ISOLATION AND CHARACTERIZATION OF A PUTATIVE ENHANCER SEQUENCE FLANKING A POLYOMAVIRUS TRANSGENE.

Ramez Fahmy

Microbiology & Immunology
McGill University
Montreal, Canada.

December 1995

A Thesis submitted to the Faculty of
Graduate Studies and Research
in partial fulfillment of the requirements of the degree of
Master's of Science

ISBN C-612-12190-9

Canada

ABSTRACT

Our laboratory has produced six lines of mice harbouring the polyomavirus Large T-antigen (PVLT) transgene. Mice of all lines express PVLT in testis and three lines develop testicular adenomas. However, mice of a single line, designated MTPVLT-8, express PVLT exclusively in testis and cardiac tissue. These mice develop cardiac hypertrophy and die prematurely. We hypothesize that expression of PVLT in cardiac cells is a concequence of trangene integration near a cardiac specific enhancer element. To support this hypothesis, DNA flanking the transgene was isolated and tested for inclusion of any possible regulatory elements. A restriction map of genomic DNA surrounding the trangene integration site was constructed as a starting point. A lambda clone containing 5-10 Kb of the 5' flanking DNA was isolated and used for a sequence of experiments. This DNA sequence did not cross-hybridize with DNA isolated from other species and appears to be mouse specific. Northern analysis demonstrate a sequence homology to a single RNA transcritp appearing to be expressed in a wide variety of mouse tissues. A gene-bank comparison on a sequenced part of the clone resulted in an 89% sequence homology match with human gualadyn gene sequence. Although a regulatory element was not evident in the preliminary work, we are now in a position to attemp cloning various parts of the clone into expression vectors containing easily assayable reporter genes. Due to lack of appropriate cell lines, in vivo experimentation must be performed.

RESUME

Six lignées de souries possédant un transgène pour l'antigène grand T du virus du polyôme(PVLT) ont été produites dans notre laboratoire. Les souries expriment le transgène dans les testicules et trois de ces lignées ont développées un adénôme testiculaire. Cependant, une lignée (MTPVLT-8) exprime le PVLT exclusivement dans les testicules et le muscle cardiaque. Ces dernières ont développées une hypertrophie du coeur et sont décédée prématurément.Nous émettons comme hypothèse que l'expression du PVLT dans les cellules cardiaques est dû à l'intégration du transgène près d'une séquence enhancer cardiaque. Afin de vérifier cette hypothèse, l'ADN entourant le transgène a été isolé pour identifier l'intégration possible d'un élément de régulation. Comme point de départ, Une carte de restriction de l'ADN génômique entourant le site d'intégration du transgène a été construite.Un clone lambda, contenant de 5 à 10 Kb de la région 5' flanquante de l'ADN, a été isolé et utilisé pour des expériences de séquence. Cette séquence d'ADN n'hybride avec aucun autre ADN isolé de différentes espèces et semble donc être spécifique à la sourie. Des analyses Northern ont démontré une homologie de séquence pour un seul transcrit qui s'exprime dans une grande variété de tissus de souries. Le séquençage d'une partie de ce clone a démontré une homologie de 89% avec le gène gualadyn humain suite à une recherche sur "GENE-BANK". Bien qu'un élément régulateur n'ait pas été mis en évidence dans ces travaux préliminaires, nous planifions cloner différentes parties de ce clone dans un vecteur d'expression qui contiennent des gènes rapporteurs. En raison d'un manque de lignées cellulaires adéquates, des expériences in-vivo devront être entreprises.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank Dr. Lorraine Chalifour for her guidance throughout this project, and appreciate the freedom she allowed me to explore my ideas no matter how crazy some of them were.

I thank all my fellow co-workers over the years for their help and wisdom, and my summer students Jack and Jessie for their help on the project. A special thanks goes to Harvey for his work in isolating Clone-6.

Finally, I thank mom and dad for their help and support throughout my educational career, and all the times they told me to just do my best when the going got rough.

CHAPTER 1

INTRODUCTION

INTRODUCTION

In recent years genetic screening and gene therapy are slowly becoming components in preventing and treating diseases. Research today is focused on understanding the mechanisms of genetic control on growth and development.

The study of transcriptional regulation is at the heart of this research. From the moment we are conceived to the time we die our bodies are constantly undergoing genetic change. Although all cells in our bodies contain identical chromosomes, the genes they express are significantly different. These differences result in the formation and maintenance of the diverse organs of the human body.

Transcriptional regulation, mediated by promoters, enhancers, and cellular binding factors, is responsible for the normal life cycle of all living cells. By identifying and understanding the principles of gene regulation we will be much closer to developing therapies for cancer and illness resulting from deregulation of normally expressed genes.

The use of transgenic animals to study transcriptional regulation has been extremely valuable in identifying transcriptional elements such as promoters, enhancers and cellular binding factors and their possible roles.

In our laboratory we developed a transgenic mouse line (MTPVLT-8) containing the polyomavirus large-T antigen (PVLT) driven by the mouse metallothionein promoter. These mice develop a cardiomyopathy resulting in premature death. This phenotype is the direct result of PVLT expression in heart tissue. Evidence suggests that expression in the heart is due to integration of the transgene, containing the PVLT, close to a cardiac-specific endogenous enhancer.

A cardiac-specific enhancer can be a very important tool used to direct expression of genes exclusively in the heart tissue, with major implications to possible gene therapies directed to cardiac abnormalities.

1.1: Formation of transgenic animals

Techniques for introducing functional genes into host organisms have proven to be powerful tools for probing complex biological processes. The use of gene transfer is invaluable for studying long lived diploid organisms where classical genetic approaches are impractical. Over the past 24 years a variety of methods have been developed to introduce foreign DNA into both somatic and germ cells of mammals. These techniques include direct microinjection of foreign DNA sequences into fertilized eggs, direct delivery of a transgene into the blastecoel cavity, and retroviral infection or transfection of transgene sequences into embryonic stem cells (Hogen et al. 1986, Jaenisch 1988). Direct microinjection of DNA into the pronucleus has been the method most extensively used. Microinjection was first attempted in the late 1970s when existing methods were adapted for microinjection of mRNA, and later DNA, into mouse eggs (Brinster et al. 1980, 1981a). The first successful microinjection of DNA leading to stable transgenic mice was published in the latter part of 1980 (Gordon et al. 1980) and within a few months four other research teams reported similar success in integrating foreign DNA into the mouse genome. It was shown that in some cases the foreign genes could be expressed (Brinster et al. 1981b, Wagner et al. 1981) and that not only were the genes incorporated into somatic tissues, but were also in the germ line (Constantini et al. 198 1,Gordon et al. 1981, Stewart et al. 1982). Moreover, the transgenes were transmitted and expressed in the offspring of founder animals in a stable manner (Palmiter et al. 1982).

The offspring constitute a transgenic line, whereby each animal carrying the transgene is identical to the founder animals in that it contains the transgene in the exact chromosomal location as its predecessor.

1.1.1: Pronuclear microinjection

Pronucliar microinjection of cloned DNA into fertilized eggs has become the most prevalent technique for introducing genes into the germline. Thousands of transgenic mice, as well as a few transgenic sheep, goats, pigs, rats, and rabbits have been produced by this method (Hammer et al. 1985). One of the great advantages of this technique is that any cloned DNA can be used. Injection of linear DNA into the pronuclei yields about 25% newborn mice which carry one or more copies of the injected DNA (Brinster et al. 1985). The term "transgenic" refers to those animals that have stable integration of a foreign gene, known as "transgenes", into their genomes. Since transgene integration often occurs prior to DNA replication in the one-cell embryo, about 70% of founder transgenic animals carry the transgene in all their cells, including the germ cells. Integration occurs after one or more rounds of replication in the remaining 30% of the mice; hence, only a fraction of the cells in the founder animals contain the transgene and the animals are chimeric.

