IDENTIFICATION OF A NOVEL GENE, TRAP1, USING A REPLICATION-COMPETENT PROMOTER-TRAP RETROVIRUS IN THE MOUSE MAMMARY GLAND

Josée Coulombe

Thesis submitted to
the School of Graduate Studies and Research
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

Department of Biochemistry
University of Ottawa,
Ottawa, Ontario
Canada

ISBN 0-612-19938-X

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

Breast cancer remains one of the leading malignancies in women of the developed world despite the extra efforts that have been deployed in the past decade or so to the study of this disease. It is well known that the degree of aggressivity of the tumor is related to the stage of differentiation. Transformation of a cell having a high proliferation potential will lead to a more aggressive tumor than transformation of a more differentiated cell. For this reason it is important to have a good understanding of the process of mammary differentiation.

The work in this thesis was directed at finding novel genes that are involved in the normal development of the breast with the hope that these may open new avenues for breast cancer therapies. The approach used was to look for genes that are expressed during breast development in the living mouse. A promoter-trap retrovirus containing an oncogene as the promoterless indicator gene was introduced locally in the mammary glands of pubescent female mice. Integration of proviruses near an active cellular promoter lead to the expression of the oncogene from a cellular promoter which was detected by the formation of a tumor. To identify genes that are specifically active during differentiation of the breast tissue we looked for the appearance of tumors after the animals were mated. Two promoter-trap retroviruses were tested. The first one (COPT) was replication deficient and although it did work *in vitro* it did not give a high enough level of infection to be used *in vivo*. The second virus (PyT) was replication competent and was found to be a more suitable promoter-trap to be used *in vivo*. This was demonstrated by the isolation of a novel murine gene named TRAP1.

ACKNOWLEDGEMENTS

I want to express my gratitude to my supervisor, Dr. Douglas Gray, first for giving me the opportunity to do a PhD in his laboratory, and for communicating his love of research to me. I greatly appreciated his patience and trust. Working in his laboratory has been a pleasant and challenging experience.
I want to extend my thanks to the other investigators of the Cancer Research Group, Drs. Bell, McBurney, and Vanderhyden for their contribution to my education and their support.
I would like to thank the many people whom I worked with during this degree for their help and good times (most of the time!). More specifically I want to thank:
Yvonne Avis, not only for her valuable technical assistance but also for her friendship. Her sense of humour kept us all a little closer to sanity (or is it the opposite?). Thank you for your attentive ear.
Manon Dubé pour son aide et sa grande patience.
Beth Mason for all the help generously provided during those numerous years, and for being a friend. Thank you so much.
Dr. Barbara Vanderhyden for her inspiration. Thank you for always keeping your door open.
Elizabeth Douville, d'abord pour m'avoir fait connaître le groupe de recherce sur le cancer, mais surtout pour sa fidèle amitié (malgré quatre années de colocation!).
Gaël Vidricaire pour ses encouragements sans borne.
Colin for his support and patience.
Enfin, un gros merci à ma famille, Paul, Colette, Isabelle, Lucie, Dominique et Martin pour leur support et encouragements tout au long de mes études. Merci aussi à Emilie et Louis-Philippe pour l'intéret porté à la santé de mes souris.

I want to thank the "Fond pour les Chercheurs et l'Aide à la Recherche" for a studentship, and the Medical Research Council and the National Cancer Institute for funding this project.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

LIST OF ABBREVIATIONS

BES	N,N-bis [2-hydroxyethyl]-2-aminoethanolsulfonic acid
bHLH	basic helix-loop-helix
CAT	Chloramphenicol acetyl transferase
DEPC	Diethylpyrocarbonate
DHFR	Dihydrofolate reductase
DNA	Deoxyribonucleic acid
ECM	Extra cellular matrix
EDTA	Ethyl diaminetetra acetic acid
ES cells	Embryonic stem cells
EST	Expressed sequence tag
HGF	Hepatocyte growth factor
IRES	Internal ribosomal entry site
G418	Geneticin
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
GNRF	Guanosine nucleotide releasing factor
KCl	Potassium chloride
KDa	Kilodalton
KH_2PO_4	Potassium dihydrogen orthophosphate
K_2HPO_4	Dipotassium hydrogen orthophosphate
LB	Luria broth
LiCl	Lithium chloride
LTR	Long terminal repeat
LZ	Leucine zipper
MEM	Minimal essential medium
$MgSO_4$	Magnesium sulfate
MMTV	Mouse Mammary Tumor Virus
MoMuLV	Moloney murine leukemia virus
MOPS	3-[N-Morpholine] propane-sulfonic acid
MT	Middle T antigen
NaCl	Sodium chloride
Na_2HPO_4	Disodium hydrogen orthophosphate
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
ONPG	O-nitrophenyl-β-D-galactopyranoside
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

pgk	Phosphoglycerate kinase
PEG	Polyethlyne glycol
PyMT	Polyomavirus middle T antigen
PyV	Polyomavirus
RNA	Ribonucleic acid
RACE	Rapid amplification of cDNA ends
SA	Splice acceptor
SDS	Sodium dodecyl sulfate
SM	Storage medium
TB	Terrific broth
TBE	Tris/borate/EDTA
TE	Tris/EDTA
TRIS	Tris (hydroxymethyl) aminomethane
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER 1

INTRODUCTION

Breast cancer

Breast cancer is affecting an ever growing number of women. It is estimated that 11% of Canadian women have a lifetime risk of developing breast cancer (Statistiques canadiennes sur le cancer 1996). Despite the progress that has been made regarding treatments, the mortality rate remains practically unchanged, a consequence of a steady rise in the incidence every year (Harris *et al.* 1992a; Statistiques canadiennes sur le cancer 1996). Breast cancer is more prevalent in the industrialized world but it is becoming more frequent in developing countries. The factors that put women from these countries at higher risk are not known. The fact that immigrants who come from countries with a low incidence of breast cancer become at higher risk when they move to a country where the incidence is higher (McMichael and Giles 1988; Buell 1973) would indicate that the environment is to blame rather then the genetic makeup of the person. The environmental factors responsible for these differences have not been identified. The only factor that could be associated with an increased risk, beside a genetic predisposition, is an elevated lifetime number of menstrual cycles (early menarche, late menopause, no child birth). This may be a reflection

of the life styles of more industrialized countries and represents a gigantic challenge for the development of prevention strategies.

Despite extensive research the etiology of breast cancer is still unknown. To begin to understand the cause of breast cancer, panels of human breast tumors have been screened for the presence of genetic alterations. A number of protooncogenes have been found to be altered or over expressed in certain mammary tumors (*neu*/Her-2, EGFR, c-Ha-*ras*, c-*myc*, TGF-α, cyclin D1, p53; van de Vijver *et al.* 1988; Slamon *et al.* 1989; King *et al.* 1985; van de Vijver *et al.* 1987; Zhou *et al.* 1987; Berger *et al.* 1988; Lundy *et al.* 1991; Fitzpatrick *et al.* 1984; Sainsbury *et al.* 1987; Sainsbury *et al.* 1985; Sainsbury *et al.* 1985; Yokota *et al.* 1986; Slamon *et al.* 1984; Escot *et al.* 1986; Mariani-Costantini *et al.* 1988; Garcia *et al.* 1989; Dickson *et al.* 1987; Zukerberg *et al.* 1995; Weinstat-Saslow *et al.* 1995; Lammie *et al.* 1991; Schuuring *et al.* 1992; Malkin *et al.* 1990; Harris *et al.* 1992a; Reik *et al.* 1985; Cornelis *et al.* 1994), and the list is likely to expand as more genes involved in cell growth and differentiation are identified. While a number of these genes have been shown to promote tumorigenesis in transgenic mice (Stewart *et al.* 1984; Leder *et al.* 1986; Sinn *et al.* 1987; Andres *et al.* 1987; Schoenenberger *et al.* 1988; Tremblay *et al.* 1989; Muller *et al.* 1988; Wang *et al.* 1994), they are not believed to initiate the formation of spontaneous human tumors. These mutations may be subsequent to transformation, due in part to the genomic instability associated with neoplastic cells, and/or may confer a growth advantage to the tumor. Unfortunately, none of these alterations are specific to breast cancer and they give very little clue as to what the early events are. There are still no specific markers available for early diagnosis of breast cancer, reducing the chances of more optimistic

prognoses. Some of these genetic alterations however, are useful markers in the staging of breast cancer which helps determine which treatment should be applied. For example, carcinomas in which the Her-2/*neu* protooncogene is overexpressed have poor prognoses. These tumors are usually negative for estrogen and progesterone receptors, and this was recently shown to be linked to the increased activity of the Her-2/neu gene product (Pietras *et al.* 1995). The presence of estrogen and progesterone receptors in a tumor is associated with a good prognosis. These tumors are often responsive to hormonal and chemotherapy.

Approximately 5% of all breast cancers (and 25% of those with onset before age 30) can be attributed to a genetic predisposition (Claus *et al.* 1991). Two genes, BRCA1 and BRCA2 (located on chromosomes 17 and 13 respectively; Hall *et al.* 1990; Wooster *et al.* 1994) have been linked to a high incidence of breast cancer in certain families. Together BRCA1 and BRCA2 account for the majority of families with early onset of breast cancer. BRCA1 is associated with 45% of families with an increased incidence of breast cancer alone but with 80% of families with a high incidence of both breast and ovarian cancer (Easton *et al.* 1993). BRCA2 is associated with a similar proportion (45%) of early onset of breast cancer alone. Both genes have been cloned, and both encode large proteins which may be related (Jensen *et al.* 1996).

When it was cloned (Miki *et al.* 1994) BRCA1 was believed to be a transcription factor, based on the presence of a zinc finger motif in its amino acid sequence, but it was recently shown to be a secreted molecule that shares some features with a family of proteins called granins (Huttner *et al.* 1990). BRCA1 germ line mutations have been identified in families with an increased incidence of early onset of breast and ovarian cancer, and tumors

from affected individuals were shown to have lost the wild type allele, suggesting that BRCA1 is a tumor suppressor gene (Neuhausen and Marshall 1994). This is supported by the demonstration that expression of BRCA1 in breast cancer cells reduced their proliferation (Holt *et al.* 1996). The effect was found to be specific to breast cancer cells. BRCA1 is expressed in a number of tissues undergoing proliferation and differentiation (Marquis *et al.* 1995), including breast epithelium, and is induced by ovarian hormones (Gudas *et al.* 1995; Marquis *et al.* 1995). Secretion of some granins has been shown to be regulated by hormones. BRCA1 may therefore be a breast epithelium-specific growth suppressor. Given the high penetrance of the BRCA1 germ line mutation (Claus *et al.* 1991; Futreal *et al.* 1994) one might expect to find a large number of sporadic tumors with BRCA1 mutations. A survey of sporadic breast tumors, looking for loss of heterozygosity and mutations in the gene, indicated that somatic BRCA1 mutations were rare (Futreal *et al.* 1994; Cropp *et al.* 1994). However, a study of the expression level showed reduced BRCA1 expression in sporadic breast tumors (Thompson *et al.* 1995) indicating that the mechanism leading to the loss of BRCA1 function may be different in sporadic and familial breast cancer. BRCA1 mutations therefore, may be more relevant to the more common (87%) late onset category of breast cancer than was first believed.

BRCA2 has only been cloned recently, and expression and functional data are not yet available. It has been postulated to be related to BRCA1 based only on sequence homology in the granin domain (Jensen *et al.* 1996). Other loci, including the p53 gene (Malkin *et al.* 1990), are thought to be involved in hereditary cases of early onset of breast cancer.

The lack of characteristic markers for breast cancer makes early diagnosis more difficult; mammography remains the major mean of detecting mammary tumors. This test however is not a routine one and is used only in higher risk groups (members of families with a genetic predisposition and women over 50 years of age, although not widely available yet). For the other women, breast cancer is often detected at later stages of the disease decreasing the chances of a good prognosis. Although treatments have improved, they only offset the growing incidence, keeping the mortality rate at a steady level (Harris *et al.* 1992a; Statistiques canadiennes sur le cancer 1996). The most efficient treatments remain the breast-conserving surgery (lumpectomy) followed by irradiation (Harris *et al.* 1992b). Despite these drastic measures, the rate of relapse is high and metastatic events are very common and occur rapidly, emphasizing again the importance of an early diagnosis.

Because prevention of breast cancer is still in its infancy it is important to improve on both early detection and treatments. Although treatments have improved, the current therapies based mostly on the removal of the tumor mass followed by irradiation, are not sufficient to reduce the number of lives lost to breast cancer every year.

As more genetic alterations are discovered in all types of malignancies, it becomes more evident that neoplasia is the result of a deregulation of the controls of cellular growth and differentiation. The new thinking in cancer therapy is toward strategies directed at changing the fate of tumor cells into a more differentiated state. It has long been recognized that the aggressivity of a tumor is related to the state of differentiation of the tumor cells , with a less differentiated cell developing into a more aggressive tumor. In recent years a wealth of information has emerged on the pathways leading to cellular proliferation or

differentiation and it appears that different tissues use different signals as well as different combinations of effector molecules to drive the cell to divide or differentiate. This means that each type of cancer may require a slightly different therapeutic approach, and it is therefore important to understand the mechanisms controlling differentiation of various organs.

The work in this thesis was directed at developing a strategy to look for novel genes involved specifically in the development of the normal murine mammary gland. It is hoped that as the genetic controls responsible for breast development are identified they may be used as tools for early detection as well as the design of novel cancer therapies.

Mouse mammary gland development

The mammary gland is unusual in that, unlike other organs, it is not fully differentiated at birth. It awaits puberty to undergo further development and reaches full differentiation during pregnancy. In the mouse embryo the mammary anlage slowly develops into a rudimentary ductal tree between day 10 and 17 (Sakakura 1987). This proliferation depends heavily on the presence of the surrounding mesenchyme. Embryonic mammary glands can be induced to proliferate *in vitro* but only in the presence of the mesenchyme. The nature of the interaction between the epithelium and the mesenchyme is not well understood but is undoubtedly reciprocal. The embryonic stage of mammary development does not appear to require any hormone. Although mammary glands can grow and proliferate upon treatment with insulin, prolactin and aldosterone in culture, these hormones don't appear to be involved in the embryo *in vivo* (Sakakura 1987). One exception

is in the male embryo where the mammary epithelium renders the mesenchymal cells responsive to androgens causing these latter to contract, crushing the epithelium and destroying the glands (Sakakura 1987).

At birth, the mouse mammary glands contain 15 to 20 branched ducts but no alveolus. From that point until puberty there is very little proliferation of the epithelium (Daniel and Silberstein 1987). The glands only grow isometrically. Maturation of the ovaries and the production of the sex hormones estrogen and progesterone at puberty trigger a burst of growth in the mammary glands. This second stage of development occurs around the fourth to the eighth week of age in mice. The mammary ductal tree expands, infiltrating the surrounding stroma. The end result is a very extensive array of branching ducts converging toward a central duct which connects with the nipple. There is still very little alveolar development at this point. The growth of the ductal tree is lead by the end buds which form at the extremity of each duct. These respond to both systemic hormones for the growth and to local signal for the direction and the extent of the growth. The end buds progress until they reach the limit of the fat pad, which is the differentiated stroma. Mammary development then pauses until the onset of pregnancy.

Elevated levels of estrogen and progesterone during pregnancy initiate the final stage of mammary development which is characterized mainly by the development of the alveoli (Daniel and Silberstein 1987). As the pregnancy progresses, alveoli start forming at the end of each duct in response to the lactational hormones prolactin and placental lactogen and the genes coding for the milk proteins are progressively activated, reaching their maximum level of expression during lactation (Rosen 1987). Later during pregnancy, cellular secretion

begins due to the presence of lactogenic and adrenal steroid hormones. At parturition the levels of estrogen and progesterone drop abruptly resulting in active lactation. The latter continues for as long as nursing is maintained, after which the glands resorb, via the process of apoptosis (Strange *et al.* 1992; Marti *et al.* 1994; Li *et al.* 1996), returning to their pre-pregnancy state.

Up until recently our knowledge of the signals controlling the development of the mammary gland was more or less limited to the hormones involved. In recent years a number of genes involved in cell growth and division have been identified and their expression pattern as well as the effect of some of these have been studied in normal breast development. A number of genes have been found to be up regulated by hormones or signal peptides. Some of the molecules that mediate the stroma/epithelium interactions have started to emerge. For example, the hepatocyte growth factor (HGF, also called scatter factor) is produced by stromal cells and the receptor, the c-*met* protooncogene, is found on mammary epithelial cells (Niranjan *et al.* 1995; Moghul *et al.* 1994). HGF was shown to have either a morphogenic or mitogenic effect on epithelial cells, depending on their type (myoepithelial vs luminal) (Niranjan *et al.* 1995). Another example is the c-kit receptor and its ligand, the stem cell factor (also called kit ligand). c-kit is normally expressed at low levels in the mammary gland epithelium (Matsuda *et al.* 1993). The role of c-kit in the mammary gland is not clear, as different groups found either a loss of c-kit expression (Natali *et al.* 1992a; Natali *et al.* 1992b), or co-expression of c-kit and the kit ligand (KL) (Hines *et al.* 1995) in breast tumors. Finally, stromal cells produce extracellular matrix components in response to signals from the epithelium. The composition of the ECM varies

according to the type and location of the epithelium (Keely *et al.* 1995). These genes however are not exclusive to mammary development and it is still not clear what are the key events leading to mammary differentiation. The aim of the work presented in this thesis was to identify genes specifically involved in breast differentiation using a novel approach: an *in situ* promoter-trap.

Promoter-trap

Promoter-trap is a strategy that was developed to identify active cellular genes. A promoterless indicator gene is inserted in the genome where integration downstream of an active promoter may lead to its expression. The disrupted transcription unit can be cloned using the transgene as a tag. Because promoter-traps need to integrate in the gene in order to be expressed, they make perfect insertional mutagens. For this reason, promoter-traps have been a very effective strategy for cloning genes which are developmentally regulated in the mouse, using ES cells. ES cells are pluripotent stem cells that are derived from the inner cell mass of a blastocyst and can be propagated in culture while keeping their full differentiation potential (Martin 1981). After manipulation, ES cells can be reintroduced into blastocysts to generate chimeric embryos, to look for developmental defects. Because the insertion is disruptive, breeding the animals to homozygousity generates a null mutant which allow detection of recessive phenotypes. By including a visual marker such as the lacZ gene in the vector, the pattern of expression of the trapped gene can be followed through embryonic development by staining embryos with the chromogen X-gal. Promoter-traps have increased considerably the efficiency of cloning by insertional mutagenesis since they

select for insertions into genes only, as opposed to inert sequences, decreasing significantly the number of ES clones to be screened for a phenotype. The enrichment was estimated to be 100 to 1000 fold (Chang *et al.* 1993). Promoter-traps also facilitate the cloning of the integration site since the transgene is located in the gene. In previous insertion mutagenesis strategies (reviewed in Gridley *et al.* 1987) the insertion could have an effect several hundreds of kilobases away from the site of integration often making the identification and cloning of the affected gene difficult.

Promoter-trap vectors used for insertional mutagenesis in ES cells can be of two types: a plasmid vector which is electroporated, or a retroviral vector. Rossant and her colleagues used a promoter-trap plasmid vector with the lacZ gene as the promoterless indicator gene and the neomycin phosphotransferase gene under the control of the human β-actin promoter to select for integrations (Gossler *et al.* 1989). To increase the chances of the lacZ gene to be expressed, a splice acceptor site was inserted in front, so that integration of the construct in an intron would allow lacZ sequences to be spliced with the endogenous transcript (also called gene-trap). They found that 1% of all integration events, as scored by G418 resistance, were also expressing lacZ (Wurst *et al.* 1995). Thirteen percent of the lacZ expressing clones that were used to generate chimeras (36 clones) showed a restricted pattern of expression at day 8.5 of embryogenesis, the period of development they chose to study. The cloning of the integration site from some transgenic lines that showed an embryonic lethal phenotype led to the identification of novel genes (Skarnes *et al.* 1992; Gasca *et al.* 1995). Other groups have used similar vectors to clone developmentally regulated genes from ES cells (Takeuchi *et al.* 1995) or F9 cells (Okazaki *et al.* 1994). The

promoter-trap strategy has also been integrated into gene targeting vectors (Vaulont *et al.* 1995). With the use of an internal ribosome entry site (IRES) both the selectable marker and lacZ are expressed from the promoter of the targeted gene as part of a bicistronic message. This greatly diminishes the number of non-homologous integration events because the selectable marker needs a cellular promoter to be expressed. If one can extrapolate from the results reported by Rossant's group where 1% of random integrations trapped a cellular promoter, the use of a promoter-trap targeting vector can possibly reduce the background of homologous recombination by two orders of magnitude.

Soriano and coworkers introduced a similar gene-trap, retroviral vector into ES cells (Friedrich and Soriano 1991). To avoid having to select for lacZ expression among G418 resistant clones, they used a fusion of the lacZ and neomycin phosphotransferase gene, called β-geo. This should ensure that all G418 resistant clones also express lac-Z. They found this to be true in 95% of their clones based on X-gal staining. Further analysis of some of these apparently lacZ negative clones showed staining during certain stages of embryogenesis indicating that the expression of lacZ in the ES cells might have been below the detection level of X-gal staining. They included a splice acceptor site in front of β-geo, and to avoid removing the viral ψ packaging sequence by splicing from a viral upstream splice acceptor sequence, they positioned their SAβ-geo cassette in the reverse orientation relative to the viral transcription (hence the name ROSAβ-geo for *r*everse *o*rientation *s*plice *a*cceptor β-geo). They reported that the addition of the splice acceptor site in front of the promoterless gene (tested in a plasmid vector) increased the frequency of the promoter-trap by 45-fold (from 0.1% to 4.6%). They also found that a retroviral vector is 2.5 times more

efficient than a plasmid vector (they compared electroporation of the retroviral vector with infection with the virus). Out of 24 transgenic lines that showed X-gal staining during embryonic development, 9 turned out to be recessive lethal (38%). One of these insertion mutants was found to have disrupted the transcriptional enhancer factor 1 gene (TEF-1; Chen *et al.* 1994). Using the same vector, another group reported the disruption of the BTF-3 general transcription factor gene (Deng and Behringer 1995).

Yet another group has designed a similar retroviral promoter-trap that facilitates the cloning of the integration site by including a pBR322 origin of replication in the vector. They used their vector to look for genes involved in neuronal differentiation using F9 cells (Okazaki *et al.* 1994).

Ruley's group designed a different type of retroviral promoter-trap vector (von Melchner and Ruley 1989). They chose to insert the promoterless reporter gene in the U3 region of the MoMuLV LTR, upstream of the viral promoter. This positioned the transgene 30 base pairs away from host DNA. They initially inserted the histidinol dehydrogenase (*his*) coding sequence (conferring resistance to histidinol) in the LTR, and the neomycin phosphotransferase gene under the transcriptional control of the Herpes Simplex Virus thymidine kinase gene promoter, in place of the viral genes. By comparing their titers (1.4 - 1.9 × 10^6 cfu/ml) to those of other MoMuLV-based vectors, they demonstrated that the *his* coding sequence (1.8 kb) in the 3' LTR did not interfere with the ability of the virus to be passaged. They were able to introduce an insert as large as 3.1kb (lac-Z) in that site without affecting the titer (Reddy *et al.* 1991). Stuhlmann and coworkers also introduced a 1.8 kb cassette containing the DHFR cDNA and obtained replication competent viruses at high

titers (Stuhlmann *et al.* 1989). The ratio of histidinol resistant over G418 resistant clones was 3.8×10^{-4}, meaning that one integration event out of 2,500 allowed histidinol dehydrogenase expression. This low frequency of promoter-trap was later demonstrated to be due to the low sensitivity of histidinol (Chang *et al.* 1993). The frequency was 1 in 400 when hygromycin was used as the promoterless gene, and 1 in 200 for either thymidine kinase or lacZ. This represents a 46 and 23-fold difference with the frequency obtained with the ROSAβ-gal vector of Soriano's group (11.6%). Unlike their ROSAβ-geo vector, the ROSAβ-gal vector has separate lacZ and neo genes (but still has a splice acceptor site in front of lacZ) so that they could calculate separately the number of integrations and the number of promoter-trap events which gave a frequency of promoter-trap of 11.6%. This 23 and 46-fold difference is in the same range as what was reported by these same authors, when comparing a vector with a splice acceptor site to one without (45-fold) indicating that the U3 vector is as efficient as the ROSA vector without a splice acceptor site.

Using various selectable markers in the U3 region, Ruley and coworkers demonstrated that their virus could trap novel genes in both NIH 3T3 fibroblasts (von Melchner *et al.* 1990) and ES cells (von Melchner *et al.* 1992), including developmentally regulated genes such as the REX-1 gene.

Retroviruses

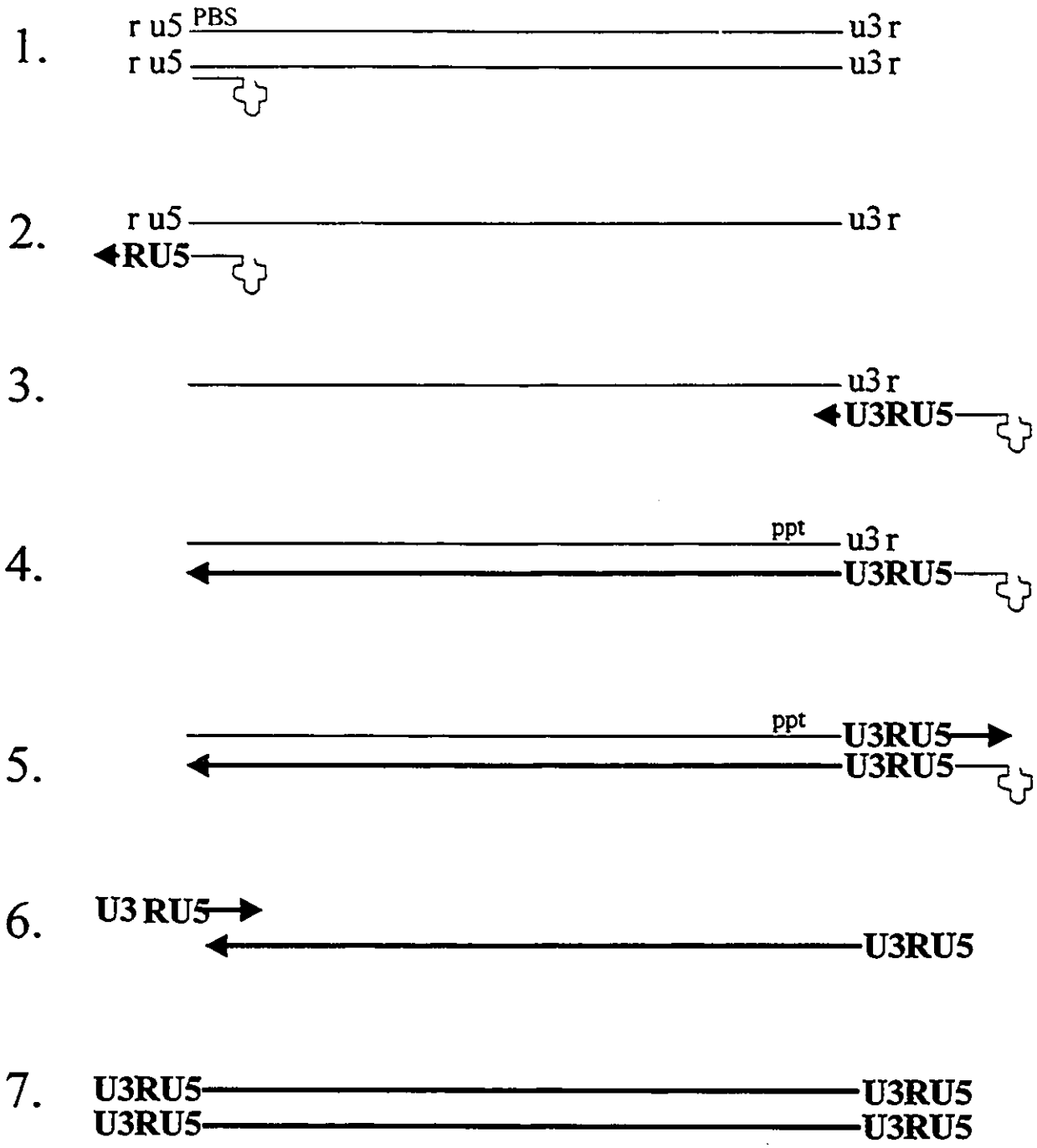
Retroviruses are positive strand RNA viruses which go through a DNA intermediate to replicate. Upon entry into the cell through specific receptors (cationic amino acid transporter for MoMuLV; Wang *et al.* 1991; Kim *et al.* 1991) the viral RNA is converted

to double stranded DNA by the viral enzyme reverse transcriptase. The DNA migrates to the nucleus and integrates in the host genome, where it is called the provirus, and is replicated as part of the cellular DNA. Provirus integration requires cellular DNA replication and occurs via recombination which is orchestrated by the viral integrase (encoded by the pol gene). In the process, 4-6 nucleotides (depending on the retrovirus) of host sequences are duplicated on each side of the integration site, the only alteration made to the host DNA. Although integration occurs through a precise recombination mechanism, it does not depend on any specific cellular sequences, meaning that integration could happen anywhere in the genome. However, there is considerable evidence indicating that integration tends to take place preferentially around DnaseI hypersensitivity areas, which often correspond to regions of promoter activity (Vijaya *et al.* 1986; Rohdewohld *et al.* 1987; Schuback and Groudine 1984; Scherdin *et al.* 1990). In the proviral form, viral sequences are transcribed to produce full length genomic RNA as well as some spliced transcripts from which the viral proteins are translated. These include structural proteins (gag and env gene products) necessary for the assembly of infectious particles, and enzymatic proteins (pol gene products) which are packaged with two copies of genomic RNA in the virion. Finally, the mature virus particles are produced by budding out of the cell to complete the life cycle.

The provirus is not a direct linear copy of the RNA genome. Two direct repeats (the LTRs) are located at each end of the provirus. These sequences contain all the transcription control elements; promoter, enhancer and polyadenylation signal. The LTRs are formed during the process of reverse transcription (Figure 1.1). The synthesis of DNA complementary to the viral RNA starts at a site called the primer binding site from a cellular

Figure 1.1 Reverse Transcription of Viral Genomic RNA

Thin lines and small letters represent RNA, thick lines and capital letters represent DNA. **1.** Schematic of two identical copies of the MoMuLV RNA genome. The "r" sequence (68pb) is found at both ends of the RNA, u5 (76bp) is unique to the 5' end, and u3 (449bp) is unique to the 3' end. The cloverleaf-like structure represents a cellular proline tRNA molecule that is partly complementary to the primer binding site (PBS). **2.** Reverse transcription is initiated by the viral polymerase (reverse transcriptase) at the primer binding site using the tRNA as a primer, and stops at the 5' end of the RNA. RNAseH (a subdomain of the polymerase) digests the RNA portion of the RNA/DNA duplex, exposing the RU5 sequence of the DNA strand. **3.** The nascent (-)strand DNA is translocated to the 3' end of the RNA template at the complementary r/R site. **4.** (-)strand DNA synthesis resumes until the end of the RNA template. **5.** RNaseH nicks the rest of the RNA template providing an RNA primer for the initiation of the second strand of DNA at the polypyrimidine tract (ppt). The (+)strand DNA synthesis uses the (-)strand DNA as template and stops in the primer binding site. **6.** The (+)strand is then translocated to the end of the (-)strand at the primer binding site and the synthesis of both DNA strands is completed. **7.** The end result is a double stranded DNA molecule with two long terminal repeats (LTRs).


1.
r u5
PBS
u3 r
r u5
u3 r
2.
r u5
u3 r
RU5
3.
u3 r
U3RU5
4.
ppt
u3 r
U3RU5
5.
ppt
U3RU5
U3RU5
6.
U3 RU5
U3RU5
7.
U3RU5
U3RU5
U3RU5
U3RU5

tRNA primer (proline tRNA in MoMuLV), and stops upon reaching the 5' end of the RNA (ru5). As DNA is synthesized, RNaseH (a function encoded by a domain of the polymerase) degrades the RNA part of the u5r RNA:DNA duplex exposing the U5R portion of the DNA. The nascent (-)stand DNA is then translocated to the 3' end of the RNA template (this may be on the same or on the other RNA molecule) at the complementary r site. Elongation of the (-)strand DNA resumes up to and including the primer binding site. During that time, the RNaseH introduced a nick in the RNA:DNA duplex at the start of the u3 and the RNA is used as a primer to initiate the (+)strand DNA synthesis which uses the (-)strand DNA as a template. The (-)strand DNA is again translocated, this time onto the (+)strand DNA molecule and joined at the primer binding site. The newly joined DNA molecules contain all the sequences necessary to complete the synthesis of both stands, forming the complete DNA provirus. At the end of the process a region unique to the 3' end of the RNA (u3) has been duplicated at the 5' end of the DNA and similarly, a region unique to the 5' end of the RNA (u5) has been duplicated at the 3' end of the DNA. Each LTR is therefore composed of a U3 sequence, an R sequence (present at both ends of the RNA), and a U5 sequence (U3RU5).

In the provirus, transcription initiates at the end of the U3 region in the 5' LTR and terminates at the end of the R region in the 3' LTR generating genomic RNA with only one U5 (at 5' end) and one U3 (at 3' end), and the r sequence at both ends. Between the two LTRs sequences are located the viral genes which encode the above mentioned viral proteins. These genes are essential for the replication of the virus. Viruses lacking part or all of these genes fail to produce infectious particles and are called replication deficient, as

opposed to replication competent viruses. Replication deficient viruses can be rescued by providing the missing genes in *trans*. This can be done in two ways. A replication competent virus (helper virus) can be used to coinfect with the replication deficient virus to provide the missing viral proteins. Alternatively, the replication deficient virus can be propagated in a cell line containing all the viral genes, a so called packaging cell line. All of the viral genes can therefore be replaced by foreign DNA making retroviruses excellent mammalian vectors. Once introduced in the cellular genome the recombinant sequences are stable and behave like host DNA. Unlike the introduction of DNA into cells by transfections methods, where the DNA often integrates in concatamers with large rearrangements of host DNA, retroviruses insert a single provirus per site of integration without any loss of host sequences. Retroviruses have also been the vector of choice when a high efficiency of DNA delivery is required. Infection is more efficient than any transfection method. Finally, retroviruses allow promoter-trap experiments to be done *in situ*, as is the case in the work presented in this thesis.

