, and K. Robertson.** 1991. A novel retrovirally induced embryonic lethal mutation in the mouse: assessment of the developmental fate of embryonic stem cells homozygous for the 413.d proviral integration. Dev. **111:**969-981.

26. **Conlon, F., K. M. Lyons, N. Takaesu, K. Berth, A. Kispert, B. Herrman, and E. Robertson.** 1994. A primary requirement for *nodal* in the formation and maintenance of the primitive streak in the mouse. Dev. **120:**1919-1928.

27. **Danos, O., and R. C. Mulligan.** 1988. Safe and efficient generation of recombinant retroviruses with amphotrophic and ecotropic host ranges. Proc. Natl. Acad. Sci. USA. **88.:**6460-6464.

28. **De Gregory, J., A. Russ, H. von Melchner , H. Rayburn, P. Priyaranjan, N.A. Jenkins, N.G. Copeland, and H.E. Ruley.** 1994. A murine homolog of the yeast RNA1 gene is required for postimplantation development. Genes Dev. **8**:265-276.

29. **Deutsch, U., G. R. Dressler, and P. Gruss.** 1988. Pax-1, a member of a paired box homologous murine gene family is expressed in segmenting structures during development. Cell. **53**:617-625.

30. **Doetschman, T. C., H. Eistette, M. Katz, W. Schmidt, and R. Kemle.** 1985. The *in vitro* development of of blastocyst- derived embryonic cell lines : formation of visceral yolk sac, blood islands and myocardium. J. Embryol Exp. Morph. **87**:27-45.

31. **Evans, M., A. Bradley, and E. Robertson.** 1985. EK cell contribution to chimeric mice: from tissue culture to sperm. *Banburry Report. 20*, p. 93-102. *In* R. J. a. F. Constantini (ed.), *Genetic Manipulation of the Early Mammalian Embryo.* Cold Spring Harbor Laboratory Press, New York.

32. **Evans, M. J. a. M. H. Kaufman.** 1981. Establishment in culture of pluripotent cells from mouse embryos. Nature. **292**::154-156.

33. **Fekete, D. M., and C.L. Cepko.** 1993. Replication-competent retroviral vectors encoding alkaline phosphatase reveal spatial restriction of viral gene expression/transduction in the chick embryo. Mol. Cell. Biol. **13**:2604-2613.

34. **Fekete, D. M., and C.L.Cepko.** 1993. Retroviral infection coupled with tissue transplantation limits gene transfer in the chicken embryo. Proc. Natl. Acad. Sci. **90**::2350-2354.

35. **Fields-Berry, S. C., A. L. Halliday, and C. L. Cepko.** 1992. A recombinant retrovirus encoding alkaline phosphatase confirms clonal boundary assignment in lineage analysis of murine retina. Proc. Natl. Acad. Sci. USA. **89**::693-697.

36. **Forrester, L., Nagy, A., Mehran, S., Watt, A., Stevenson, L., Bernstein, A., Joyner,A. and Wurst, W.** 1996. An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid *in vitro*. Proc. Natl. Acad. Sci. USA. **93**:1677-1682.

37. **Friedrich, G., P. Soriano.** 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Gen. and Dev. **5:**1513-1523

38. **Gasca, S., Hill, D., Klingensmith, J. and Rossant, J.** 1995. Characterization of a Gene Trap Insertion Into a Novel Gene, *cordon-bleu* , Expressed in Axial Structures of the Gastrulating Mouse Embryo. Dev. Gen. **17:**141-154.

39. **Gogos, J. A., R. Thompson , W. Lowry , B. F. Sloanne , H. Weintraub , and M. Horwitz.** 1996. Gene Trapping in Differentiating Cell Lines: Regulation of the Lysosomal Protease Cathepsin B in Skeletal Myoblast Growth and Fusion. The Journal of Cell Biology. **134:**837-847.