The exact mechanism of integration of microinjected DNA is still unknown. However, the injected DNA often integrates as multiple copies in a tandem head-to-tail array at a single chromosomal site (Gridley et al. 1987). However, other arrangements have been documented (Hammer et al. 1985, Palmiter et al. 1983a). There is no evidence for site directed integration (Brinster et al. 1985) and many believe that the integration event occurs randomly in the chromosomes. Integration occurs on many different autosomes (Krumlauf et al. 1985, Lacy et al. 1983) as well as sex linked chromosomes (Lo et al. 1986). One hypothesis for the integration mechanism (Brinster et al. 1985) is that randomly generated chromosomal breaks are the rate-limiting step for integration which may explain the predominance of single integration sites. At these breaks the injected DNA molecules initiate integration which may account for the observed fivefold increase in integration efficiency with linear vs. circular DNA. Many believe that homologous recombination may occur among the injected molecules explaining the prevalence of tandem arrays, however, at present no evidence of homologous

recombination between injected and host DNA sequences has been found, possibly as a result of differences in chromatin structures between the injected and chromosomal DNAs. Although this is a plausible mechanism, no substantial evidence has been found to support it.

Due to the random nature of DNA integration, different lines of mice isolated from the same injection experiment, and injected with the same foreign DNA fragment, are genotypically and often phenotypically different. In general, the DNA element integrates at different chromosomal locations and in various copy numbers. The integration complexes differ not only in copy number, but some may contain inverted repeats or different sorts of fusions. These rearangements can include internal portions of two copies of the foreign DNA (rearrangements). Junction points between the array and the chromosomes generally involve sequences internal to the foreign DNA (Covarrubias et al. 1986, 1987; Wilkie et al. 1987). In some cases arrays may contain DNA fragments from elsewhere in the genome or from unknown sources (Wilkie et al. 1987). Finally, small mutations such as deletions or nucleotide substitutions my occur during the integration process. In general, once the DNA has successfully integrated, stable transmition for many generations is possible with no evidence of rearrangements; although in a few cases, partial deletions or amplifications were reported (Wilkie et al. 1987)

1.2: Expression of foreign genes in transgenic animals

The transfer of foreign DNA to produce transgenic animals has been extensively documented. In the majority of gene transfers in mammals the foreign DNA is expressed in the host. A large number of diverse transgenes have been expressed in mice including DNA from chickens, *Drosophila*, humans, pigs, rats, rabbits as well as viral and bacterial DNA. Some examples of transgene expression in mice include the rat α-actin gene in skeletal and cardiac muscle, the human and rabbit beta-globin gene expressed in erythroid cells, and the human insulin gene expressed in β-cells (Ottesen et al. 1994). In the

majority of cases these exogenous DNAs are slightly modified to facilitate their identification from their endogenous counterparts. The expression of such genes have also shown diversity in their pattern of expression. They are either expressed in many tissues, localized tissues or in cells of a specific lineage.

By producing transgenic mice containing genes of interest and then assaying the various tissues for mRNA or protein, it becomes possible to elucidate the necessary genetic elements for tissue-specific gene expression. It is, however, important to modify the integrated gene to identify its products over the endogenous counterpart. Modifications may include use of genes from other species, or modification where a few nucleotides have been inserted or deleted.

The use of "reporter" genes is also very common. In this technique hybrid genes are constructed whereby control elements of genes of interest are linked to a reporter gene and regulate its expression. It is then possible to assay for the reporter gene that codes for an easily identifiable enzyme which is not endogenous to mice, a hormone which produces pronounced physiological effects, a polypeptide that is easily identifiable immunologically, or an oncogene that could lead to tumors in the cell types where it is expressed.

1.2.1: Gene expression in cells and tissues

Some genes are expressed in many cell types. Examples include genes transcribed from the mouse metalothionein (MT-1), collagen, and various viral promoters. Genes which are normally expressed in several cell types have also been tested in transgenic mice. Examples include the alpha-fetoprotein gene which is initially induced in yolk sac, fetal liver, and gastrointestinal cells but later expression declines in all of these cell types. The transferrin gene, expressed predominantly in liver, testis, and brain is another example of a gene thats expression is restricted to a few cell types (McKnight et al. 1983). In both cases these genes have been expressed in their appropriate cells in transgenic

mice. Other genes normally expressed in a single cell type have been studied. The immunoglobulin K gene (Brinster et al. 1983, Storb et al. 1984) and an elastase-I gene (Swift et al. 1984) were the first examples of tissue-specific gene expression; subsequently, other globin genes such as the myosin light-chain-2, α-actin, α-crystallin, insulin, as well as several others have been expressed in a tissue-specific manner. Transgenic mice have been produced which express these genes predominantly, if not exclusively, in the cell types specific for their endogenous counterparts.

As a general rule, it appears that if a tissue-specific transgene is expressed, it is expressed in the appropriate tissue, apparently independent of its chromosomal integration site. This implies that trans-acting proteins responsible for tissue-specific expression are capable of finding their target sequences at almost any chromosomal location and initiate transcription. However, expression levels of a particular gene among founder animals varies greatly (Chada et al. 1986, Ornitz et al. 1985, Townes et al. 1985, Swift et al. 1984) suggesting that chromosomal position can influence accessibility of the genes to transcription factors. Some mice may not express the gene at all which, among other possible factors, may be due to its integration into heterochromatin domains.

Since transgenes integrate in multiple copies, usually in a tandem array, there is no way of knowing how many of them may be functional templates for transcription. Moreover, there is a very poor correlation between gene expression and copy number (Ornitz et al. 1985, Swift et al. 1984) which suggests that only a subset of the integrated genes is expressed or that the entire array is very sensitive to chromosomal position. For several different transgenes the level of expression in some mice approaches or exceeds that of the endogenous genes, which indicates that optimal expression may not depend on the normal chromosomal position or that the transgene construct omitted regulatory control elements. It is worth noting that the signals for tissue-specific expression, and their recognition appear to be evolutionary conserved between species, since many human genes for example have been appropriately expressed in mice.

1.3: Factors that influence transgene expression

It is very unlikely that any two founder mice injected with the same transgenic construct will exhibit identical expression patterns. Though some will have similar localization of expression and perhaps relatively similar levels of expression, others will have very noticeable differences. These differences may be inappropriate expression in one or more tissues, overexpression, underexpression or even a total lack of expression. These differences are almost certainly a result of mutations that may occur during the integration process rendering the gene inoperative, integration into chromosomal locations that adversely affect transcription, or possible integration near endogenous enhancers that can lead to overexpression or inappropriate expression.

1.3.1: Insertional mutagenesis

The insertion site of foreign DNA into the host chromosome is probably random, although it may be possible that integration occurs preferentially into "open" chromatin. It is therefore probable that occasionally the foreign DNA will disrupt the function of an endogenous gene or its control elements. If the foreign DNA contains an enhancer/promoter, it is also probable that it may lead to the expression of endogenous genes which are near by. Most insertional mutations are recessive and, thus, their effect can only be noticed following inbreeding of the transgenic offspring. The overall frequency of insertional mutations including visible mutations and prenatal lethality in transgenic lines has been estimated to be between 5% and 10% (Palmiter et al. 1986, Weiher et al. 1990).