Thesis work

The goal of this thesis was to design a promoter-trap strategy that would allow us to identify genes involved in the development of the normal murine mammary gland. Given that cancer is the result of a deregulation of cellular growth and differentiation pathways, our efforts at understanding, detecting and developing better therapies for breast cancer can only benefit from a better knowledge of the genetics of breast development. It is our belief that there exist genes that are specific to the differentiating mammary epithelium and we decided

on a promoter-trap approach to identify such genes. The study of breast development is greatly facilitated by the fact that most of it takes place after birth. This, and the fact that unlike most organs, the mammary gland is easily accessible in the animal prompted us to do the promoter-trap *in situ*, by delivering the promoter-trap retrovirus through the mammary central duct of living mice. In a study on the role of *ras* in mammary neoplasia, Wang *et. al.* demonstrated that it is feasible to infect rat mammary epithelium specifically by injecting the retrovirus through the nipple (Wang *et al.* 1991). They first used the lacZ visual marker to test the efficiency of the infection. By staining sections of infected mammary glands with X-gal, the authors estimated their infection efficiency to be 0.3% (with a titer of 5×10^7 cfu/ml). Assuming that a rat has a total of 10^8 mammary cells, they suggest that about 10^6 cells would be infected provided that all twelve glands were infected. This technique would allow us to target the epithelium of the end buds specifically, which are the site of active cellular proliferation (Daniel and Silberstein 1987).

The two promoter-trap retroviruses used in this thesis were constructed based on the viruses presented by Ruley (von Melchner and Ruley 1989) , with the promoterless gene in the MoMuLV LTR. Because the promoter-trap is done in living animals, it is not feasible to use cytotoxic drugs to select for transgene expression. Instead, we used oncogenes as reporter genes so that expression could be indicated by the formation of a tumor, which can easily be detected in the animal.

The *ras* gene was the first cellular oncogene to be identified, from a human bladder carcinoma (Goldfarb *et al.* 1982). In mammals, there are three highly homologous *ras* genes which all code for a very similar plasma membrane-associated 21 kDa protein

(Barbieid 1987). Ras is a member of a large family of G proteins and is involved in signal transduction (reviewed in Satoh *et al.* 1992). These proteins bind and hydrolyze GTP. In its active form (GTP bound) Ras interacts with downstream effector molecules, one of which (*raf*) initiates a phosphorylation cascade by a series of kinases which ultimately leads to gene activation. By hydrolyzing GTP to GDP, Ras becomes inactive. The level of Ras is relatively constant in the cell. Its activity however is regulated by activators (guanine nucleotide releasing factors; GNRF) which catalyse the release of GDP thereby favouring the binding of GTP, and inhibitors (GTPase activating proteins; GAP) that catalyse the hydrolysis of GTP (Haubruck and McCormick 1991). Activating mutants of ras can no longer respond to inhibitors of its activity. One of the most common activating mutations is the point mutation at codon 12 that was found in the ras gene of the Harvey murine sarcoma virus (v-Ha-ras). Ras is a downstream effector of a number of transmembrane tyrosine kinase receptors (Satoh *et al.* 1992). In the case of the epidermal growth factor receptor, for example, activation and autophosphorylation of the receptor allow its interaction with an adaptor molecule (Grb2) that is found associated with the Ras activator mSOS (Satoh *et al.* 1992). Activation of the receptor therefore has the effect of bringing the GNRF (mSOS) in contact with Ras at the plasma membrane. Being the converging point of a number of signalling pathways involved in cellular proliferation, it is not surprising that an activated form of ras (oncogenic) leads to the transformation of various cell types. Overexpression of activated *ras* has been demonstrated to be transforming in established cells (Land *et al.* 1983) as well as in primary cells either in transgenic mice (Andres *et al.* 1987; Tremblay *et al.* 1989; Sinn *et al.* 1987) or infected embryos (Compere *et al.* 1989b).

The transformation efficiency of ras can be enhanced by a number of oncogenes including c-*myc*. c-*myc* is a member of a small family of highly related nuclear phosphoproteins (c-*myc*, N-*myc* and L-*myc*). c-*myc* (from now on referred to as *myc*) has been the most studied because it is widely expressed during embryonic development and in adult tissues. It is one of a few transcription factors that contain both a leucine zipper (LZ) and a basic helix-loop-helix domain (bHLH), both involved in protein dimerization and DNA binding (Jones 1990). *myc* has been shown to bind to a specific DNA sequence (Blackwell *et al.* 1990; Prendergast and Ziff 1991) when associated with the closely related human protein *max* (or the murine homologue *myn* ; Prendergast *et al.* 1991). Heterodimers of *Myc* and *Max* (or *Myn*) are transcriptional activators. *Max* can also form homodimers and bind to the specific sequence, acting as a repressor (*Max* contains both LZ and a bHLH motifs but does not have a transactivating domain; Reddy *et al.* 1992; Kato *et al.* 1992). The activity of *myc* is also regulated by another protein called *Mad* which dimerises with *myn*, preventing its association with *myc*. The role of myc has been associated with DNA synthesis and cell survival. Levels of *myc* are high in proliferating cells and decline during differentiation. Its effect may therefore be mediated through the up regulation of genes involved in cellular proliferation. In addition to being regulated at the level of its activity, *myc* expression is tightly regulated, both positively and negatively, at the level of transcription and translation. It is not too surprising then to find that the transforming activity of *myc* is due to overexpression rather than activating mutations in the coding sequence. *myc* overexpression has been associated with a number of cases of Burkitt lymphoma and B-cell acute lymphocytic leukemia in human, as well as murine plasma cell

tumors, where the *myc* locus was found translocated next to an immunoglobulin gene locus (Taub *et al.* 1982; Erickson *et al.* 1983), with no alteration to its coding sequence. It seems that in each case the new surrounding sequences either stabilize the *myc* transcript or eliminate some negative transcriptional control. *myc* was identified as the cellular homologue of the transforming gene (v-*myc*) found in the avian myelocytomatosis virus MC29, where the intact coding sequence of *myc* was fused to part of the gag retroviral protein (Alitalo *et al.* 1983). *myc* overexpression can cause the immortalization of primary cells but is rather inefficient at fully transforming. Transgenic mice expressing *myc* in the mammary gland (MMTV-*myc*) developed adenocarcinomas of the breast after a long latency (Stewart *et al.* 1984; Leder *et al.* 1986; Sinn *et al.* 1987). Embryos infected with a retrovirus expressing *myc* also develop tumors (Compere *et al.* 1989a).

The combination of *myc* and *ras* however is more efficient then either oncogene alone at transforming primary cells. Leder's group showed that double transgenics carrying c-*myc* and v-Ha-*ras* under the MMTV LTR promoter developed mammary adenocarcinomas with better efficiency and a much shorter latency than either of the single transgenics (Sinn *et al.* 1987). Similarly, Jaenisch's group infected mouse embryos with retroviruses expressing c-*myc* or v-Ha-*ras* and found that the combination of the two oncogenes was more powerful at causing neoplasia, including carcinomas (Compere *et al.* 1989a). These studies made it possible to expect a promoter-trap carrying these two oncogenes to cause carcinomas of the breast in mice.

The first virus that we constructed (COPT, Figure 1.2A) was replication deficient and contained a promoterless activated viral *ras* (v-Ha-*ras*) inserted in the LTR, 95bp away from

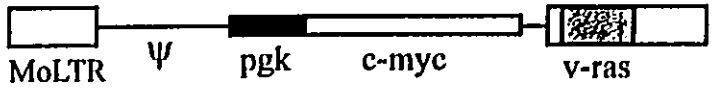
the 5' end. The c-*myc* coding sequence, under the control of the ubiquitously expressed pgk gene promoter, was introduced in place of the viral genes. Therefore, *myc* is constitutively espressed but *ras* is only expressed when the provirus integrated near an active cellular promoter. The virus was called COPT for *c*ooperating *o*ncogenes *p*romoter-*t*rap.

The second retrovirus was constructed with a single oncogene at the same position in the LTR and was kept replication competent (PyT, Figure 1.2B). The oncogene used was the Polyomavirus middle T antigen. The promoter-trap retrovirus was named PyT for *P*olyoma *t*rap. PyMT has been demonstrated to be one of the genes responsible for the transformation of cells by the Polyomavirus (PyV) (Treisman *et al.* 1981). This DNA virus belongs to the papovavirus class of DNA tumor viruses. In the natural host (the mouse), the PyV replicates inside the cell until it reaches a certain density and lyses the cell. In cells which are not permissive for the replication of the virus, the fate of the host cell may be transformation. It has been shown that PyMT, one of three immediate early gene products of the virus, was causing transformation by associating with, and activating the cellular protooncogene *src*, a membrane associated protein tyrosine kinase involved in signal transduction. Although in cells transformed by the PyV, the large T antigen was also shown to be necessary, in other systems it was demonstrated that PyMT can transform primary cells on its own. One example is where PyMT was shown to transform mammary epithelium when driven by the MMTV promoter in transgenic mice (Guy *et al.* 1992). A number of

Figure 1.2 Promoter-Trap Retroviral Vectors

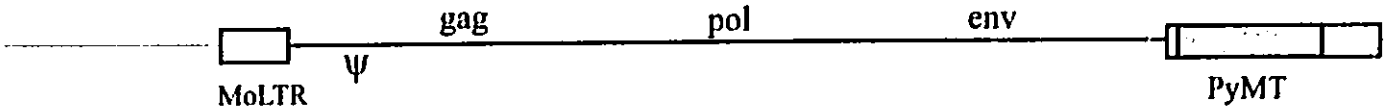
Thick lines represent MoMuLV sequences and thin lines represent plasmid sequences. Open boxes represent the LTRs. In the COPT vector, a pgk-c-*myc* cassette replaced most of the gag, pol and env genes. The v-Ha-*ras* sequence was inserted in the 3' LTR so that it could be duplicated in the 5' LTR during reverse transcription. The PyT vector had the PyMT sequence inserted in the 3' LTR. The packaging signal (ψ) was intact in both vectors to allow the viral RNA genome to be packaged into virions.

COPT



1kb

PyT



1kb

transgenic lines were generated and in all of them multifocal breast tumors formed, indicating that PyMT was quite efficient at transforming mammary epithelium requiring very little help. The activation of the PyMT gene by cellular promoters should therefore be expected to generate mammary tumors in the animals.

Performing the promoter-trap experiment in the living animal had the great advantage of offering a biological screen to identify genes that are specifically active during mammary differentiation. Mice were injected with the virus at puberty because cellular proliferation is a requirement for proviral integration. The first tumors to appear would be the result of proviruses integrating in ubiquitously expressed genes. The interesting tumors would be the ones appearing after the mice were mated. The cloning of the integration site of one such tumor was part of the work presented in this thesis.

Some of the strategies that have been used in the past to identify cell type specific loci include the use of subtractive libraries (Davis *et al.* 1987; Basset *et al.* 1990; Yaswen *et al.* 1990). This approach involves the isolation of messenger RNA from differentiated and undifferentiated cells separately and subtracting one from the other by hybridization. Although there has been a number of variations and improvements in methods to isolate genes that are differentially expressed between two cell types or following treatments, all of these methods are based on the use of mRNA as the starting material and are unavoidably biased towards more highly transcribed genes. Here we were interested in regulatory genes that are expressed in the normal epithelium and are most likely to be expressed at low levels. Also, the expression of these genes would be restricted to a small proportion of the epithelium, that is the cells of the terminal end buds. This would dilute the mRNAs even

more and would involve isolating a large number of these end buds to obtain starting material. Although modern amplification methods would facilitate the task, because we don't know the time at which and for how long key regulatory genes are active during mammary differentiation, this would require that the procedure be performed for various stages of mammary development. Promoter-traps have the great advantage of targeting the DNA rather than the RNA, which eliminates problems of copy number bias. The strength of this approach is illustrated by the number of reports in recent years where developmentally regulated genes have been cloned using promoter-traps (Joyner and Martin 1987; Friedrich and Soriano 1991; Takeuchi *et al.* 1995; Gasca *et al.* 1995; Okazaki *et al.* 1994). Interestingly, a number of these genes turned out to be transcription factors, indicating that promoter-trap is a good strategy to identify genes whose expression may be low or temporally restricted.

CHAPTER 2

MATERIALS AND METHODS

Bacterial transformation

Bacteria were transformed by electroporation. 10 µl of DNA from a ligation mixture (or 5µl of TE containing 25ng of plasmid) were mixed on ice with 50µl of competent bacteria (TG1 or DH5α) and incubated for five minutes. The DNA/ bacteria solution was transferred to a 2mm cuvette and electroporated using either a BioRad (25µF, 2.5kV, 200Ω) or a BTX (2.45kV, 129Ω) gene puler. 1ml of LB (10g peptone, 5g yeast extract, 5g sodium chloride in 1L) was added promptly and mixed well by pipetting. The bacteria were allowed to recover at 37°C for 30 to 60 minutes with agitation. An aliquot of 10-200µl of bacteria (in a total volume of 200µl) was spread onto a 100mm bacterial dishes coated with 1.5% agar (15g of agar in 1L of LB) containing 50µg/ml ampicillin (Boehringer Mannheim). The plates were incubated, inverted, at 37°C overnight.

Small scale plasmid preparation

Bacteria were grown in 5ml of LB or TB (12g bacto-tryptone, 24g yeast extract, 4ml glycerol, and 10ml of a 170mM KH_2PO_4 and 720mM K_2HPO_4 solution, in 1L) containing

50µg/ml of ampicillin, at 37°C with agitation overnight. Two times 1.5ml of bacterial culture was centrifuged in a 1.5ml microfuge tube for 5 seconds and the supernatant was removed. For subcloning or restriction analysis purposes, plasmid DNA was isolated using the following boiling method (Holmes and Quigley 1981). The bacterial pellet was resuspended in 350µl of STET solution (8% sucrose, 0.5% Triton X-100, 50mM EDTA pH 8.8, 10 mM Tris-HCl pH 8.0) by vortex. 250µg of lysozyme (Boehringer Mannheim; 10mg/ml stock in 10mM Tris pH 8.0) was added and mixed by vortex. The bacterial lysate was boiled for four minutes and centrifuged at room temperature for five minutes. The white genomic DNA was removed using a sterile tooth pick. 350µl of isopropanol was added and mixed by inversion. The DNA was recovered by centrifugation at 4°C for five minutes. The DNA pellet was rinsed with 1ml of cold 70% ethanol and centrifuged again for five minutes. The pellet was dried under vacuum and resuspended in 200µl of TE (10mM Tris pH 8.0, 1mM EDTA).

For sequencing purposes the DNA was prepared using the following alkaline lysis method (Birnboim 1983). The bacterial pellet was resuspended in 100µl of an ice cold solution containing 50mM glucose, 10mM EDTA and 25mM Tris-HCl pH8.0, by vortexing. After five minutes at room temperature, 200µl of a freshly prepared solution containing 0.2N NaOH and 1%SDS was added. Samples were mixed by inversion and incubated on ice for five minutes. 150µl of an ice cold solution of potassium acetate pH4.8 (7ml of glacial acetic acid and 15ml of a 5M potassium acetate solution in 25ml total) was added. Samples were mixed briefly by vortexing and incubated on ice for five minutes. Following centrifugation at 4°C for five minutes, the supernatant was transferred to a clean tube and centrifuged for

another five minutes. The supernatant was again transferred to a clean tube and RnaseA (Boehringer Mannheim; 10mg/ml) was added to 50µg/ml. Samples were incubated at 37°C for 30 minutes. An equal volume of phenol/chloroform (1:1) was added, mixed by vortexing for one minute and centrifuged at room temperature for five minutes. The supernatant was transferred to a clean tube and 1ml of cold 95% ethanol was added. After mixing by vortexing, the samples were incubated on dry ice for 20 minutes. The samples were centrifuged at 4°C for five minutes. The DNA pellet was rinsed with cold 70% ethanol and centrifuged at 4°C for five minutes. The pellet was dried under vacuum and resuspended in 20µl of water. The DNA concentration was determined by diluting 1-2µl of DNA in a total of 50µl of water and reading the absorbance at 260nm on a Beckman DU® 640 spectrophotometer using a microcuvette.

Large scale plasmid preparation

Bacteria were grown (from a 5ml overnight culture) in 500ml of Luria Broth (LB) containing 50µg/ml of ampicillin at 37°C with agitation, overnight. Plasmid DNA was isolated using the same alkaline lysis method described above on a larger scale. The bacterial culture was centrifuged at 5,500g (6,000rpm in a Beckman JA-14 rotor) for 10 minutes at 4°C. The bacterial pellet was resuspended in 4ml of a solution containing 50mM glucose, 10mM EDTA and 25mM Tris-HCl pH8.0 and mixed by vortex. 2mg of lysozyme was added and samples were mixed by vortexing. After a 10 minute incubation at room temperature 9ml of a freshly made solution containing 0.2N NaOH and 1%SDS was added. The samples were mixed by inversion and incubated on ice for five minutes. 4.5ml of a cold

solution of potassium acetate pH4.8 was added and mixed briefly by vortexing. After a 45 minutes incubation on ice the samples were centrifuged at 12,000xg (10,000rpm in a Beckman JA-20 rotor) for 10 minutes. The supernatant was transferred to a clean tube and 11ml of isopropanol was added and mixed by vortexing. After a 10 minute incubation at room temperature the samples were centrifuged at 12,000xg for 10 minutes. The supernatant was discarded and the tube was inverted to drain excess liquid. The nucleic acid pellet was resuspended in 1ml of TE. RnaseA was added to 100µg/ml and the samples were incubated at 37°C for 30 minutes. 20µl of a 10%SDS solution and was added and mixed by vortex. Proteinase K (Boehringer Mannheim; 5mg/ml stock solution in water) was added to 200µg/ml and incubated at 37°C for one hour. An equal volume of phenol/chloroform (1:1) was added and mixed by vortexing for one minute. After centrifugation for five minutes at room temperature the supernatant was transferred to a clean tube and an equal volume of butanol was added and mixed by inversion. After a three minutes centrifugation at room temperature (clinical centrifuge) the bottom aqueous phase was transferred to a clean 15ml conical tube (Falcon). The volume was adjusted to 2ml with TE. Exactly 2.3g of cesium chloride was added and dissolved by inversion. 100µl of a 10mg/ml solution of ethidium bromide was added and mixed well. The DNA/CsCl solution was transferred to a quick sealed tube (Beckman; 11x32mm), sealed, balanced and centrifuged in a TLV-100 rotor in a Beckman TL-100 ultracentrifuge at 60,000rpm overnight. The DNA was recovered by inserting a needle at the top of the tube to allow air in. A second needle attached to a 3ml syringe was inserted in the tube at right angle below the DNA band, and approximately 0.5ml of solution was drawn in the syringe. The needle was removed and the DNA

transferred to a 4ml polystyrene push cap tube. The ethidium bromide was extracted by two washes of equal volume of water-saturated butanol. The bottom aqueous phase was transferred to a clean tube and two volumes (1-2ml) of water were added and mixed well. Three volumes (3-6ml) of 95% ethanol were added, mixed and the samples were incubated at -20°C for 30 minutes. The DNA was recovered by centrifugation at 12,000xg for 30 minutes at 4°C. The air dried pellet was resuspended in 1ml of TE. The DNA concentration was determined by diluting 10µl of DNA in 1ml of water and reading the absorbance at 260nm.

Large scale plasmid DNA was also prepared using the maxi protocol of QIAGEN.

Plasmid vectors

All restriction enzymes were from either Bethesda Research Laboratories (BRL), Boehringer Mannheim Canada (BMC) or New England Biolabs (NEB). To make pLTRras, the coding sequence of the v-Ha-*ras* gene (position 338 to 910) was amplified by PCR from the HB-11 plasmid (ATCC 41013), using oligonucleotides containing a BamHI restriction site (5'-TTATAGGGATCCGCGATGACAGAATACAAG-3' and 5'-TTCAGTGGATCCTCAGGACAGCACACACTT-3'). The PCR reaction (50µl) contained 25ng of plasmid DNA, 0.2mM each deoxinucleotide triphosphate (dNTP) (Pharmacia), 15pmol each oligonucleotide, 1x PCR buffer (Stratagene), 1U Pfu polymerase (Stratagene). The amplification was performed in a COY thermocycler by 30 cycles consisting of 30 seconds at 94°C, 30 seconds at 55°C and 1 minute at 72°C. The PCR product was precipitated with 0.3M sodium acetate and 2.5 volumes of 95% ethanol. The DNA was

recovered by centrifugation at 4°C for 15 minutes, rinsed in 70% ethanol and resuspended in an appropriate volume of water. The DNA was digested with BamHI and the 580bp fragment was purified by standard agarose gel electrophoresis (Sambrook *et al.* 1989), using 1% SeaKem agarose (Sigma). The DNA was recovered from the agarose using glass milk (GeneClean). The ras sequence was inserted in the MoMuLV-LTR using the pLTR12 plasmid (Reik *et al.* 1985) which contains one LTR with an engineered BamHI site in place of the Sau3IA site at position 95.

The pUCRas plasmid was constructed by replacing the 3'LTR of the pUCN2 vector (Armentano *et al.* 1987) with the v-Ha-*ras*-containing LTR from the pLTRras plasmid described above. The pUCN2 vector was first digested with ClaI and then partially digested with SstI. The proper dilution of SstI necessary to give a partial digest was first determined by digesting small amounts of pUCN2 plasmid (3µg/50µl) with increasing dilutions of SstI (0.2-0.025U) for one hour at 37°C. A concentration of 0.1U/3µg of SstI was used to cut ClaI-digested pUCN2. The 6.9kb vector fragment was gel purified as described before and ligated to the purified 1140bp ClaI-SstI fragment containing the LTR(ras) from the pLTRras vector.

To make the COPT vector, a pgk-c-*myc* cassette (EcoRI-BamHI) from the pPGKmycX vector (Dr.McBurney; it contains part of the first exon and all of the second and third exons of c-*myc*, linked to the pgk promoter) was inserted in place of the neomycin resistance gene (EcoRI insert) in pUCRas. This was done by ligating the EcoRI site first, gel purifying the ligated product as described above, then making the other end blunt using the Klenow fragment of DNA polymeraseI (Sambrook *et al.* 1989) and ligating again.

The PyT vetor was constructed by inserting the Polyoma Virus middle T antigen (PyMT) sequence, a HindIII to EcoRI fragment from the MMTVMT vector (a gift from Dr. W. Muller), in the BamHI site of pLTR12 by blunt ligation. The remainder of the Moloney provirus was ligated to the modified pLTR12 vector (pLTRMT) as an EcoRI-ClaI fragment from pMov3 (gift of Dr. R. Jaenisch). This positioned the PyMT coding sequence in the 3' LTR.

The pTRAP-CAT vector was constructed by ligating the 2kb NheI genomic fragment of TRAP1 into the XbaI site of the pCAT-Enhancer vector (Promega). Clones containing the insert in each orientation were isolated.

Cell lines and transfection

All cell lines (NIH3T3, P19, GP+E86 and XC) were grown in α-MEM (GIBCO BRL) supplemented with 10% bovine serum (5% fetal and 5% newborn), at 37°C with 5% CO_2. Transfections were done using calcium phosphate (Graham and van der Eb 1973) and BES buffer (Cheng and Okayama 1987). For stable transfections, NIH3T3 and GP+E86 cells were seeded at $5x10^5$ cells per 100mm tissue culture dish. The following day 10μg of retroviral vector DNA and 1μg of selectable marker DNA were mixed in a solution of calcium chloride to a final concentration of 62.5mM in 0.5ml. While vortexing very gently, 0.5ml of a BES buffered solution (50mM BES, 280mM NaCl, 1.5mM $Na_2HPO_4 \cdot 2H_2O$, pH 6.96) was added dropwise. Calcium phosphate precipitates were allowed to form at room temperature for 5-10 minutes. The 1ml of $CaPO_4$/DNA solution was added to the cells

dropwise. The cells were rinsed with PBS (8g NaCl, 0.2g KCl, 1.44g KH_2PO_4, 0.24g Na_2HPO_4, pH 7.4 in one liter) and fed fresh medium 8 hours later. Two days post transfection, the cells were replated at a density of 10^5 per 100mm dish with 400µg ml of G418 (Gibco BRL) or 2µg/ml of puromycin (Sigma). Resistant colonies were isolated 10-14 days later and expanded.

For transient transfections, P19 (McBurney 1993) or 293 cells were plated at a density of $5x10^5$ cells per 60mm tissue culture dish. Five micrograms of each plasmid (pCAT and pgk-β-gal) were used to prepare the calcium phosphate precipitate as described above. The cells were rinsed with PBS and feed fresh medium after 6 hours, and were harvested 24 hours post transfection.

Virus production and infection

The pUCRas vector was transfected alone, and the COPT vector was cotransfected with the pgk-neo vector (McBurney *et al.* 1991), in the GP+E86 packaging cell line. Twelve (pUCRas) and 24 (COPT) colonies were isolated and expanded. The PyT vector was cotransfected with the pgk-puromycin vector (from Dr.McBurney's laboratory) in NIH 3T3 cells and 24 colonies were isolated, expanded and tested for titer by XC assay (see below).

To determine the titer of the pUCRas producer clones, NIH 3T3 cells were seeded at $2x10^4$ cells per 60mm tissue culture dish. The producer cells were seeded at $2.5x10^6$ cells per 60mm dish. The next day, the supernatant from the producer clones was cleared of cellular debris by centrifugation in a clinical centrifuge for 3 minutes at maximum setting. Dilutions of the virus (10^{-3} -10^{-5}) were prepared on ice in α-MEM supplemented with 2%

newborn serum. The medium was removed from the NIH 3T3 cells and these were infected, in duplicates, with 1ml of the various virus dilutions containing 8µg/ml of polybrene (Sigma; 10mg/ml stock solution in PBS). Cells were incubated at 37°C for 30 minutes and 4ml of medium was added and the cells were returned to 37°C. The following day the cells were fed fresh medium with 400 µg/ml of G418. This was repeated after 4-5 days. When colonies were visible (10-14 days post infection) the medium was removed and cells were soaked in a solution of methyleneblue (2% in 95% ethanol) for 15-20 minutes. The staining solution was removed and the cells were rinsed in copious amount of water. The plates were inverted to dry and titers were calculated by counting the number of colonies and correcting for the dilution factor.

To determine the titer of COPT producer clones (or verify the structure of the PyT virus) by Southern blot analysis, NIH 3T3 cells were seeded at 1-2x10^5 cells per 60mm tissue culture dish and producer cells were seeded at 2.5x10^6 cells per 60mm dish. The following day, the supernatant from the producer clones (5ml) was cleared of cellular debris by centrifugation for 5 minutes (maximal speed in a clinical centrifuge) and was added to the NIH 3T3 cells with 8µg/ml of polybrene as described before. The cells were fed fresh medium the next day and harvested for DNA extraction two days post infection. For serial passages of the virus (PyT), the supernatnant from the infected cells (5ml) was transfered to fresh NIH3T3 cells (seeded at 1-2x10^5 cells per 60mm tissue culture dish the day before) in the same manner as above.

For injection in the nude mice, a total of 8x10^6 NIH3T3 cells were seeded at 10^6cells per 100mm tissue culture dish. The COPT producer cells were seeded in eight 100mm tissue

culture dishes at $2.5x10^6$ cells per dish. The next day, the supernatant from the producer cells was filtered through a 0.2µ filter and transfered to the NIH3T3 cells with 8µg ml of polybrene (multiplicity of infection of 1). The following day, the infected cells were trypsinized, rinsed and resuspended in PBS ($4x10^6$ cells/ml) . The cells were injected subcutaneously at four sites in a nude mouse (two shoulders and two hips). The same number of uninfected NIH3T3 cells were injected in the same way in a second nude mouse.

For injection of the COPT or PyT virus in the mammary glands, the producer cells were seeded at $2.5x10^6$ cells per 60mm tissue culture dish. The following day, the supernatant was filtered through a 0.2µ filter and polybrene was added to 80µg/ml. The virus/polybrene solution was kept on ice until ready to be injected in the mice.

Plaque (XC) assay

For the XC assay (Rowe *et al.* 1970), NIH 3T3 cells were plated at a density of $2x10^4$ cells per well in a 6-well tissue culture dish with 8µg/ml of polybrene. The following day, fresh virus from producer clones was filtered through a 2µ filter and diluted on ice in α-minimum essential medium (α-MEM) supplemented with 2% fetal calf serum. The medium was removed from the NIH 3T3 cells and 800µl of the various virus dilutions were added in duplicates. The infected cells were incubated at 37°C for 30 minutes, after which 2 ml of medium were added and the cells were returned to 37°C. Cells were fed fresh medium on the second and forth day. After 5-6 days, or when the cells were confluent, the supernatant was removed and the cells were UV-irradiated with 20mJ using a BioRad UV chamber. XC cells were added at a density of $8x10^5$ cells per well in 2ml of medium. The

cells were fed fresh medium every second day until they reached confluency. Plaques were visualized by methyleneblue staining: the medium was removed and approximately 2ml of a solution of 2%methyleneblue in 95%ethanol was added for 15-20 minutes. The staining solution was removed and the dishes were rinsed in copious amount of water. After allowing the dishes to dry inverted, the number of plaques were counted. The titer was calculated taking into account the dilution factor.

Nipple injections

Nipple injections were performed using a modification of the method published by Wang *et al.* (Wang *et al.* 1991). Four week old Balb/C female mice (Charles River) were anaesthetized using Avertin (10g of Tribromoethanol [Aldrich] in 10ml of tert.amylalcohol [BDH], diluted in saline) at 0.4 mg per gram of body weight. Freshly isolated virus, to which 80 μg/ml of polybrene (Sigma) was added, was injected directly in the nipple using a 30 gauge needle. Approximately 0.05 - 0.1 ml (10^4-$2x10^4$ cfu for COPT and $5x10^4$-10^5 cfu for PyT) was injected into each nipple.

X-gal staining of mammary glands

Following extraction (three days post-infection), the mammary glands were fixed in 0.2% gluteraldehyde overnight at 4°C, or in 2% gluteraldehyde for several hours at 4°C. The glands were rinsed twice in PBS and stained in X-gal staining solution (10ml: 35mg of K^+-ferrocyanide[8mM], 27.3mg of K^+-ferricyanide[8mM], 9.7ml of 100mM PBS pH7.4, 10µl of 1.0M $MgSO_4$, 10mg of X-gal (Sigma) dissolved in 300µl of dimethylsulfoxide) at 37°C

overnight. The stained glands were rinsed 2-3 times in PBS and 2-3 times in water, and were cleared of fat by soaking in a 1%HCl solution for 5-10 minutes. They were rinsed in a 25% then 50% glycerol solution for few hours each, and finally preserved in 100% glycerol.

Southern blot analysis

DNA was isolated by digesting the cells or the tissue in 1ml of DNA extraction buffer (75mM NaCl, 25mM EDTA, 10mM Tris pH7.6, 1%SDS, 400 µg/ml proteinase K) at 37°C overnight on a rotator. An equal volume of phenol:chloroform (1:1) was added and mixed by inversion. After centrigugation at 3,000xg at room temperature, sodium acetate was added (0.3M final concentration pH5.2) to the DNA solution and 2.5 volumes of 95% ethanol was added. The DNA was recovered by centrifugation at 12,000xg (Beckman JA-20 rotor) at 4°C for 30 minutes, air dried and resuspended in 1ml of TE. The concentration was estimated by reading the absorbance at 260nm. 10µg of DNA was digested with 70U of the appropriate enzyme in a total volume of 100µl at 37 °C overnight. The DNA was precipitated as before and resuspended in a small volume of water. The DNA was electrophoresed on 1% agarose (Sigma) using standard procedures. The gel was soaked in a staining solution of ethidium bromide (~1µg/ml) for 15-20 minutes and photographed. The gel was then soaked in denaturing solution (1.5M NaCl, 0.5N NaOH) for 45 minutes and in neutralizing solution (1.5M NaCl, 1M Tris pH8.0) for 45 minutes. The DNA was transfered to a Hybond N membrane (Amersham) following standard protocol (Sambrook *et al.* 1989). Following UV cross-linking (100mJ in a Biorad UV chamber), the membrane was

prehybridized in a solution containing 5x SSC (20x SSC: 175.3g NaCl, 1.88.2g sodium citrate in 1L, pH7.0),1x Denhartd's solution (50x: 5g ficoll 400, 5g polyvinylpyrrolidone, 5g bovine serum albumin fraction V, in 500 ml of water), 50% deionized formamide, 1%SDS and 100µg/ml Herring sperm DNA, at 42°C for a few hours to overnight. 25ng of DNA fragment was radioactively labelled using ^{32}P-dCTP(Amersham) and random oligonucleotides primers (Boehringer Mannheim). The probe was separated from unincorporated nucleotides using a G-50 Sephadex (Pharmacia) column. The probe was denatured by boiling for 4-5 minutes and added to the membrane. Hybridization was carried out at 42°C overnight with agitation. The membrane was rinsed twice in 2xSSC at room temperature and washed in 0.1x SSC and 0.1%SDS at 65°C for one hour with agitation. The membrane was exposed to X-ray film (Dupon NEN).

Northern blot analysis

RNA was extracted using lithium chloride/urea (Auffray and Rougeon 1980). Cells or tissue (frozen in liquid nitrogen) were placed in 5ml of LiCl/urea solution (3M lithium chloride and 6M urea) and subjected to a polytron pulse for 2 minutes. The lysed cells/tissue was left on ice overnight. The RNA was centrifuged at 11,000xg (Beckman JA-17 rotor) at 4°C for 30 minutes. It was rinsed twice in 3M LiCl and centrifuged as before. The RNA pellet was air dried and resuspended in 500µl of DEPC treated water (1ml of DEPC [BDH] in 1L of water, stirring for a few hours and autoclaved to inactivate the DEPC). The concentration was determined by reading the absorbance at 260nm. PolyA+ RNA was isolated according to the protocol of Jacobson (Jacobson 1987) (BDH columns, BRL resin),

with the exception that the bound RNA was eluted four times with 0.5ml of elution buffer.

For northen blot analysis, 10µg of total RNA or 3µg of PolyA$^+$ RNA was precipitated with sodium acetate (0.3M pH5.2) and 95% ethanol, and resuspended in 20 µl of RNA loading buffer (50% formamide, 1x MOPS buffer, 17.5% formaldehyde, and loading dye[10x: 0.4% bromophenolblue in 50% glycerol and 1mM EDTA]). The samples were heated at 68°C for 10 minutes and cooled down on ice for a few minutes. Ethidium bromide was added to 20µg/ml and the samples were run on a 1% formaldehyde-agarose gel (50%formaldehyde, 1x MOPS buffer [200mM MOPS, 50mM $NaOAc{\cdot}3H_2O$, 10mM EDTA, pH7.0]). The RNA was transfered to a Hybond-N membrane and hybridized as described for Southern blots.