40. **Gogos, J. A., W. Lowry, and M. Karayiorgou.** 1997. Selection for Retroviral Insertions into Regulated Genes. Journal of Virology. **71:**1644-1650.

41. **Gossler, A., A.L. Joyner, J. Rossant, and W.C. Skarnes.** 1989. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science:463-465.

42. **Gridley, T., Soriano, P., and Jaenisch R.** 1987. Insertional Mutagenesis in Mice. TIG. **3:**162-167.

43. **Gridley, T., D.A. Gray, T. Orr-Weaver, P. Soriano, D.E Barton, U. Franke, and R. Jaenisch.** 1990. Molecular analysis of the *Mov 34* mutation: transcript disrupted by proviral integration is conserved in *Drosophila*. Development. **109:**235-242.

44. **Grossniklaus, U., H.J. Bellen, C. Wilson, and W.J. Gehring.** 1989. P-element-medited enhancer detection applied to the study of oogenenis in *Drosophila*. Development. **107:**189-200.

45. **Hahnel, A., D. A. Rappolee, J. L. Millan, T. Manes, C. A. Ziomek, N. G. Theodosiou, Z. Werb, R. A. Pedersen, and G. A. Schutz.** 1990. Two alkaline phosphatase genes are expressed during early development in the mouse embryo. Development. **110:**555-564..

46. **Hogan, B., Beddington, R., Constantini, F., and Lacy, Elizabeth.** 1994. Manipulating the Mouse Embryo.

47. **Holland, P. W. H., and B. L. M. Hogan.** 1988. Expression of homeobox genes during mouse development: a review. Genes & Dev. **2::**773-782.

48. **Jaenisch, R., D. Jähner, P. Nobis, I. Simon, J. Löhler, K.Harbers, and D. Grotkopp.** 1981. Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell. **24::**519-529.

49. **Jaenisch, R.** 1980. Retroviruses and embryogenesis: microinjectionof Moloney murine leukemia virus into midgestation mouse embryos. Cell. **19::**181-188.

50. **Jaenisch, R.** 1988. Transgenic animals. Science. **240::**1468-1474.

51. **Jähner, D., H. Stuhlmann, C.L. Stewart, K. Harbers, J. Löhler, I. Simon, and R. Jaenisch.** 1982. *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. Nature. **298::**623-628.

52. **Johnson, P., and F. T.** 1990. Limited bidirectional activity of two housekeeping gene promoters: human HPRT and PGK. Gene. **88:**207-213.

53. **Jonsson, J., Q. Wu, K. Nilsson, and R. A. Phillips.** 1996. Use of a promoter-trap retrovirus to identify and isolate genes involved in differentiation of a myeloid progenitor cell line in vitro. Blood. **87:**1771-1779.

54. **Joyner, A.** 1991. Gene Targeting and Gene Trap Screens Using Embryonic Stem Cells: New Approaches to Mamalian Development. . **13:**649-655.

55. **Keller, G., K. Kennedy, T. Papayannopoulou, and M. V. Wiles.** 1993. Hemptopoietic Commitment during Embryonic Stem Cell Differentiation in Culture. Mol. and Cell Biol.:473-486.

56. **Keller, G. M.** 1995. In Vitro differentiation of embryonic stem cells. Current Biology. **7:**862-869.

57. **Kerr, W. G., G.P. Nolan, A.T. Serafini, and L.A. Herzenberg.** 1989. Transcriptionally defective retroviruses containing *lacZ* for the in situ detection of endogenous genes and developmentally regulated chromatin., p. *LIV*: 767-776. Cold Spring Harbor Symp. Quant. Biol.

58. **Kessel, M., and P. Gruss.** 1990. Murine developmental control genes. Science. **249::** 374-379.

59. **King, W., M. Patel, L. Lober, S. Goff, and M. Nguyen-Huu.** 1985. Insertion mutagenesis of embryonal carcinoma cells by retroviruses. Science. **228::**554-558.