It is often easy to identify insertional mutations since the affected host genes flank the transgenic sequence and therefore can be isolated. After sequencing the host flanking region, investigators are able to determine whether the mutation is a result of disruption of a functional endogenous gene or possible mutations which occur to the transgene itself upon integration. Many reports of deletions, duplication and rearrangements around the

site of integration have been described (Wilkie et al. 1987). In fact, the frequency of insertional mutations is higher in transgenic mice that were produced using microinjection techniques as compared with those poduced using retroviral insertion. This difference is most likely the result of large regions of chromosomal DNA being deleted or duplicated during integration of microinjected DNA, whereas retroviruses insert properly, with duplication of only a few base pairs of host DNA at the integration site.

1.3.2: Chromosomal position effect

Chromosomal position effect refers to altered transgene expression from that anticipated due to the insertional site and the influence of the DNA around it. Al-Shawi et al. (1990), found that after foreign DNA was injected into mice and 5 lines established, 4 lines expressed the transgene in the target tissue, whereas one line did not express the gene. To determine whether this lack of expression was due to mutations rendering the gene nonfunctional, or the result of the localization of the transgene, researchers recovered the foreign gene from the primary transgenic mouse and reinjected it into new mouse embryos by pronuclear microinjection to produce secondary transgenic mouse lines. If the lack of expression in primary animals resulted from mutations affecting the transgene, then the same non-expressional pattern in the secondary lines would be expected. If, on the other hand, the lack of expression resulted from the localization of the transgene in the host DNA, then expression in the secondary mice similar to the original findings, most expressing the transgene in tissue, would be expected. The results indicated that most mice expressed the trangene, accourdingly the authors demonstrated that reintegration of the recovered trangene DNA into new founder animals restored expression to that originally expected. Integration of the reisolated DNA into the secondary mice occurred randomly and the probability of reinsertion in the same location as the original founder animal was extremely low. Furthermore, in these experiments flanking DNA was also introduced into the secondary mice, thus the positional effect must have been due to influences originating from more distant parts of the chromosome.

Chromosomal position effect may also be attributed to inappropriate or overexpression of genes in various tissues. In this case the transgene may integrate close to an endogenous enhancer/promoter element which controls the foreign gene expression (Hamada 1986, Gossler et al. 1989).

1.4: Transcriptional enhancers

1.4.1: Definition and function

Regulation of transcription in prokaryotes as well as eukaryotes is controlled by short promoter sequences usually 10-35 bases upstream of the transcriptional start site. Many eukaryotic promoters consist of a conserved sequence, TATAAA, -20 to -30 base pairs(bp) upstream known as the TATA box. Some also contain an additional upstream element GG(T or C)CAATCT known as the CAAT box. These short sequences serve as contact points for the binding of RNA polymerase and help position the polymerase on the transcriptional start site (reviewed by Rosenberg and Court 1979). In prokaryotes deletion of these promoter sequences abolishes transcription, however this is not the case in many eukaryotic systems. Early studies using the histone H2A gene showed that deletion of the TATA box did not necessarily abolish expression (Grosschedl and Birnstiel 1980a), and that DNA sequences positioned several hundred base pairs away from the transcriptional start site could positively alter transcription (Grosschedl and Birnsteil 1980b). Moreover, these sequences were functional even in an inverted orientation.

Experiments with the simian virus 40 (SV40) early region demonstrated that the TATA box was not required for early gene expression (Benoist & Chambon 1980), however deletion of a sequence about 150 bp upstream of the transcriptional start site adversely affected gene expression. Thus, DNA sequences located considerably far from the transcriptional start site can significantly affect transcription. Not only did these sequences affect transcription of their indigenous genes but, in the case of SV40, they

were also able to affect transcription of other genes. Transcription of the β-globin gene increased by 200 fold when linked to a 366 bp DNA segment of the SV40 promoter region (Banerji et al. 1981). Transcriptional enhancement occurred only when the SV40 sequence was present in *cis*, did not depend upon DNA replication, and could occur when the sequences were placed in both orientations either 1400 bp 5' or 3300 bp 3' of the gene.

Enhancers are defined by the observed properties of these viral transcriptional control elements, specifically the ability to: (a) increase transcription of *cis*-linked promoters. (b) operate in an orientation- independent manner, (c) exert an effect over large distances independent of position, and (d) enhance the expression of heterologous promoters (Atchison 1988).

Following identification of these enhancer elements in viral genomes, cellular enhancers were also discovered. Some of the most well characterized cellular enhancers are those of the immunoglobulin genes (Banerji et al. 1983, Gillies et al. 1983, Mercola et al. 1983). Banerji et al. (1983) demonstrated enhanced expression of a -globin promoter linked to a several hundred bp segment derived from the Ig heavy chain locus which functioned in an orientation and position independent manner. The enhancement occurred only when the sequence was linked in *cis* and did not depend on DNA replication. Furthermore, it was shown that enhancement occurred only in lymphoid cells suggesting a tissue-specific mechanism.

1.4.2: Organization of enhancer elements

As studies into enhancer activity progressed it became obvious that enhancer activity was not dependent on a single DNA sequence motif, but rather on a combination of motifs each with specific function. This concept of an "enhancer effect" is a product of several sequence motifs (modules), each of which contribute to the overall activity of the enhancer (Serfling et al. 1985).

These sequence motifs are generally binding sites for trans-acting nuclear factors. Thus, enhancer activity is not dependent only on the presence of DNA sequence motifs but on the binding factors which regulate activity as well. Furthermore, some sequence motifs are able to bind several different trans-acting factors which may be present in only particular cell types. Some tissue-specific factors may be present in one or several cell types in an inactive form and later become active after post-transcriptional modifications. The transformation of trans-acting factors into their active state by cellular agents is most likely the basis of tissue-specific and developmental control of transcription.

Similarly, it is also possible to have negatively acting sequences whereby the binding of trans-acting elements prevents expression of some genes in certain cell types. Therefore, mutations to these sequence motifs leading to inability to bind trans-acting factors, or mutations to the trans-acting factors normally bound to enhancer motifs rendering them unable to bind, may lead to altered expression. Enhancer activity is therefore dependent on the presence of cis-dependent DNA sequence motifs, the presence and state of trans-acting cellular factors, and the interaction between them.

1.4.3: Enhancer localization (enhancer trap)

Transgenic mice can be used to effectively identify cis-acting control elements responsible for tissue specific expression. They are also the only means of:

1> Localizing control elements in cells that cannot be cultured.

2> Demonstrating exclusive expression in the appropriate cell type.

3> Observing their expression during normal development.

Tissue-specific control elements (enhancers) have been localized to areas adjacent to their promoter, many kilobases upstream or downstream of the promoter, or even within the genes they regulate. Some cases of promoters which appear to behave in a tissue-specific manner have also been reported (Sternberg et al. 1988).

Use of transgenic mice to identify possible tissue-specific elements is straight forward, but often tricky. First a transgene consisting of a promoter-reporter gene construct is used to produce transgenic mice. Tissues and cells are then assayed for unusual expression of the reporter gene. If for example, expression is significantly higher in muscle cells compared to other cell types it may be possible that the transgene had integrated close to a muscle specific enhancer. The second step is to isolate the transgene and the endogenous DNA flanking it. By using the isolated DNA to produce new mice or transfecting into cells that can be cultured, it is possible to determine the presence of a possible enhancer. If transcriptional levels in transgenic animals, or cultured cells, are approximately equivalent to those expected, then the majority of all the necessary *cis*-acting elements are probably present. If levels are low or absent, then either one or more important transcriptional elements may be absent either by deletions, mutation, or chromosomal position effect may may be of influence. If normal expression is achieved, then strategies using mutational analysis can be used to locate the cis-acting elements important for tissue-specific expression. Mutational analysis may work well for identifying single elements, but is often difficult to implement if transcription is influenced by multiple factors working in a cooperative manner.

To prove that a tissue-specific enhancer element was isolated, it must be shown that it appropriately controls a heterologous promoter regardless of orientation (Serfling et al. 1985).