PCR analysis

To isolate DNA flanking the provirus, 1-2 µg of genomic DNA from the tumor was digested with 10U of either SmaI (25°C) or TaiI (50°C) for at least 2 hours. The DNA was diluted to a concentration of 10ng/µl to favor self ligation and ligated at 16°C overnight using 1U of T4 DNA ligase (BRL). 500ng of circularized DNA was used in the PCR reaction which included 15 pmol of each primer (oligo1: 5'-ACAGGGATCCCGATCTGAAC-3', oligo2: 5'-CCAATCAGTTCGCTTCTCGC-3') (see Figure 3.8 for location of the oligonucleotides), 0.2mM of each deoxinucleotide triphosphate (dNTP)(Parmacia) and the manufacturer's buffer in a 50µl volume. Reaction samples were prepared by mixing all the components except the enzyme. The samples were preheated at

96°C for 1 minute and annealed at 62°C for 5 minutes in a COY thermocycler. The tubes were centrifuged briefly to bring down condensation and 4U of Vent polymerase (exo⁻ ; NEB) were added and the reactions were covered with parafilm oil and put through 50-60 cycles of 3 minutes at 72°C , 1 minute at 96°C and 1 minute at 62°C. The last cycle was followed by a 10 minute incubation at 78°C. A 10µl aliquot of the product was analyzed on a 1.2% agarose gel.

Genomic cloning

To build the genomic library , 50µg of DNA from the tumor was digested with EcoRI and 12-20kb fragments were isolated using a 5-20% potassium acetate gradient. The digested DNA was layered on top of the gradient and spun at 40,000rpm for 7 hours at room temperature in an SW41 rotor (Beckman). The DNA was recovered in fractions of 1.5ml through a needle (20 gauge) inserted at the bottom of the tube. An aliquot of each fraction was analyzed on a 0.8% agarose gel, and the fractions containing the desired size DNA fragments were pooled and dialysed in TE. 1µg of EcoRI fragments was ligated to 2µg of λDASHII arms digested with EcoRI (Stratagene), in a volume of 10µl with 1U of T4 DNA ligase (BRL) at 16°C overnight. The DNA was packaged into phage particles using a Gigapack extract (Stratagene) acording to the manufacturer's instructions. To determine the titer of the library, 10µl of appropriate dilutlions (10^{-3}-10^{-5}) of the phage lysate was mixed with 100µl of an overnight culture of K802 bacteria, grown in LB (with 10mM $MgSO_4$), and incubated at 37°C for 15 minutes. 3-4ml of warm (55°C) 0.7% agarose (with $MgSO_4$) was added, mixed quickly by inversion and poured onto a layer of solidified 1.5% agar in

100mm bacterial petri dishes. The dishes were incubated, inverted, at 37°C for 10-16 hours. The number of plaque forming units (pfu) was calculated from the number of plaques taking into account the dilution factor.

To screen the library, 500 000 pfu were plated on 10 150mm bacterial petri dishes (50 000pfu per dish) in the same way as described for the determination of the titer, except that the phage lysate was mixed with 200 µl of bacteria and 7ml of 7% agarose was added. After an overnight incubation at 37°C the plates were cooled down at 4°C. Plaque lifts were done in duplicates on Hybond-N membrane (Amersham). The membrane was laid on the plate for a few minutes and lifted gently. The membrane was marked with pensil and a corresponding mark was placed on the dish to locate positive plaques later. The DNA was denatured by placing the filters between Whatman papers and autoclaving them at 100°C for one minute. The DNA was UV cross-linked (100mJ) using a BioRad UV chamber. Filters were hybridized with a radioactively labelled PyMT sequence and washed in the same way as described for a Southern blot.

Positive plaques were removed from the agar using the large end of a Pasteur pipette, placed in a 1.5ml eppendorf tube with 1ml of SM medium (5.8g NaCl, 2g $MgSO_4 \cdot H_2O$, 50ml of 1M Tris-Cl pH7.5, and 2% gelatin in 1L) and two drops of chloroform, and vortexed briefly. The phages were allowed to diffuse out of the agarose at 4°C for at least four hours, or at room temperature for two hours. Samples were centrifuged briefly to bring down the chloroform. Clones were plaque purified by two more rounds of plating and screening as described above (on 100mm petri dishes) using progressively higher dilutions until a single plaque could be isolated. Usually, final dilutions of 10^{-4} and 10^{-5} were used for these two

rounds of purification. The final single plaque was placed in a 10ml polypropylene tube with a drop of overnight culture of bacteria. An extra tube contained bacteria only and served as a control for cell lysis by the phage. After a 10 minute incubation at 37°C, 2ml of LB with magnesium was added and the samples were incubated with shaking at 37°C for 4-6 hours (or until culture is clear compared to the control). Two drops of chloroform was added and mixed by vortexing. Samples were centrifuged at 3,000xg for 5 minutes.

To prepare phage DNA, 15µl of the above lysate was mixed with 1ml of an overnight culture of bacteria and incubated at 37°C for 10 minutes. This starter was added to 50ml of LB (with $MgSO_4$), in a 250ml conical flask, and incubated at 37°C with vigorous shaking overnight. Ten drops of chloroform were added and mixed by swirling. 1.5g of NaCl was added and dissolved completely. Samples were centrifuged at 2,000xg for 5 minutes. The supernatant was transfered to a new tube and 5g of PEG 8000 (BDH) was added and mixed by inversion. The samples were placed on ice for one hour then centrifuged at 2,000xg for 30 minutes at 4°C. The supernatant was discarded and the tube inverted to drain all liquid. The pellet was dissolved in 600µl of SM by pipetting up and down. The suspension was transfered to a 1.5ml eppendorf tube, 25 µg of Dnasel (BMC) was added and inubated at 37°C for 30 minutes. 2.5µl of diethylpyrocarbonate (DEPC)(BDH) was added and mixed well to inactivate the enzyme. The samples were divided evenly into two tubes and a solution containing 1M Tris pH 8.5, 1.5% SDS and 0.1M EDTA was added. After mixing, the samples were heated at 65°C for 5 minutes. 300µl of a 5M potassium acetate solution was added and samples were incubated on ice for 30 minutes. After centrifugation at 4°C at maximum speed, the supernatant was transfered

to a new tube and the tube was filled up with isopropanol. The DNA was recovered by centrifugation at 4°C for 10 minutes. The dried DNA pellet was resuspended in 200µl of TE. 10 µl of DNA was digested with the appropriate enzyme(s) and run on a 0.8% agarose gel. The insert DNA was subcloned into the pBluescript SK+ vector (Stratagene) as an EcoRI-KpnI fragment as described before.

cDNA cloning

A murine lung cDNA library obtained from Stratagene was screened, and clones isolated in the same way as it was described for the genomic library with the exception that XL1-Blue bacteria (Stratagene) were used and 200,000 plaques were screened (20,000 plaques per 150mm dish). The cDNA inserts were recovered in the form of plasmids using a helper phage (ExAssist) as per manufacturer's protocol (Stratagene).

Sequencing

Plasmid DNA was prepared on a small scale using the alkaline lysis method described earlier. For manual sequencing the reactions were prepared using a Phamacia T7 polymerase sequencing kit and ^{35}S-αdATP from Amersham. The reactions were run on a 6 or 8% polyacrylamide (19:1 acrylamide:bisacrylamide) gel with 50% urea in TBE buffer (5xTBE: 54g Tris-HCl, 27.5g boric acid, 0.3g EDTA in 1L). The gel was fixed in 10%methanol and 10% acetic acid for 15-20 minutes, transfered to a Watman paper and dried at 80°C under vacuum using a BioRad gel drier. The gel was subjected to autoradiography using Kodak BMR1 film. Sequencing was also done on an ABI automated

sequencer using the ABI Collection Prism 377 software. Sequence analysis was done using DNASIS software.

CAT assay

Transiently transfected cells were washed in cold PBS and harvested in a solution of 0.25M Tris pH 7.8. The cells were scraped off the plate using a rubber policeman and transfered to a 1.5ml microfuge tube on ice. The samples were lysed through two rounds of freeze/thaw on dry ice and at 37°C. Half of the sample was saved for β-galactosidase assay (see below) and the other half was heated at 65°C for 10 minutes and centrifuged in a microcentrifuge at maximum speed at 4°C. The supernatant was transfered to a new tube and assayed for protein concentration using a BioRad micro assay. 5µg of protein was used for the CAT assay using a fluorochrome based kit (FMBIO® TANGERINE FAST CAT® from HITACHI). An aliquot (5-15µl) of the sample was subjected to thin layer chromatography using TLC aluminium sheets (silica gel 60 without fluorescent indicator) (MERCK) in a mixture of 15% methanol/85% chloroform until the two forms of the substrate were properly separated. The results were analysed on a Hitachi fluoroimager using the FMBIO(V4.04) software. The results are presented as persent converted substrate (amount of converted substrate x 100 divided by the total amount of converted plus unconverted substrate).

β-galactosidase assay

Samples for the β-galactosidase assay (from transient transfections) were processed together with the ones for the CAT assay up to the point where the samples are divided in half. The half of the sample that was used for the β-galactosidase assay was directly centrifuged at 4°C in a microcentrifuge for 10 minutes after the two rounds of freeze/thaw. The supernatant was transfered to a new tube and assayed for protein concentration using a BioRad micro assay. To assay for β-galactosidase activity, 10μg of protein was mixed with 800μl of assay buffer (100mM $NaPO_4$ pH7, 1mM $MgSO$ and 10mMβ-mercaptoethanol) and 200μl of ONPG (4mg/ml)(BDH). The absorbance was measured at 420nm on a Beckman DU*640 spectrophotometer, and the data analysed using the manufacturer's "Kenetic/Time" program. The slope was used to compare the efficiency of transfection between samples.

CHAPTER 3

RESULTS

A) The COPT promoter-trap retrovirus

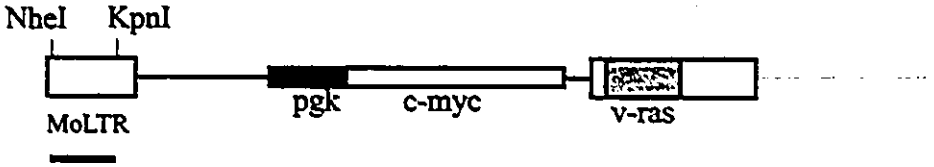
Construction and production of the COPT retrovirus

The COPT retroviral vector contained the v-Ha-*ras* oncogene coding sequence in the U3 region of the 3' LTR of the MoMuLV, upstream of the promoter (Figure 3.1A). Other examples of MoMuLV vectors containing foreign sequences in the LTR showed no effect on the replication of the virus, with titers being comparable to those of viruses with wild type LTRs (Stuhlmann *et al.* 1989; Reik *et al.* 1985; Lobel *et al.* 1985; von Melchner and Ruley 1989; Reddy *et al.* 1991). The *ras* sequence amplified by PCR included the complete translated sequence, starting four base pairs before the initiating AUG, and ending at the termination codon. The COPT vector also contained the c-*myc* oncogene under the control of the ubiquitous pgk promoter (McBurney *et al.* 1991). This cassette was inserted in place of the retroviral genes between position 1484 and 8125 of the retroviral vector. The vector contained an intact packaging signal sequence (ψ) so that it could be packaged into infectious virions. Upon infection, the *ras* sequence is duplicated in the 5' LTR, through the process of reverse transcription (see Figure 1.1), positioning the oncogene 95bp away from

Figure 3.1 COPT Promoter-Trap Retrovirus

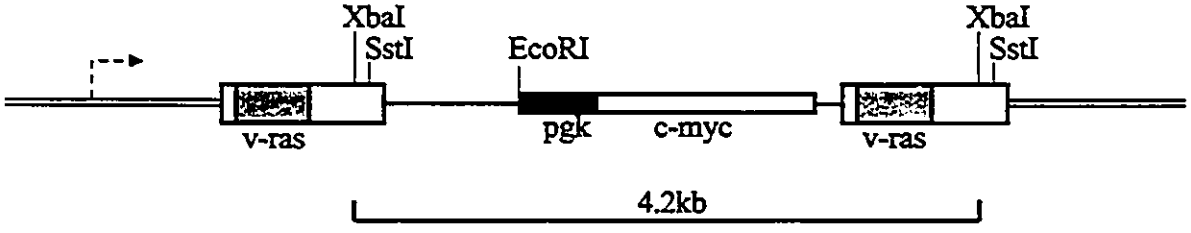
Thick lines represent MoMuLV sequences, thin lines represent plasmid sequences, and the double lines represent cellular DNA. Open boxes represent the LTRs. The large shaded boxes represent the v-Ha-*ras* sequence and the thinner grey boxes represent the c-myc sequence. Black boxes represent the pgk promoter.
A. The COPT retroviral vector showing the ras sequence inserted in the 3' LTR and the pgk-c-*myc* cassette replacing most of the viral genes. The dark line under the 5' LTR shows the NheI-KpnI fragment used for probes.
B. The COPT provirus showing the *ras* sequence duplicated in the 5' LTR. The broken line arrow indicates a potential active cellular promoter. Any restriction enzyme present in the LTRs only generate a 4.2kb proviral fragment. The only EcoRI site is indicated.

A


NheI
KpnI
MoLTR
pgk
c-myc
v-ras

1kb

B


XbaI
SstI
EcoRI
XbaI
SstI
v-ras
pgk
c-myc
v-ras
4.2kb

cellular DNA (Figure 3.1B). In the event that the provirus integrated near an active cellular promoter, *ras* may be expressed and since *myc* is constitutively expressed, the presence of the two oncogenes should lead to the formation of a tumor. The virus was given the name COPT for *c*ooperating *o*ncogenes *p*romoter-*t*rap.

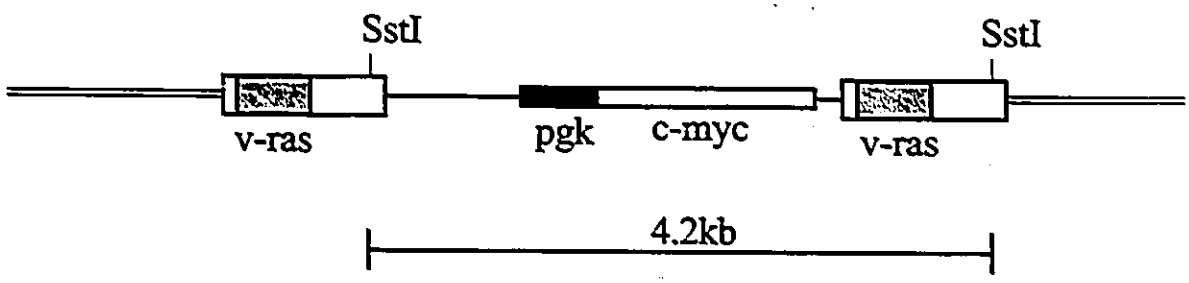
Because the virus is replication deficient, the GP+E86 packaging cell line (Markowitz *et al.* 1988) was used to produce the COPT virus. This cell line, derived from NIH 3T3 cells, encodes the three viral genes from two separate plasmids, thus eliminating the possibility of recombination events that could produce wild-type virus. The COPT retroviral vector was cotransfected in this cell line along with a vector conferring neomycin resistance (pgk-neo) (McBurney *et al.* 1991). Twenty-four G418 resistant colonies were isolated, expanded and tested for virus production.

To test that the COPT virus was intact, NIH 3T3 fibroblasts were infected with COPT virus from the various producer clones and the DNA was analysed by Southern blot analysis 48 hours post-infection. By digesting the DNA with SstI which cuts once in each LTR we expected to see one band of 4.2kb using *ras* as a probe (Figure 3.2A). Figure 3.2B shows the results from six representative clones. Different producer clones showed varying amounts of virus with some being below detectable levels. The proviral band was of the expected size, indicating that the COPT vector was producing a functional intact virus. Clone 3 had the strongest signal and was selected as the source of COPT virus.

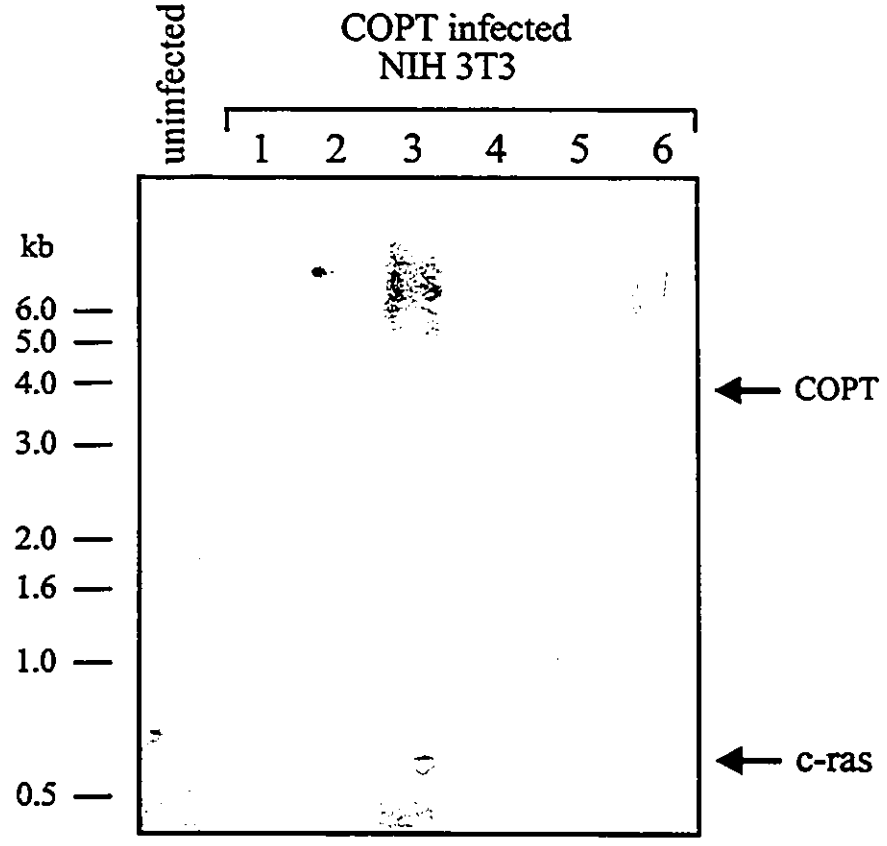
Figure 3.2 Southern Blot Analysis of the COPT Virus

A. The COPT provirus. Thick lines represent MoMuLV sequences and double lines represent cellular DNA. Open boxes represent the LTRs and the shaded boxes represent the v-Ha-*ras* sequence. SstI digestion generates a 4.2kb fragment.
B. Southern blot analysis of NIH 3T3 cells infected with COPT virus from different producer clones (1-6). The DNA was digested with SstI and the blot was probed with the ras sequence. Markers are the 1kb ladder (BRL).

A


SstI
SstI
v-ras
pgk
c-myc
v-ras
4.2kb

B

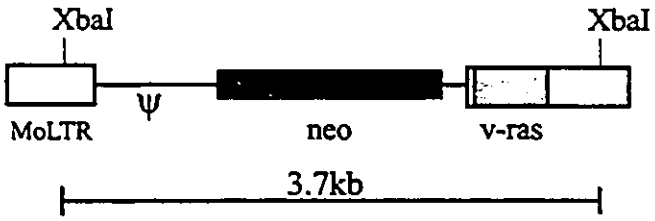

uninfected
COPT infected
NIH 3T3
1
2
3
4
5
6
kb
6.0
5.0
4.0
3.0
2.0
1.6
1.0
0.5
COPT
c-ras

Because COPT is replication deficient and lacks a selectable marker, the titer could not be assessed directly. Instead, an estimate of the COPT titer was obtained using the same type of Southern blot analysis. A retrovirus very similar to COPT was used to demonstrate that the amount of virus produced can be reflected by the intensity of the signal from infected cells on a Southern blot. The pUCRas vector (Figure 3.3A) is identical to the COPT vector, except for the neomycin phosphotransferase gene replacing the pgk-c-*myc* cassette. The neo gene is expressed from the LTR promoter. The pUCRas virus was produced in GP+E86 cells in the same way as described for COPT (except that it was transfected alone since it encodes its own drug resistance gene). The titer of 6 pUCRas producing clones was established by drug selection of NIH 3T3 cells infected with dilutions of the virus. The titers were in the same range as those obtained from the COPT producer clones (10^5 cfu/ml). To ensure that this approach was valid, NIH 3T3 fibroblasts were first infected with undiluted as well as a two fold dilution of the virus and the signal on a Southern blot was compared. Figure 3.3B shows that the intensity of the signal from infected cells reflected the titer of the various pUCRas producer clones. It also shows that for every clone tested, a two fold dilution of the virus gave a signal which was half the intensity of that from undiluted virus. When comparing the signal from cells infected with the COPT virus to the signal from cells infected with viruses of known titer (pUCRas) it was estimated that the titer of clone 3 was approximately $2x10^5$ cfu/ml.

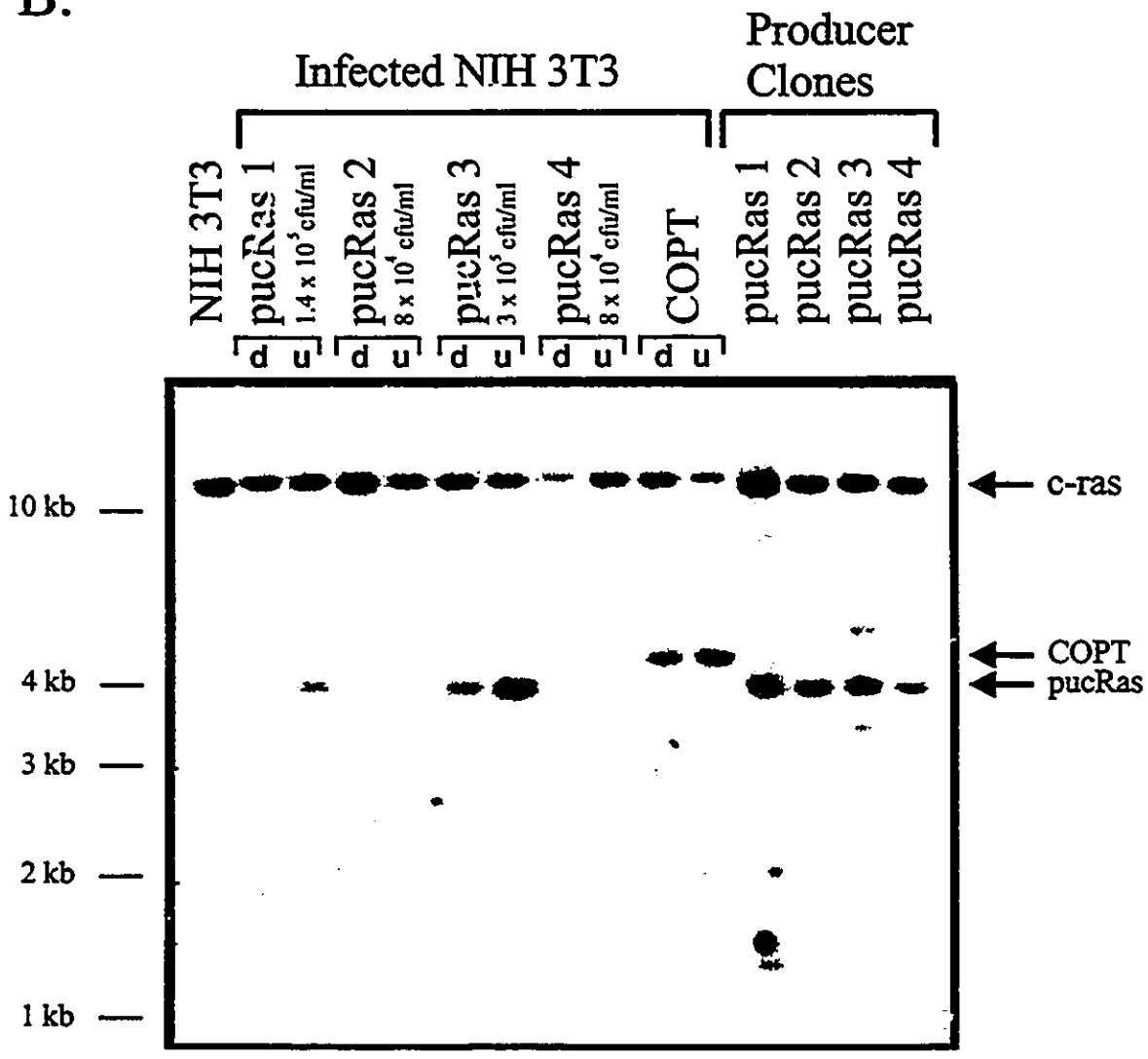
Figure 3.3 Titer Determination by Southern Blot Analysis

A. The pUCRas retroviral vector. Thick lines represent MoMuLV sequences and double lines represent cellular DNA. Open boxes represent the LTRs and the light shaded box represents the v-Ha-*ras* sequence. The dark shaded box represents the neomycin prosphotransferase gene (neo), expressed from the LTR promoter. ψ represents the packaging signal. XbaI digestion generates a 3.7kb fragment.
B. Southern blot analysis of NIH 3T3 cells infected with pUCRas virus from different producer clones (1-4), or infected with the COPT virus from clone 3. For each clone, cells were infected with undiluted (u) as well as diluted (d) (two-fold) virus. The DNA was digested with XbaI and the blot was probed with the ras sequence. The titer of the pUCRas clones is indicated. The titer of the COPT clone 3 was estimated at 2 x 10^5 cfu/ml. Also shown is DNA from the four pUCRas producer clones. Markers are the 1kb ladder (BRL).

A.


XbaI
XbaI
ψ
MoLTR
neo
v-ras
3.7kb

B.


Infected NIH 3T3
Producer Clones
NIH 3T3
pucRas 1
1.4 x 10⁵ cfu/ml
pucRas 2
8 x 10⁴ cfu/ml
pucRas 3
3 x 10⁵ cfu/ml
pucRas 4
8 x 10⁴ cfu/ml
COPT
pucRas 1
pucRas 2
pucRas 3
pucRas 4
d u
10 kb
4 kb
3 kb
2 kb
1 kb
c-ras
COPT
pucRas

Testing of the promoter-trap in vitro

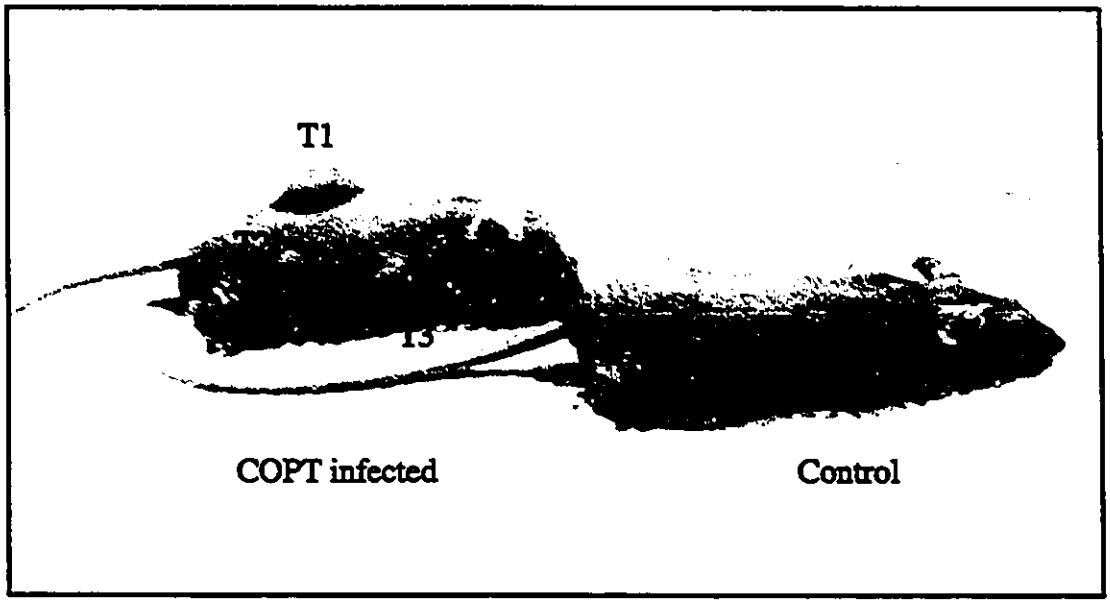
The promoter-trap capability of the COPT virus was tested in nude mice. NIH 3T3 cells were infected with the COPT virus from the producer clone 3 at a multiplicity of infection (m.o.i.) near one. The next day, $6x10^6$ infected cells were injected subcutaneously in one mouse in four different sites: above the two shoulders and above the two hips. As a control for the potential of the cell line to form tumors, a second mouse was injected with $6x10^6$ uninfected NIH 3T3 cells. Three tumors formed in the mouse injected with the COPT infected cells with a latency of two to three weeks (Figure 3.4). The control mouse did not develop any tumor during the same period of time. The mice were sacrificed at four weeks post injection, and the tumors were isolated. The tumors were reimplanted in a nude mouse by mincing a small portion of the tumor tissue with blades and breaking it down further by passing it through a 20 gauge needle to obtain a single cell suspension. Each tumor cell suspension was injected in two sites per mouse. Only tumor 1 developed a secondary tumor, which was isolated. Tumor 3 was injected in the same mouse as tumor 1 and may not have had time to grow a secondary tumor before the mouse had to be killed (three weeks later) because of the large size of the first tumor. The mouse injected with tumor 2 never developed tumors.

Analysis of the tumors

For each tumor (including the secondary tumor 1) a small portion was used to establish a culture. The tumor tissue was minced with blades and passed through a 20 gauge needle as described above. The tumor cells were grown in the same medium used to grow

Figure 3.4 The COPT Promoter-Trap in Nude Mice

The mouse on the left was injected subcutaneously with COPT infected NIH 3T3 cells. T1 appeared after 2½ weeks, and T2 and T3 appeared after three weeks. The mouse on the right was injected with uninfected NIH 3T3 cells.


T1
T3
COPT infected
Control

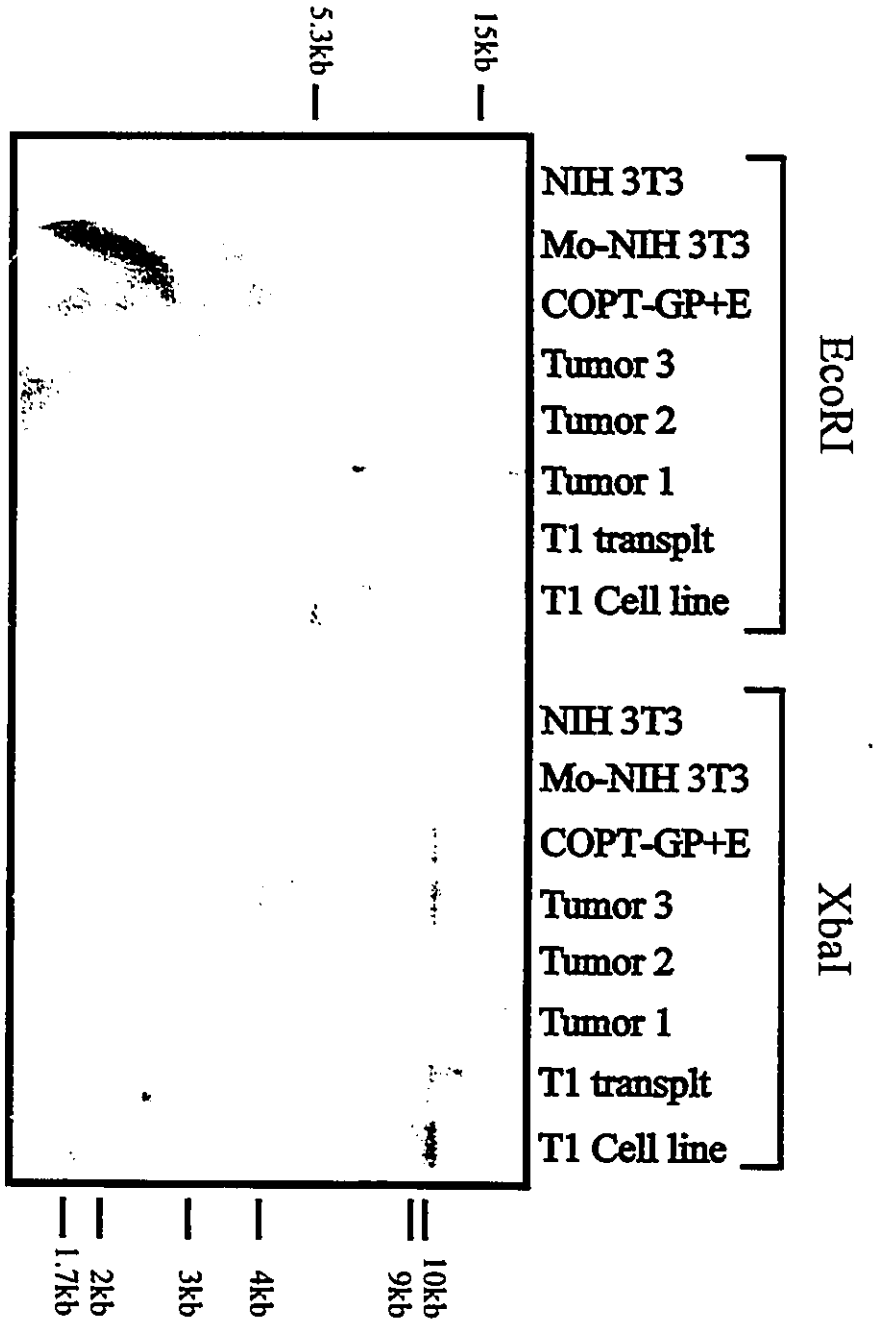
NIH 3T3 cells (α-MEM with 10%serum). Tumors 1 and 3 grew well but tumor 2 failed to grow past a few generations. Part of each tumor was used to prepare DNA for Southern blot analysis and PCR amplification. For tumor 2, this represented the rest of the tumor because it was quite small. For tumors 1 and 3, as well as the secondary tumor 1, the rest of the tumor was used to prepare RNA for northern blot analysis. RNA was also prepared from the established tumor cell lines.

Southern blot analysis

The tumors were first tested for the presence of the COPT provirus by Southern blot analysis to establish a correlation. The tumor DNA was digested with XbaI which should give a 4.2kb band when probed with the *ras* sequence (Figure 3.1B). Figure 3.5 shows the results of this analysis. Control samples included untreated NIH 3T3 cells (lane 1) and COPT producer clone 3 (lane 3). The second lane was DNA from NIH 3T3 cells that were infected with MoMuLV. The band seen at 10kb represents the endogenous ras gene. All three tumors were positive for the provirus but two of them showed rearrangements. Tumor 3 had a shorter proviral band and tumor 1 showed two bands (~1.7kb and 9kb). To investigate the nature of these rearrangements, the same blot was probed again with myc and then the pgk promoter to try to identify the missing sequences (data not shown). The provirus in tumor 1 was negative for the presence of both myc and pgk which would indicate that a large portion of the provirus was lost. The provirus in tumor 3 was positive for both myc and pgk promoter sequences. Further restriction enzyme analysis indicated that the internal deletion in tumor 3, estimated at 500bp, is most likely

Figure 3.5 Southern Blot Analysis of COPT-induced Tumors in Nude Mice

Southern blot containing DNA from the three tumors obtained in the nude mouse (Fig 3.4). The DNA was digested with EcoRI or XbaI as indicated, and the blot was probed with the ras sequence. Lane 1 contained DNA from uninfected NIH 3T3 cells. Lane 2 contained DNA from NIH 3T3 cells infected with a wild type MoMuLV. Lane 3 contained DNA from the COPT producer clone. Tumors 1,2 and 3 correspond to T1, T2 and T3 in Figure 3.4. Tumor 1 was propagated in another nude mouse (transplant) and in culture (cell line). Markers are the 1kb ladder (BRL), and high molecular weight markers (BRL).