60. **Korn, R. S., M; Neuhaus, H; Hensiling, U; Soininen, R; Zachgo, J; Gossler, A.** 1992. Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimaera embryos with high frequency and detect endogenous genes. Mech Dev. **39:**95-109.

61. **Kuehn, M. R., A. Bradley, E. Robertson, and M. Evans.** 1987. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. Nature. **326::**295-298.

62. **Lazo, P. A., V. Prasad, and P.N. Tsichlis.** 1987. Splice acceptor site for the *env* message of Moloney murine leukemia virus. J. Virol. **61::**2038-2041.

63. **Li, E., T.H. Bestor, and R. Jaenisch.** 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. **69::**915-926.

64. **Lindenbaum, M. H., and F. Grosveld.** 1990. An in vitro globin gene switching model based on differentiated embryonic stem cells. Genes & Dev. **4::**2075-2085.

65. **MacGregor, G. R., A.E. Mogg, J.F. Burke, and C.T. Caskey.** 1987. Histochemical staining of clonal mammalian cell lines expressing E.coli b-galactosidase indicates heterogeneous expression of the bacterial gene. Somatic Cell Mol. Genet. **13::**253-265.

66. **Mann, R., R. C. Mulligan, and D. Baltimore.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell. **33::**159-159.

67. **Martin, G. R.** 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. USA. **78::**7634-7638.

68. **McBurney, M., W. A. Staines, K. Boekelheide, D. Parry, K. Jardine and L. Pickavance.** 1994. Murine PGK-1 promoter drives widespread but not uniform expression in transgenic mice. Dev.dyn. **200:**278-293.

69. **McClannahan, T., S. Darlymple , M. Barkett and F. Lee.** 1993. Hematopoietic growth factor receptor genes as markers of lineage commitment during *in vitro* developement of hematopoietic cells. Blood. **81:**2903-2915.

70. **Meisler, M.** 1992. Insertional mutation of 'classical' and and novel genes in transgenic mice. TIG. **8**.

71. **Miller-Hance, W., M. Lacorbier, S. J. Fuller, S. M. Evans, G. Lyons, Schmidt C, Robbins J, and K. R. Chie.** 1993. *In vitro* chamber specification during embryonic stem cell development. J. Biol. Chem. **268:**25244-25252.

72. **Nagy, A., J. Rossant, R. Nagy, W. Nerwerly-Abramow, and J. Roder.** 1993. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc. Natl. Acad. Sci. USA. **90:**8424-8428.

73. **Nakano, T., H. Kodama, and T. Honjo.** 1994. Generation of lymphohematopoietic cells from embryonic stem cells in culture. Science. **265:**1098-1101.

74. **Neale, F., J. Clubb, D. Hotchkis, and S. Posen.** 1965. Heat stability of human placental alkaline phosphatase. J. clin. Path. **18:**359-363.

75. **Nieto, M. A., P. Gilardi-Hebenstreit, P. Charnay and D.G. Wilkinson.** 1992. A receptor protein tyrosine kinase implicated in the segmental patterning of the hindbrain and mesoderm. Development. **116:**1137-1150.

76. **Niswander, L., and G. Martin.** 1992. Fgf-4 expression during gastrulation, myogenesis, limb and tooth development in the mouse. Dev. **114:**755-768.

77. **Nusse, R., and H. E. Varmus.** 1992. Wnt genes. Cell. **69::**1073-1087.

78. **O'Kane, C. J., and W. Gehring.** 1987. Detection in situ of genomic regulatory elements in Drosophila. Proc. Natl. Acad. Sci. USA. **84:**9123-9127.

79. **Peckham, I., S. Sobel, J. Comer, R. Jaenisch, and E.Barklis.** 1989. Retrovirus activation in embryonal stem carcinoma cells by cellular promoters. Genes & Dev. **3:**2062-2071.

80. **Poirer, F., C.-T. J. Chan, P. M. Timmons, E. J. Robertson, M. J. Evans, and P. W. J. Rigby.** 1991. The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. Development. **113:**1105-1114.