1.4.4: Muscle-specific enhancers

An underlying principle of cellular differentiation is specific gene expression. The preferential expression of cellular genes allows not only differentiation, but also cellular regulation and development. Preferential expression involves cell or tissue-specific binding factors which often bind tissue-specific enhancers allowing the transcription of specific cellular genes. In skeletal muscle, a host of muscle specific factors such as MyoD (Davis et al 1987), myogenin, myf-4 (Edmondson & Olson 1988), myf-5 (Braun et al.

1988), and MRF-4/herculin/myf-6 (Miner & Wold 1990) have been identified. These basic helix-loop-helix proteins are exclusively expressed in skeletal muscle, and they are believed to control each others expression as well as the expression of muscle structural proteins.

Many muscle specific genes normally activated during myogenesis have been identified, some include the chick α-skeletal actin gene (Gricnik et al. 1986), troponin I genes (Konieczny et al. 1985), the myosin heavy chain gene (Medlcr et al. 1983) among others.

One of the best characterized tissue-specific enhancers are those controlling expression of the muscle creatine kinase (MCK) gene (Johnson et al. 1989). Two skeletal muscle-specific enhancers, have been identified and characterized one located 1026-1256bp upstream of the MCK start site and another within the first intron of the MCK gene. In addition, an enhancer which upregulates both skeletal and cardiac muscle gene expression has been identified.

Identification of the 5' enhancers was achieved by performing deletion analysis on a large portion of the 5' flanking region of the MCK gene fused to the reporter gene CAT. CAT expression was then measured in different cell types in *vitro* as well as in transgenic mice (Johnson et al. 1989). Such analysis identified a major MCK muscle specific enhancer located 1020-1256bp upstream of the transcriptional start site. This enhancer upregulated MCK expression 200 fold when present in the correct orientation, and a 20 fold increase when the enhancer was inverted or placed at variable distances as compared to expression with the promoter alone. Moreover, high levels of expression exclusive to muscle cell types was found in both in *vitro* and in *vivo*. Furthermore, if this element was fused to a heterologous promoter it again conferrd elevated expression predominantly in muscle cell types.

1.4.5 Known cardiac-specific enhancers.

The identification of a cardiac specific enhancer was published shortly after the undertaking of the present project (Parmacek et al. 1992). The slow/cardiac troponin C gene (cTnC) is continuously expressed in embryonic and adult cardiac myocytes and only transiently in embryonic fast skeletal myotubes. Analysis has demonstrated the presence of two enhancer sequences, one controlling cardiac specific expression and the other controlling expression in skeletal myotubes. The function of the cardiac enhancer located between -124 to -32 bp of the cTnC gene is independent of the skeletal muscle enhancer located within the first intron of the gene (bp 997 to 1141).

Identification of the cTnC cardiac specific enhancer was carried out using deletion analysis. An initial construct containing 2.2 Kb of 5' cTnC sequence linked to the CAT reporter gene driven by the cTnC promoter was used to transfect primary cultures of rat neonatal cardiac myocytes. Results obtained showed an 80-100 fold increase in CAT activity over the promoterless CAT vector. Inclusion of additional intragenic and 3' flanking sequence showed no increase in CAT activity, suggesting that all regulatory elements required for high expression was included within the 2.2 Kb fragment. Deletions up to base pair -124bp had no effect on transcriptional levels. However, further deletion up to -79bp resulted in a 90% drop in CAT activity.

To determine whether this particular region contained a cardiac specific enhancer rather than a cardiac specific promoter element, a 69bp fragment (bp -124 to -56) was isolated. Two new constructs, one containing the cTnC promoter-CAT construct without the fragment and the same construct containing the fragment 3' of the CAT gene, were again transfected into cells. CAT activity in those cells that did not contain the 69bp fragment was just above baseline, but expression was restored to 80-100% above baseline in those cells that contained the fragment 3' of the CAT gene. Moreover, this fragment increased expression independent of position and orientation, characteristic of a true enhancer element.

Constructs containing the 69bp fragment 3' of the CAT gene driven by either the minimal SV40 and herpes simplex virus thymidine kinase promoter elements were used to test whether expression could be achieved using heterologous promoters. In both cases the cardiac enhancer failed to increase transcription from both promoters in cardiac myocytes. It thus appears that this enhancer requires the endogenous cTnC promoter.

To test the specificity of the isolated promoter/enhancer the construct containing bp -124-CAT and a promoterless control were transfected into various cell types. Comparisons showed an increase of CAT activity by 90 fold over the promoterless control plasmid in primary cultures of neonatal cardiac myoctyes. No activity was observed in C2C12 skeletal myoblasts, C2C12 skeletal myotubes, and all non-muscle cell lines tested including primary cardiac fibroblasts.

Cardiac myocytes have the ability to take up and express exogenous DNA injected directly into the left ventricular wall in vivo (Acsadi et al. 1991, Lin et al. 1990). To determine whether this enhancer element has comparable functionality in vivo, several cTnC-CAT plasmid constructs were injected into the left ventricular walls of 6-week-old Sprague-Dawley rats. Five days post injection, the animals were sacrificed and cardiac homogenates were assayed for CAT activity. Results similar to those in vitro were obtained. The 156bp 5' flanking cardiac-specific promoter/enhancer increased CAT activity about 70-fold in vivo. Deletion from base -124 to -79 abolished transcriptional activity. Again, introduction of the 69bp fragment (-124 to -79) downstream of the CAT gene restored transcriptional levels in vivo.

These observations confirm the identity of DNA sequence fragment immediately upstream of the cTnC gene which posses all the required characteristics of an enhancer element and appears to be genuinely cardiac specific.

1.5: Physiological consequence of transgene expression

It is now possible to redirect genes encoding regulatory proteins to novel cell types. Redirecting expression of secreted proteins to organs such as liver, can dramatically increase the levels of protein in the circulation. In the new target organ this expression may lack any normal regulatory feedback control mechanisms. The ability to express genes in heterologous tissues, overexpress certain products, or express them prematurely is a valuable tool in understanding the functions of complex systems that may not be possible in *vitro*.

1.5.1:Growth control

Several studies have been conducted on transgenes encoding growth control genes. For example, rat, bovine, or human growth hormone (GH) under the control of a heterologous promoter (usually MT) to allow elevated and chronic production has produced mice up to twice the normal size.

1.5.2: Cancer

The introduction of several oncogenic constructs into mice was shown to produce tumorigenic or other pathological changes. For example, transgenic mice with an integrated SV40 enhancer and early region (coding for large and small T antigen) often develop choroid plexus tumors (Brinster et al. 1984, Palmiter et al. 1985). In some cases only the SV40 large T antigen is sufficient to cause tumorigenesis. If, however, the SV40 enhancer suffers from deletions the incidence of tumors is greatly reduced and those that develop are rarely of choroid plexus origin. Furthermore, if promoter/enhancer elements of tissue-specific genes are fused to the SV40 T-antigen, tumors develop in the cell types specified by the enhancer. For example, the promoter/enhancer region from the insulin gene directs tumors exclusively to the beta-cells of the pancreatic islets (Hanahan 1985), whereas the promoter/enhancer region of the elastase-I gene directs tumors to the acinar

cells of the pancreas (Ornitz et al. 1985b). In similar experiments, tumorigenesis has been directed to lactotroph cells, somatotroph cells, and the lens.

1.6: Metalothionein transcribed genes.