5.3kb
15kb
NIH 3T3
Mo-NIH 3T3
COPT-GP+E
Tumor 3
Tumor 2
Tumor 1
T1 transplt
T1 Cell line
EcoRI
NIH 3T3
Mo-NIH 3T3
COPT-GP+E
Tumor 3
Tumor 2
Tumor 1
T1 transplt
T1 Cell line
XbaI
1.7kb
2kb
3kb
4kb
9kb
10kb

around the 3' end of the myc sequence. This conclusion was supported by northern blot analysis (see below).

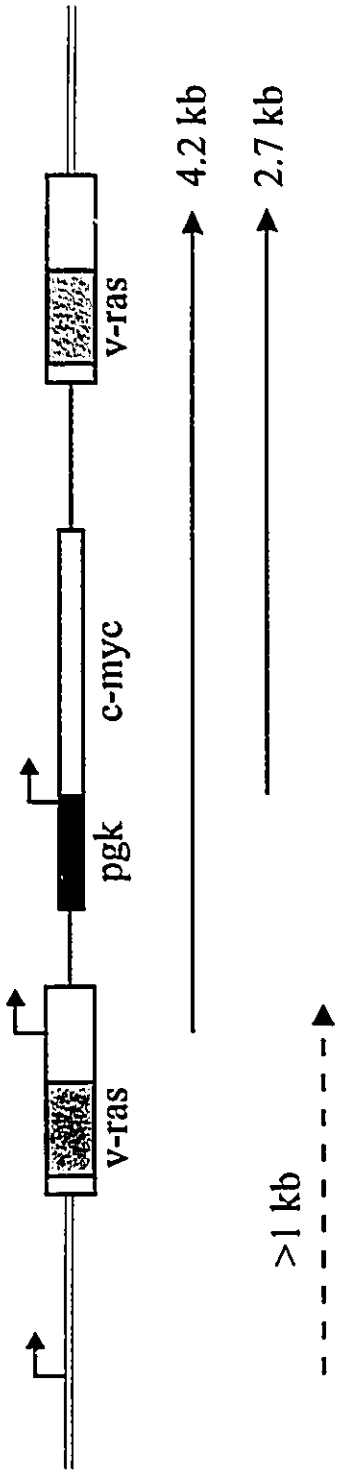
To get an idea of how many viruses infected the original cell that gave rise to the tumors, genomic DNA was digested with an enzyme which cuts only once in the provirus. EcoRI was chosen because it cuts in the middle of the provirus (Figure 3.1B). Each integrated provirus should give two bands (when probed with *ras* or the LTR) corresponding to the proviral-cellular junction fragments, the size of which will depend on the site of integration. Figure 3.5 shows the results of this analysis using the *ras* sequence as the probe. Tumor 3 showed two bands (3.5kb and ~7kb) that are large enough to be junction fragments. Tumor 2 showed only one band (~8kb). The second band may be too faint to see or simply too large to have transferred efficiently. It is difficult to know what to expect from tumor 1 due to the large rearrangement. Only one band (5.5kb) was detected which may mean that the internal EcoRI site was part of the deletion and the band represent a genomic fragment harboring the truncated provirus. The fact that no more than two junction fragments were seen indicated that each tumor contained a single provirus.

Northern blot analysis

As illustrated in Figure 3.6, two transcripts are expected from the COPT provirus. The longer one is generated by the LTR promoter and should be 4.2kb in length. The second, shorter transcript is produced by the pgk promoter and should be 2.7kb long. If a tumor was caused by a promoter-trap event, an extra transcript initiated by an active cellular promoter and terminated in the 5' LTR should be detected (indicated by the dotted line). This transcript would be responsible for the expression of *ras*. The length of this

Figure 3.6 COPT Proviral Transcript

The COPT provirus. Thick lines represent retroviral sequences and double lines represent cellular DNA. Open boxes represent the LTRs and the shaded boxes represent the *ras* sequence. The arrows represent a cellular promoter and the two proviral promoters. Every provirus should produce two transcripts, one initiating in the 5' LTR (4.7kb), and one initiating at the pgk promoter (2.7kb). The dotted line represents a hybrid transcript initiating in the cellular sequence and terminating in the 5' LTR. This transcript would have a minimum size of 1kb and would be present only if the provirus integrated near an active cellular promoter.


v-ras
c-myc
pgk
v-ras
4.2 kb
2.7 kb
>1 kb

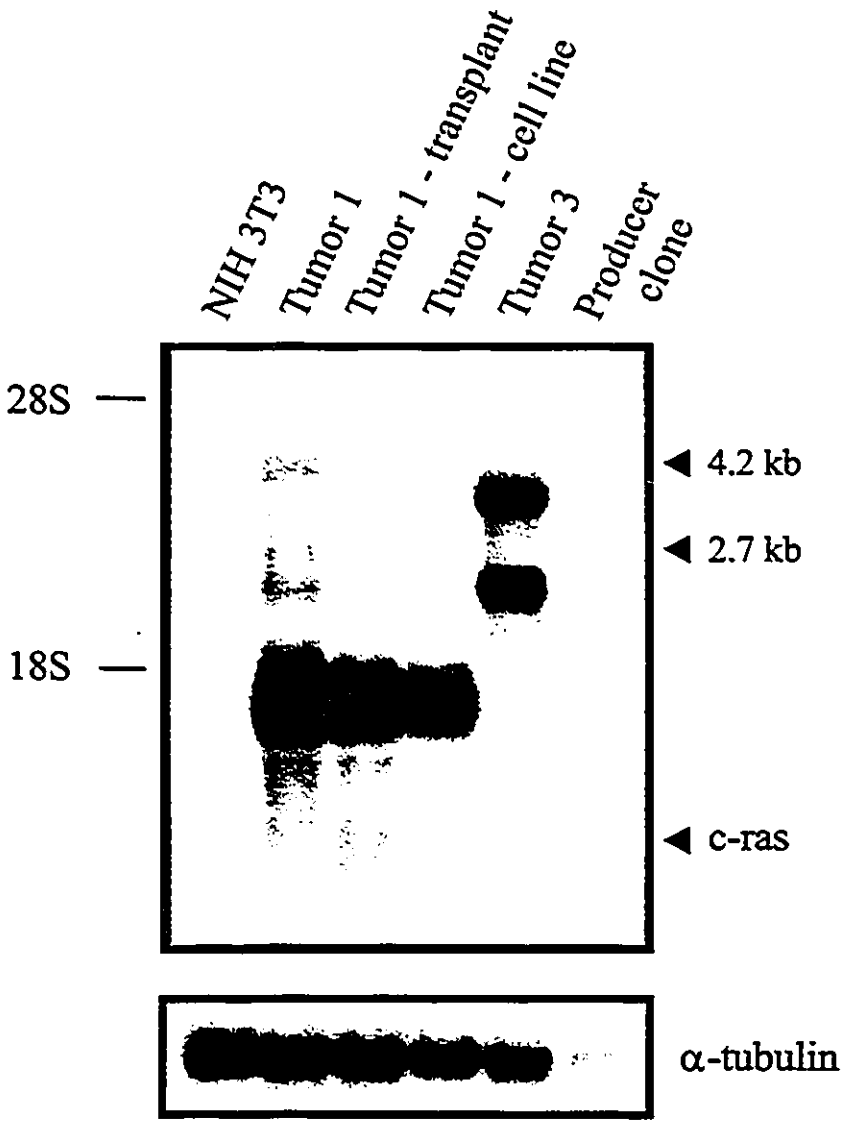
hybrid transcript would depend on the distance between the start of transcription in the cellular gene, and the start of the provirus, but should be at least 1kb long, that is the distance from the 5' end of the provirus to the polyadenylation site in the 5' LTR. Transcripts initiating at the cellular promoter and terminating in the 3' LTR may be produced but are not expected to allow *ras* protein expression (von Melchner and Ruley 1989; Reddy *et al.* 1991).

Total RNA (10μg) was used for northern blot analysis and the *ras* sequence was used as the probe (Figure 3.7). As a negative control, the first lane contained RNA from NIH 3T3 cells. As a positive control for the two proviral transcripts, the last lane contained RNA from the COPT producer clone 3. As shown by the loading control (α-tubulin) this lane contained much less RNA, which explains the weaker signal. The primary tumor 1 showed a number of transcripts, some of which correspond to the two expected proviral transcripts (4.2kb and 2.7kb). These transcripts seem to have disappeared in the tumor 1 cell line and in the transplanted tumor. The main transcript in all of the tumor 1 samples is approximately 1.5kb long. This is consistent with a hybrid genomic-provirus transcript.

Tumor 3 showed two transcripts which are both shorter than the expected size for the two proviral transcripts by about 500bp. This was consistent with the Southern blot analysis which indicated an internal deletion of about the same size somewhere in the *myc* sequence. These two transcripts can also be detected by probing with *myc* or pgk promoter sequences. No additional transcript was detected in this tumor. As mentioned earlier, RNA from tumor 2 was not available.

Figure 3.7 Northern Blot Analysis of COPT-Induced Tumors in Nude Mice

Northern blot analysis using 10 µg of total RNA extracted from two of the three tumors obtained in the nude mouse (Figure 3.4). RNA from uninfected NIH 3T3 cells and from the COPT producer clone 3 were included (first and last lanes). The blot was probed with the ras sequence. α-tubulin was used as the loading control. The position of the 18S and 28S ribosomal RNA is indicated on the left. The two expected proviral transcripts are indicated by arrow heads. Tumor 1 was propagated in another nude mouse (transplant) and in culture (cell line).


NIH 3T3
Tumor 1
Tumor 1 - transplant
Tumor 1 - cell line
Tumor 3
Producer clone
28S
18S
4.2 kb
2.7 kb
c-ras
α-tubulin

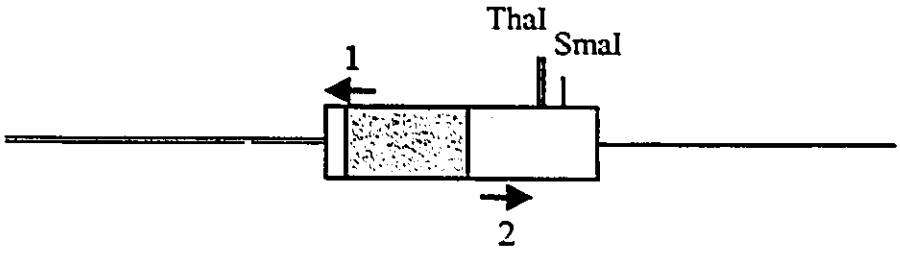
PCR amplification and sequencing

To clone the site of integration, genomic DNA adjacent to the 5' end of the provirus was amplified by the polymerase chain reaction. Figure 3.8 illustrates the reverse PCR strategy employed (Silver and Keerikatte 1989). Genomic DNA was digested with either ThaI or SmaI. These sites are found in the LTRs at position 322 324 and 404 respectively. Because ThaI has a four base pair recognition site (CGCG) its site should normally be found in genomic DNA at an average frequency of once every 256bp. SmaI has a six base pair recognition site (CCCGGG) and is therefore found less frequently than ThaI in genomic DNA (approximately once every 4kb). However, CG dinucleotides are found in DNA at a lower frequency than expected (four times less frequent; Cross and Bird 1995) making ThaI and SmaI sites less frequent than estimated. On the other hand, CpG dinucleotides are more frequent around promoter regions (Cross and Bird 1995) than in random DNA sequence, and restriction sites containing CpG dinucleotides are often found in clusters near promoters. Since our promoter-trap requires that proviruses integrate near cellular promoters, ThaI and SmaI sites may be present upstream of the provirus at a greater frequency than expected in random DNA sequences. Digestion with these enzymes should release a DNA fragment containing a portion of the LTR, including *ras*, and some genomic sequence. The length of the restriction fragment depends on where the enzyme(s) cut in the genomic DNA. Self-ligation of these junction fragments allows the genomic sequences to be amplified using oligonucleotides in the LTR with opposite direction of polymerisation (going away from each other, see Figure 3.8).

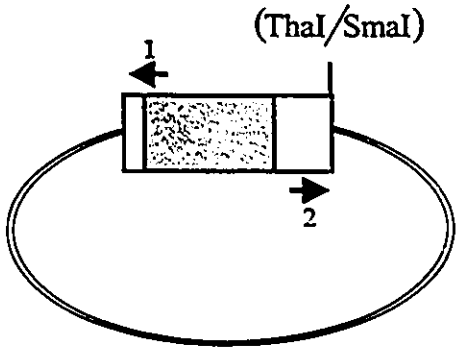
Figure 3.8 Reverse PCR Strategy

PCR strategy used to amplified genomic DNA 5' of the proviral integration site.
1. Illustrated is the 5' LTR of the COPT provirus. The thick line represents retroviral *gag* sequences and the double line represents cellular DNA. The open box represents the LTR and the shaded box represents the v-Ha-*ras* sequence. Genomic DNA from the tumors or from NIH 3T3 cells was digested with either ThaI or SmaI.
2. The junction fragments were allowed to ligate on their own, forming a close circular molecule. Using oligonucleotides as shown (1 and 2), the genomic sequences were amplified by PCR.
3. The resulting PCR products contained cellular DNA flanked by portions of the LTR.

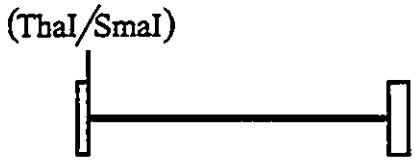
1.



2.



3.



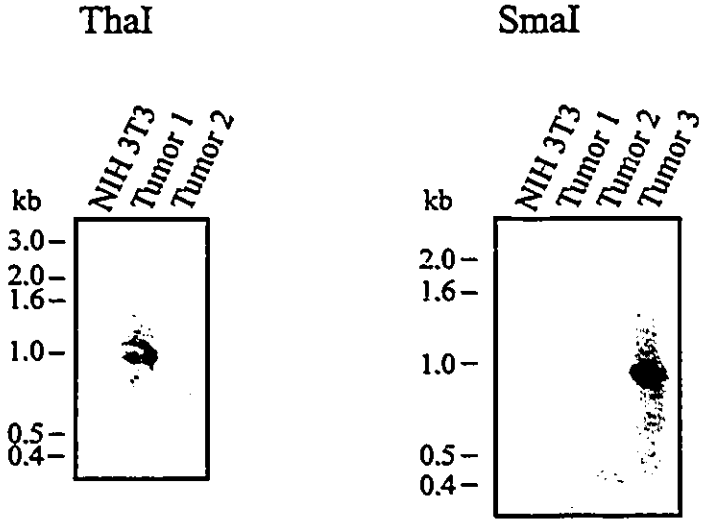
To make the amplification specific to the COPT provirus, as opposed to endogenous ecotropic proviruses, one oligonucleotide (oligo1) spanned the junction between the 5' end of the *ras* and the LTR. The second oligonucleotide (oligo2) was positioned in the LTR between the 3' end of *ras* and the ThaI restriction site. These oligonucleotides should give an amplification product with a minimum size of 139bp.

Using ThaI to digest the genomic DNA, a fragment of approximately 1kb was amplified from tumor 1. To confirm that this fragment contained part of the LTR, Southern blot analysis was performed on the fragment using a portion of the LTR as a probe (NheI-KpnI fragment shown in Figure 3.1A). Figure 3.9A shows that the 1kb PCR product did contain part of the LTR and hopefully included genomic DNA 5' of the provirus. One could not however exclude the possibility that the amplified fragment represented the 3'LTR. From an intact provirus the product from the 3' LTR should produce a 1.1kb fragment. Because of the truncated provirus in tumor 1, this product may not be of the expected size. The question would have to be answered by sequencing. The 1kb fragment was then subcloned in the SmaI site of the pGEM4 vector (Promega). Four clones were used to obtain sequence from both ends of the insert, using the T7 and Sp6 primers (Pharmacia). From the sequence it was observed that all clones were truncated versions of the 1kb fragment. This was most likely caused by an over active MungBean nuclease used to blunt the end of the amplified fragment before subcloning. The various segments of sequence obtained contained either some of the 3'LTR (including part of the env gene), some of the 5' LTR (including some of the downstream sequences) or sequences downstream of the 5' LTR. Although these sequences were found on separate pieces of DNA from different subclones

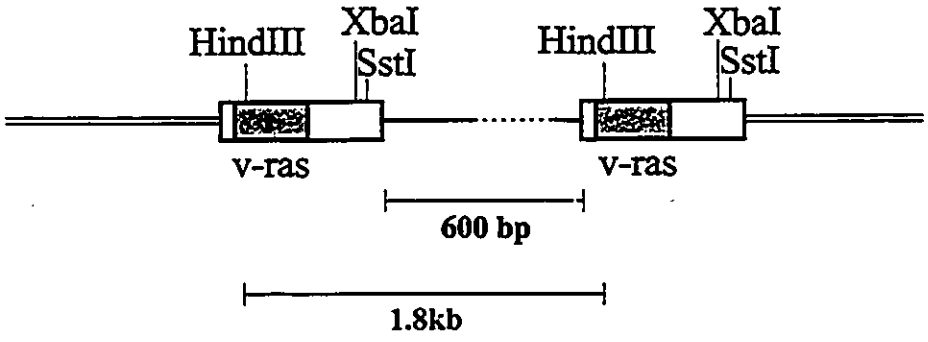
Figure 3.9 PCR Amplification Products

A. Southern blot analysis of fragments amplified from tumor DNA by the polymerase chain reaction. The left panel shows the amplification products from the indicated tumors when genomic DNA was digested with ThaI. The right panel shows the amplification products from the indicated tumors when the DNA was digested with SmaI. The probe used was the LTR fragment shown in Figure 3.1A. NIH 3T3 DNA served as a negative control. Markers are the 1kb ladder (BRL).
B. Structure of the rearranged COPT provirus from tumor 1 as deduced from Southern and northern blot analyses and sequencing. The LTRs are shown by the open boxes and the v-Ha-*ras* sequence by the shaded boxes. The double lines represent cellular DNA. The distance between the LTRs was estimated at 600bp.

A


ThaI
SmaI
NIH 3T3
Tumor 1
Tumor 2
Tumor 3
kb
3.0
2.0
1.6
1.0
0.5
0.4

B


HindIII
XbaI
SstI
HindIII
XbaI
SstI
v-ras
v-ras
600 bp
1.8kb

of the 1kb fragment, it strongly suggested that the provirus in tumor 1 had lost a large piece of DNA between the two LTRs. Because the subclones were severely truncated by the nuclease it was not possible to map the break points of the deletion. Unfortunately, the 5' genomic junction did not amplify, and the possibility of a promoter-trap could not be confirmed.

No fragment could be amplified from TaqI digested DNA from tumor 3. SmaI however, gave an amplification product of approximately 900bp. This amplification product was also shown to contain LTR sequences by Southern blot analysis (Figure 3.9A), but could never be successfully subcloned. It was therefore not possible to prove or disprove the possibility of a promoter-trap event in that tumor either.

Flanking genomic DNA from tumor 2 could never be amplified using either TaqI or SmaI.

Mammary gland injections

To first test the efficiency of virus delivery to the target cells via nipple injections, pubescent mice were injected with a MoMuLV-based virus expressing the β-galactosidase gene from the LTR promoter. The titer of the so called BAG virus (Price *et al.* 1987) was in the same range as that of the COPT virus ($2x10^5$ cfu/ml). Staining of whole glands with X-gal , two to three days post injections, revealed blue staining in the cells of the end buds (Figure 3.10). The intensity of the stain decreased as the distance from the end buds to the nipple increased. No staining was detected in uninfected glands from the same mouse. It

Figure 3.10 X-gal Staining of Mammary Glands

Whole mount of three mammary glands stained with X-gal (see materials and methods). The bottom gland was isolated from a transgenic mouse carrying the lacZ gene under the control of the pgk promoter (Dr. McBurney). The middle gland was injected *in situ* with the BAG virus (encodes lacZ). The top gland came from the same animal as the middle gland but was not injected with the virus.

Non-injected

BAG-Injected

Transgenic

was not possible to estimate the efficiency of infection from whole mounts due to the low resolution.

To test the COPT virus *in vivo*, 24 four week old Balb/C female mice were used. The mice were anaesthetized and a fine needle was used to inject the virus in the gland through the nipple. A volume of 0.05 to 0.1 ml of fresh virus was injected into each of the ten glands whenever possible. That represented approximately 10^4- $2x10^4$ infectious particles per gland. The mice were monitored every four to five days for the presence of palpable tumors. After 8 weeks, none of the mice had developed tumors and they were all mated. After a total of 6 months, no tumor had formed and the mice were sacrificed. One mouse was used to extract RNA from infected mammary glands to see if COPT transcripts could be detected. Using polyA$^+$ selected RNA and c-*myc* as a probe, no proviral transcripts were seen despite a good signal for the endogenous *myc* transcript (data not shown).

B) The PyT promoter-trap retrovirus

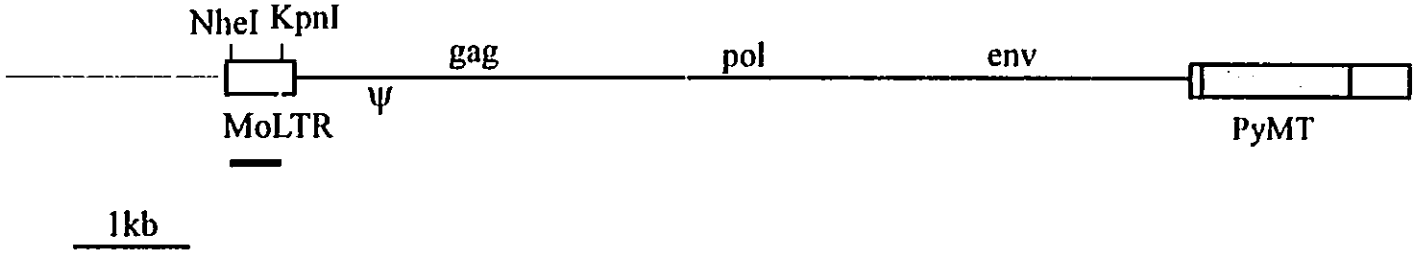
Construction and production of the PyT retrovirus

The PyT retroviral vector, also a MoMuLV based vector, contained the Polyomavirus middle T antigen (PyMT) oncogene inserted in the U3 region of the 3' LTR (Figure 3.11A). The rest of the retroviral sequences were kept intact making the PyT retrovirus replication competent. Upon integration of the provirus in host DNA PyMT is positioned 95bp away from cellular sequences and may be expressed from a cellular promoter (Figure 3.11B).

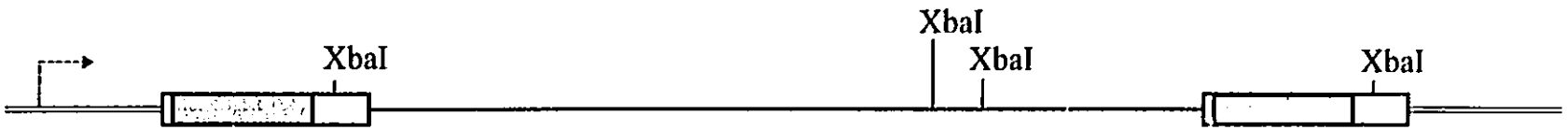
Figure 3.11 PyT Promoter-Trap Retrovirus

Thick lines represent MoMuLV sequences, thin lines represent plasmid sequences, and double lines represent cellular DNA. Open boxes represent the LTRs. The Polyomavirus Middle T antigen (PyMT) is represented by a shaded box.
A. Schematic of the PyT retroviral vector with the PyMT inserted in the 3' LTR. ψ represents the packaging signal. The dark line under the 5' LTR indicates the NheI-KpnI fragment used for probes.
B. Schematic of the PyT provirus. The PyMT sequence has been duplicated in the 5' LTR during reverse transcription. The broken line arrow indicates a potential cellular promoter. XbaI restriction sites are indicated.

A



B



67-A

The PyT vector was cotransfected in NIH 3T3 cells with a vector conferring puromycin resistance (pgk-puro, from the laboratory of Dr. McBurney). Twenty-four puromycin resistant colonies were isolated, expanded and tested for PyT virus production. This was done by infecting NIH 3T3 fibroblasts with PyT virus from different producer clones. The DNA from the infected cells was isolated 48 hours post-infection and analysed by Southern blot. The DNA was digested with XbaI which cuts four times in the provirus so that two bands of 5.5kb and 3.6kb were expected using a portion of the LTR as a probe. Figure 3.12A shows a schematic of the PyT provirus and the expected bands when the portion of the LTR indicated in figure 3.11A (Nhe1-KpnI) was used as a probe. Figure 3.12B shows that all of the clones were producing virus.

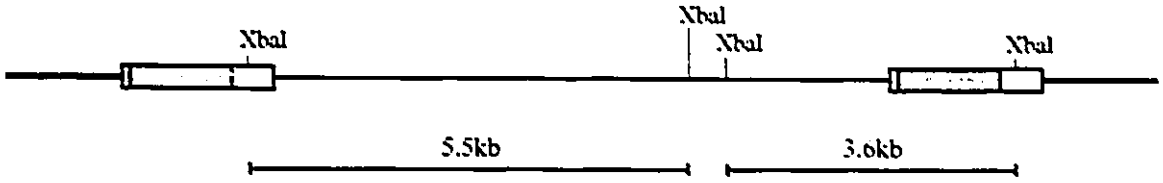
The titer from the different producer clones was determined using an XC assay which is based on the fact that XC cells, which are tumor cells from a rat sarcoma induced by the Rous Sarcoma Virus, fuse with Moloney virus-infected cells to form large syncytia. NIH 3T3 cells were infected with dilutions of the virus. Once the cells reached confluence, they were UV irradiated and XC cells were added. Any colony of infected NIH 3T3 cells would fused with the XC cells to form large syncytia which can be visualized as clear plaques by methyleneblue staining. The titers ranged from 1 to $5x10^6$ cfu/ml.

From Southern blot analysis of PyT-infected NIH 3T3 cells, like the one shown in Figure 3.12B, the observation was made that for most producer clones an extra band was present (2.3kb) (3' LTR-PyMT). The size of the band could correspond to the size expected of a wild type 3' LTR. To test that idea, the same Southern blot was reprobed with the PyMT sequence. As shown in Figure 3.13, only the 3.6kb band corresponding to the 3'LTR

Figure 3.12 Southern blot analysis of the PyT virus

A. Schematic of the PyT provirus and the expected XbaI fragments. Thick lines represent MoMuLV sequences and double lines represent cellular DNA. The open boxes represent the LTRs and the shaded boxes the Poliomavirus Middle T antigen.
B. Southern blot analysis of NIH 3T3 cells infected with PyT virus from different producer clones. The DNA was digested with XbaI and the blot was probed with the MoMuLV LTR (see Fig 3.11A). Markers are the 1kb ladder (BRL). The first lane contained DNA from uninfected NIH 3T3 cells. The numbers across the top indicate the different producer clones. The larger band represents the fragment of the provirus containing part of the 5' LTR (see A. above). The middle size band represents the fragment containing most of the 3' LTR. The smaller band represent a wild type 3' LTR (see text). Also shown at the bottom is the titer of each producer clone.

A


XbaI
XbaI
XbaI
XbaI
5.5kb
3.6kb

B

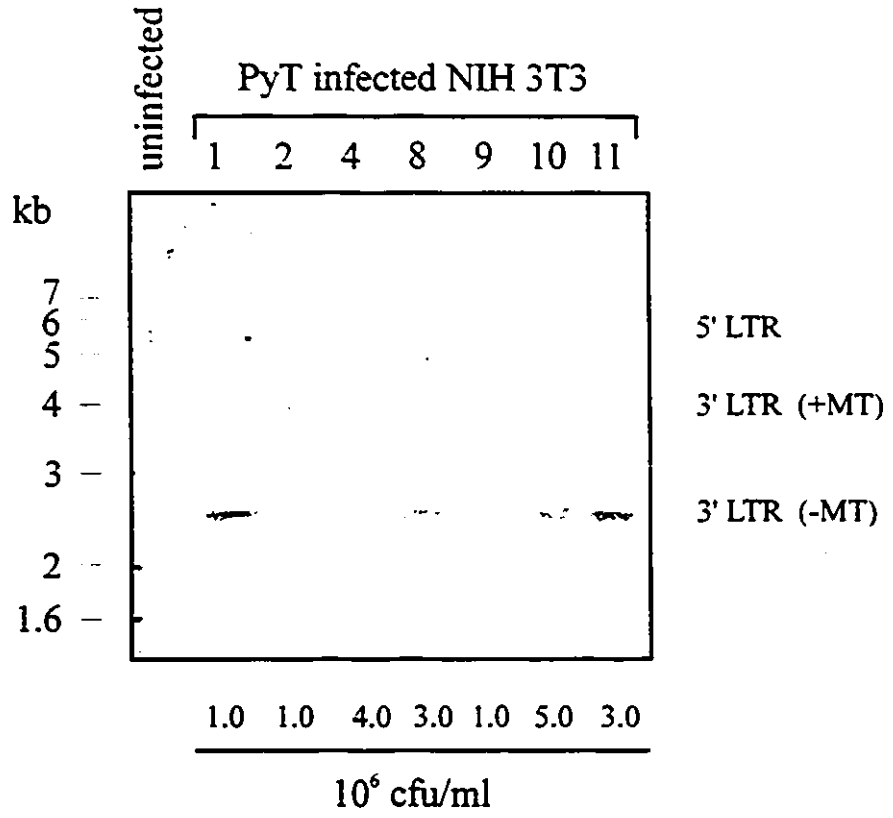
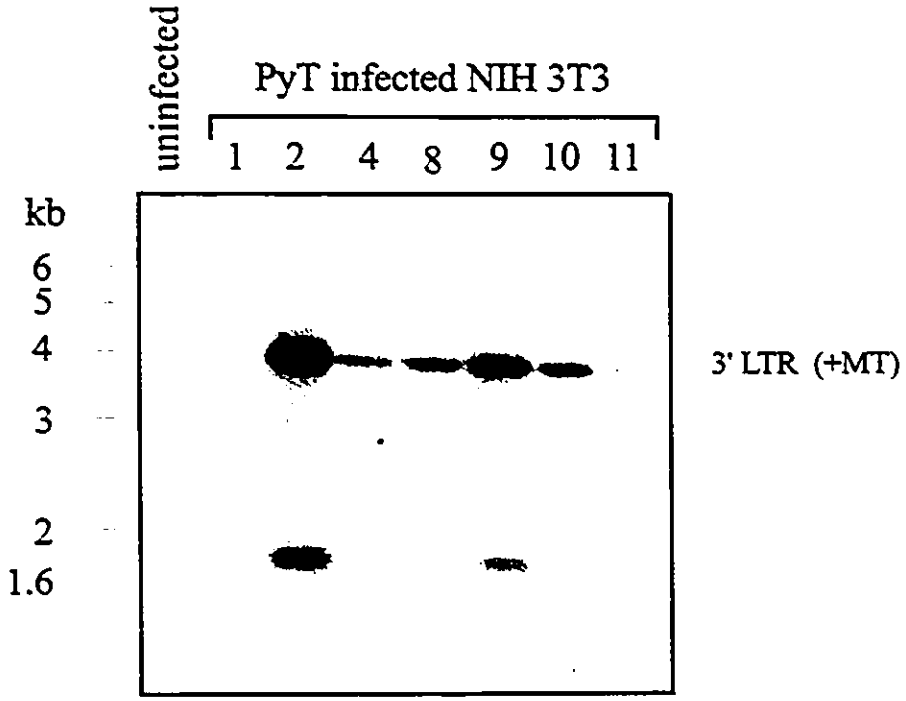

uninfected
PyT infected NIH 3T3
1 2 4 8 9 10 11
kb
7
6
5
4
3
2
1.6
5' LTR
3' LTR (+MT)
3' LTR (-MT)
1.0 1.0 4.0 3.0 1.0 5.0 3.0
10^6 cfu/ml

Figure 3.13 Southern blot analysis of the PyT virus using PyMT as a probe.

The Southern blot is the same as the one shown in Figure 3.12 but was probed with the PyMT sequence. Lanes are labelled the same as in Figure 3.12.


uninfected
PyT infected NIH 3T3
1 2 4 8 9 10 11
kb
6
5
4
3
2
1.6
3' LTR (+MT)

of the PyT provirus is detected by PyMT, and not the 2.3kb band. This pointed to the possibility that the PyMT sequence was somewhat unstable and got deleted at a frequency which varied among different producer clones. Since clone 2 showed a more stable provirus than the other clones (Figure 3.12B), it was chosen for future experiments.

First, to investigate the stability of the PyMT sequence in clone 2, an experiment was performed where the virus was passaged serially. NIH 3T3 fibroblasts were infected with PyT virus and every 48 hours, the supernatant was transferred onto fresh NIH 3T3 cells and the infected cells were harvested for Southern blot analysis. Once again, XbaI was used to digest the DNA and the probe was the same fragment of LTR as before. Figure 3.14 shows how the virus became unstable after three rounds of infection and had completely lost the PyMT sequence by the fifth passage.

Mammary glands injections

The mammary gland injections were done in the same manner as for the COPT virus. Twenty-four Balb/C pubescent females were anaesthetized and injected with the PyT virus in the mammary glands through the nipple (0.5-1 x 10^5 cfu per gland). All ten glands were injected whenever possible. The mice were monitored every three to four days for the appearance of palpable tumors. Three mice out of the nineteen survivors developed tumors at three different times of mammary gland development. This is summarized in Figure 3.15.

The first tumor was detected in an inguinal gland six weeks post-infection (nulliparous mouse). The same mouse eventually developed a few other tumors which appeared to be of mammary origin. A total of five apparently separate tumor masses were

Figure 3.14 Serial Passage of the PyT virus

Southern blot analysis of NIH 3T3 cells in which the PyT virus was passaged. The DNA was digested with XbaI and the blot was probed with the MoMuLV LTR (see Fig 3.11). The first lane contained DNA from the PyT producer clone 2. The second lane contained DNA from cells that were infected with the PyT virus from clone 2. The next lane represents cells that were infected with PyT virus collected from infected NIH 3T3 cells (1°infection), and so on. Forty-eight hours separated each infection. The origin of the three bands was described in Fig 3.12. Markers are the 1kb ladder (BRL).