81. **Price, J., D. Turner, and C. Cepko.** 1987. Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. Proc. Natl. Acad. Sci. USA. **84::**156-160.

82. **Quertermous, E. E., H. Hidai, M. A. Blanar, and T. Quertermous.** 1994. Cloning and Characterization of a basic helix-loop-helix protein expressed in early mesoderm and the developing somites. Proc. Natl. Acad, Sci. USA. **91:**7066-7070.

83. **Reddy, S., H. Rayburn, H. von Melchner, and H.E. Ruley** . 1992. Fluorescence-activated sorting of totipotent embryonic stem cells expressing developmentally regulated *lacZ* fusion genes. Proc. Natl. Acad. USA. **89:**6721-6725.

84. **Reddy, S., J.V. DeGregorgi, H. von Melchner, and H.E.Ruley.** 1991. Retrovirus promoter-trap vector to induce lacZ gene fusions in mammlian cells. J. Virol. **65:**1507-1515.

85. **Reik, W., H. Weiher, and R. Jaenisch.** 1985. Replication-competent Moloney murine leukemia virus carrying a bacterial

suppressor tRNA gene: Selective cloning of proviral and flanking host sequence. Proc. Natl. Acad. Sci. USA. **82:**1141-1145.

86. **Risau, W., H. Sariola , H. G. Zerwes, J. Sasse, P. Ekblom, R. Kemler, and T. Doetschman.** 1988. Vasculogenesis and angiogenesis in embryonic-stem-cell-derived embryoid bodies. Development. **102:**471-478.

87. **Robertson, E., A. Bradley, M. Kuehn, and M. Evans.** 1986. Germ-line transmission of genes introduced into cultured pluripotent cells by retroviral vector. Nature. **323::**445-448.

88. **Robertson, E. J., M.H. Kaufman, A. Bradley, and M.J. Evans.** 1983. Isolation, properties, and karyotype analysis of pluripotent (EK) cell lines from normal and parthenogenetic embryos, p. 647-663. *Cold Spring Harbor Conferences on Cell Proliferation.* Cold Spring Harbor Laboratory Press, New York.

89. **Robertson, E. J.** 1987. Teratocarcinomas and embryonic stem cells a practical approach. IRL Press, Washington DC.

90. **Rohdewohld, H., H. Weiher, W. Reik, R. Jaenisch, and M. Breindl.** 1987. Retrovirus integration and chromatin structuire: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites. J. Virol. **61::** 336-343.

91. **Rohwedel, J., V. Malysev, E. Bober, H. Arnold, J. Hescheler, and A. M. Wobus.** 1994. Muscle Cell Differentiation of Embryonic Stem Cells eflects Myogenesis in Vivo: Developmentally Regulated Expression of Myogenic Determination Genes and Functional Expression of Ionic Currents. Developmental Biology. **164:**87-101.

92. **Rossant, J., and R.A.Pedersen.** 1986. *Experimental approaches to mammalian development.* Cambridge University Press, Cambridge, England.

93. **Rossant, J., and N. Hopkins.** 1992. Of fin and fur: mutational analysis of vertebrate embryonic development. Genes & Dev. **6::**1-13.

94. **Russ, A., C. Friedel , K. Ballas, U. Kalina, D. Zahn, K. Strebhardt, and von H. Melchner.** 1996. Identification of genes

induced by factor deprivation in hematopoietic cells undergoing apoptosis using gene-trap mutagenesis and site-specific recombination. Proc. Natl. Acad. Sci. USA. 93:15279-15284

94a. **Sam M; Wurst W; Forrester L; Vauti F; Heng H and A. Bernstein.** 1996. A novel family of repeat sequences in the mouse genome responsive to retinoic acid. Mamm Genome 7: 741-748.

95. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

96. **Sanes, J. R., J.L.R. Rubenstein, and J.-F. Nicolas.** 1986. Use of a recombinant retrovirus to study post-implantation cell lineage in mouse embryos. EMBO J. **5::** 3133-3142.