Expression of the endogenous mouse MT-1 gene occurs in virtually every cell, although the level of expression may vary. Expression may also be affected by a variety of inducers such as glucocorticoids, metals, inflammatory signals and interferon, but again response varies among different tissues (Palmiter et al. 1986). MT-1 fusion genes are expected to be expressed in all cells as long as the required cis-acting regulatory elements are present, however, this is seldom the case (Palmiter et al 1983). The expression of MT-1 fusion genes is genrally high in liver, intestines, kidney, heart, pancreas, and testis, but the expression varies widely among tissues and among founder animals.

Occasionally distorted expression occurs in some mice relative to the endogenous gene, for example, very high expression in kidney or pancreas with virtually no expression in liver (Palmiter et al. 1983). Whether the normal tissue distribution of MT-1 gene expression is a consequence of the various inducable elements (such as the metal regulatory elements) or whether an enhancer acts independently of these elements is unclear (Palmiter 1986). The variability of expression among mice and tissues may suggest that the MT-1 promoter is very sensitive to chromosomal position. Thus, if an enhancer is tissue-specific the location will only affect the levels of expression in that cell type, whereas with an enhancer/promoter that functions in many cell types, the influence of chromosomal position may vary in the different cell types, and from one animal to the other depending on the integration site. While many genes express their effects normally in transgenic animals, in some genes the effects are not expressed or are expressed in negligible levels. Metallothionein fusion genes are the best documented for lack of expression. Many transgenic mice generated carrying the mouse MT-1 promoter fused to beta-galactosidase, human factor IX, or rat alpha-tumor growth factor failed to express any of the genes in the liver. However, in each case the gene constructs were shown to be

functional when introduced into tissue culture cells. It was shown that the MT-beta-galactosidase construct was very active following microinjection into fertilized eggs, and activity persisted for several days of development. In the case of MT-hepatitis B surface-antigen gene, expression is very low, but upon substitution of the mouse albumin promoter/enhancer for the mouse MT-1 sequence, normal expression was observed. Therefore, a plausible conclusion would be that during development something occurs which renders these genes inexpressable in the adult liver.

1.7: Polyomavirus large T antigen

The polyomavirus Large T antigen (PVLT) is one of three early genes transcribed by the polyomavirus (PV), a small double stranded virus belonging to the papovavirus family. PVLT is a DNA-binding phosphoprotein with ATPase and helicase activity. The presence of PVLT is required for successful viral DNA replication and transcriptional regulation.

PVLT has the ability to disrupt normal cell growth and differentiation. It is a known immortalizing gene able to confer an immortalized phenotype in rodent primary cells (Glaichenthaus et al. 1985, Rassoulzadegan et al. 1982, 1983). Many different cell types become deregulated following PV infection where PVLT may have a significant role (Dawe et al. 1987, Eddy et al. 1982, Freund et al. 1987, Gross 1970, Rockford et al. 1990). The precise mechanism by which the polyomavirus alters cellular phenotypes is not yet completely known, however, PVLT is known to bind the product of the retinoblastoma (Rb) gene, a protein involved in cellular control of proliferation, possibly altering its function and allowing altered cellular control and growth.

Expression of PVLT may lead to latent tumor formation (Bautch et al. 1987), inhibition of differentiation (Giep et al. 1989), cardiomyocyte hypertrophy (Chalifour et al. 1990) and hyperplasia of Leydig cells (Chalifour et al. 1992).

1.8: Line 8 transgenic animals used in this study

The mouse transgenic line designated MTPVLT-8 was primarily used in this study. C57 Bl/6 x Balb/c mouse embryos were injected with a mouse metalothionein promoter (MT) element linked to the polyomavirus large T antigen (PVLT) gene and a SV40 poly-A tail coding region. The embryos were then reimplanted in pseudopregnant CD-1 females. The resulting offspring were assayed for transgene incorporation and positive founder animals were latter crossed with normal C57 bl/6 x Balb/c mice to produce F2 hybrids. Six lines were selected for this study of which one, designated MTPVLT-8, produced mice phenotypically different from all others. Line 8 mice were shown to develop testicular adenomas similar to that observed in two other lines, MTPVLT-10 and MTPVLT-11, but were the sole transgenic mouse line to also develop cardiac hypertrophy. Typically, line-8 mice develop cardiac hypertrophy and die prematurely within 6 to 12 months as a result of cardiac failure. Post mortem examination revealed heart sizes 3 to 5 fold larger than normal with both ventricular and artrial hypertrophy observed. At the cellular level, it was reported that the individual cardiomyocytes enlarge and that there is no significant increase in cell number (Holder et al. 1995). Furthermore, at the molecular level PVLT mRNA was detected in heart tissues of line 8 transgenic mice while none was detected in heart tissues of any other MTPVLT lines (Holder et al. 1995).

1.9: Aims

We hypothesized that the observed cardiac hypertrophy in line 8 transgenic mice is due to the expression of the MTPVLT transgene in heart tissue. This expression may be controlled by chromosomal position effect since no other transgenic mice harboring the exact trangenic sequence express the gene in heart. Transgene integration in the vicinity of a cardiac specific enhancer is the most plausible explanation of the observed heart specific nature of expression. Tissue specific transgene expression due to integration near enhancers have been well documented (Hamada 1986, Sternberg 1988).

The aim of this project was to isolate, characterize, and identify any potential cardiac specific enhancers flanking the integration site from line-8 mice.

CHAPTER 2

MATERIALS

MATERIALS

2.1: Restriction and modifying enzymes

Restriction enzymes and modifying enzymes were purchased from Pharmacia with the exception of Nsi I purchased from New England Biolabs.

2.2: Thermostable polymerases and helper products.

Taq was purchased from Promega, PFU from Stratagene, and Vent DNA polymerase from New England Biolabs. Taq Extender, and Perfect Match DNA polymerase enhancer were purchased from Stratagene.

2.3: PCR and random priming reagents

Random primers and nucleotides were all purchased from Pharmacia. Klenow was purchased from either Pharmacia or Promega. All PCR primers were synthesised by Sheldon Biotechnologies (Montreal, QC) with the exception of the anchor primer, external and internal primers which were synthesized by National Biosciences (Plymouth, MN).

2.4: Membranes and blotting reagents.

GeneScreen Plus and colony lift membranes were purchased from NEN/Dupont.

2.5: Radioactivity & autoradiography

Radioactive [^{32}P]α-dCTP and [^{32}P]γ-ATP were purchased from either Amercham or NEN/Dupont. Film and screens used for autoradiography were either X-OMAT film from Kodak or NEN/Dupont Reflections autoradiography film.

2.6: Kits

Sequencing was performed using the fmol DNA Sequencing System from Promega. The Sephaglass bandPrep kit from Pharmacia was used to purify extracted DNA from agarose gels. The Epicurian Coli kit from Stratagene and TA cloning kit from Invitrogen were used in the bacterial minilibrary and clone-6 subcloning respectively.

2.7: Other reagents

Agar, agarose, NaCL, SDS, borate, were purchased from ICN. Dextran Sulfate was purchased from Pharmacia, Tris base from Boehringer Mannheim and guanadinium isothiocyanate from BRL. All other reagents were purchased either from BDH or Anachemia.