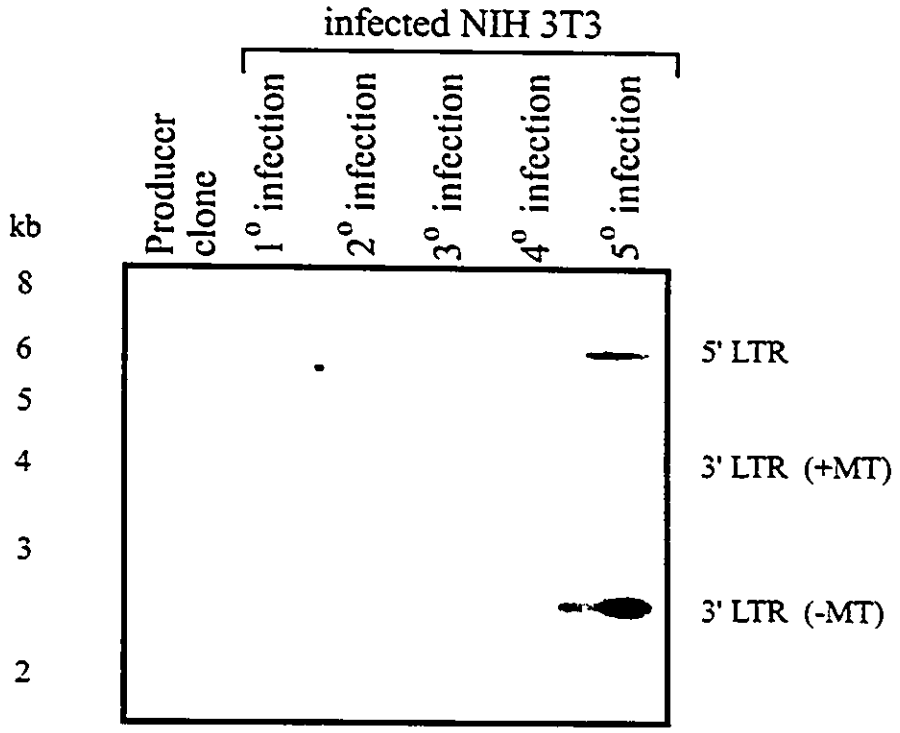

infected NIH 3T3
Producer clone
1° infection
2° infection
3° infection
4° infection
5° infection
kb
8
6
5
4
3
2
5' LTR
3' LTR (+MT)
3' LTR (-MT)

Figure 3.15 Summary of tumors obtained with the PyT virus

Summarized are the number of tumors (mammary and others) obtained in three animals injected with the PyT virus, with the time of detection post-infection (and post-mating). The mice were labelled according to the stage of mammary development at which they were at the time of detection.

MOUSE	NUMBER OF MAMM. TUMOR(S)	TIME OF DETECTION (wk post-infection)	OTHER TUMOR
Nulliparous	4 (5)	6	–
Monoparous	1	9 (3 wk post-mating)	spleen
Weaning	2	13 (7 wk post-mating)	colon

isolated. One was the first tumor detected and was isolated from the other ones. The second, third and fourth tumors were all in the inguinal region (opposite to the first tumor) and very close to each other. Although one was of a darker colour (red as opposed to whitish) they were difficult to isolate from each other. The fifth tumor was quite small and was located near the shoulder. The liver and a kidney were isolated to be used as negative controls. After detection of the first tumor the rest of the animals were mated. Three weeks later a second mouse (monoparous) developed a tumor. This was near the end of the pregnancy. That tumor grew slowly and remained small as the mouse was monitored daily until it was sacrificed 4 days later. Once again the tumor seemed to be in the mammary tissue. While isolating a kidney to serve as a negative control, a lump was found on the surface of the spleen and was saved for analysis.

Finally, a third mouse (weaning) developed a tumor seven weeks after the matings, which was a week after weaning the pups. The animal was sacrificed a week later and a second mammary tumor was found, as well as a small tumor on the colon which was saved for analysis. A kidney was isolated to serve as a negative control. The rest of the animals were killed four months post infection when no further tumors formed. Many animals showed symptoms of viraemia, including an enlarged thymus (Kozak and Ruscetti 1992).

Analysis of the PyT tumors

For some of the tumors an effort was made to grow them in culture. As described before, a small portion of the tumor mass was minced with blades and passed through a 20 gauge needle in a small volume of PBS. The cells were grown in α-MEM supplemented

with 10% fetal bovine serum, insulin (10μg ml) and dexamethazone (10^{-7}M). Unfortunately the tumor cells did not grow past a few generations. Most of the tumor tissue was used to extract DNA for Southern blot analysis. From tumors that were large enough RNA was also extracted.

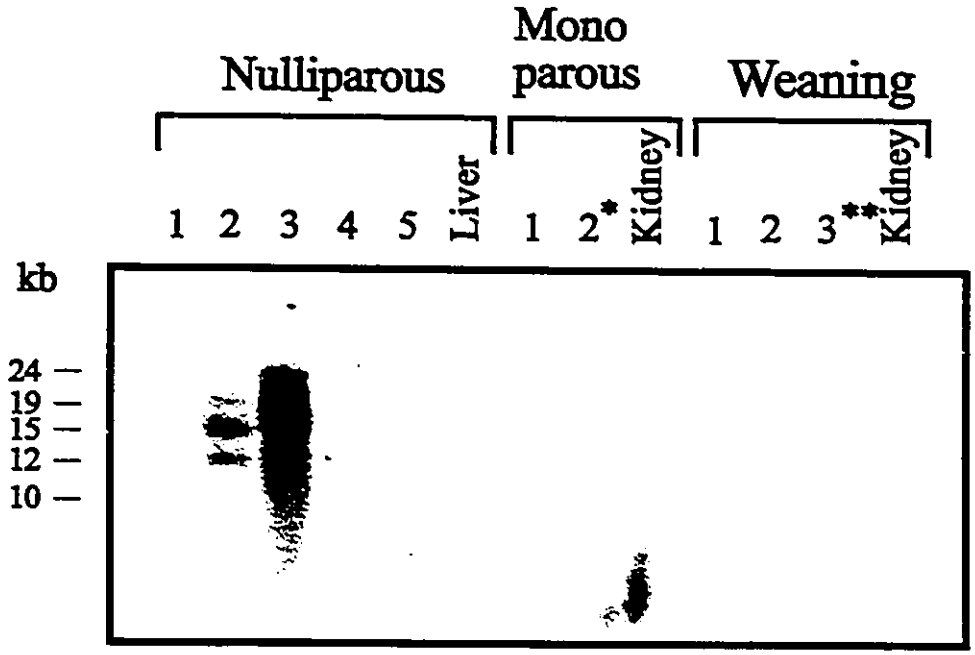
Southern blot analysis

The tumors were first tested for the presence of the PyT provirus by Southern blot analysis. The DNA was digested with EcoRI because this enzyme does not cut anywhere in the provirus so that each integrated provirus will generate an independent band. This would tell us how many proviruses are in each tumor. It can also give information on whether two tumors are related in terms of their origin.

Figure 3.16 shows the Southern blot for the EcoRI digest of tumor DNA probed with PyMT. The first mouse (nulliparous) showed the presence of the PyT provirus in all of the tumors maybe with the exception of the last one (tumor 5). The first tumor had one integrated provirus (~19kb). Tumors 2,3 and 4 had a few integrated proviruses, some might have been the same indicating maybe that there was some contamination with each other during dissection (on this particular gel tumor 4 DNA did not digest but other blots showed that some of the bands are the same as in tumor 2 or 3). Tumor 5 was small and the proviral signal might have been diluted with surrounding normal tissue. The liver was negative for the presence of the virus as expected. The second mouse (pregnant or monoparous) had only one mammary tumor which was positive for the virus, with two integrated proviruses. The fact that the bands are equally intense indicated that there was simultaneous infection of the

Figure 3.16 Southern Blot Analysis of PyT-Induced Tumors

Southern blot analysis (0.8% agarose gel) of DNA extracted from tumors induced by the PyT virus in mice. The DNA was digested with EcoRI and the blot was probed with the Polyomavirus Middle T antigen sequence. Mammary tumors from the three animals are indicated by numbers. 2* was a spleen tumor, and 3** was a colon tumor. Liver or kidney were used as negative controls. Nulliparous, monoparous and weaning refer to the stage of mammary development at which the mice were at the time the tumors were detected. Markers are high molecular weight DNA markers (BRL).


Nulliparous
Mono parous
Weaning
1 2 3 4 5 Liver
1 2* Kidney
1 2 3** Kidney
kb
24
19
15
12
10

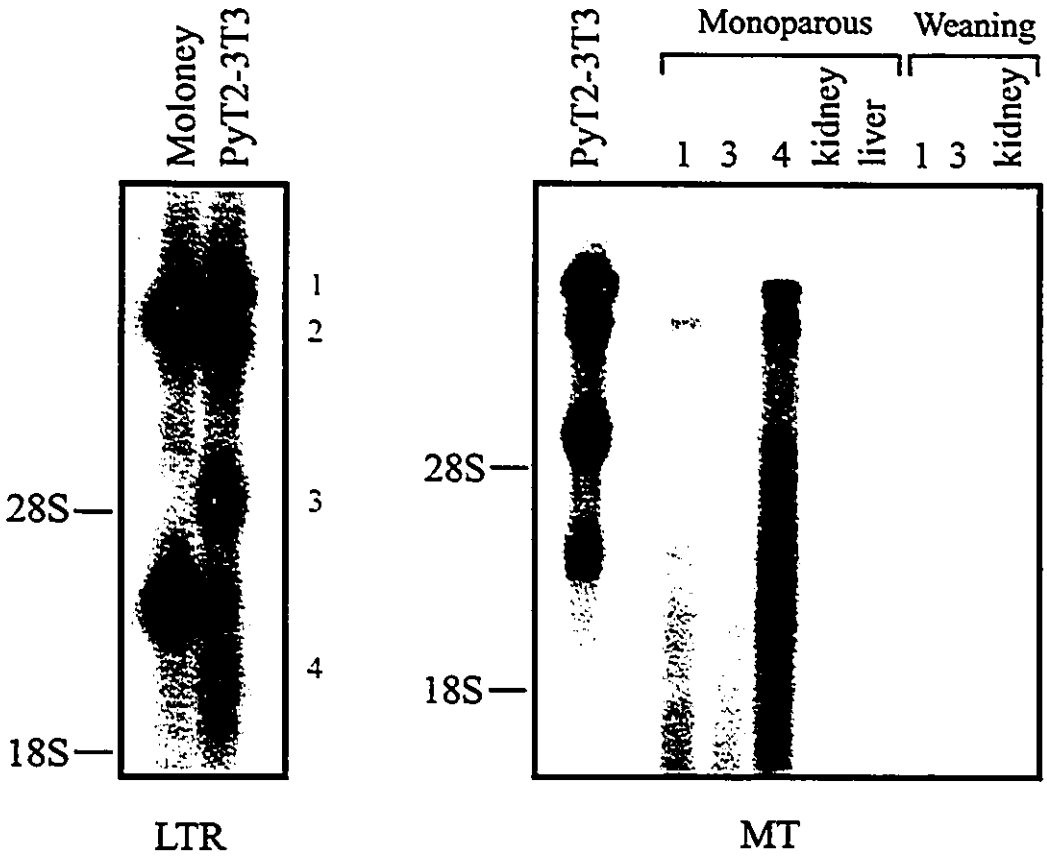
original cell by two viruses. The kidney was negative, as expected. The spleen tumor also showed the presence of the PyT provirus. There was the question of whether the tumor found on the spleen was a metastasis of the mammary tumor or a primary tumor from an independent infection. The fact that the spleen tumor had only one integrated virus and that the size of the EcoRI fragment did not match either one seen in the mammary tumor indicated that the spleen tumor originated from a separate event. This also indicated that the PyT virus was replicating in the mouse and made its way to the spleen possibly through the lymphatic system or through the blood circulation. For the third mouse (weaning) there was no detectable signal in any of the tumors.

Northern blot analysis

Most of the tumors were too small to be able to extract both DNA and RNA. Figure 3.17 shows the results of a Northern blot analysis on a number of tumors using total RNA. The quality of the RNA recovered from the tumors was poor and made conclusions difficult. The left panel shows transcripts from cells infected with wild type Moloney or recombinant PyT virus. When probing with the LTR, two mRNA were seen from the wild type virus as expected. In the PyT infected cells, four main transcripts were detected by the LTR. Bands labelled 1 and 3 represent the full length transcripts from an intact PyT provirus. These were also detected by the PyMT probe (lane 1 in the right panel). Bands 2 and 4 also were detected by the PyMT probe and appeared to be truncated versions of transcripts labelled 1 and 3, as they were both proportionally shorter. Also seen with the LTR probe were two transcripts of the same size as the wild type transcripts, that were not

Figure 3.17 Northern Blot Analysis of PyT-Induced Tumors

Northern blot analysis of tumors induced by the PyT virus in mice. The left panel shows a northern blot containing RNA extracted from NIH 3T3 cells infected with either a wild type MoMuLV (Moloney) or with the PyT virus from clone 2 (PyT2-3T3). The blot was probed with the LTR. The right panel shows a northern blot containing RNA from a number of PyT-induced tumors. The blot was probed with the Polyomavirus Middle T antigen sequence. The first lane contained RNA from NIH 3T3 cells infected with the PyT virus from the producer clone 2 (same as in left panel). Numbers at the top correspond to mammary tumors from the indicated animals. Liver and kidney served as negative controls. Monoparous and weaning refer to the stage of mammary development at which the mice were at the time the tumors were detected. The position of the 18S and 28S ribosomal RNA is indicated.


Moloney
PyT2-3T3
1
2
3
4
28S
18S
LTR
PyT2-3T3
Monoparous
Weaning
1 3 4 kidney liver
1 3 kidney
28S
18S
MT

detected by PyMT sequences (the smaller one, between bands 3 and 4, is more visible than the larger one, between bands 1 and 2). This is consistent with the Southern analysis which showed a loss of the PyMT sequence in a subpopulation of the infected cells (Figure 3.12B). Some of the expected (larger) proviral transcripts were present in the tumor samples probed with PyMT (right panel) but the quality of the RNA did not allow smaller or novel transcripts, that would be indicative of a promoter-trap, to be detected. Consistent with the Southern analysis, the tumors from the third mouse (weaning) were negative for the presence of any PyMT-positive transcript. It was not tested whether any transcript could be detected with an LTR probe.

Cloning of the integration site from a mammary tumor

An attempt at isolating genomic DNA flanking the integration site from the various tumors using the PCR strategy described earlier was unsuccessful. An alternative PCR approach that should have overcome the limitations of the reverse PCR strategy was also tried. The new method, published by Sorenson and coworkers (Sorensen *et al.* 1993), eliminated the need to digest and ligate the DNA prior to amplification. Unfortunately, that method did not work out either. Instead, a more conventional approach was used to clone the integration site. A genomic library was constructed from tumor DNA and the fragment containing the provirus was isolated by hybridization to the PyMT sequence. Because each tumor would need its own library it was not feasible to work with all the tumors at once. The tumor which formed in the pregnant mouse was selected for a couple of reasons. First, we had no convincing evidence that any of the tumors from the third mouse were positive

for the provirus. In addition, the amount of DNA necessary to build the library was a limiting factor for the mammary tumors of that mouse. Second, according to our biological screening strategy, the tumors found in the nulliparous mouse would be the result of integrations in "housekeeping genes". For the same reasons, the mammary tumors from the pregnant mouse was more interesting as we suspected it to be the result of an integration in a gene that was newly activated when mammary development was induced.

The DNA was digested with EcoRI so that the two genomic fragments corresponding to the two integration sites (in Figure 3.16) could be isolated. Knowing that the size of these fragments was in the range of 12-20kb the library was enriched for larger size fragments using a potassium acetate gradient. Four clones were originally isolated and two of them were considered as candidates for the two proviral integration sites based on mapping study (Figure 3.18). The size of clones 1 and 4 were compatible with the size of the EcoRI proviral fragments detected by Southern blot (Figure 3.16). For further analysis, the genomic sequences flanking both sides of each provirus were subcloned in the Bluescript vector using EcoRI and KpnI sites.

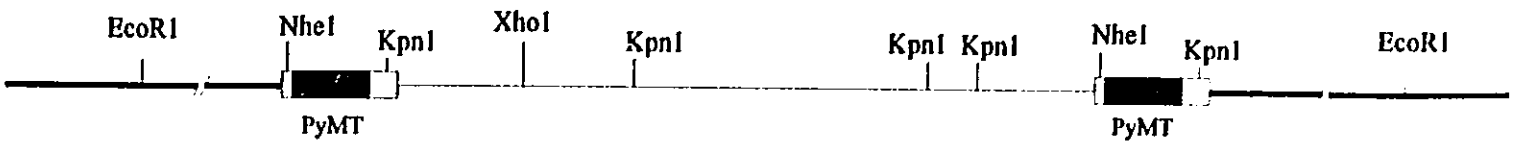
Southern blot analysis of genomic flanking sequences

To confirm that the two isolated genomic clones corresponded to the two bands seen in the Southern blot shown in Figure 3.16, the genomic sequence from each clones was used to reprobe that same Southern. Each 5' fragment identified one of the two bands first detected with PyMT (data not shown). To see whether these genomic sequences were part of a single copy gene as opposed to repetitive sequences, they were used to probe a Southern

Figure 3.18 Map of genomic clones corresponding to two proviral integration sites from one PyT-Induced tumor.

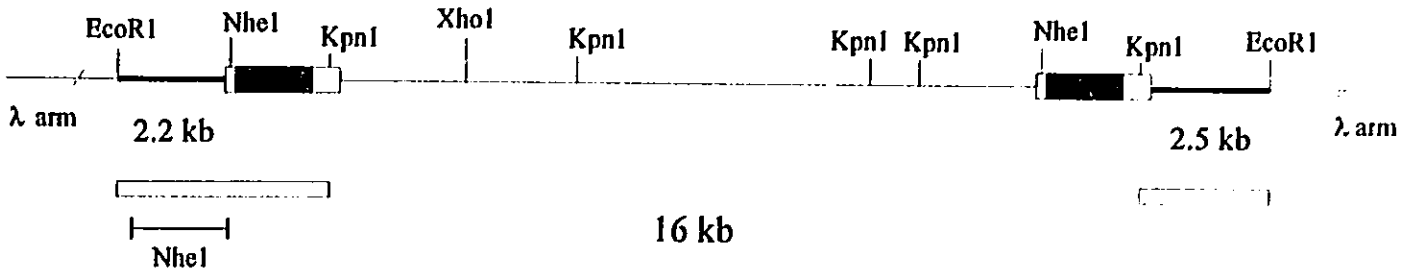
At the top is depicted the genomic structure of the PyT provirus. The thick lines represent cellular sequences and the thin lines represent proviral sequences. The open boxes represent the LTRs and the shaded boxes represent the PyMT sequences. Restriction sites used in the mapping or subcloning of the flanking genomic sequences are indicated. The distance of the EcoRI sites from the LTRs will differ for each provirus. λ DAHII clones 1 and 4 show the structure of two PyT proviruses found in the mammary tumor from the monoparous mouse. The thin broken lines represent lamda phage sequences. The thick lines represent genomic DNA. 16kb and 15kb indicate the size of the genomic inserts. The size of the flanking genomic sequences is indicated. The junction fragments were subcloned into a plasmid vector as EcoRI to KpnI fragments (indicated by light shaded boxes). The NheI fragment shown below the 5'-genomic sequence of clone 1 was used as a probe for Southern and northern blot analysis, as well as for the screening of a cDNA library.

genomic

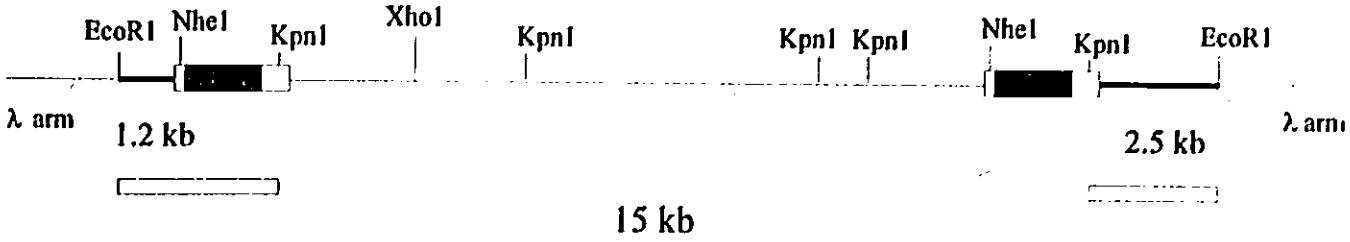


λ DASH II

1



4



SI-A

blot containing DNA from NIH 3T3 cells digested with three different enzymes. Also included on one of the blots was genomic DNA from the tumor digested with EcoRI (Figure 3.19A). Figure 3.19 shows that both clones gave patterns that are consistent with single copy gene. As expected, clone1 identified an extra band in tumor DNA when compared to NIH 3T3 DNA (EcoRI digest), corresponding to the disrupted allele.

Northern blot analysis of genomic flanking sequences

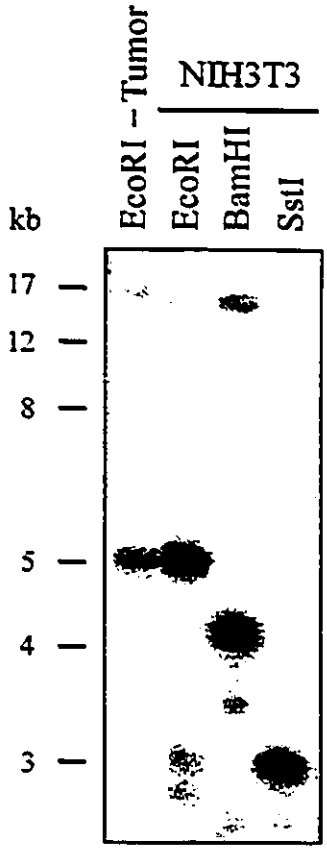
Northern blot analysis was used to see whether expression of the trapped gene was specific to the mammary gland. With the hope that the isolated genomic sequences from the genomic clones 1 and 4 could detect a transcript, they were used to probe a northern blot containing RNA from mammary tissue from a virgin and a pregnant mouse. A first attempt with total RNA was negative for both clones (data not shown). Using polyA$^+$ selected RNA clone 1 detected a band corresponding to a transcript of approximately 4.5kb. The intensity of the signal however was equivalent in the virgin and the pregnant mouse mammary RNA samples, indicating that expression was not restricted to, nor enhanced during mammary development. Clone 4 did not detect any transcript with polyA$^+$ RNA. Clone 1 was then selected as the putative promoter-trap and was named TRAP1.

To study the pattern of expression of TRAP1 in more detail, northern blot analysis was repeated with a wider variety of tissues (Figure 3.20). The gene was expressed at comparable levels in all tissue tested with the exception of testes which showed higher expression. Testes also showed an extra band in the range of 1-2kb but this was not investigated further.

Figure 3.19 Southern Blot Analysis of NIH 3T3 DNA using genomic flanking sequences as probes

Southern blots containing DNA from NIH 3T3 cells, probed with the 5' flanking sequence of clone 1 (A) and clone 4 (B). The DNA was digested with the indicated enzymes. In addition to NIH 3T3 cell DNA, panel A contained DNA from the mammary tumor digested with EcoRI. Markers were the high molecular weight DNA markers (BRL) in panel A, and the 1kb ladder (BRL) in panels A and B.

A


EcoRI – Tumor
NIH3T3
EcoRI
BamHI
SstI
kb
17
12
8
5
4
3

B

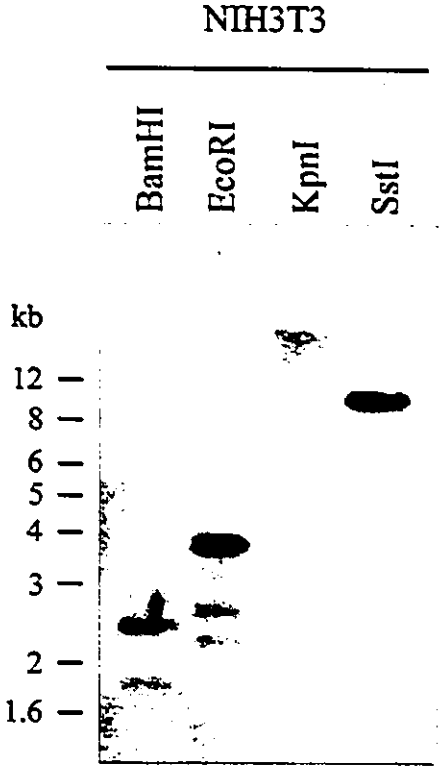
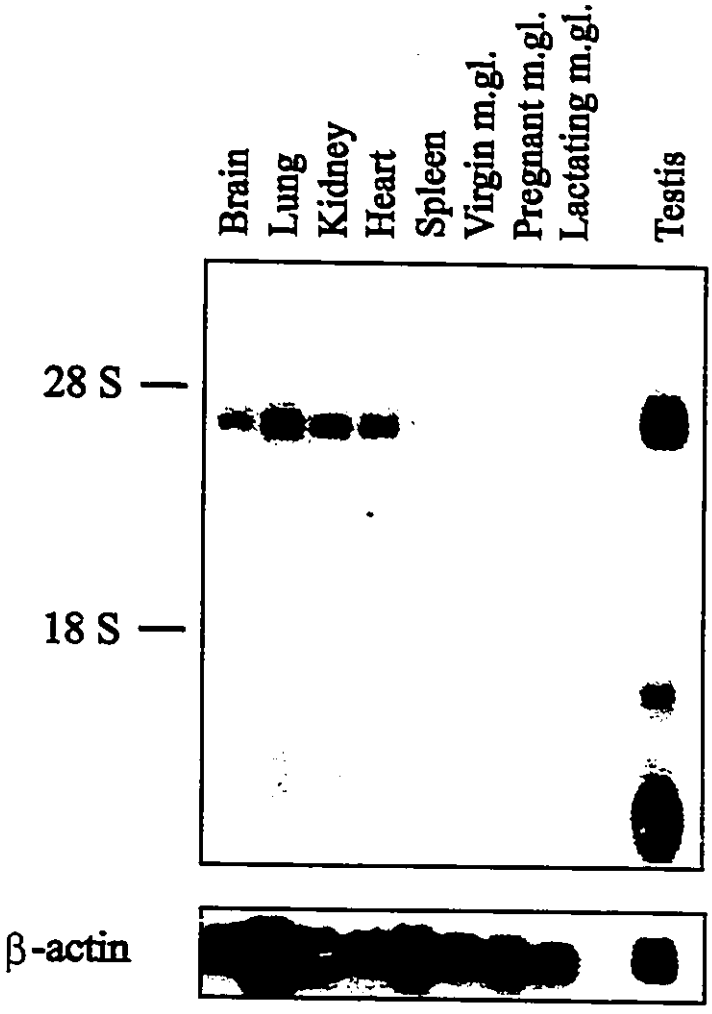

NIH3T3
BamHI
EcoRI
KpnI
SstI
kb
12
8
6
5
4
3
2
1.6

Figure 3.20 Northern Blot Analysis of TRAP1 in Adult Mouse Tissues

Northern blot analysis of TRAP1 using 3μg of polyA $^+$ selected RNA from the indicated adult mouse tissues. The blot was probed with the 5' flanking sequence from clone 1 (TRAP1) (NheI fragment shown in Figure 3.18). The position of the 18S and 28S ribosomal RNA is indicated. β-actin was used as a loading control.


Brain
Lung
Kidney
Heart
Spleen
Virgin m.gl.
Pregnant m.gl.
Lactating m.gl.
Testis
28 S
18 S
β-actin

Sequence analysis of genomic DNA

In order to determine whether TRAP1 was a known gene, partial sequence was obtained from the ends of the genomic clones. In the 2.2kb fragment flanking the 5' LTR, clone1 had a B1-like element in the first 200bp next to the EcoRI site. This, as expected, showed homology to a great number of genes containing B1 or B1-like repetitive elements when compared to the Genbank database. That end of the fragment contained an NheI restriction site about 200bp into the sequence so that what was first believed to be an EcoRI-NheI fragment was in fact an NheI fragment that excluded the B1 like element. The last 220bp at the 3' end of the same genomic fragment (next to the LTR) were more informative (Figure 3.21B). That sequence showed a high degree of identity (85%) to a human partial cDNA (EST22025), one of 600 expressed sequence tags isolated from the human brain (Adams *et al.* 1991) (Figure 3.21C). This is in agreement with our clone being expressed in the brain. Together with the northern blot analysis the sequence data suggested that the PyT provirus in clone 1 integrated in the transcribed region of a novel murine gene.

Sequence data from clone 4 did not revealed significant homology to any known sequence. The evidence was quite strong in favour of clone 1 being the promoter-trap.

Cloning of the TRAP1 cDNA

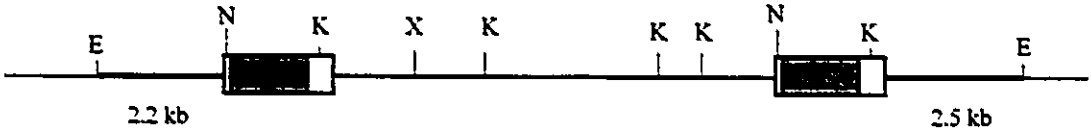
Because TRAP1 was not exclusively expressed in the breast but rather appeared to be expressed in a wide range of tissues, an available mouse lung cDNA library (Stratagene) was used to clone the cDNA of TRAP1. The same 5' flanking genomic fragment (NheI insert) used for Southern and norther blot analysis was used as the probe. Three clones of 600, 800

Figure 3.21 Sequence Homology of the TRAP1 Genomic Clone

A. Schematic of the TRAP1 genomic insert clone. Thin lines represent plasmid sequences, thick lines represent genomic DNA, and the intermediate size lines represent proviral sequences. Open boxes represent LTR sequences and the shaded boxes represent the Polyomavirus Middle T antigen sequence. The size of the 5' and 3' genomic sequences are indicated.
B. Sequence from the 5' junction fragment. Capital letters represent genomic sequence and the small letters represent the LTR sequence.
C. Sequence homology between the TRAP1 genomic sequence and the human expressed sequence tag EST02205. The identity is 85%.

A

TRAP1 genomic clone



B

```
CCACTCAATCTCACTCATCCAGCCATCGTTGACCACACAGACTTTTTTTT  50
TTCAACGACCCCCTCTCCCTAAATGGAGTTGGGCGGGGGGGAAATGAAAA 100
CGAGTTGGTCTTTACTTTTTTAAAAGAGACTTTTTGATCCAATGAGAGCC 150
CCACAAGTAACTGAGTTTGGGTCCTGGTTGGTTGCTTTATTTTTTCCCT  200
CCGGGGAAAAGCCCCTCCCCCtgaaagaccccacctgtaggt...       221
```

C

```
 60 CCCCTCTCCCTAAATGGAGTTGGGCGGGGGGGAAATGAAAAC-GAGTTGGTCTTTACTTT 118  TRAP1
    |||| |||||||||||||||||| |||||||||||||| || ||||||| | ||  |||
144 CCCCCCTCCCTAAATGGAGTTGGGTGGGGGGGAAATGAATACTGAGTTGGCC-TTTATTT 202  EST02205

119 TTTAAAAGAGACTTTTTGATCCAATGAGAGCCCCACAAGTAACTGAG-TTTGGGTCCTGG 177  TRAP1
    |||||||||  ||||||||||||||||  ||||  || ||| |||| ||||||||||||
203 TTTAAAAGA--CTTTTTGATCCAATGAGGCCCCCTAAA-TAATTGAGTTTTGGGTCCTGG 259  EST02205

178 TTGGTTGCTTTATTTTTTT 196  TRAP1
    ||||||  |||  |||| |
260 TTGGTTTGTTTTATTTTGT 278  EST02205
```

and 900bp were isolated and sequences in both directions. The two smaller clones were internal to, and completely homologous to the larger one. The largest clone was used as a probe to obtain more clones. The second round of screening identified relatively short cDNA inserts (0.9- 1.6kb) and a different approach was taken to obtain the full length TRAP1 cDNA. Having a partial sequence of the cDNA it was possible to use a RACE strategy to obtain the 5' and 3' portions of the TRAP1 cDNA independently. By selecting TRAP1 specific primers so that they amplify an overlapping portion of the cDNA, the two ends of the cDNA can be joined to generate the complete cDNA. cDNA was made from lung polyA+ selected RNA (reagents from Clonetech) and adaptors optimised for RACE (Clontech) were ligated on. Each end of the cDNA were amplified in separate reactions using one TRAP1 specific oligonucleotide and one oligonucleotide from the linker (Clontech). The 5' product gave a fragment of 900bp. Because the TRAP1-specific oligonucleotide was positioned 750bp away from the start of the partial cDNA, a 900bp fragment indicated that the 5' end of the partial clone was 150bp away from the start of the transcript. Knowing this and the fact that the size of the TRAP1 transcript is 4.5kb, it was deduced that a 4.2kb fragment should be generated from the 3' RACE reaction. Unfortunately, every attempt at amplifying the 3' end of the cDNA failed to give the predicted size band.

Sequence of the partial TRAP1 cDNA

The 900bp sequence compiled from the three initial cDNA clones (from the Stratagene lung cDNA library) was compared to the Genbank database. In addition to the previous

expressed sequence tag (EST22025) identified with the genomic sequence, the TRAP1 cDNA showed a high degree of identity (80-90%) to a number of more recently isolated human ESTs. Again, no murine sequence had any significant homology to TRAP1.

The TRAP1 sequence was translated to search for the presence of an open reading frame (Figure 3.22). An ORF started at position 337 of the partial cDNA and extended to the 3' end. The putative initiating methionine was positioned in a perfect Kozak consensus sequence (Kozak 1987). The fact that the 5' RACE reaction identified only an extra 150bp upstream of the partial cDNA reinforces the possibility that the AUG codon at position 337 corresponds to the initiating methionine. A database search (Nblast from NCBI) did not reveal any significant homology to any known protein. The best homology found was to the general transcription factor TFIIS but the homology (46% similarity over 97bp) was not located in any of the known functional domains of the TFIIS factor and was not believed to be pertinent.

A comparison of the genomic and the cDNA sequence revealed that the provirus integrated 47bp upstream of the putative initiating methionine (Figure 3.22).