97. **Sasaki, H., B. L. M. Hogan.** 1993. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation. Development. **118:**47-59.

98. **Sassoon, D., and N. Rosenthal.** 1993. Detection of Messenger RNA by in Situ Hybridization. *Methods in Enzym.* 225:384-404.

99. **Scherdin, U., K. Rhodes, and M. Breindl.** 1990. Transcriptionally active genome Regions are preferred targets for retrovirus integration. *J. Virol.* 64::907-912.

100. **Scherer, C., J. Chen , A. Nachabeh , N. Hopkins , and H. E. Ruley.** 1996. Transcriptional specificity of the pluripotent embryonic stem cell. Cell Growth Differ. **7:**1393-1401.

102. **Schnieke, A., K. Harbers, and R. Jaenisch.** 1983. Embryonic lethal mutation in mice induced by retrovirus insertion into the a1(I) collagen gene. *Nature.* 304:: 315-320.

103. **Schöler, H. R. S., N. Ruppert, K. Suzuki, Chowdhury, and P. Gruss.** 1990. New type of POU domain in germ line-specific protein Oct-4. *Nature.* **344**:435-439.

104. **Shih, C.-C., J.P. Stoye, and J.M. Coffin.** 1988. Highly preferred targets for retrovirus integration. *Cell.* 53::531-537.

105. **Singh, G., D. M. Supp, C. Schreiner, J. McNeish, H. J. Merker, N. G.Copeland, N. A. Jenkins, S. S. Potter, and W. Scott.** 1991. *legless* insertional mutation: morphological, molecular, and genetic characterization. Genes Dev. **5::** 2245-55.

106. **Skarnes, W., Moss, J., Hurtley, S. and Beddington, R.** 1995. Capturing genes encoding membrane and secreted proteins important for mouse development. Proc. Natl. Acad. Sci. **92:**6592-6596.

107. **Skarnes, W. C.** 1990. Entrapment vectors: A new tool for mammalian genetics. *Biotechnology.* 8::827-831.

108. **Skarnes, W. C., B. A. Aeurbach, and A. Joyner.** 1992. A gene trap approach in mouse embryonic stem cells : the lac Z reporter is activated by splicing , reflects endogenous gene expression and is mutagenic in mice. Genes and Dev. **6:**903-918.

109. **Sorge, J., A.E. Cutting, V.D. Erdman, and J.W. Gautsch.** 1984. Integration-specific retrovirus expression in embryonal carcinoma cells. *Proc. Natl. Acad. Sci. USA.* 81::6627-6633.

110. **Soriano, P., G. Friedrich, and P. Lawinger.** 1991. Promoter interactions in retrovirus vectors introduced into fibroblasts and embryonic stem cells. *J. Virol.* 65::2314-2319.

111. **Soriano, P., T. Gridley, and R. Jaenisch.** 1987. Retroviruses and insertional mutagenesis in mice: proviral integration at the Mov 34 locus leads to early embryonic death. Genes and Dev. **1:**366-375.

112. **Sosic, D., B. Brand-Saberi, C. Schmidt, B. Christ, and E. N. Olson.** 1997. Regulation of *paraxis* Expression and Somite Formation by Ectoderm- and Neural Tube-Derived Signal. Developmental Biology. **185:**229-243.

113. **Spence, S. E., D.J. Gilbert, D.A. Swing, N.G., Copeland, andN.A. Jenkins.** 1989. Spontaneous germ line virus infection and retroviral insertional mutagenesis in eighteeen transgenic *Srev* lines of mice. *Mol. Cell. Biol.* 9::177-184.

114. **Stuhlmann, H., Cone, R., Mulligan, R. C., and Jaenisch, R.** 1984. Introduction of a selectable gene into different animal tissue

by a retrovirus recombinant vector. *Proc. Natl. Acad. Sci. USA.* 81:: 7151 - 7155.