2.8: Equipment

The following equipment was used in most of the research. Sequencing was conducted using the BRL Model 52 sequencing apparatus. Agarose gels were electrophosed using a BioRad electrophoresis apparatus. Power supplies used were the Pharmacia GPS 200/400 and ECPS 3000/150. Densitometry and gel analysis was

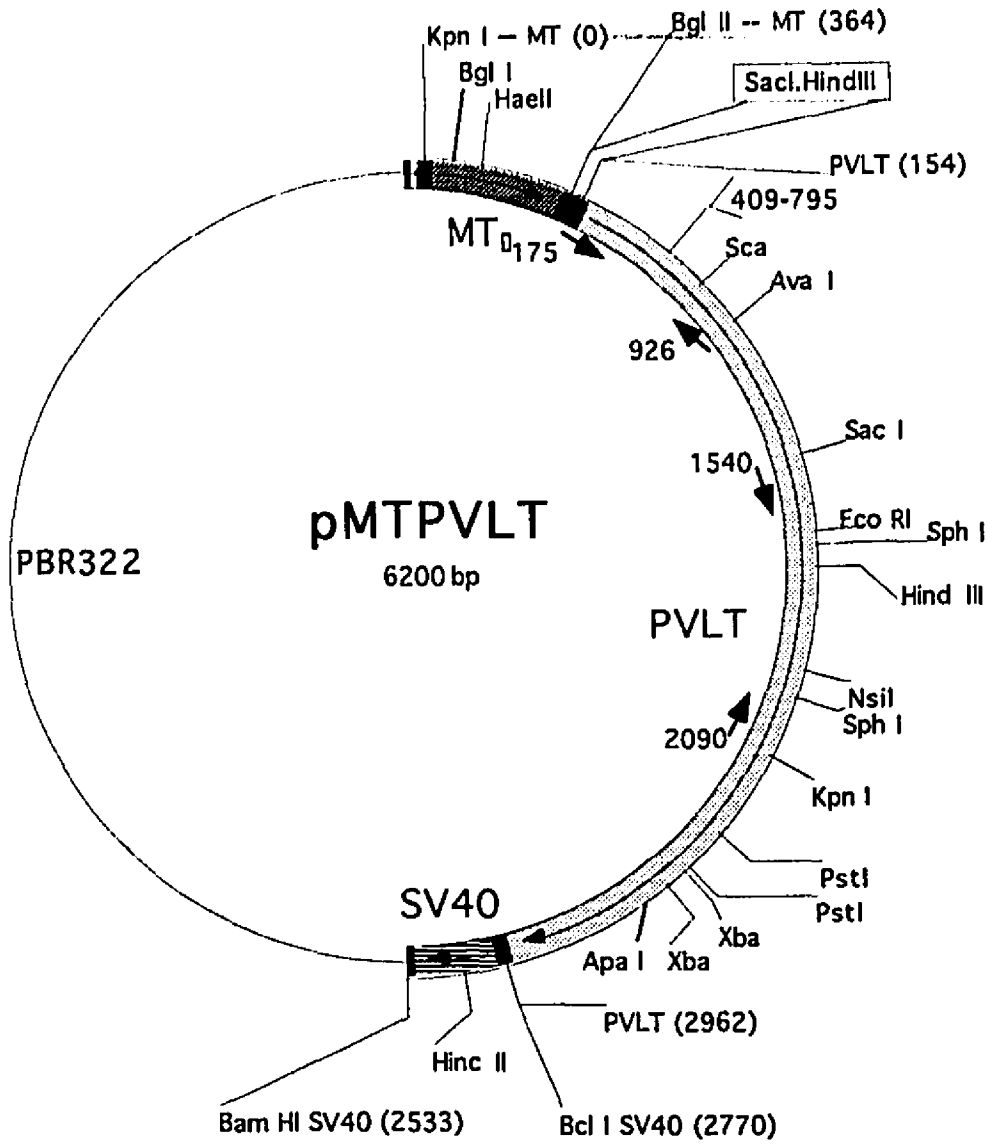
performed using either the Molecular Dynamics' PhosporImager SI and/or the Personal Densitometer SI.

2.8: pMTPVLT

The transgene in MTPVLT-8 mice originates from the pMTPVLT clone. Most PVLT specific probes were made using this plasmid as template. A schematic of the plasmid and some of the important restriction sites as shown in figure 1.

Figure 1

A schematic of the pMTPVLT. Restriction sites within the transgene are indicated. PCR primers used to produce probes to the transgene are indicated by arrows.



Arrows represent PCR primers. Primer numbers refer to the location within the PVLT sequence coding sequence and not within this plasmid.

Figure 1

CHAPTER 3

METHODS

METHODS

3.1: Isolation of genomic DNA

Tissue was extracted and immediately frozen on dry ice. Genomic DNA was isolated from a transgenic mouse (Designation 81215311x1-1) exhibiting both the enlarged heart and seminal vesicle/testis phenotype. Lung tissue (0.5445 g) was ground using a mortar and pestle and suspended in 1.2 ml Digestion Buffer (100mM NaCl; 10 mM Tris-Cl, pH 8; 25 mM EDTA, pH8; 0.5% SDS) per 100 mg of tissue. Proteinase K was added at 0.1mg/ml for 6 hours and again for 17 hours at 55°C. The solution was then extracted twice with phenol/chloroform and once with chloroform. RNase was added to the supernatant to a final concentration of 10ug/ml for 30 minutes at 37°C. One more phenol extraction was performed followed by a single chloroform extraction. The DNA was precipitated using 2.2 volumes of 95% ethanol, spooled, washed in 70% ethanol and allowed to air dry for 20 minutes. It was then resuspended in distilled water overnight at 4°C.

3.2: Digestion of genomic DNA

A total of 20ug of MTPVLT Line-8 positive transgenic and CD-1 non-transgenic mouse genomic DNA was digested with various restriction enzymes. DNA was digested in a 2ml eppendorf tubes consisting of 20ug genomic DNA, the appropriate volume of 10X buffer, 60ul of spermidine (10mM) and distilled water in a final volume of 600ul. The reaction mix was heated to 65°C for 30 minutes and allowed to cool to room temperature. To each tube 100 Units of each restriction enzyme was added and incubation proceeded overnight at 37°C. An additional 100 units was added the following day and allowed to react for 2-4 hours. The digested DNA was then precipitated with 1/10 volume of 3M ammonium acetate (pH 5.2) and 2.2 volumes of ethanol overnight at 20°C. The DNA was recovered by centrifugation then rinsed in 75% ethanol. The pellet was vacuum

dried and resuspended in 40ul of distilled water. A 10ul alliquot of each digest was run on a small agarose gel to determine if complete digestion was achieved.

3.3: Random priming of DNA and RNA probes

Radioactive probe was generated using a standard random priming reaction (Asubel et al. 1993). Approximately 100-200ng of template DNA was mixed with 1ul of random primers (pdN6, 5.0 ug/ul) and the volume was brought up to 14ul with distilled water. The reaction tube was heated at 95°C for 3 minutes then placed immediately on ice for 2 minutes. A complete reaction mix contained 2.5ul of a dNTPs (0.5mM of dATP, dGTP, and dTTP), 2.5 ul Klenow Buffer (0.5M Tris-Cl pH 7.5; 0.1M $MgCl_2$; 10mM DTT; 0.5mg/ml BSA), 5ul $^{32}P\alpha$-dCTP (3000ci/mMol Amersham) and 1ul Klenow (5u/ul, Promega) in a final reaction volume of 25ul. The labelling reaction was carried out at 37°C for 3 hours. The reaction was stopped by the addition of 75ul of TE (10mM Tris-Cl; 1.0 mM EDTA) and the unincorporated radioactive dCTP was removed by centrifugation through a Sephadex G50 spin column. The DNA probe was denatured by adding 0.1 volumes of 1 N NaOH for 5 minutes followed by neutralization using an equal volume of 1 N HCl. A 5ul aliquot of the final product was subjected to scintillation counting to measure the amount of incorported nucleotides.

3.4: Southern blotting

3.4.1: Gel preparation and electrophoresis

Approximately 15ug of digested genomic DNA was loaded into a 0.8% agarose gel (0.8g agarose/100ml 1xTBE) and electrophoresis was performed overnight at 25 volts.