Promoter activity in the 5' genomic fragment

Once it was realized that the provirus integrated relatively close to the start of transcription, the question arose as to whether the promoter for TRAP1 was included in the 2.2kb of 5' upstream genomic sequence. To answer that question, the 2kb NheI fragment was subcloned in front of the bacterial CAT gene in the pCAT-Enhancer vector (TRAP-

Figure 3.22 Nucleotide and Deduced Amino Acid Sequence of the Partial TRAP1 cDNA Clone

Nucleotide numbers are shown on the left and amino acid numbers are shown on the right. Five stop codons upstream of the putative initiating methionine (position 337) are shown in bold. The site of integration of the PyT provirus is indicated by the black triangle.

```
  1  GGC CGA AGG ACC GTC ACC ACA TAA CTT CAA AAA TAA TCA ACC ACC TTC CCT GCC
 55  CAA ACC TAC CCA ATC TCA CTC ATC CAG CAT TTC CGT TTT GGA ATC CAC TCA GAC
109  TTT TTT TTT CAA CGA CCC CCT CTC CCT AAA TGG AGT TGG GCG GGG GGG AAA TGA
163  AAA CTG AGT TGG TCT TTA CTT TTT TTA AAA GAG ACT TTT TGA TCC AAT GAG GCC
217  CCA CAA GTA ACT GAG TTT GGG TCC TGG TTG GTT GCT TTA TTT TTT TTC CTC CAA
                                 ▼
271  AAT TTT ACC CCT CCC CCC TGA GCC CGA GGT GCT GAC GTC GCA AAA AAA TTG GAT
325  AAA GCC ACC ATC ATG GGT TCA GGT CCC ATA GAC CCC AAA GAA CTG CTC AAG GGC
                   M   G   S   G   P   I   D   P   K   E   L   L   K   G   14
379  CTG GAT AGC TTC CTT ACC CGG GAC GGA GAA GTG AAG AGT GTG GAT GGA ATT TCC
      L   D   S   F   L   T   R   D   G   E   V   K   S   V   D   G   I   S  32
433  AAG ATC TTC AGT CTA ATG AAG GAG GCA CGA AAG ATG GTG AGT CGG TGC ACG TAC
      K   I   F   S   L   M   K   E   A   R   K   M   V   S   R   C   T   Y  50
487  TTG AAC ATA ATC CTG CAG ACC CGC GCT CCA GAA GTG CTC GTC AAG TTT ATT GAT
      L   N   I   I   L   Q   T   R   A   P   E   V   L   V   K   F   I   D  68
541  GTC GGT GGC TAC AAG CTT CTG AAC AAC TGG CTA ACA TAT TCA AAG ACC ACC AAC
      V   G   G   Y   K   L   L   N   N   W   L   T   Y   S   K   T   T   N  86
595  AAC ATT CCT CTC CTA CAG CAG ATT CTG CTG ACT CTG CAG CAC CTC CCG CTC ACT
      N   I   P   L   L   Q   Q   I   L   L   T   L   Q   H   L   P   L   T 104
649  GTG GAC CAT CTC AAG CAG AAC AAC ACA GCA AAA CTA GTG AAG CAG CTC AGC AAG
      V   D   H   L   K   Q   N   N   T   A   K   L   V   K   Q   L   S   K 122
703  TCA AGT GAG GAT GAA GAG CTC CGG AAA TTG GCA TCA GTC CTT GTC AGT GAC TGG
      S   S   E   D   E   E   L   R   K   L   A   S   V   L   V   S   D   W 140
757  ATG GCT GTC ATC CGC TCC CAG AGT AGC ACC CAG CCT GCA GAG AAA GAT AAG AAG
      M   A   V   I   R   S   Q   S   S   T   Q   P   A   E   K   D   K   K 158
811  AAA AGG AAG GAA GAG GGG AAA AGT CGA ACC ACG CTT CCT GAG CGG CCT TTG ACT
      K   R   K   E   E   G   K   S   R   T   T   L   P   E   R   P   L   T 176
865  GAA GTG AAG GCT GAG ACC CGG GCT GAG GAA GCC CGA GAG AAG AAG AAA GAA AAG
      E   V   K   A   E   T   R   A   E   E   A   R   E   K   K   K   E   K 194
919  CG
```

CAT, Figure 3.23). This vector contains the SV40 enhancer which may help to detect expression form weaker promoters. The insert was subcloned into the XbaI site in either orientation to use the opposite orientation as a negative control. As a positive control a vector containing CAT under the control of the MoMuLV-LTR was used (gift of Ninan Abraham, Dr. Bell's laboratory). Each of the three vectors, as well as the original pCAT vector, were cotransfected transiently with a β-galactosidase expressing plasmid (pgk-β-gal; McBurney *et al.* 1991) in either the human 293 or mouse P19 cells. These cell lines were chosen for their high efficiency of transient transfection (McBurney 1993). The cells were harvested 24 hours post transfection and protein extracts were assayed for CAT and β-galactosidase activity. Results from 293 cells are shown in Figure 3.24. P19 cells gave the same results. β-galactosidase activity was used to compare the efficiency of transfection. A difference of no more than two-fold was observed between the transfections. The genomic fragment in the same orientation as the provirus transcription could drive high expression of CAT (76% converted substrate) whereas the opposite orientation, or the pCAT-Enhancer vector alone could not. This provided very strong evidence that the tumor which arose in the pregnant mouse was caused by a promoter-trap event and that clone 1 represented the gene where the virus integrated.

Figure 3.23 TRAP-CAT Vector

The TRAP-CAT vector contains the ampicillin resistance gene, the TRAP1 5' genomic sequence in front of the chloramphenicol acetyl transferase (CAT) gene, and the SV40 enhancer. The TRAP1 sequence (2kb NheI fragment) was inserted in the XbaI site of the pCAT-Enhancer vector (Promega). Clones containing the TRAP1 sequence in either orientation were isolated.

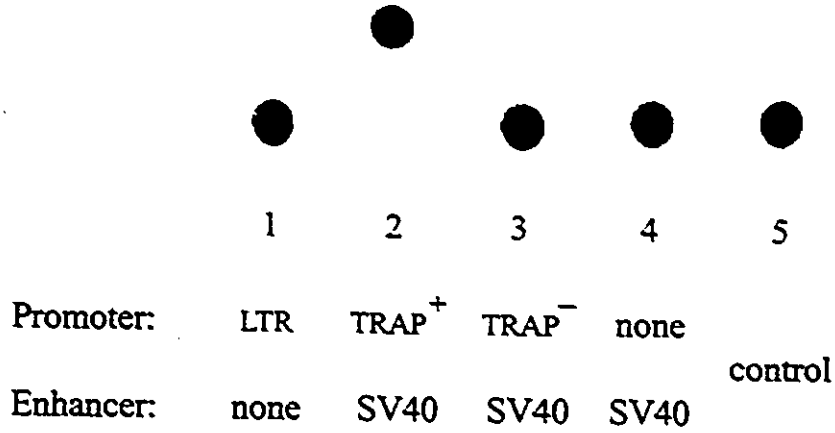
Amp r
EcoRI
SV40
Enhancer
HindIII
PstI
SalI
AccI
(XbaI/NheI)
TRAP-CAT
SV40
5'TRAP1
CAT
(XbaI/NheI)
AUG
EcoRI

Figure 3.24 TRAP1 Promoter Assay

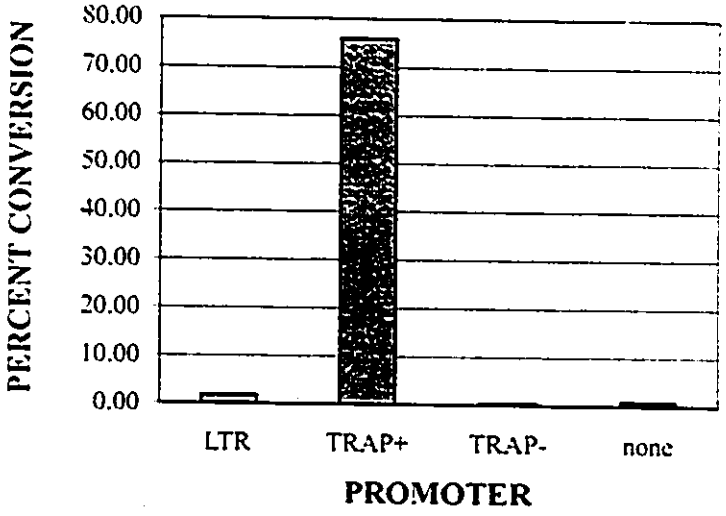
A. Example of a typical CAT assay using the TRAP-CAT vector in a transient transfection of human 293 cells. The lower spot represents unconverted chloramphenicol substrate and the upper spot represents mono-acetylated chloramphenicol (see materials and methods for details of the assay). The vector used in lane 1 had the CAT gene expressed from the MoMuLV LTR (no SV40 enhancer). This served as a positive control. Lanes 2 and 3 show the CAT activity from TRAP-CAT vectors with the TRAP1 sequence in the positive or negative orientation respectively. Lane 3 shows the CAT activity from the pCAT enhancer vector. Lane 5 was a negative control for the CAT assay and contained no protein in the reaction.

B. CAT activity from above, expressed as percent of converted substrate. Samples are the same as in A. This represents the average of three experiments in human 293 cells.

A


1
2
3
4
5
Promoter:
LTR
TRAP+
TRAP−
none
control
Enhancer:
none
SV40
SV40
SV40

B


PERCENT CONVERSION
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
LTR
TRAP+
TRAP-
none
PROMOTER

CHAPTER 4

DISCUSSION

A promoter-trap retrovirus was used to try to identify novel genes that are specifically expressed during murine mammary differentiation. A promoterless oncogene was inserted in the LTR of a MoMuLV vector so that upon integration of the provirus into the host genome the reporter gene may be expressed from a cellular promoter. Two viruses were constructed, the first one being replication deficient (COPT) and the second one being replication competent (PyT). Each virus was used to infect murine mammary glands by injection of the virus *in situ* through the nipple. An oncogene was used as the promoterless reporter gene in order to be able to detect expression in the animal, where activation of the oncogene should form a tumor. Doing the experiment *in vivo* provided a natural screen for genes whose expression is activated during differentiation.

The replication deficient retrovirus (COPT) was not successful at trapping any gene in the mouse mammary gland but the second virus (PyT) did trap at least one novel murine gene.

A) The COPT promoter-trap retrovirus

The COPT virus

The COPT promoter-trap retrovirus had a promoterless v-Ha-*ras* oncogene inserted in the MoMuLV LTR, and the c-*myc* oncogene driven from the ubiquitous pgk promoter. The *myc* oncogene should be expressed at all times whereas the *ras* gene will be expressed only if the provirus integrated near an active cellular promoter.

Being replication deficient, the COPT virus was produced in the GP+E86 packaging cell line which provided the viral gene products necessary for the assembly of infectious virions. The titer of the producer clones could not be established by conventional drug selection due to the lack of a selectable marker in the retrovirus. Instead, titers were estimated in an indirect fashion using Southern blot analysis. This method was judged adequate based on the demonstration that a two-fold difference in titer could be detected by Southern blot using a very similar virus of known titer. Other methods for estimating the titer of replication deficient retroviruses lacking selectable markers include the assay for reverse transcriptase activity in the virus-containing supernatant (Goff *et al.* 1981), or slot blot analysis of virion RNA (Huszar *et al.* 1989). In our hands, a two to five-fold dilution of the virus could not be reflected reproducibly by slot blot analysis. Our Southern blot approach had the advantage of assessing for infectious particles only and had the convenience of verifying the structure of the provirus at the same time. The fact that the COPT virus had the expected structure was also confirmed by northern blot analysis (Figure 3.7, last lane) where predicted size transcripts from the LTR (4.2kb) and pgk (2.7kb) promoters were detected.

Not every clone could be shown to produce COPT virus. In retrospect, these earlier infections were performed in such a way that, for most clones, the multiplicity of infection (m.o.i.) would have been well below one. The infections for Figure 3.3 were done with a multiplicity of infection closer to 2. This is supported by a ratio of proviral to endogenous ras signal close to one. Also, XbaI (Figure 3.3) was found to give a more complete digestion and therefore a better signal than SstI. At the time, clone 3 was found to produce the most virus and was used for testing the promoter-trap. The titer of clone 3 ($2x10^5$ cfu/ml) was found to be in the same range as those for the pUCRas clones and was believed to be representative of the other COPT producer clones. The titers obtained from the GP+E86 cell line, in the range of 10^5 cfu/ml, were consistent with those published using the same (Markowitz *et al.* 1988) or similar packaging cell lines (Cepko *et al.* 1984; Mann *et al.* 1996).

The COPT promoter-trap

The ability of the COPT retrovirus to work as a promoter-trap was tested using nude mice. This was believed to be a better test than the ability to form foci in soft agar because cells that are considered transformed by this assay do not always form tumors in nude mice. It was also important to demonstrate that our indicator oncogene could form tumors in the mouse. COPT-infected NIH 3T3 cells produced three tumors after injection in mice, with a latency of two to three weeks, whereas uninfected cells did not produce any tumor during the same period of time. Knowing how many cells were infected at a multiplicity of infection near one, and assuming that the three tumors arose from a promoter-trap event, this

would represent one promoter-trap for every two million integrations. This is two to four orders of magnitude lower (depending on the reporter gene) than what has been reported for similar promoter-trap retroviruses (von Melchner and Ruley 1989; Chang *et al.* 1993). This however is not a fair estimate of the efficiency of the COPT virus to act as a promoter-trap because by killing the mouse to extract the tumors we discarded any other potential tumors which may have appeared later on. Promoter-trap events are usually scored by drug selection in culture where every potential positive is allow to develop a resistant colony. Even taking this into account, the efficiency of the promoter-trap remains low, which can be a reflection of the difference in sensitivity between drug selection in culture and tumor formation *in vivo*. Whereas every cell expressing a drug resistance gene should be resistant to that particular drug, not every cell expressing *myc* and *ras* will go on and form a tumor. This is illustrated in double transgenic mice expressing both *myc* and *ras*, each under the control of the MMTV promoter (Sinn *et al.* 1987). Tumors arose in a stochastic fashion indicating that other events are required in addition to *myc* and *ras* activation.

Southern blot analysis confirmed that the three tumors from the nude mouse came from COPT infected cells. All three tumors were shown to contain proviral DNA, although two of them had undergone rearrangement. It was also determined that each tumor contained a single retrovirus based on the restriction pattern of the proviruses using enzymes that would reveal the junction fragments. This was expected because the cells injected in the mouse were infected at a multiplicity of infection near one.

Tumor 2 was the only tumor with an apparent intact COPT provirus. Unfortunately, there was no RNA available from that tumor because of its small size. This tumor could not be

propagated in culture like the two other tumors, maybe because it required certain growth factors that were only available in the animal. It was therefore not possible to obtain any evidence of a promoter-trap event either by northern blot analysis or sequence data because sequences flanking the integration site could not be isolated by PCR amplification.

Tumor 3 did not show the presence of any additional transcript on a northern blot, which would be indicative of a promoter-trap. However, northern blot analysis using total RNA, may not be sensitive enough to detect expression from certain cellular promoters. It is possible that the rearrangement observed in that tumor somehow led to the expression of *ras* from either the pgk or the 5' LTR promoter. To address this question the nature of the rearrangement was investigated further. Using Southern blot analysis with different restriction enzymes, the small internal deletion was estimated to be approximately 500bp and localized near the 3' end of the *myc* sequence because the provirus was still positive for both pgk and *myc*. This was supported by the northern blot which showed that both proviral transcripts were shorter by about 500bp, which would be expected if the deletion is in a region common to both transcription units. Given that the two transcripts are of the expected size, taking into account the deletion, it appears unlikely that the rearrangement caused some kind of novel splicing event which would allow *ras* to be expressed from one of the two internal promoters. Alternatively, the deletion could have fused the *myc* and the *ras* sequences together, but this would require that the reading frame be conserved and that *ras* can work as a fusion protein with *Myc*. The possibility remains that the tumor was caused by an insertion event and not by the expression of *ras* from a cellular promoter. The amount of *ras* protein was never investigated in any of the tumors. Amplification and sequencing

of flanking genomic DNA would have helped clarify the question, but no sequence data could be obtained from this tumor.

Tumor 1 was determined to have lost most of the provirus because it was negative for the presence of both pgk and myc sequences. It may either have lost all of the provirus but one LTR, or have lost the middle part of the provirus. It was difficult to decipher between these two possibilities using Southern blot analysis because the sequence of the two LTRs is identical. If both LTR's were still present, it raised the possibility that the 3' *ras* was expressed from the 5' LTR. The northern analysis helped clarify the question. As seen in Figure 3.7, tumor 1 showed some level of expression of the two expected proviral transcripts which are seen in the producer clone (last lane), as well as two transcripts that are proportionally shorter. This would indicate that the deletion (which may have taken place in two steps) occurred after the initiation of the tumor because at some point tumor cells were producing transcripts from a full length provirus. This would argue that a promoter-trap was the cause of the tumor and not the rearrangement. These transcripts are present at even lower levels in the transplanted tumor and in the cell line from tumor 1 (lanes 3 and 4 in Figure 3.7). The loss of part of the provirus may have somehow provided a growth advantage to the cell, and these cells may have outgrown the original ones in the tumor. As the tumor was passaged, either in the mouse or in culture, fewer and fewer original cells remained which would explain the lower levels of these transcripts in the transplanted tumor and the derived cell line. If this were true it would be expected that as the cell line was passaged further these transcripts would eventually disappear. One may speculate that the loss of part of the provirus brought the 3'LTR closer to the trapped cellular promoter,

providing a second enhancer for the cellular promoter. This could be the growth advantage of the tumor cells containing the deletion. Alternatively, an additional genetic alteration may have happened in a tumor cell following the rearrangement of the provirus, giving that cell a growth advantage. The size of the additional transcript seen in tumor 1 and its derivatives (1.6kb) is consistent with a fusion transcript initiating 600bp upstream of the start of the provirus, and terminating in the 5' LTR. This transcript could be responsible for *ras* expression. The failure to amplify the genomic sequence upstream of the integration site could not confirm this, but the amplification of the 3' LTR added some weight to the internal deletion theory. The fact that an amplified fragment of approximately 1kb contained sequences from both LTRs (this is known because of the associated sequences 5' or 3' of the LTR sequence) as well as some sequence downstream of the 5' LTR, suggested that the two LTRs were intact and had been brought closer together by a large deletion. From the size of the amplified fragment the distance between the LTRs was estimated at 600bp. If the two LTRs are indeed intact and separated by 600bp, an XbaI digestion of the provirus would predict a restriction fragment of about 1.8kb, which is what was observed in Figure 3.5. In addition, every restriction digests that cut in each LTR gave the same size band. The estimated size of a transcript extending from one LTR to the other would be about 1.8kb. This is slightly larger than the size of the transcript seen in Figure 3.7 (1.6kb) which would argue in favour of a the latter being a fusion transcript. This could be resolved by probing the northern blot (Figure 3.7) with either gag or env sequences. The prediction would be that if *ras* is driven off the 5'LTR, gag and env sequences would be present in the transcript, whereas they would not if *ras* was expressed from a cellular promoter.

The rearrangements observed in two of the tumors raised a concern as to the stability of the COPT provirus. In general, recombinant proviruses are very stable (Smith *et al.* 1991). In this *in vitro* test system, the loss of any part of the provirus except the 5' LTR should not affect tumor growth, but in a primary cell where *myc* expression needs to be maintained as well as *ras* expression, any tumor losing a functional *myc* would no longer grow. Therefore, the fact that *myc* is dispensable for the growth of the tumor in this assay may have allowed these two rare events to be scored whereas they would not have been in a primary cell. The number of integrations analysed was too small to be able to extrapolate these findings to all COPT proviruses.

The PCR amplification did not work as well as anticipated. One of the reasons might be that the limited number of enzymes that can be used may decrease the odds of getting a fragment that is of a suitable length for amplification. One needs enzymes that will cut at a reasonable frequency in genomic DNA and at convenient sites in the LTR. If the restriction site is too close to the provirus the fragment may be too small to religate efficiently on its own. Alternatively, a restriction site located too far from the provirus may hamper the amplification. The two enzymes available to use, SmaI and ThaI, may not have been optimal for this application due to their CpG content.

The COPT virus as an in vivo promoter-trap

Although infection of the mammary glands with the BAG retrovirus showed that the virus could reach the target cells, the COPT retrovirus was not successful at trapping promoters *in vivo*. No tumors were detected over a six months period that included a pregnancy.

Although the combination of *myc* and *ras* is capable of transforming primary cells in the animal, the frequency at which it does so is too low to make the COPT virus an efficient promoter-trap *in vivo*.

Conclusion

The COPT promoter-trap retroviral vector produced infectious virus with the predicted structure. Southern blot analysis showed that the *ras* sequence was stable in the provirus, both in tissue culture and in the tumors. Although rearrangements of the provirus were observed in the tumors, they did not involve the *ras* sequence but rather internal sequences. The GP+E86 packaging cell line produced COPT virus with the expected titers ($2x10^5$ cfu/ml). Although it was not formally proven that the three tumors were caused by a promoter-trap event, there was strong evidence, at least from tumor 1, that the COPT virus was able to act as a promoter-trap *in vitro*. The efficiency was somewhat lower than that reported with other promoter-trap retroviruses containing a promoterless gene in the LTR. This was most likely due to the low sensitivity of the reporter gene *in vivo*. The formation of a tumor by oncogenes is more demanding on a cell than conferring resistance to a cytotoxic drug in culture. The former depends on a number of events whereas the latter relies on the expression of one gene. This most likely explains why no tumors were observed in Balb/c mice following injection of the virus in the mammary glands. For the combination of the *ras* and *myc* oncogenes to work as a promoter-trap in the animal, a much larger number of target cells would need to be infected, numbers that could not be attained with the current titer of COPT producing clones.

B) The Pyt promoter-trap retrovirus

The PyT virus

The PyT promoter-trap retrovirus utilised the Polyomavirus Middle T antigen (PyMT) as the reporter gene. This oncogene was felt to be a more sensitive reporter gene for an *in vivo* promoter-trap because it was shown to be a very strong transforming gene in primary mammary epithelium using transgenic mice. It has a shorter latency than the *myc/ras* combination of oncogenes and a higher efficiency. The great advantage of this oncogene is that it can transform primary mammary epithelium efficiently on its own which meant that the virus could be kept replication competent, increasing the level of infection in the mammary gland. It also meant that because the promoter-trap did not rely on any other part of the provirus, it should not be as affected by proviral rearrangements, as was the COPT virus.

The PyT retrovirus contained the PyMT coding sequence (1.4kb) in the LTR of a MoMuLV. Larger recombinant sequences have been introduced in the LTR of MoMuLV vectors without affecting the titers. Because the virus was replication competent it did not require a packaging cell line and was produced in NIH 3T3 fibroblasts. The titers obtained from the various producer clones (10^6 cfu/ml) were within the range of what has been reported for replication competent retroviruses (Reik *et al.* 1985). It was verified that the virus produced had the expected structure using Southern blot analysis of PyT infected NIH 3T3 fibroblasts. This lead to the observation that although the predicted virus was produced, it was not completely stable and the PyMT sequence was lost at a frequencies which varied

among the different producer clones. This was investigated further by following the stability of the virus through serial passaging. The observation was that most of the virus had deleted part of its genome after three passages, and by the fifth passage no intact virus could be detected. The exact nature of the rearrangement is not known but it was consistent between the different clones, in that the size of the truncated 3' LTR was always the same. The fact that the new 2.3kb band corresponds to the size of a wild type LTR and was not detected by the PyMT sequence suggested that homologous recombination happened between a PyMT-containing and a wild type LTR. The only place where both recombinant and wild type LTRs are present is in the producer cells where the vector has a wild type 5' LTR. It is therefore possible that the LTRs have undergone recombination either in the preintegrated state of the vector during transfection, or in the integrated form. As was presented in Figure 3.12, different clones produce different ratios of recombinant and wild type virus. The proportion of intact recombinant virus generated by a given clone (for example clone 2) appeared to be reproducible. This would suggest that the recombination took place during transfection and that the integrated recombinant vector is stable. Northern blot analysis of different producer clones suggested that this may be the case. As was seen in Figure 3.17 for NIH 3T3 cells infected with PyT virus from clone 2, six different transcripts were generated. In addition to the two expected recombinant transcripts, two mRNA species were seen that comigrated with wild type MoMuLV transcripts, and were not detected by the PyMT sequence. The two other transcripts were detected by both LTR and PyMT sequences but were shorter than the wild type transcripts. The origin of these is unknown. They could be the result of a splicing event or premature termination of transcription in the PyMT

sequence. The implication of recombination between LTRs in the producer clones would be that a mixture of wild type and recombinant virus would be produced and one may expect the wild type virus to have a growth advantage over the recombinant virus so that as the virus is passaged the recombinant virus would eventually be diluted out. This is exactly what was observed during the passaging of the virus (Figure 3.14). The implication *in vivo* would be that some cells would be infected with wild type virus and some others with the PyT virus. Yet another subset of cells would be infected with virions containing one wild type and one recombinant RNA molecule. The type of virus produced by these cells would depend on which of the two RNA species was used to generate the U3 region of the provirus (see Figure1.1). In the end, both types of viruses will be produced by infected mammary cells, thereby diluting the titer of the recombinant virus. To avoid such homologous recombination Stuhlmann and coworkers (Stuhlmann *et al.* 1989) obtained their virus from transiently transfected NIH 3T3 cells (48 hours). Clone 2 was chosen as the source of PyT virus because it produces the most intact recombinant virus.

The in vivo promoter-trap

Out of 19 surviving PyT-injected mice, three eventually developed tumors. It was difficult to know what kind of latency to expect for tumors to appear. Out of five strains of MMTVMT mice generated in the laboratory of Dr. Muller (Guy *et al.* 1992), two developed tumors around the fifth week of age without any pregnancy. The three other strains required at least two pregnancies before they developed tumors (average of 135 days). In the first case, the hormonal control of the MMTV promoter seemed to have been overridden by

adjacent cellular sequences. This may give us an idea of the latency to expect from "housekeeping" gene promoters. Interestingly, the first tumor appeared after six weeks. At that time the rest of the animals were mated thinking that most active promoters would have had their effect by then. Two mice developed tumors three and seven weeks later, the first one at the end of a pregnancy and the second one a week after weaning its litter. Idealistically, these would be the result of the activation of genes expressed during pregnancy and lactation respectively. Because the virus is replication competent one cannot exclude the possibility that either or both of these tumors may have been the result of the trapping of a "housekeeping" gene during a later infection. This could only be determined from the expression pattern of these genes in the developing breast. By waiting for a certain period of time before activating the differentiation specific genes, one can screen out trapped "housekeeping" promoters from the few initial rounds of infection. It is practically impossible to estimate the efficiency of the PyT promoter-trap, first because the virus is replicating and one doesn't know how many cells eventually get infected with how many viruses, and secondly, because of homologous recombination a certain proportion of all infections will be from wild type virus.

Analysis of the tumors

For at least two animals, Southern blot analysis showed that the tumors correlated with the presence of the PyT virus. The reason why the tumors from the third mouse failed to show the presence of the provirus is not known. These tumors may have been caused by the integration of wild type viruses. Probing the Southern blot with the LTR sequence should

answer this question but because of the background signal caused by endogenous viruses, it is often difficult to see novel proviruses, especially in the higher molecular weight range. Alternatively, failure to see PyT proviruses in these tumors may be a problem of detection. These tumors (as well as tumor 5 in the first mouse) were small and surrounding normal tissue may be diluting the proviral signal.

Southern blot analysis also showed that some tumors had more than one integrated provirus. For tumors 2,3 and 4 in the first mouse (nulliparous) there was most likely cross contamination between the tumors when they were isolated. They were very close to each other and were difficult to separate. From an EcoRI digest, it was quite obvious that the spleen tumor was an independent event from that of the mammary tumor because the site of integration of the provirus did not match either of the two in the mammary tumor. This indicated that the virus was replicating and travelling to distant sites.

The northern blot analysis was undertaken with the hope that novel transcripts would be visible in tumors, identifying a promoter-trap. Unfortunately nothing could be concluded from this analysis, else than the fact that some retroviral transcripts are present as expected. The RNA obtained from some of the tumors was not of good enough quality to detect transcripts below the 28S marker. It was hoped that the tumors could be propagated in culture so that RNA would be available from all of the tumors but none of the tumors survived in the dish. Northern blot analysis may not be sensitive enough to detect such novel transcripts. RNase protection would be more sensitive and should not be as affected by moderate level of degradation. Unfortunately, the proper conditions to amplify genomic

sequences flanking the site of integration could not be worked out and it was therefore not possible to determine how many of the tumors resulted from a promoter-trap.

Cloning of the TRAP1 gene

A different approach was chosen to analyse the integration site of one particular tumor. This tumor was chosen on the basis of the time at which it appeared relative to mammary differentiation, and the fact that the presence of the provirus could be established clearly by Southern blot analysis. Both integration sites from the mammary tumor from the pregnant mouse (monoparous) were cloned using a genomic library. The 5' flanking sequence from the two clones was analysed to determine which of the two integrations led to the promoter-trap. It was thought to be rather unlikely that both integrations would be a promoter-trap. As expected, only one clone was positive for the presence of transcribed sequences next to the 5' LTR. This was concluded from both northern blot analysis and sequencing. Northern blot data indicated that the gene in which the provirus had integrated was expressed at moderate levels. This may explain why the tumor was slow growing.

Sequence data provided enough information to confirmed that the genomic fragment upstream of the 5' LTR contained some transcribed sequence by showing a high degree of identity to a partial human cDNA, cloned as an expressed sequence tag (EST2205). The human clone was too small (278bp) to provide any clues as to the function of that gene. One can only speculate that it would be well conserved between the two species based on the high degree of identity at the nucleotide level (85%). No known mouse sequence was found

to have homology to the TRAP1 genomic sequence indicating that the murine gene had not been cloned.

A partial cDNA was first obtained from a commercial lung cDNA library. Further screening of the library using the cDNA was moderately successful and a different strategy was employed to obtain the full length cDNA. A technique known as rapid amplification of cDNA ends (RACE) (Frohman *et al.* 1988) made used of the already known partial cDNA sequence of TRAP1 to amplify the complete cDNA in two segments using the polymerase chain reaction. Whereas the 5' end of the cDNA could be amplified, the 3' end presented a bigger challenge. Seeing as the 5' end of the cDNA could be amplified from that new library which was primed with a poly dT oligonucleotide, the 3' part of the cDNA had to be represented also. That was confirmed by Southern blot analysis of an aliquot of the library where the partial TRAP1 clone detected a band representing the full length cDNA around 4.5kb. One may speculate that the TRAP1 mRNA contained some secondary structure which was difficult to bypass by both PCR amplification and reverse transcription with the conditions used in the commercial library (made using both random and poly dT primers) but that the structure was successfully bypassed by the conditions used to make the library for the RACE (Clontech) .

By comparing genomic and cDNA sequence data it was determined that the provirus had integrated in what is believed to be the 5' untranslated region of the TRAP1 gene, 47 nucleotides upstream of the putative initiating methionine. The fact that only 250bp of extra sequence was found at the 5' end of the cDNA using RACE supported the idea that the partial cDNA contains part of the 5' untranslated region. This would make the untranslated

region approximately 500bp which is larger than the average leader region (20-100 nucleotides) (Kozak 1987). Interestingly, longer 5' untranslated regions are often seen in protooncogenes (Kozak 1987). In light of this it seemed reasonable to expect the 2.2kb fragment of genomic DNA upstream of the provirus to contain the TRAP1 promoter. When subcloned in front of the CAT gene, this genomic DNA showed promoter activity in both mouse embryonal carcinoma and human embryonic kidney cell lines. TRAP1 expression was detected in both adult tissues (northern blot) and in cell lines of embryonic origin (CAT assay), indicating that the expression of TRAP1 is ubiquitous. The strength of the TRAP1 promoter could not be evaluated from this assay due to the presence of the strong SV40 enhancer. From northern blot analysis however, TRAP1 was found to be expressed at moderate levels in adult tissues. The demonstration of an active promoter upstream of the provirus provided very strong evidence that the tumor was caused by a promoter-trap.

Conclusion

The PyT virus was produced with the expected structure but producer clones were found to generate a mixture of recombinant and wild type viruses most likely due to homologous recombination during the transfection process. The overall amount of virus produced however, was as expected. It was shown that the PyT virus could cause the formation of tumors in animals. Because the genomic flanking sequences could not be isolated for all the tumors we could not verify that all of the tumors arose from promoter-trap events. However, most tumors showed the presence of the provirus indicating at least a correlation. For one tumor it was possible to show that the PyMT oncogene was expressed from the promoter of

a novel cellular gene and thus demonstrating that the PyT virus can be used *in vivo* to clone novel genes. The fact that TRAP1 was found to be expressed at low levels demonstrated that PyT is sensitive enough to allow detection of weak promoters.

The main advantage of using an *in vivo* promoter-trap approach is that "housekeeping" genes can be separated from mammary specific genes by mating the animals. The fact that TRAP1 expression was not specific to the developing gland suggested that the waiting period before mating the animals could be extended to account for weaker promoters. The identification of the genes of interest may even require a few pregnancies before their activity can be detected.

By reducing the contamination from wild type viruses and optimizing a rapid way of isolating flanking genomic sequences, a great number of integration events in potentially developmentally regulated genes could be screened.

C) General Conclusions

The work presented in this thesis demonstrated that it is possible to use a replication competent promoter-trap *in vivo* to clone novel genes active in the murine mammary gland. A first replication deficient virus failed to work efficiently as a promoter-trap *in vivo*. This was attributed to a low level of *in situ* infection and a reporter gene that lacked in sensitivity. A second virus was constructed to improve on both of these points. The replication competent promoter-trap virus PyT relied on a single strong oncogene to detect active cellular promoters *in vivo*. The potential of this virus was illustrated by the cloning of a

novel murine gene from the mammary gland of a living animal. The application of the PyT virus could be extended to any other developmental system provided that the PyMT oncogene can transform the tissue of interest.

APPENDIX

A Replication-Competent Promoter-Trap Retrovirus.

Josée Coulombe[1,3] Yvonne Avis[3], and Douglas A. Gray[1,2,3†]

Departments of [1]Biochemistry and [2]Medicine, University of Ottawa,
[3]Ottawa Regional Cancer Centre, 501 Smyth Road, Ottawa, Canada K1H 8L6

Running Title: PyT retrovirus

† corresponding author

ABSTRACT

A promoter-trap retrovirus has been constructed in which a promoterless Polyoma virus middle T antigen gene was inserted in the U3 region of the LTR of a replication-competent Moloney murine leukemia virus. The resulting virus, designated PyT, was used to infect mouse mammary glands *in situ*. As expected, mammary tumors appeared in some infected animals. These tumors were found to contain PyT proviruses of the predicted structure. From one such tumor the PyT provirus and surrounding sequences from the integration site were cloned. The provirus was found to have integrated adjacent to the promoter of a novel mouse gene (TRAP1) that was expressed at low levels in various mouse tissues. These data show that the PyT retrovirus provides a sensitive means of detecting active promoters *in vivo*.

INTRODUCTION

Promoter-trap strategies have proven to be very effective for cloning genes that are active at various stages of development. Frequently this has been done in embryonic stem (ES) cells, which when reintroduced into blastocysts generate chimeric embryos that can be bred to identify developmental defects (2,5,6,15,17). By using a promoter-trap strategy one can considerably increase the efficiency of insertional mutagenesis because there is selection for insertions into genes as opposed to "junk" DNA. This enrichment decreases the number of ES clones to be screened for a phenotype. The enrichment of the promoter-trap has been estimated to be 100 to 1000 fold (2). Promoter-traps also facilitate the cloning of the integration site in that the transgene is located within the gene to be cloned.