115. **Stuhlmann, H., Jaenisch, R., and Mulligan, R. C.** 1989. Construction and properties of replication-competent murine retroviral vectors encoding methotrexate resistance. *Mol. Cell. Biol.* 9::100 - 108.

116. **Stuhlmann, H., Jaenisch, R., and Mulligan, R. C.** 1989. Transfer of a mutant dihydrofolate reductase gene into pre- and postimplantation mouse embryos by a replication-competent retrovirus vector. *J. Virol.* 63::4857 - 4865.

117. **Takada, S., K. L. Stark, M. J. Shea, G. Vassileva, J. A. McMahon, and A. P. McMahon.** 1993. *Wnt-3a* regulates somite and tailbud formation in the mouse embryo. Genes & Development. **8:**174-189

118. **Takeuchi, T., Y. Yamazaki, Y. Katoh-fukui, R. Tsuchiya, S. Kondo, J. Motoyama,T. Higashinakagawa.** 1995. Gene trap capture of a novel mouse gene, *jumonji*, required for neural tube formation. Genes Dev. **9:**1211-1222.

119. **Tam, P. P. L. and R. S. P. Beddington.** 1987. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development. **99:**109-126.

120. **Tam, P. P. L., S. Meier ; A.G. Jacobson.** 1982. Differentiation of the metameric pattern in the embryonic axis of the mouse. II. Somitomeric organization of the presomitic mesoderm. Differentiation. **21:**109-122.

120b. **Tam P.P.L. and P.A. Trainor.** 1994. Specification and segmentation of the paraxial mesoderm. *Anat. Embryol (Berl).* **189:** 275-305.

121. **Thomas, K. R. and M. R. Capecchi.** 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. **51::**503-512

122. **Torres, M., A. Stoykova, O. Huber, K. Chowdhury, P. Bonaldo, A. Mansouri, S. Butz, R. Kemler, P. Gruss.** 1997. An

alpha-E-catenin gene trap mutation defines its function in preimplantation development. Proc. Natl. Acad. Sci. USA. **94:**901-906.

123. **Varmus, H.** 1988. Retroviruses. *Science*. 240::1427-1435.

124. **Varmus, H. E., N. Quintrell, and S. Ortiy**. 1981. Retroviruses as mutagens: insertion and excision of a non-transforming provirus alters expresion of a resident transforming provirus. *Cell*. 25::23-36.

125. **Varmus, H. E., and S.R. Swanstrom**. 1984. Replication of retroviruses. I*n* RNA tumor viruses. Vol. 2. N. T. R. Weiss, H. Varmus, and R. Coffin, editor. Cold Spring Harbor Laboratory,, Cold Spring Harbor, NY. 75-134.

126. **Vijaya, D. L. S., and H.L. Robinson**. 1986. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. *J. Virol.* 60::683-692.

127. **von Melchner, H., S. Reddy, and H.E. Ruley**. 1990. Isolation of cellular promoters by using a retrovirus promoter trap. *Proc. Natl. Acad. Sci. USA*. 87:3733-3737.

128. **von Melchner, H., J.V. DeGregorgi,, H. Rayburn, S. Reddy, C. Friedel, and H.E. Ruley.** 1992. Selective disruption of genes expressed in totipotent embryonal stem cells. Genes & Dev. **6:**919-927.

129. **von Melchner, H. and H. E. Ruley.**1989. Identification of cellular promoters by using a retrovirus promoter trap. J. Virol. **63:**3227-3233

130. **Wagner, E. F.** 1990. Mouse genetics meets molecular biology at Cold Spring Harbor. *New Biologist*. 2::1071-1074.

131. **Wang, R. C., R Bautch, V.** 1992. Embryonic stem cell-derived cystic embryoid bodies form vascular channels: an *in vitro* model of blood vessel development. Development. **114:**303-316.

132. **Weiher, H., T. Noda, D.A. Gray, A.H. Sharpe, and R.Jaenisch**. 1990. Transgenic mouse model of kidney disease: insertional inactivation of ubiquitously expressed gene leads to nephrotic syndrome. *Cell*. 62::425-434.