3.4.2: Southern transfer

The gel was submerged in a 0.25N HCl solution for approximately 20-30 minutes and then neutralized in 0.4N NaOH for 20 minutes. The DNA was transferred onto a Gene Screen Plus membrane using a downward capillary protocol (Ausubel et al. 1994) with 0.4N NaOH transfer buffer for 3 hours. The membrane was then neutralized in 2xSSC and air dried.

3.4.3: Prehybridization and hybridization

The filter was prehybridized in 40 ml of prehybridization solution (10% Dextran Sulfate; 1% SDS; 1M NaCl) and 100ug/ml Herring sperm DNA for two hours. The prehybridization solution was discarded and 25 ml of hybridization solution (prehybridization solution containing $1x10^5$ cpm/ml of probe) was added. Hybridization was carried out overnight in a 65°C rotating oven.

3.4.4: Washing and autoradiography

The filter was washed with 2xSSC twice for 5 minutes followed by two 30 minute washes at 65°C with 1xSSC; 1% SDS. A final wash was carried out for 10 minutes using 0.1xSSC. The filter was then wrapped and placed for autoradiograph exposure.

3.4.5: Stripping filters

Filters were stripped using 0.1xSSC; 1% SDS. The SSC-SDS solution was heated to boiling and added to the filters. The filters were shaken on an orbital shaker for 15 minutes. The procedure was repeated twice. The membranes were then exposed for 48hrs and checked for any remaining signal.

3.5 Direct PCR cloning of flanking DNA

3.5.1 Digestion and purification of genomic DNA.

Line-8 mouse genomic DNA, 10-15 ug, was digested with a Nsi I restriction endonuclease to create 3' overhang DNA fragments. The DNA was size separated on a 0.8% agarose gel in TAE. DNA fragments ranging in size from 4-9Kb, the expected size of the 5' flanking sequence, were cut out of the gel and purified by spinning through glass wool.

3.5.2 Ligation of anchor primer

Ligation of the anchor primer to the size fractionated DNA and 5ug total genomic DNA was performed in a 10ul reaction mix (1ul ligation buffer; 1ul 10mM ATP; 50uM anchor primer; 8 Weiss units of T4 DNA ligase). The DNA was heated to 65°C for 15 minutes prior to adding the ligase to dissociate any annealed ends. Ligation was performed at 15°C overnight. The ligation mix was then brought up to a final volume of 50ul, heated at 65°C for 15 minutes, and spun through a Sephadex G50 column to remove unligated primers.

3.5.3 PCR amplification

PCR was performed using Taq DNA polymerase, PFU polymerase , or Vent DNA polymerase with their provided 10x polymerase buffer and 10x $MgCl_2$ solutions. Reactions were performed in either 50 or 100 ul final volume. Primer concentrations were fixed at 50uM while dNTP concentration varied from 2-5mM and $MgCl_2$ concentrations varied from 0.5-2X. Where applicable 1ul of either Perfect Match DNA polymerase enhancer or Taq extender was added along with the appropriate buffer. PCR was performed using a Perkin Elmer Cetus DNA Thermal Cycler. The reaction mix was

heated at 95°C for 10 minutes prior to addition of the polymerase to denature the DNA. Cycling was performed for 31-38 cycles at 95°C for 30 seconds, 55-58°C for 30 seconds, and 72°C for 1 minute per Kb amplification product.

3.6 Subcloning of genomic DNA into pUC18

Total MTPVLT-8 genomic DNA was digested with Eco RI and separated on an agarose gel. To minimize unwanted DNA fragments, DNA within 6-8 Kb was excised from the agarose and isolated by spinning through glass wool (See Figure 8). This DNA was then cloned into a PUC18 cloning vector and transformed into competent bacterial cells. Bacteria were grown on plate containing 50ug/ml ampicilin at 37°C overnight. Colony lifts were performed, and probed using a PVLT specific probe.

3.7: RNA isolation

RNA was isolated from CD1 mouse tissues. Between 0.4-0.5 grams of each tissue was homogenized in 5ml denaturing solution (4M guanidinium thiocyanate; 2.5mM sodium citrate, pH 7.0; 0.5% sarcosyl; 0.1M beta-mercaptoethanol) in a 50ml conical tube (Falcon) for 30 seconds. To the homogenate the following was added: 0.5ml 2M sodium acetate (pH 4.0), 5ml water saturated phenol, and 1ml chloroform/isoamyl alcohol (49:1). The suspension was vortexed for 10-15 seconds and ice cooled for 15 minutes followed by centrifugation for 20 minutes at 7500 rpm at 4°C. The aqueous phase was collected and transferred into a 13ml screwcap tube (Starsted) followed by the addition of 5ml of 100% isopropanol. The RNA was precipitated for 1 hour at -20°C and recovered by centrifugation at 7500 rpm at 4°C for another 20 minutes. The supernatant was removed and the RNA pellet resuspended in 3ml of denaturing solution. The RNA was again precipitated using 1 times volume 100% isopropanol at -20°C for an hour followed by a 20 minute spin as above. The pellet was washed in 75% ethanol for 10-15 minutes and subjected to a final 20 minutes spin. The pellet was then allowed to air dry for 1-2 hours

and was resuspended in 1 ml TE (10mM tris-Cl; 0.1mM EDTA, pH7.4). The RNA suspension was stored at 80°C until needed.

3.8: Northern blotting

3.8.1: Loading and electrophoresis

Total RNA, 10ug, was suspended in a final volume of 20ul in water. Loading dye (Bromophenol Blue, Xylene Cyanol) was added and the samples were subjected to electrophoresis at 107 volts in a 1% denaturing agarose gel consisting of 1X MOPS solution (0.02M MOPS; 5mM sodium acetate; 1mM EDTA, pH7.0) and 10% formaldehyde for 3 hours. The gel was then stained with ethidium bromide to verify the presence of the 28S and 18S RNA species.

3.8.2: Northern transfer

The gel was then washed in water for 20 minutes prior to vacuum blotting (BioRad model 785) using a 50mM sodium hydroxide transfer solution for 1.5 to 2.0 hours.

3.8.3: Membrane staining

The membrane was removed from the blotting apparatus and neutralized in a 2X SSC solution for 5 minutes. It was then subjected to an RNA staining solution (0.02% methylene blue; 0.5M sodium acetate) for 30 minutes to visualize the 28S and 18S RNA species. The bands were marked and a photograph taken. The filter was then destained for 15 minutes in Wash A (1xSSPE; 20% ethanol) and 15 minutes in Wash B (0.2xSSPE; 1%SDS) until the filter was free of any dye.

3.8.4: Probing northern blot

The membrane was prehybridized at 63°C in prehybridization solution (5xSSPE; 5xDenhart's; 0.5% SDS; 100ug/ml herring sperm DNA) for 3 hours. To the solution 2-2.5 x 10^5 cpms probe/ml was added. Hybridization was performed at 63°C for 19 hours.

3.8.5: Washing

Following hybridization the filter was washed in 2xSSC; 0.1% SDS twice for 15 minutes at room temperature followed by two 15 minute washes at 63°C in 1xSSC; 0.1% SDS. The filters were then subjected to autoradiography. If necessary additional washings were performed at 63°C in 0.1xSSC; 0.1% SDS for 15-30 minutes depending on the background signal.

3.9: Sequencing

3.9.1:End Labeling of Oligonucleotide Sequencing Primer

Sequencing oligonucleotides were synthesized by Sheldon Biotechnology. The lyophilized primer was resuspended to final concentration of 1ug/ul. Labeling was carried out in a 10 ul reaction containing 3.5 ul Primer (20ng/ul), 1ul PNK buffer (Supplied), 4.5 ul ^{32}P-dATP (>5000uCi MMol), and 1ul of PNK 8U/ul at 37°C for 1 hour. The enzyme was deactivated by heat at 95°C for 2 minutes.