Two types of promoter-trap vectors have been used for insertional mutagenesis in ES cells. The first has a promoterless reporter gene in a plasmid vector which is electroporated into cells (6,15). The second type makes use of retroviral vectors (5,17,18), which have a number of advantages over plasmid vectors. The integration of proviral DNA into the genome is less disruptive than the integration of naked DNA and therefore eliminates the possibility of rearrangements which can inactivate the promoter. Retroviruses also have a certain affinity for DNase hypersensitive sites which are found near promoter regions (13,16), therefore increasing the chances of integration within a gene. Finally, as we have shown here, retroviral vectors allow promoter-traps to be performed *in vivo*.

Different variations of a retroviral promoter-trap vector have been used previously (5,17,18). In one experiment, a promoterless drug resistance gene was inserted in the U3

region of the LTR so that upon proviral integration the indicator gene was positioned 30 base pairs away from genomic sequence (18). The presence of foreign DNA in this region of the LTR did not alter the ability of the virus to replicate (18). This strategy had the advantage that the virus was replication competent. A second group introduced a reporter gene preceded by a splice acceptor site, in place of the viral structural genes (5). This increased the probability of the reporter gene to be expressed (expression of the reporter occurs upon insertion in the gene anywhere after the first splice donor - such strategies are more accurately described as "gene traps") but rendered the virus incapable of replicating. Because the comparatively low titer of defective viruses would limit the number of insertion events that we could screen *in vivo*, we modelled our strategy on the replication-competent promoter-trap virus (18).

We chose to identify promoters in the mouse mammary gland. This gland is unusual in that most of its development takes place postnatally (the breast remains rudimentary until puberty whereupon the ductal structure of the gland expands: during pregnancy the mammary gland reaches full development (4)) and it is likely that there is a program of gene expression that accompanies development. The mammary gland is also ideal for a retroviral promoter-trap strategy in that its epithelium is topologically continuous with the exterior of the animal, and can be infected by introduction of virus through the nipple.

Obviously it is not possible to perform *in vivo* experiments using retroviruses that incorporate genes conferring resistance to cytotoxic drugs that would kill the animal. We therefore chose to include a promoterless oncogene as the marker gene in our experiments. Mouse mammary epithelium has been shown to be very susceptible to transformation by the

Polyoma virus middle T antigen (PyMT). Transgenic mice carrying this viral gene under the control of the mouse mammary tumor virus promoter (MMTV-LTR) develop multifocal mammary tumors around the age of two months (8). The apparent single-hit kinetics of PyMT transformation of mouse mammary epithelium suggested that PyMT would be an appropriate reporter gene for an *in vivo* promoter-trap retrovirus. Because retroviruses require that the cells be in a replicative stage to integrate (3), we chose to initiate infections at puberty, when the mammary gland undergoes ductal expansion (4). Here we report the construction of PyT, a replication-competent retrovirus incorporating a promoterless PyMT gene, its use in infection of pubertal mouse mammary glands, and the subsequent identification of a cellular promoter active in that tissue.

MATERIALS AND METHODS

Retroviral vector and virus production: The Polyoma Virus middle T antigen (PyMT) sequence, a HindIII to EcoRI fragment from the MMTVMT vector (a gift from Dr. W. Muller; 8) was inserted in pLTR12 (12) which contains one Moloney LTR with an engineered BamHI site in place of the Sau3A site at position 95. The insert was ligated to the vector by blunt ligation. The remainder of the Moloney provirus was ligated to the modified pLTR12 vector as an EcoRI-ClaI fragment from pMov3 (gift of Dr. R. Jaenisch). This positioned the PyMT coding sequence in the 3' LTR.

The PyT construct was cotransfected with a neomycin resistance encoding plasmid (pgk-neo; 10) in NIH 3T3 cells by the $CaPO_4$ method (modification of Graham et al; 7) and clones were selected in G418. Producer clones were tested for viral titer using an XC assay (14). For the nipple injections, the supernatant from a plate of confluent cells was cleared of cells by passing it through a 0.2 μ filter. For cell infection, supernatants were prepared as above and were transfered to NIH 3T3 or NMuMG cells with 8μg/ml of polybrene (Sigma).

Middle T expression: 293 cells (plated at $5x10^5$ per 60mm dish the day before) were transfected transiently, using $CaPO_4$, with various amounts of a plasmid encoding the PyMT sequence under the control of the pgk promoter. Cells were harvested twenty-four hours later in sample buffer (4% SDS, 20% glycerol, 120mM Tris pH6.8, 1mg/ml bromophenolblue).

Approximately five million NMuMG cells were infected with PyT virus at a multiplicity of infection near two. Forty-eight hours later the cells were plated in 0.3% agarose, and

colonies were isolated after four weeks. Proteins were extracted from confluent 60mm dishes in 200µl of sample buffer.

For western analysis, one twentieth of the protein lysate was loaded on a 10% SDS polyacrylamide gel and transferred to a nitrocellulose membrane (Costar). Middle T antibody was the monoclonal Ab815.

Nipple injections: Nipple injections were performed using a modification of the method of Wang *et al.* (19). Four week old Balb/C female mice were anaesthetized using Avertin (10g of Tribromoethanol [Aldrich] in 10ml of tert.amylalcohol [BDH], diluted in saline) at 0.4 mg per gram of body weight. Freshly isolated virus, to which 80 μg/ml of polybrene (Sigma) was added, was injected directly in the nipple using a 30 gauge needle. Approximately 0.05 - 0.1 ml ($5x10^4$-10^5 cfu) was injected into each nipple.

Tumor analysis and cloning of integration sites: Tumors were surgically removed and DNA was extracted with 400 μg/ml of proteinase K in SDS buffer (75 mM NaCl, 25 mM EDTA, 10 mM Tris pH 7.6, 1% SDS) at 37 °C overnight. Following two phenol/chloroform extractions and ethanol precipitation, 10μg of DNA was digested with EcoRI and was run on a 1% agarose gel. The DNA was transferred to Hybond-N membrane (Amersham). The PyMT sequence was radioactively labelled using random oligonucleotide primers (Boehringer). For the construction of the library, 50 μg of DNA was digested with EcoR1 and 12-20 kb fragments were isolated on a 5-20% potassium acetate gradient. The size selected DNA was ligated to λDASHII arms digested with EcoRI (Stratagene) and packaged using a Gigapack extract (Stratagene). 500 000 plaques were screened and six positives

were plaque purified. Phage DNA was digested with various enzymes and was run on 1% agarose gel for Southern blot analysis.

Subcloning and sequencing: The genomic DNA flanking the provirus was subcloned into pBLUESCRIPT (Stratagene) using EcoRI and KpnI sites. DNA was sequenced using the Pharmacia T7 sequencing kit with either their Universal primer or a primer located in the LTR. The sequence was analysed using the DNASTAR program.

Northern analysis: Total RNA was isolated using LiCl and messenger RNA was prepared using oligo dT columns (BDH column, BRL resin). 3 μg of poly A$^+$ RNA was run on 1% agarose formaldehyde gel and transferred to a Hybond-N membrane (Amersham). An EcoRI-NheI genomic fragment flanking the 5' LTR was labelled as before.

Promoter assay: An NheI fragment which included all of the cloned 5' genomic sequence (except for 220bp at the 5' end) and the first 31bp of the 5'LTR, was subcloned into the XbaI site of the pCAT-enhancer vector (Promega). 5µg of the so-called TRAP-CAT vector was cotransfected in 293 cells, together with an equal amount of a plasmid encoding lacZ (pgk-lacZ; gift of Dr. McBurney) using the $CaPO_4$ method. Twenty four hours later the cells were harvested in 0.25M Tris pH 7.8 and lysed with two rounds of freeze/thaw (dry ice/37°C). The samples were divided into two, one for each enzyme assay, and one half (for the CAT assay) was heated at 65°C for 10 minutes. Samples for both assays were centrifuged in a refrigerated microfuge for 10 minutes. The supernatants were assayed for protein concentration using a micro-assay from BIO-RAD. For the β-galactosidase assay, 10μg of protein was mixed with 800μl of assay buffer (100mM $NaPO_4$ pH 7, 1mM $MgSO_4$ and 10mM β-mercaptoethanol) and 200μl of ONPG (4mg/ml; BDH). Samples were

analysed on a Beckman DU® 640 spectrophotometer using the manufacturer's "kinetics/Time" program, at 420nm. 5μg of protein was used to assay for CAT activity using a fluorochrome-based kit (FMBIO® TANGERINE FAST CAT® kit; Hitachi). The products were separated on thin layer chromatography and analysed on a fluoroimager (Hitachi) using the manufacturer's FMBIO (V4.04) software.

RESULTS:

In vitro analysis:

The Polyoma middle T antigen (PyMT) sequence was inserted in the 3' LTR of a Moloney murine leukemia retroviral vector (Fig 1A). If viral replication proceeded normally, upon integration of the provirus the PyMT sequence would be duplicated in the 5' LTR and would be positioned 95bp 3' to genomic sequences.

Before proceeding to *in vivo* experiments, we tested the replication of the virus in cell cultures. The PyT vector was transfected into NIH 3T3 cells and 24 producer clones were isolated. The best clones gave a titer of 10^6 cfu/ml as determined by XC assay (14). The integrity of the provirus was confirmed by Southern blot analysis of PyT virus-infected NIH 3T3 cells. Bands of the predicted size were detected in digests with SstI (not shown) and XbaI (Fig 1B). The stability of the virus was tested for two selected producer clones, again by Southern blot analysis of PyT infected cells. The virus was found to be stable for three rounds of infection when passaged through NIH 3T3 cells at 48 hours intervals. After several passages, the PyMT sequences were lost (Fig 1B, 5° infection), most likely by a recombination event.

The expression of Middle T antigen was also verified in culture. PyMT was expressed transiently in 293 from the pgk promoter and the protein was detected by western blot using whole lysate (Figure 2A). The amount of PyMT detected correlated with the amount of plasmid introduced. Expression of PyMT by promoter-trap was tested by infecting mammary epithelial cells (NMuMG) with the PyT virus and selecting for growth in soft agarose. Expression of Middle T by 12 random clones was verified by western blot

(Figure 2B). Most clones (ten out of twelve) showed detectable levels of the PyMT protein. Clones 3 and 8 either did not express PyMT and grew in soft agarose as the result of unrelated events, or expressed PyMT at levels that were sufficient for the phenotypic change, but below the level of detection of the western blot.

Mammary gland infections:

Four week old Balb/C female mice were used for the mammary injections because this is a period where the mammary epithelium proliferates to form the ducts (4), proliferation being a requirement for retroviral infection (3). A summary of the tumors obtained is shown in table 1. A total of twenty animals were injected, most in all ten glands. The first tumor was detected in an animal six weeks post-infection. This animal eventually developed two additional visible tumors and was sacrificed (*nulliparous* in Fig 3). One tumor mass was quite large and turned out to be a mixture of two or three tumors (as evidenced by blot analysis, Fig 3). To ensure activation of mammary specific promoters, animals that were free of tumor at six weeks post infection were mated. One mouse developed a tumor three weeks later. The tumor stopped growing after about a week and the animal was sacrificed (*monoparous* in Fig 3). A second tumor was observed on the spleen and was saved for analysis. A third mouse developed a tumor a week after weaning her first litter (seven weeks post-mating). The animal was sacrificed a week later and a second tumor was found on the colon which was also saved for analysis. No other tumors were detected among the rest of the injected animals and they were sacrificed after five months.

A second set of mice (19) were infected in the same way with a replication competent Moloney retrovirus containing a different DNA sequence in place of PyMT. No tumors were detected after 10 weeks.

Analysis of the tumors:

All tumors were analysed for the presence of the PyT provirus by Southern blot analysis (Fig 3). EcoRI was chosen because it did not have any recognition site in the provirus and could therefore reveal the number of integrated proviruses (one EcoRI band equals one integrated provirus). The presence of the virus was confirmed in all but the three tumors from the mouse who developed tumors after weaning (Fig 3). The spleen tumor found in the second mouse (2* in Fig 3) was independent of the mammary tumor as indicated by their different pattern of provirus integration site. This suggests that the PyT virus was replicating and getting into the circulation. This assertion was also supported by the observation that the tumor free mice eventually showed symptoms of Moloney infection (9) including enlarged thymuses.

Cloning of the integration site and analysis of the genomic flanking sequence:

We chose to focus on the mammary tumor that arose in the pregnant mouse (*monoparous* in Fig 3). A genomic library was constructed from size selected EcoRI digested tumor DNA and probed with the LTR sequence. Two different clones were isolated corresponding to the two integrations sites. This was confirmed by reprobing the Southern blot shown in Figure 3 with genomic sequences from these two clones sequentially. Each clone showed one band corresponding to one of the two integration sites seen on the initial probing with PyMT (data not shown). Figure 6A shows the map

of the genomic clone corresponding to the promoter-trap. This was deduced from both sequencing and northern blot analysis data (see below). Also shown is the size of the 5' and 3' genomic flanks as mapped by Southern blot analysis using EcoRI in combination with either KpnI or XhoI (Fig 6A).

Genomic sequence from the 5' end of the integration site was used to probe a Southern blot containing NIH 3T3 cell DNA digested with three different enzymes (Fig 4). The restriction enzyme pattern of our clone is consistent with a single copy gene, as evidenced by the single intense band detected in EcoRI or SstI digests (Fig 4). In addition, the genomic probe detected an extra band in an EcoRI digest of genomic DNA from the tumor corresponding to the disrupted allele. These data confirmed that the cloned sequences were from the site of proviral integration.

Northern blot analysis of various mouse tissues was used to test whether the genomic DNA flanking the provirus contained transcribed sequences. Using RNA from breast tissue of virgin and pregnant mice we could at the same time determine whether the trapped gene was expressed specifically during mammary differentiation. Using poly A^+ RNA and the same genomic probe as above, we detected a low abundance transcript of approximately 4.5 kb which was present in all tissues tested (Fig 5). Other transcripts of lower abundance were also detected. The major transcript was present in equal amounts in pregnant and virgin breast tissue. Testis contained high levels of this transcript together with two additional transcripts not detected in the other tissues.

Sequence was obtained from both ends of the genomic clone. A 220bp of sequence adjacent to the 5'LTR (Fig 6B) showed no similarity to any mouse sequence in the

GenBank database but revealed strong similarity to a partial human cDNA sequence, previously isolated from human brain as expressed sequence tag (EST) (1) (Fig 6C). The sequence identity between TRAP1 and EST02205 was 85% over 136 bases ending at the 3' limit of the 278 base EST02205. It is possible that further similarity exists with the remainder of the mRNA from which EST02205 was derived.

The TRAP1 promoter:

Because retroviruses tend to integrate near the 5' end of genes we decided to test for the presence of the TRAP1 promoter in the 2.2kb genomic fragment 5' of the integration site. This fragment was subcloned in front of the bacterial chloramphenicol acetyl transferease (CAT) gene in a vector containing the SV40 enhancer. The vector was assayed for CAT activity by transient transfection in the human 293 cell line. Figure 7 shows an example of the thin layer chromatograph (A) and the CAT activity (B) reported as percent of the substrate converted (the percent conversion was calculated from non-saturating amounts of samples). There was a 75-fold difference between the percent conversion of the putative promoter (TRAP+) and that of the vector alone (average of three experiments). The transfection efficiency did not vary more than twofold as calculated by β-galactosidase activity .

To demonstrate that the TRAP1 promoter did drive the expression of PyMT in the tumor, the 5' junction fragment of the provirus (which included the genomic sequence used in the CAT assay above and the start of the provirus up to the Xho1 site in the gag sequence) was transfected transiently into 293 cells and PyMT was detected by western bot

of whole lysates (Figure 8). The amount of PyMT detected paralleled the amount of plasmid added, indicating that the PyMT sequence in the LTR was expressed from the TRAP1 promoter.

DISCUSSION:

From twenty animals that were infected in the mammary glands with the PyT virus we collected a total of ten tumors, eight of which appeared at sites and with gross morphology consistent with a mammary origin. Six tumors out of ten were found to contain PyT proviral DNA. One animal developed three distinct tumors that were found to be negative for the PyT provirus. We currently cannot explain the origin of these tumors, which arose in a strain (BALB/c) of low spontaneous tumor incidence (11). In contrast, no tumor formed in animals infected with a similar replication-competent virus containing irrelevant and nonfunctional sequences (Coulombe and Gray, unpublished).

The fact that a number of tumors have more than one integrated provirus indicates that the efficiency of mammary infection by nipple injection can be very high. It is unlikely that multiple EcoRI bands of roughly equal intensity represent re-infection events during the growth of the tumor - rather, these bands represent proviral content in the cell that initiated tumor formation. It is difficult to estimate the efficiency of the promoter-trap *in vivo* because one cannot accurately estimate number of target cells in the mammary gland (cycling cells to which the virus has access), or the number of infection events for the PyT retrovirus prior to the loss of the marker PyMT sequences.

The isolation and characterization of one of the promoter-trap integration sites allowed us to clone what appears to be a novel mouse gene that we called TRAP1. We base this conclusion on several lines of evidence. First, a fragment (2.2kb) from this gene hybridized to a 4.5kb transcript on a northern blot of various mouse tissues. This gene was

expressed at low levels, as the transcript was only detected in poly A^+ RNA. Second, the same fragment of genomic DNA was used to probe a Southern blot of normal DNA digested with several enzymes (Figure 4), and the pattern obtained was consistent with a single copy gene. These data make it unlikely that the transcripts detected with the TRAP1 genomic probe are derived from some repetitive element present in the probe. Finally, genomic sequence adjacent to the 5' LTR showed 85% identity over 136 bases to the previously identified human brain partial cDNA EST02205. Transcripts were detected by the TRAP1 probe in the mouse brain (Fig 5), our results are therefore consistent with TRAP1 genomic sequences corresponding to the murine homolog of the gene from which the human EST02205 was derived.

Northern blot analysis showed that TRAP1 was expressed at comparable levels in all adult tissues tested. The tumor from which TRAP1 was isolated was detected during pregnancy, but its early appearance (at three weeks post-mating) exceeded expectations based on previous transgenic mouse studies in which PyMT was driven by the MMTV promoter (8). In these studies, PyMT-induced mammary tumors had a four week latency. For this reason, and in light of the low abundance of the TRAP1 mRNA transcript, we favor the hypothesis wherein relatively low levels of PyMT were expressed from the TRAP1 promoter. This level of expression resulted in the appearance of a tumor nine weeks post-infection, which coincidentally corresponded to three weeks post-mating. There is no evidence that the TRAP1 gene is transcriptionally regulated during gestation, or that the gene is involved in mammary differentiation.

Northen blot analysis also showed that the TRAP1 promoter is active in normal mammary tissue and was not activated from a silent state by the retroviral enhancer. To avoid the possible activation of silent promoters through juxtaposition of the Moloney enhancer these elements could have been eliminated from the virus by deletion or mutation, but such alterations would likely have resulted in lower viral titers. We chose to make as few alterations to the Moloney sequences as possible in order that the titer (and hence the number of insertion events) be optimal for *in situ* infections, where the volume of the inoculum is necessarily limited.

It is an expectation of our promoter-trap strategy that integration of the PyT provirus should occur immediately downstream of an active promoter: intervening sequences beyond a certain length would be expected to contain AUG codons that would impede the translation of the PyT marker gene. We therefore thought it likely that the 2.2 kb 5' genomic fragment cloned from the integration site should contain a promoter element. This prediction was confirmed in experiments where the 5' genomic sequences were attached to a CAT reporter gene (Fig 7). The activity of this promoter was found to be relatively high when compared to the Moloney murine leukemia virus LTR, but the comparison has the caveat that the TRAP1 promoter was aided by a distal SV40 enhancer element, whereas the MoMuLV LTR was not (it did, however, contain nascent LTR enhancer elements). Our expectation (based on northern analysis) was that the TRAP1 promoter would be weak. Either its activity is greatly elevated by the SV40 enhancer, or the 2.2 kb fragment is devoid of neighbouring negative elements that exist in the mouse

genome (or both). We have also shown that the TRAP1 promoter could drive the expression of PyMT (Figure 8).

We have shown that PyT, a replication-competent retrovirus, can be used to introduce a promoterless oncogene into the genomes of mouse cells in the intact animal. The data presented herein demonstrate that this virus will infect a sufficient number of cells in a target tissue to cause the formation of a tumor. Further, from cloning a provirus from one such tumor we have demonstrated the presence of a cellular promoter immediately adjacent to the proviral LTR as predicted. We are therefore confident that the PyT virus can be used to screen for active promoters in murine tissues or cells, either *in vivo* or *in vitro*. By producing the virus in an amphotropic cell line, it may also be possible to infect nonmurine cells or tissues, although we have yet to test this prediction.

ACKNOWLEDGEMENTS:

This work was supported by a grant from the Medical Research Council of Canada to D.A.G. J.C. was the recipient of a Fonds pour la Formation de Chercheurs et l'Aide à la Recherche studentship.

REFERENCES

1. Adams, M. D., M. Dubnick, A. R. Kerlavage, R. Moreno, J. M. Kelley, T. R. Utterback, J. W. Nagle, C. Fields, and J. C. Venter. 1992. Sequence identification of 2,375 human brain genes. Nature 355:632-634.

2. Chang, W., S. C. Hubbard, C. Friedel, and H. E. Ruley. 1993. Enrichment of insertional mutants following retrovirus gene trap selection. 1993. Virol.193:737-747.

3. Brown, P. O. 1990. Integration of retroviral DNA. In "Retroviruses Strategies of Replication", Springer-Verlag, Berlin, Heidelberg.

4. Daniel, C. W., and G. B. Silberstein. 1987. Development of the rodent mammary gland. In "The Mammary Gland", Plenum, New York, NY.

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

6. 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 244:463-465.

7. Graham, G. L., and A. J. van der Eb. 1973. A new technique for the assay of infectivity of human ademovirus 5 DNA. Virol.52:456-467.

8. Guy, C.T., R. D. Cardiff, and W. J. Muller. 1992. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model. Mol.Cell.Biol. **12**:954-961

9. Kozak, C.A., and S. Ruscetti. 1992. Retroviruses in Rodents. In "The Retroviridae". Plenum, New York, NY.

10. McBurney, M. W., Sutherland, L. C., Adra, C. N., Leclair, B., Rudnicki, M. A., and Jardine, K. 1991. The mouse Pgk-1 gene promoter contains an upstream activator sequence. Nucleic Acids Res. **19**:5755-5761.

11. Nandy, S., and C.M. McGrath. 1973. Mammary neoplasia in mice. Adv. Cancer Res. **17**:353-414.

12. 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 sequences. Proc.Natl.Acad.Sci. USA **82**:1141-1145.

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

14. Rowe, W. P., W. E. Pugh, J. W. Hartley. 1970. Plaque assay techniques for murine leukemia viruses. Virol. **42**:1136-1139.

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

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

17. von Melchner, H., J. V. DeGregori, 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.

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

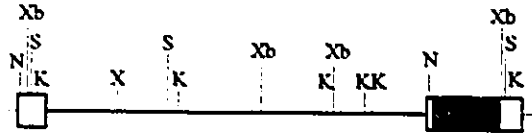
19. Wang, B., W. S. Kennan, J. Yasukawa-Barnes, M. J. Lindstrom, and M. N. Gould. 1991. Carcinoma induction following direct *in situ* transfer of v-Ha *ras* into rat mammary epithelial cells using replication-defective retrovirus vector. Cancer Res. **51**:2642-2648.

FIGURE 1 A. Schematic representation of the PyT retroviral promoter-trap DNA vector. The open box represents the LTR and the cross-hatched box the Polyoma virus middle T antigen sequence. Thin lines represent plasmid sequence. K:KpnI; N:NheI; S:SstI; Xb:XbaI; X:XhoI.

B. Southern blot of a PyT producer clone and PyT virus-infected NIH 3T3 cells. NIH 3T3 cells were infected with the PyT virus and were fed fresh medium 24h later. Forty eight hours post infection, the supernatant of the infected cells was used to infect fresh NIH 3T3 cells. This cycle was repeated for several passages. DNA was isolated from infected cells (48h post infection) and digested with XbaI for Southern blot analysis.

A

PyT vector



B

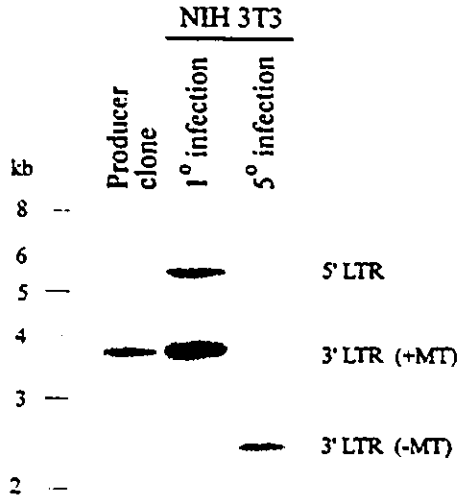
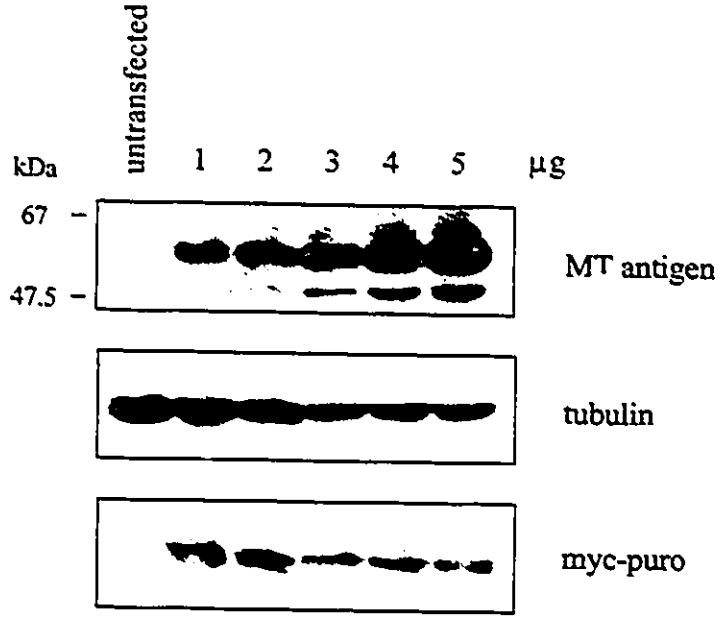


FIGURE 2 Western blot analysis of PyMT expression.

A. Transient expression of PyMT in human 293T cells. The amount of DNA used is as indicated. Tubulin was used as a loading control. Myc-puro is a fusion between six copies of a human myc epitope and the puromycin protein, and served as an internal control for the transfections. B. Expression of PyMT by promoter-trap. pgk-PyMT is a stable clone used as a positive control for PyMT expression. NMuMG is the parental cell line for the promoter-trap clones (negative control). The promoter-trap clones 1-12 were derived from colonies that grew in soft agarose after infection with the PyT virus. The loading control is tubulin.

A


untransfected
kDa
1
2
3
4
5
µg
67
47.5
MT antigen
tubulin
myc-puro

B

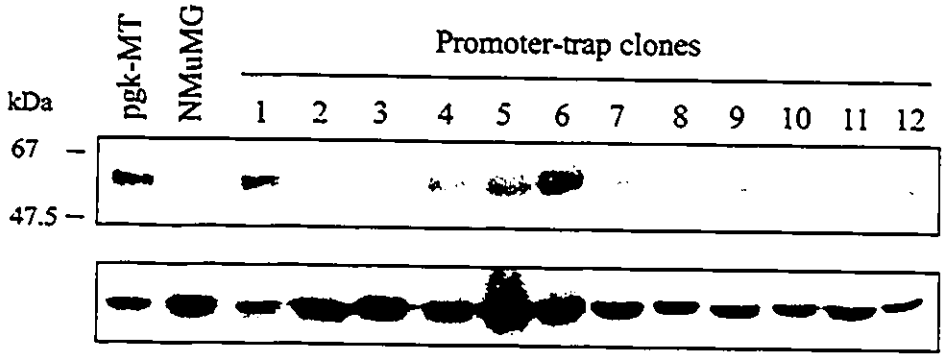

pgk-MT
NMuMG
Promoter-trap clones
kDa
1
2
3
4
5
6
7
8
9
10
11
12
67
47.5

FIGURE 3 Southern blot of EcoRI digested tumor DNA, probed with the PyMT sequence. Each band represents one proviral integration site. Tumors from two different animals are shown (nulliparous and monoparous) with normal liver or kidney DNA as controls. Each number represents a different mammary tumor except for 2° which is a spleen tumor.

Nulliparous
Mono parous
1 2 3 4 5 Liver
1 2* Kidney
kb
24
19
15
12
10

FIGURE 4 Southern blot of tumor DNA digested with EcoRI, and NIH 3T3 cell DNA digested with EcoRI, BamHI or SstI. The probe was the genomic fragment 5' of the proviral integration site. Comparison of tumor DNA and NIH 3T3 DNA digested with EcoRI confirmed that the probe was derived from the integration site.

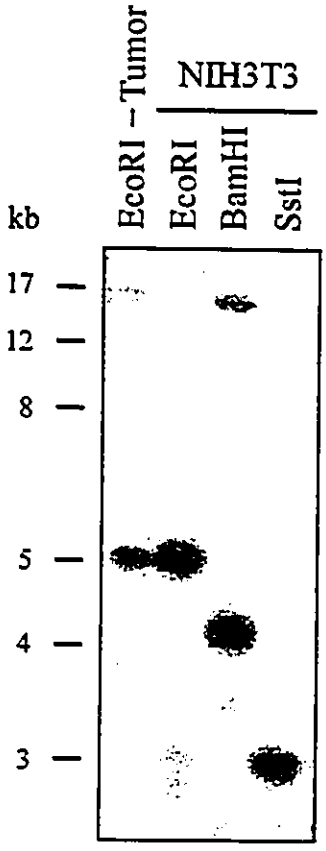

Tumor
NIH3T3
EcoRI
EcoRI
BamHI
SstI
kb
17
12
8
5
4
3

FIGURE 5 Northern blot (poly A^+ RNA) of various adult mouse tissues. The probe used was the genomic fragment 5' of the retroviral integration site. β-actin was used as a loading control.

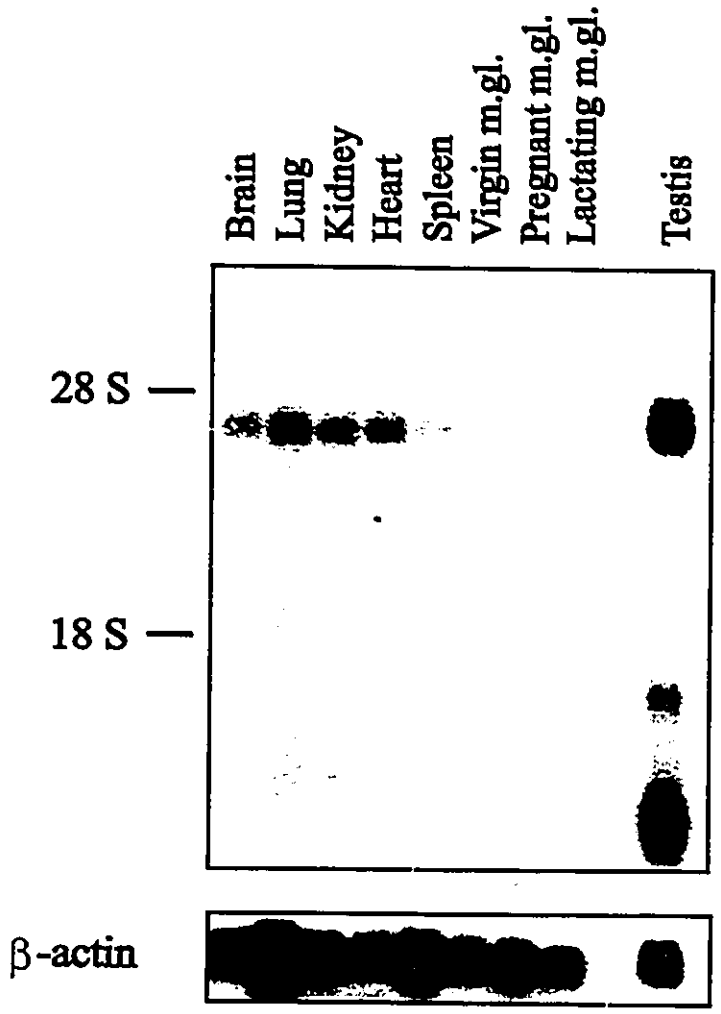

Brain
Lung
Kidney
Heart
Spleen
Virgin m.gl.
Pregnant m.gl.
Lactating m.gl.
Testis
28 S
18 S
β-actin

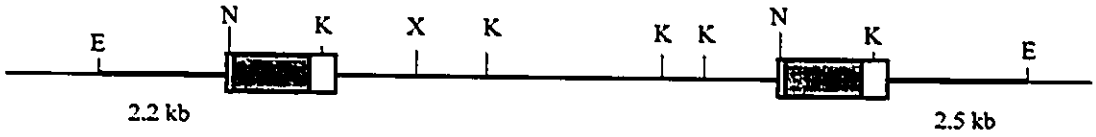
FIGURE 6 A. Schematic representation of the TRAP1 genomic clone. The open box represents the LTR and the cross-hatched box the Polyoma virus middle T antigen sequence. The thick line represents genomic sequence and the thin line the phage arms. E:EcoRI; K:KpnI; N:NheI; X:XhoI.

B. Sequence data from the genomic DNA proximal to the 5'LTR. Numbers refer to the genomic sequence (capital letters). Lowercase letters represent LTR sequence.