133. **Weiss, R., N. Teich , H. Varmus, and J. Coffin.**1985. RNA tumor viruses: appendices, 2nd ed, vol. 2, Cold Spring Harbour.

134. **Wiles, M. V., and G. Keller**. 1991. Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. *Development*. 111::259-267.

135. **Wilkins, A. S.** 1992. Mus musculus: From oocyte to implanting blastocyst. I*n Genetic analysis of animal development.* Ed. Wiley-Liss, 161-196. NY.

136. **Williams, B. A. and P. Ordahl.** 1994. Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. Development. **120:**785-796.

137. **Wilson, C., R.K. Pearson, H.J. Bellen, C.J. O'Kane, U. Grossniklaus, and W.J. Gehring**. 1989. P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in *Drosophila. Genes & Dev.* 3;:1301-1313.

138. **Wilson, J. M., D.M. Jefferson, J.R. Chowdhury, P.M. Novikoff, D.E. Johnston, and R.C. Mulligan**. 1988. Retrovirus-mediated transduction of adult hepatocytes. *Proc. Natl. Acad. Sci. USA*. 85::3014-3018.

139. **Wolf, C. T., C.; Stoetzel, C.; Thisse, B.; Gerlinger, P.; Perrin-Schmitt, F.** 1991. The *M-twist* gene of *Mus* is expressed in subsets of mesodermal cells and is closely related to the *Xenopus X-twi* and *Drosophila twist*. Dev. Biol. **143:**363-373.

140. **Wood, S. A., W.S. Pascoe, C Schmidt, R. Kemler, M.J. Evans,and N.D. Allen**. 1993. Simple and efficient production of embryonic stem cell-embryo chimeras by coculture. *Proc. Natl. Acad. Sci. USA*. 90:: 4582-4585.

141. **Woychik, R. P., R.L. Maas, R. Zeller, T.F. Vogt, and P. Leder**. 1990. 'Formins': proteins deduced from the alternative transcripts of the *limb deformity* gene. *Nature*. 346::850-853.

142. **Wurst, W., Rossant, J., Priedeaux, V., Kownacka, M., Joyner, A., Hill, D., Guillemot, F., Gasca, S., Cado, D.,**

Auerbach, A. and Ang, S. 1995. A Large Scale Gene Trap Screen for Insertional Mutants in Developmentally Regulated Genes in Mice. **139:**889-899.

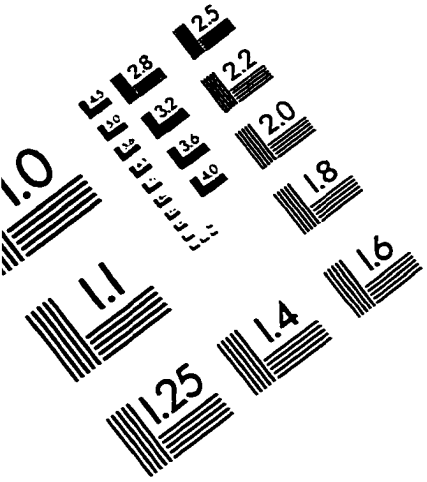
143. **Xiong, J.-W., R. Battaglino, and H. Stuhlmann.** 1997. A Large Genetic Screen for Developmental Genes in ES Cells and Embryoid Bodies Using AP Entrapment Vectors. (submitted).

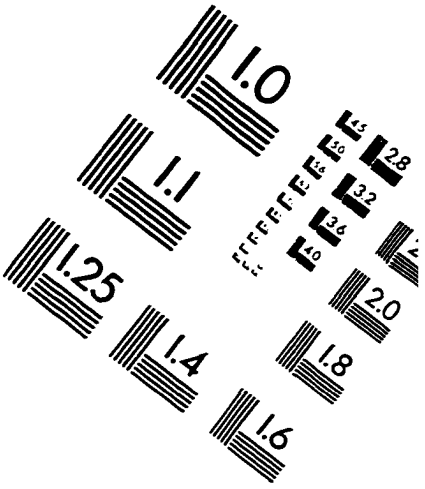
144. **Yamada, G., C. Kioussi, F. R. Shubert, Y. Eto, K. Chowdhury, F. Pituello, and P. Gruss.** 1994. Regulated Expression Of Brachyury(T), NKX1.1 And Pax Genes in Embryoid Bodies. Biochemical and Biophysical Research Communications. **199:**552-563.

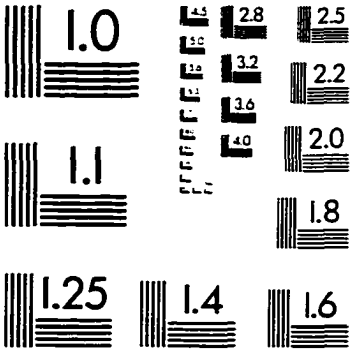
145. **Yokoyama, T., N. G. Copeland, N. A. Jenkins, C. A. Montgomery, F. F. Elder, P. A. Overbeek.** 1993. Reversal of left-right asymmetry: a situs inversus mutation [see comments]. Science. **260::**679-82.

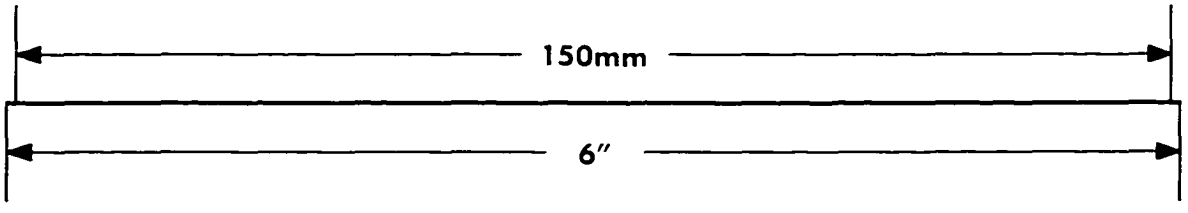
146. **Zhou, X., H. Sasaki, L. Lowe, B. L. M. Hogan, and M. Kuehn.** 1993. *Nodal* is a novel TGF-B-like gene expressed in the mouse node during gastrulation. Nature. **361:**543-547.

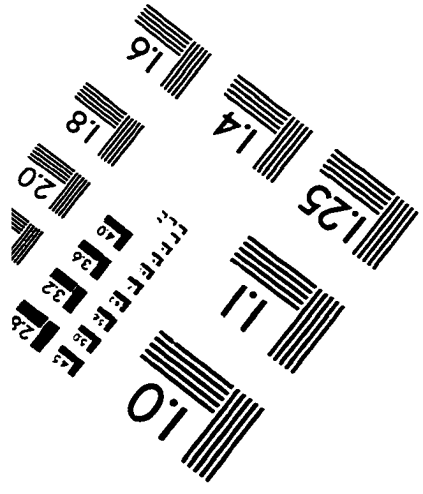
IMAGE EVALUATION TEST TARGET (QA-3)

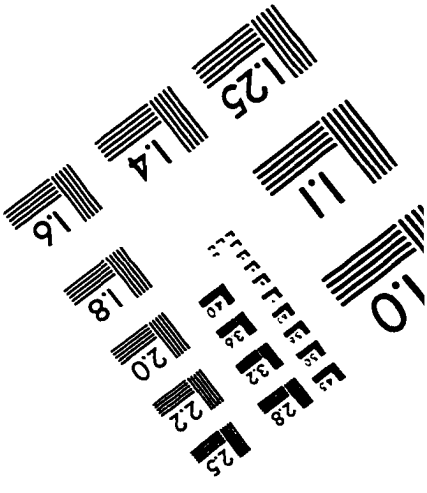












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989