C. Alignment of TRAP1 and EST02205 as obtained from GENBANK. There is 85% identity between these sequences.

A

TRAP1 genomic clone



B

```
CCACTCAATCTCACTCATCCAGCCATCGTTGACCACACAGACTTTTTTTT  50
TTCAACGACCCCCTCTCCCTAAATGGAGTTGGGCGGGGGGGAAATGAAAA 100
CGAGTTGGTCTTTACTTTTTTAAAAGAGACTTTTTGATCCAATGAGAGCC 150
CCACAAGTAACTGAGTTTGGGTCCTGGTTGGTTGCTTTATTTTTTTCCCT 200
CCGGGGAAAAGCCCCTCCCCCtgaaagaccccacctgtaggt...       221
```

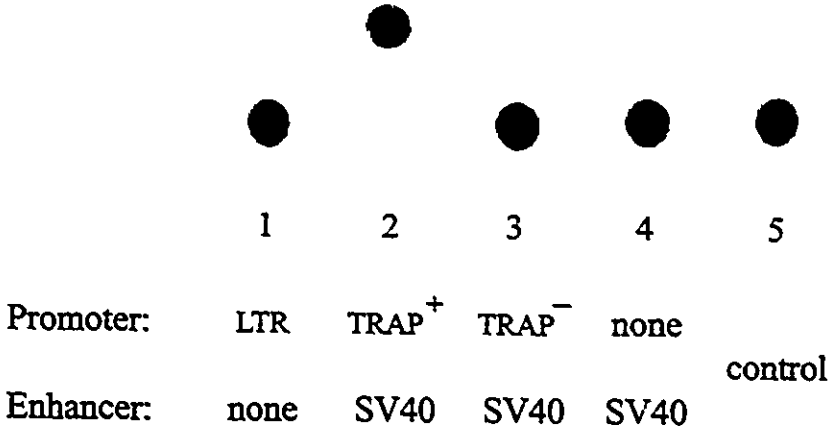
C

```
 60 CCCCTCTCCCTAAATGGAGTTGGGCGGGGGGGAAATGAAAAC-GAGTTGGTCTTTACTTT 118  TRAP1
    |||| |||||||||||||||||| |||||||||||||| || ||||||| | ||  |||
144 CCCCCCTCCCTAAATGGAGTTGGGTGGGGGGGAAATGAATACTGAGTTGGCC-TTTATTT 202  EST02205

119 TTTAAAAGAGACTTTTTGATCCAATGAGAGCCCCACAAGTAACTGAG-TTTGGGTCCTGG 177  TRAP1
    |||||||||  |||||||||||||||||  ||||  || ||| |||| ||||||||||||
203 TTTAAAAGA--CTTTTTGATCCAATGAGGCCCCCTAAA-TAATTGAGTTTTGGGTCCTGG 259  EST02205

178 TTGGTTGCTTTATTTTTTT 196  TRAP1
    ||||||  |||  |||| |
260 TTGGTTTGTTTTATTTTGT 278  EST02205
```

FIGURE 7 A) Thin layer chromatograph of a typical CAT assay (the modified fluorescent substrate generates only one acetylated product). The first lane is a positive control: a plasmid encoding CAT under the control of the MoMLV-LTR (gift of Ninan Abraham; this plasmid does not contain the SV40 enhancer). $TRAP^+$ and $TRAP^-$ are plasmids containing the TRAP1 genomic fragment subcloned in front of CAT in the positive and negative orientation respectively. The fourth lane is the vector alone, with the SV40 enhancer but no promoter. The last lane had no protein added to the assay. The amount of converted substrate in lane 5 was used as the background for the calculation of the conversion.

B) CAT activity shown as percent of converted substrate. Samples are the same as in A.

A


1
2
3
4
5
Promoter:
LTR
TRAP+
TRAP−
none
control
Enhancer:
none
SV40
SV40
SV40

B

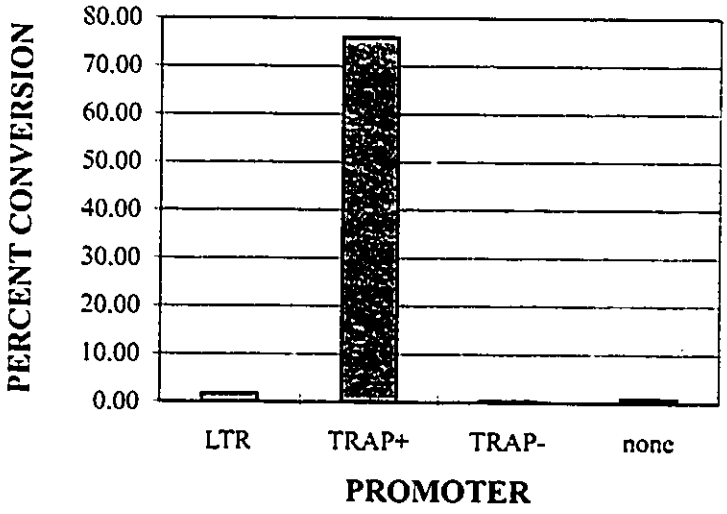

80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
PERCENT CONVERSION
LTR
TRAP+
TRAP-
none
PROMOTER

FIGURE 8 PyMT expression from the TRAP1 promoter. Western blot showing transient expression of PyMT from the TRAP1 promoter, in 293 cells. The amount of DNA used is as indicated. Tubulin was used as a loading control. Myc-puro is a fusion between six copies of a human myc epitope and the puromycin protein, and served as an internal control for the transfections.

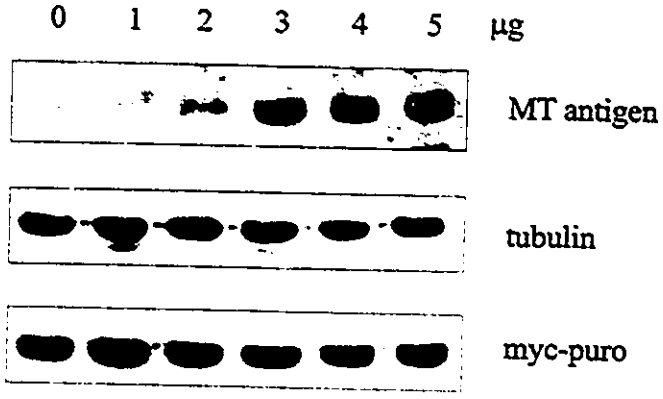

0
1
2
3
4
5
µg
MT antigen
tubulin
myc-puro

Table 1 Summary of tumors found in PyT infected animals.

MOUSE	NUMBER OF MAMM. TUMOR(S)	TIME OF DETECTION (wk post-infection)	OTHER TUMOR
Nulliparous	4 (5)	6	–
Monoparous	1	9 (3 wk post-mating)	spleen
Weaning	2	13 (7 wk post-mating)	colon

BIBLIOGRAPHY

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, et al. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project.. Science 252:1651-6.

Adams MD, Dubnick M, Kerlavage AR, Moreno R, Kelley JM, Utterback TR, et al. 1992. Sequence identification of 2,375 human brain genes.. Nature 355:632-4.

Alitalo K, Bishop JM, Smith DH, Chen EY, Colby WW,Levinson AD. 1983. Nucleotide sequence of the v-myc oncogene of avian retrovirus MC29. Proc Natl Acad Sc 80:100-4.

Andres AC, Schonenberger C, Groner B, Gennighausen L, LeMeur M,Gerlinger P. 1987. Ha-ras oncogene expression directed by a milk protein gene promoter: Tissue specificity, hormonal regulation, and tumor induction in transgenic mice. Proc Natl Acad Sci 84:1299-303.

Armentano D, Yu S-F, Kantoff PW, von Ruden R, Anderson WF,Gilboa E. 1987. Effect of internal viral sequences on the utility of retroviral vectors. J Virology 61(5):1647-50.

Auffray C, Rougeon F. 1980. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur J Biochem 107:303-14.

Barbicid M. 1987. ras genes. Ann Rev Biochem 56:779-827.

Basset P, Bellocq JP, Wolf C, Stoll I, Hutin P, Limacher JM, et al. 1990. A novel metalloproteinase gene specifically expressed in stromal cells of breast carcinomas.. Nature 348:699-704.

Berger MS, Locher GW, Saurer S, Gullick WJ, Waterfield MD, Groner B, et al. 1988. Correlation of c-erbB2 gene amplification and protein expression in human breast carcinoma with nodal status and nuclear grading.. Cancer Res 48:1238-43.

Birnboim HC. 1983. A rapid alkaline extraction method for the isolation of plasmids DNA. Methods Enzymol 100:243-55.

Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN,Weintraub H. 1990. Sequence-specific DNA binding by the c-myc protein. Science 250:1149-51.

Bradbury JM, Sykes H,Edwards PAW. 1991. Induction of mouse mammary tumors in a transplantation system by the sequential introduction of the myc and ras oncogenes. Int J Cancer 48:908-15.

Buell P. 1973. Changing Incidence of breast cancer in Japanese-American women. J Natl Cancer Inst 51(5):1479-83.

Cepko CL, Roberts BE,Mulligan RC. 1984. Construction and application of a highly transmissible murine retrovirus shuttle vector. Cell 37:1053-62.

Chang W, Hubbard SC, Friedel C,Ruley HE. 1993. Enrichment of insertional mutants following retrovirus gene trap selection. Virology 193:737-47.

Chen C, Okayama H. 1987. High effecieney transformation of mammalian cells by plasmid DNA. Mol Cell Biol 7:2745-2752.

Chen Z, Friedrich GA,Soriano P. 1994. Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. Genes & Dev 8:2293-301.

Claus EB, Risch N,Thompson WD. 1991. Genetic analysis of breast cancer in the cancer and steroid hormone study. Am J Hum Genet 48:232-42.

Compere SJ, Baldacci P, Sharpe AH, Thompson T, Land H,Jaenisch R. 1989a. The ras and myc oncogenes cooperate in tumor induction in many tissues when introduced into midgestation mouse embryos by retroviral vectors. Proc Natl Acad Sci 86:2224-8.

Compere SJ, Baldacci PA, Sharpe AH,Jaenisch R. 1989b. Retroviral transduction of the human c-Ha-ras-1 oncogene into midgestation mouse embryos promotes rapid epithelial hyperplasia. Mol Cell Biol 9(1):6-14.

Cornelis RS, van Vliet M, Vos CBJ, Cleton-Jansen A-M, van de Vijver MJ, Peterse JL, et al. 1994. Evidence for a gene on 17p13.3, distal to TP53, as a target for allele loss in breast tumors without p53 mutations. Cancer Res 54:4200-6.

Cropp CS, Nevanlinna HA, Pyrhonen S, Stenman U-H, Salmikangas P, Albertsen H, et al. 1994. Evidence for involvement of BRCA1 in sporadic breast carcinomas. Cancer Res 54:2548-51.

Cross SH, Bird AP. 1995. CpG islands and genes. Curr Biol 5:309-14.

Daniel CW, Silberstein GB. Postnatal development of the rodent mammary gland. The Mammary Gland. ; Neville MC, Daniel CW, editors.New York, NY: Plenum, 1987;

Davis RL, Weintraub H,Lassar AB. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51:987-1000.

Deng JM, Behringer RR. 1995. An insertional mutation in the BTF3 transcription factor gene. Transgenic Res 4:264-9.

Dickson RB, Kasid A, Huff KK, Bates SE, Knabbe C, Bronzert D, et al. 1987. Activation of growth factor secretion in tumorigenic states of breast cancer induced by 17 -estradiol or v-Ha-ras oncogene.. Proc Natl Acad Sci 84:837-41.

Easton DF, Bishop DT, Ford D,Crockford GP. 1993. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consorium. Am J Hum Genet 52:678-701.

Erickson J, ar-Rushdi A, Drwinga HL, Nowell PC,Croce CM. 1983. Transcriptional activation of the translocated c-myc oncogene in Burkitt lymphoma. Proc Natl Acad Sci 80:820-4.

Escot C, Theillet C, Lidereau R, Spyratos F, Champeme M-H, Gest J, et al. 1986. Genetic alteration of the c-myc protooncogene (MYC) in human primary breast carcinomas.. Proc Natl Acad Sci 83:4834-8.

Fitzpatrick SL, Brightwell J, Wittliff JL, Barrows GH,Schultz GS. 1984. Epidermal growth factor binding by breast tumor biopsies and relationship to estrogen receptor and progestin receptor levels. Cancer Res 44:3448-53.

Friedrich G, Soriano P. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes & Dev 5:1513-23.

Frohman MA, Dush MK,Martin GR. 1988. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85:8998-9002.

Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, et al. 1994. BRCA1 mutations in primary breast and ovarian carcinomas. Science 266:120-2.

Garcia I, Dietrich P-Y, Aapro M, Vauthier G, Vadas L,Engel E. 1989. Genetic alterations of c-myc,c-erbB2, and c-Ha-ras protooncogenes and clinical associations in human breast carcinomas. Cancer Res 49:6675-9.

Gasca S, Hill DP, Klingensmith J,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 Genet 17:141-54.

Goff S, Traktman P,Baltimore D. 1981. Isolation and properties of Moloney Murine Leukemia Virus mutants: use of a rapid assay for release of virion reverse transcriptase. J Virol 38(1):239-48.

Goldfarb M, Shimizu K, Perucho M,Wigler M. 1982. Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. Nature 296:404-9.

Gossler A, Doetschman T, Korn R, Serfling E,Kemler R. 1986. Transgenesis by means of blastocyst-derived embryonic stem cell lines. Proc Natl Acad Sci USA 83:9065-9.

Gossler A, Joyner AL, Rossant J,Skarnes WC. 1989. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science 244:463-5.

Graham GL, van der Eb AJ. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52:456-67.

Gridley T, Soriano P,Jaenisch R. 1987. Insertional mutagenesis in mice. TIG 3(6):162-6.

Gridley T, Gray DA, Orr-Weaver T, Soriano P, Barton DE, Francke U, et al. 1990. Molecular analysis of the Mov 34 mutation: transcript disrupted by proviral integration in mice is conserved in Drosophila. Development 109:235-42.

Gudas JM, Nguyen H, Li T,Cowan KH. 1995. Hormone-dependent regulation of BRCA1 in human cancer cells. Cancer Res 55:4561-5.

Guy CT, Cardiff RD,Muller WJ. 1992. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol Cell Biol 12(3):954-61.

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, et al. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21.. Science 250:1684-9.

Hanna Z, Jankowski M, Tremblay P, Jiang X, Milatovich A, Francke U, et al. 1993. The Vin-1 gene, identified by provirus insertional mutagenesis, is the cyclin D2. Oncogene 8:1661-6.

Harbers K, Kuehn M, Delius H,Jaenisch R. 1984. Insertion of retrovirus into the first intron of 1(I) colagen gene leads to embryonic lethal mutation in mice. Proc Natl Acad Sci 81:1504-8.

Harris JR, Lipman ME, Veronesi U,Willett W. 1992a. Breast Cancer (first of three parts). N Engl J Med 327(5):319-28.

Harris JR, Lippman ME, Veronesi U,Willett W. 1992b. Breast cancer (second of three parts). N Engl J Med 327(7):390-8.

Harris JR, Lippman ME, Veronesi U,Willett W. 1992c. Breast cancer (third of three parts). N Engl J Med 327(7):473-9.

Haubruck H, McCormick F. 1991. Ras p21: effects and regulation. Biochem Biophys Acta 1072:215-29.

Hines SJ, Organ C, Kornstein MJ,Krystal GW. 1995. Coexpression of the c-kit and stem cell factor genes in breast carcinomas. Cell Growth & Differ 6:769-79.

Hollstein M, Sidransky D, Vogelstein B,Harris CC. 1991. p53 mutations in human cancers. Science 253:49-53.

Holmes DS, Quigley M. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal Biochem 114:193-7.

Holt JT, Thompson ME, Szabo C, Robinson-Benion C, Arteaga CL, King M-C, et al. 1996. Growth retardation and tumor inhibition by BRCA1. Nat Genet 12:298-302.

Huszar D, Ellis J,Bernstein A. 1989. A rapid and accurate method for determining the titer of retrovirus vectors lacking a selectable marker. Technique 1(1):28-32.

Huttner WB, Gerdes H-H,Rosa P. 1990. The granin (chromogranin secretogranin) family. TIBS 16:27-30.

Jacobson A. Purification and fractionation of Poly(A)+ RNA. Guide to Molecular Cloning Techniques. : Berger SL, editors.London: Academic Press, 1987:

Jensen RA, Thompson ME, Jetton TL, Szabo CI, van der Meer R, Helou B, et al. 1996. BRCA1 is secreted and exhibits properties of a granin. Nat Genet 12:303-8.

Jones N. 1990. Transcriptional regulation by dimerization: two sides to an incestuous relationship. Cell 61:9-11.

Joyner AL, Martin GR. 1987. En-1 and En-2, two mouse genes with sequence homology to the Drosophila engrailed gene: expression during embryogenesis. Genes & Dev 1:29-38.

Kato GJ, Lee WMF, Chen L,Dang CV. 1992. Max: functional domains and interaction with c-Myc. Genes & Dev 6:81-92.

Keely PJ, Wu JE,Santoro SA. 1995. The spatial and temporal expression of the 2 1 integrin and its ligands, collagen I, collagen IV, and laminin, suggest important roles in mouse mammary morphogenesis. Differentiation 59:1-13.

Kim JW, Closs EI, Albritton LM,Cunningham JM. 1991. Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. Nature 352:725-8.

King CR, Kraus MH,Aaronson SA. 1985. Amplification of a novel v-erbB-related gene in a humam mammary carcinoma. Science 229:974-6.

Kozak CA, Ruscetti S. Retroviruses in Rodents. The Retroviridae. AnonymousNew York, NY, Plenum, 1992;

Kozak M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucl Acids Res 15(20):8125

Lammie GA, Fantl V, Smith R, Schuuring E, Brookes S, Michalides R, et al. 1991. D11S287, a putative oncogene on chromosome 11q13, is amplified and expressed in squamous cell and mammary carcinomas and linked to BCL-1.. Oncogene 6:439-44.

Land H, Parada LF,Weinberg RA. 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature 304:596-602.

Leder A, Pattengale PK, Kuo A, Stewart TA,Leder P. 1986. Consequences of widespread deregulation of the c-myc gene in transgenic mice: multiple neoplasms and normal development. Cell 45:485-95.

Li M, Hu J, Heermeier K, Hennighausen L,Furth PA. 1996. Apoptosis and remodeling of mammary gland tissue during involution proceeds through p53-independent pathways. Cell Growth & Differ 7:13-20.

Liang P, Pardee AB. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967-71.

Lobel LI, Patel M, King W, Nguyen-Huu MC,Goff SP. 1985. Construction and recovery of viable retroviral genomes carrying a bacterial suppressor transfer RNA gene. Science 228:329-32.

Lundy J, Schuss A, Stanick D, McCormack ES, Kramer S,Sorvillo JM. 1991. Expression of neu protein, epidermal growth factor receptor, and transforming growth factor alpha in breast cancer: Correlation with clinicopathologic parameters. Am J Pathol 138(6):1527-34.

Malkin D, Li FP, Strong LC, Fraumeni JF, Nelson CE, Kim DH, et al. 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233-8.

Mann R, Mulligan RC,Baltimore D. 1996. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33:153-9.

Mariani-Costantini R, Escot C, Theillet C, Gentile A, Merlo G, Lidereau R, et al. 1988. In situ c-myc expression and genomic status of the c-myc locus in infiltrating ductal carcinomas of the breast.. Cancer Res 48:199-205.

Markowitz D, Goff S,Bank A. 1988. A safe packaging line for gene transfer: separating viral genes on two different plasmids. J Virol 62(4):1120-4.

Marquis ST, Rajant JV, Wynshaw-Boris A, Xu J, Yin G-Y., Abel KJ, et al. 1995. The developmental pattern of Brca1 expression implies a role in differentiation of the breast and other tissues.. Nat Genet 11:17-26.

Marti A, Jehn B, Costello E, Keon N, Ke G, Martin F, et al. 1994. Protein kinase A and AP-1 (c-Fos/JunD) are induced during apoptosis of mouse mammary epithelial cells.. Oncogene 9:1213-23.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci 78(12):7634-8.

Matsuda R, Rakahashi T, Nakamura S, Sekido Y, Nishida K, Seto M, et al. 1993. Expression of the c-kit protein in human solid tumors and in corresponding fetal and adult normal tissues.. Am J Pathol 142(1):339-46.

McBurney MW, Sutherland LC, Adra CN, Leclair B, Rudnicki MA,Jardine K. 1991. The mouse Pgk-1 gene promoter contains an upstream activator sequence. Nucl Acids Res 19(20):5755-61.

McBurney MW. 1993. P19 embryonal carcinoma cells. Int J Dev Biol 37:135-40.

McMichael AJ, Giles GG. 1988. Cancer in migrants to Australia: Extending the descriptive epidemiological data. Cancer Res 48:751-6.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.. Science 266:66-71.

Moghul A, Lin L, Beedle A, Kanbour-Shakir A, DeFrances MC, Liu Y, et al. 1994. Modulation of c-MET proto-oncogene (HGF receptor) mRNA abundance by cytokines and hormones: evidence for rapid decay of the 8 kb c-MET transcript.. Oncogene 9:2045-52.

Muller WJ, Sinn E, Pattengale PK, Wallace R,Leder P. 1988. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell 54:105-15.

Natali PG, Nicotra MR, Sures I, Mottolese M, Botti C,Ullrich A. 1992a. Breast cancer is associated with loss of the c-kit oncogene product. Int J Cancer 52:713-7.

Natali PG, Nicotra MR, Sures I, Santoro R, Bigotti A,Ullrich A. 1992b. Expression of c-kit receptor in normal and transformed human nonlymphoid tissues. Cancer Res 52:6139-43.

Neuhausen SL, Marshall CJ. 1994. Loss of heterozygosity in familial tumors from three BRCA1-linked kindreds. Cancer Res 54:6069-72.

Niranjan B, Buluwela L, Yant J, Perusinghe N, Atherton A, Phippard D, et al. 1995. HGF/SF: a potent cytokine for mammary growth, morphogenesis and development.. Development 121:2897-908.

Okazaki S, Tanase S, Choudhury BK, Setoyama K, Miura R, Ogawa M, et al. 1994. A novel nuclear protein with zinc fingers down-regulated during early mammalian cell differentiation. J Biol Chem 269:6900-7.

Pietras RJ, Arboleda J, Reese DM, Wongvipat N, Pegram MD, Ramos L, et al. 1995. HER-2 tyrosine kinase pathway targets estrogen receptor and promotes hormone-independent growth in human breast cancer cells. Oncogene 10:2435-46.

Prendergast GC, Lawe D,Ziff EB. 1991. Association of myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell 65:395-407.

Prendergast GC, Ziff EB. 1991. Methylation-sensitive sequence-specific DNA binding by the c-myc basic region. Science 251:186-9.

Price J, Turner D,Cepko C. 1987. Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. Proc Natl Acad Sci 84:156-60.

Reddy CD, Dagupta P, Saikumar P, Dudek H, Rauscher III FJ,Reddy EP. 1992. Mutational analysis of Max: role of basic, helix-loop-helix/leucine zipper domains in DNA binding, dimerization and regulation of Myc-mediated transcriptional activation. Oncogene 7:2085-92.

Reddy S, DeGregori JV, von Melchner H,Ruley HE. 1991. Retrovirus promoter-trap vector to induce lacZ gene fusions in mammalian cells. J Virology 65(3):1507-15.

Reik W, Weiher H,Jaenisch R. 1985. Replication-competent Moloney murine leukemia virus carrying a bacterial suppressor tRNA gene: Selective cloning of proviral and flanking host sequences. Proc Natl Acad Sci 82:1141-5.

Robinson GW, McKnight RA, Smith GH,Hennighausen L. 1995. Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. Development 121:2079-90.

Rohdewohld H, Weiher H, Reik W, Jaenisch R,Breindl M. 1987. Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites. J Virology 61(2):336-43.

Rosen JM. Milk protein gene structure and expression. The Mammary Gland. : Neville MC, Daniel CW, editors.New York, NY: Plenum Press, 1987;

Rowe WP, Pugh WE,Hartley JW. 1970. Plaque Assay Techniques for Murine Leukemia Viruses. Virology 42:1136-9.

Sablitzky F, Jonsson J-I, Cohen BL,Phillips RA. 1993. High frequency expression of integrated proviruses derived from enhancer trap retroviruses. Cell Growth and Differ 4:451-9.

Sainsbury JEC, Malcolm AJ, Appleton DR, Farndon JR,Harris AL. 1985. Presence of epidermal growth factor receptor as an indicator of poor prognosis in patients with breast cancer. J Clin Pathol 38:1225-8.

Sainsbury JRC, Farndon JR, Sherbet GV,Harris AL. 1985. Epidermal-growth-factor receptors and oestrogen receptors in human breast cancer. Lancet :364-6.

Sainsbury JRC, Fardon JR, Needham GK, Malcolm AJ,Harris AL. 1987. Epidermal-growth-factor receptor status as predictor of early recurrence of and death from breast cancer. Lancet :1398-402.

Sakakura T. Mammary embryogenesis. The Mammary Gland. ; Neville MC, Daniel CW, editors.New York, NY: Plenum, 1987;

Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A laboratory manual. second ed.; Nolan C. editor. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1989.

Satoh T, Nakafuku M,Kaziro Y. 1992. Function of ras as a molecular switch in signal transduction. J Biol Chem 267(34):24149-52.

Scherdin U, Rhodes K,Breindl M. 1990. Transcriptionally active genome regions are prefered targets for retrovirus integration. J Virol 64(2):907-12.

Schoenenberger C, Andres A, Groner B, van der Valk M, LeMeur M,Gerlinger P. 1988. Targeted c-myc gene expression in mammary glands of transgenic mice induces mammary tumors with constitutive milk protein gene transcription. EMBO Journal 7(1):169-75.

Schubach W, Groudine M. 1984. Alteration of c-myc chromatin structure by avian leukosis virus integration. Nature 307:702-8.

Schuuring E, Verhoeven E, Mooi WJ,Michalides RJAM. 1992. Identification and cloning of two overexpressed genes, U21B31/PRAD1 and EMS1, within the amplified chromosome 11q13 region in human carcinomas. Oncogene 7:355-61.

Silver J, Keerikatte V. 1989. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. J Virology 63(5):1924-8.

Sinn E, Muller W, Pattengale P, Tepler I, Wallace R,Leder P. 1987. Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. Cell 49:465-75.

Skarnes WC, Auerbach BA,Joyner AL. 1992. A gene trap approach in mouse embryonic stem cells: the lacZ reporter is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. Genes & Dev 6:903-18.

Slamon DJ, deKernion JB, Verma IM,Cline MJ. 1984. Expression of cellular oncogenes in human malignancies. Science 224:256-62.

Slamon DJ, Godolphin.W., Jones LA, Holt JA, Wong SG, Keith DE, et al. 1989. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science 244:707-12.

Smith GH, Gallahan D, Zwiebel JA, Freeman SM, Bassin RH,Callahan R. 1991. Long-Term In Vivo Expression of Genes Introduced by Retrovirus-Mediated Transfer into Mammary Epithelial Cells. J Virol 65(11):6365-70.

Sorensen AB, Duch M, Jorgensen P,Pedersen FS. 1993. Amplification and sequence analysis of DNA flanking integrated proviruses by a simple two-step polymerase chain reaction method. J Virology 67(12):7118-24.

Statistiques canadiennes sur le cancer. Institut National du Cancer du Canada. 1996.

Stewart TA, Pattengale PK,Leder P. 1984. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MMTV myc fusion genes. Cell 38:627-37.

Strange R, Li F, Saurer S, Burkhardt A,Friis RR. 1992. Apoptotic cell death and tissue remodeling during mouse mammary gland involution. Development (115):49-58.

Stuhlmann H, Jaenisch R,Mulligan RC. 1989. Construction and Properties of Replication-Competent Murine Retroviral Vectors Encoding Methotrexane Resistance. Mol Cell Biol 9(1):100-8.

Takeuchi T, Yamazaki Y, Katoh-Fukui Y, Tsuchiya R, Kondo S, Motoyama J, et al. 1995. Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation.. Genes & Dev 9:1211-22.

Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, et al. 1982. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells.. Proc Natl Acad Sci 79:7834-41.

Tavtigian SVea. 1996. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet 12:333-7.

Teixeirs C, Reed JC,Pratt MAC. 1995. Estrogen promotes chemotherapeutic drug resistance by a mechanism involving Bcl-2 proto-oncogene expression in human breast cancer cells. Cancer Res 55:3902-7.

Thompson ME, Jensen RA, Obermiller PS, Page DL,Holt JT. 1995. Decreased expression of BRCA1 accelerates growth and is often present during sporadic breast cancer progression. Nat Genet 9:444-50.

Treisman R, Novak U, Favaloro J,Kamen R. 1981. Transformation of rat cells by an altered polyoma virus genome expressing only the middle-T protein. Nature 292:595-600.

Tremblay JP, Pothier F, Hoang T, Tremblay G, Brownstein S, Liszauer A, et al. 1989. Transgenic mice carrying the mouse mammary tumor virus ras fusion gene: distinct effects in various tissues.. Mol Cell Biol 9(2):854-9.

van de Vijver M, van de Bersselaar R, Devilee P, Cornelisse C, Peterse J,Nusse R. 1987. Amplification of the neu (c-erbB-2) oncogene in human mammary tumors is relatively frequent and is often accompanied by amplification of the linked c-erbA oncogene. Mol Cell Biol 7(5):2019-23.

van de Vijver MJ, Peterse JL, Mooi WJ, Wisman P, Lomans J, Dalesio O, et al. 1988. Neu-protein overexpression in breast cancer: Association with comedo-type ductal carcinoma in situ and limited prognostic falue in Stage II breast cancer. N Engl J Med 319(19):1239-45.

Varley JM, Swallow JE, Brammar WJ, Whittaker JL,Walker RA. 1987. Alterations to either c-erbB2 (neu) or c-myc proto-oncogenes in breast carcinomas correlate with poor short-term prognosis. Oncogene 1:423-30.

Vaulont S, Daines S,Evans M. 1995. Disruption of the adenosine deaminase (ADA) gene using a dicistronic promoterless construct: production of an ADA-deficient homozygote ES cell line. Transgenic Res 4:247-55.

Vijaya S, Steffen DL,Robinson HL. 1986. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. J Virol 60(2):683-92.

von Melchner H, Reddy S,Ruley HE. 1990. Isolation of cellular promoters by using a retrovirus promoter trap. Proc Natl Acad Sci 87:3733-7.

von Melchner H, DeGregori JV, Rayburn H, Reddy S, Friedel C,Ruley HE. 1992. Selective disruption of genes expressed in totipotent embryonal stem cells. Genes & Dev 6:919-27.

von Melchner H, Ruley HE. 1989. Identification of cellular promoters by using a retrovirus promoter trap. J Virol 63(8):3227-33.

Wang B, Kennan WS, Yasukawa-Barnes J, Lindstrom MJ,Gould MN. 1991. Carcinoma induction following direct in situ transfer of v-Ha-ras into rat mammary epithelial cells using replication-defective retrovirus vectors. Cancer Res 51:2642-8.

Wang H, Kavanaugh MP, North RA,Kabat D. 1991. Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. Nature 352:729-31.

Wang TC, Cardiff RD, Zukerberg L, Lees E, Arnold A,Schmidt EV. 1994. Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. Nature 369:669-71.

Weinstat-Saslow D, Merino MJ, Manrow RE, Lawrence JA, Bluth RF, Wittenbel KD, et al. 1995. Overexpression of cyclin D mRNA distinguishes invasive and in situ breast carcinomas from non-malignant lesions.. Nat Med 1:1257-60.

Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, et al. 1994. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.. Science 265:2088-90.

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. 1995. Identification of the breast cancer susceptibility gene BRCA2. Nature 378:789-92.

Wurst W, Rossant J, Prideaux V, Kownacka M, Joyner A, Hill DP, et al. 1995. A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice.. Genetics 139:889-99.

Yaswen P, Smoll A, Peehl DM, Trask DK, Sager R,Stampfer MR. 1990. Down-regulation of a calmodulin-related gene during transformation of human mammary epithelial cells. Proc Natl Acad Sci 87:7360-4.

Yokota J, Tsunetsugu-Yokota Y, Battifora H, Le Fevre C,Cline MJ. 1986. Alterations of myc,myb, and ras^{Ha} proto-oncogenes in cancers are frequent and show clinical correlation. Science 231:261-5.

Zhou D, Battifora H, Yokota J, Yamamoto T,Cline MJ. 1987. Association of multiple copies of the c-erbB-2 oncogene with spread of breast cancer. Cancer Res 47:6123-5.

Zukerberg LR, Yang WI, Gadd M, Thor AD, Koerner FC, Schmidt EV, et al. 1995. Cyclin D1 (PRAD1) protein expression in breast cancer: approximately one-third of infiltrating mammary carcinomas show overexpression if the cyclin D1 oncogene.. Mod Pathol 8:560-7.

CURRICULUM VITAE

NAME: Josée Coulombe
DATE OF BIRTH: October 11, 1964
NATIONALITY: Canadian
LANGUAGES: French, English

EDUCATION

1984-1987 McGill University, B.Sc. Major in Biochemistry.

1982-1984 CEGEP Bois-de-Boulogne, DEC in Health Sciences.

SCIENTIFIC BACKGROUND

1990-present University of Ottawa, PhD Biochemistry: Identification of a novel gene, TRAP1, using a replication-competent promoter-trap retrovirus in the mouse mammary gland, laboratory of Dr. Douglas A. Gray.

1987-1990 University of Guelph, M.Sc. Molecular Biology and Genetics, with honour. Transcriptional activation of Simian Virus 40 by large T antigen, laboratory of Dr. Alan G. Wildeman.

AWARDS

1991-1993 Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR).

1991-1993 University of Ottawa, supplement to scholarship.

PUBLICATIONS

Josée Coulombe, Yvonne Avis, and Douglas A. Gray. (1996) A Replication Competent Promoter-trap Retrovirus. [J. Virol., in press].

Josee Coulombe, Manon Dube, and Douglas A. Gray. (1994) Replication of a Moloney Leukemia Virus Mutant Lacking the LTR Binding Site. Virology 203:362-368.

Heather L. Davis, Barbara A. Demeneix, Beatrice Quantin, Josee Coulombe, and Robert H Whalen. (1993) Plasmid DNA is Superior to Viral Vectors for Direct Gene Transfer into Adult Mouse Skeletal Muscle. Human Gene Therapy 4:733-740.

Josée Coulombe, Lloyd Berger, Donald B. Smith, Robert K. Hehl, and Alan G. Wildeman (1992) Activation of Simian Virus 40 Transcription In Vitro by T Antigen. J. Virol 66:4591-4596.

ABSTRACTS

Josée Coulombe and Douglas A. Gray. Identification of Active Promoters in Vivo Using a Replication-competent Promoter-trap Retrovirus. (1995) Fifteen Annual Great Lakes Mammalian Development Meeting, March 31-April 2, Mount Sinai, Toronto. Speaker.

Josée Coulombe and Douglas A. Gray. A Mammary Promoter-trap. (1994) Molecular Genetics of Cancer, June 1-8, Cold Spring Harbor Laboratory, NY. Poster.

Josée Coulombe and Douglas A. Gray. Identification of Genes Involved in Mammary Development Using a Promoter-Trap Retrovirus. (1993) New Directions in Cancer Research: A Symposium in Honour of Peter Scholefield, September 19-22, Honey Harbor, Ontario. Poster.

Josée Coulombe and Douglas A. Gray. Identification of Genes Involved in Mammary Development Using a Promoter-Trap Retrovirus. (1992) Genetics and Molecular Biology of Breast Cancer, September 2-6, Cold Spring Harbor Laboratory, NY. Poster.

Josée Coulombe and Douglas A. Gray. Identification of Genes Involved in Mammary Development Using a Promoter-Trap Retrovirus. (1992) Twelfth Annual Great Lakes Mammalian Development Meeting, April 10-12, Mount Sinai, Toronto. Poster

Josée Coulombe and Douglas A. Gray. Identification of Genes Involved in Mammary Development Using a Promoter-Trap Retrovirus. (1991) The 2nd Eastern Canadian Conference on Development and Cancer, September 23-25, Montreal, Quebec. Poster

Josée Coulombe and Alan G. Wildeman. Activation of Simian Virus 40 Transcription In Vitro by T Antigen. (1990) Molecular Biology of SV40, Polyoma, and Adenoviruses meeting, August 15-19, Cold Spring Harbor Laboratory, NY. Presentation.