3.9.2: Sequence primer extension, electrophoresis and transfer.

Sequence primer extension was carried out as per instructions using the fmol DNA Sequencing System. Following a 2 minute denaturation step at 95°C, the samples were subjected to 30 cycles at 95°C for 30 seconds, 42°C for 30 seconds, and 70°C for 1

minute. The samples were loaded onto a 6% acrylamide gel and two later loadings were carried out. The gel was then transferred to Whatman paper, dried under vacuum and heat, and exposed by autoradiography for 2-3 days. All sequence acquisition was done manually and confirmed by 2 or more persons.

CHAPTER 4

RESULTS

RESULTS

The use of transgenes to identify enhnacers is well documented (Hamada 1986, Gossler et al. 1989). The transgene itself serves as a locator in the host genome. To identify an enhancer the genomic DNA flanking the integration site must be cloned and characterized. Southern blotting is useful in creating a restriction map of the DNA surrounding the transgene of known sequence. Information obtained from Southern blots can then be used to clone the flanking sequences. Northern blotting can be used to identify possible endogenous genes in the area. Finally, sequencing the DNA flanking the transgene integration site would be essential to identify potential enhancers or promotors.

4.1: Restriction mapping.

Because transgene integration occurs at random the genomic DNA surrounding the transgene following stable integration is often unknown. The use of Southern blots is an effective and easy way to identify the arrangement of the transgenic DNA in terms of copy number, orientations and whether major mutations had occurred. An appropriate procedure is to digest genomic DNA from transgenic mice with various restriction enzymes and use a probe specific to the transgenic sequence (Figure 2). Bands corresponding to repeated fragments are expected to appear, due to the tandem arrangement of inserted transgenes, as well as bands corresponding to the 3' and 5' flanking fragments.

Total genomic DNA from a line-8 trasgenic mouse was digested with various rexstriction enzymes and a Southern membrane prepared. The membrane was probed using radiaoactive DNA sequences specific to the entire transgene (Figure 2A). The banding patterns observed correspond to 5' and 3' flanking DNA fragments, transgene

repeat elements and possible banding to the the endogenous mouse MT gene. To identify the 5' flanking DNA fragments (Figure 4B), the Southern filter was stripped and rehybridized with a MT precific radioactive probe (Figure 2B). The banding patterns observed correspond to 5' end flanking fragments, repeat elements that contain parts of the MT sequence, and possible banding to endogenous MT DNA. To elemintate the any possible banding to endogenous MT DNA and still identify 5' flanking DNA fragments the filter was stripped and rehybridized with a PCR probe specific to base pairs 175-926 of the PVLT cDNA (Figure 4B). The banding pattern observed is thus representative of 5' flanking fragments and possible repeat elements (Figure 2C).

In a similar fashion the 3' flanking sequences can be determined. The filter was again stripped and reprobed with a radioactive probe specific to a Apa I - Hinc II fragment located at the 3' end of the transgene (Figure 4B). The resulting banding pattern corresponds to 3' end flanking fragments and possible repeat elements (Figure 2D).

With the information provided by this differential probing technique it is possible to identify the 5', 3' and repeat elements, as in the example for Nsi I shown in Figure 3.

Due to the tandem nature of transgene integration, intense banding corresponing to hybridization to multiple copies of the transgene repeat element is expected and indeed observed. By quantifying the intesity of the bands corresponding to the repeat elements to those of the flanking fragments, a crude estimate of the transgene copy number can be assertained (Figure 4C). Densitometric analysis suggests that line-8 mice have integrated 2 to 3 copies of the transgene (Figure 4D).

Figure 2

Digested genomic MTPVLT-8 DNA probed with MTPVLT random primed probe.

Total genomic DNA isolated from a MTPVLT line 8 transgenic mouse was digested with various restriction enzymes. The DNA was subjected to electrophoresis on a 0.8% agarose gel and transferred to GeneScreen plus membrane.

Panel A: The membrane probed using a randomly primed radioactive probe to the entire transgenic sequence.

Panel B: The membrane probed using metallothionein as probe.

Panel C: The membrane probed using a PCR probe specific to PVLT from positions 175-926 (see figure 4B).

Panel D: The membrane probed using 3' end Apa1/HincII fragment of the transgene (see figure 4B).

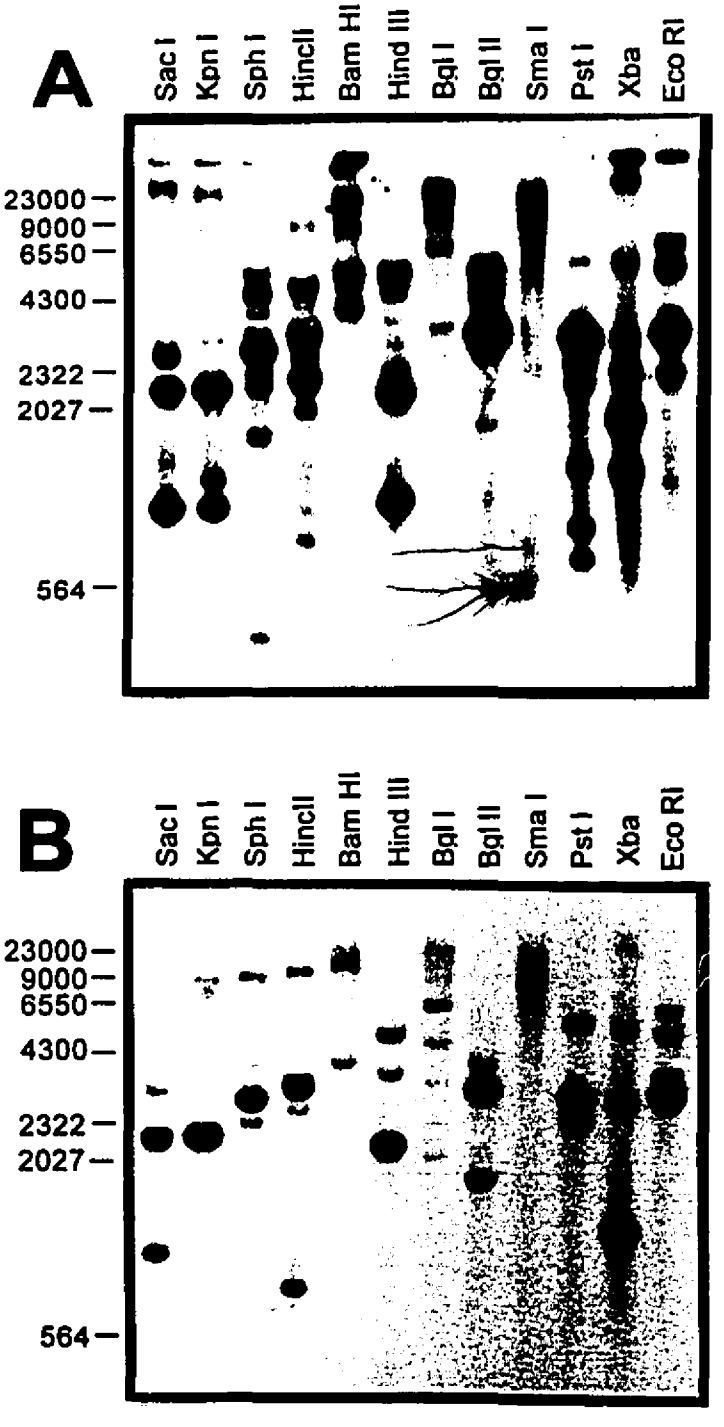


Figure 2

Digested genomic MTPVLT-8 DNA probed with MTPVLT specific probes.

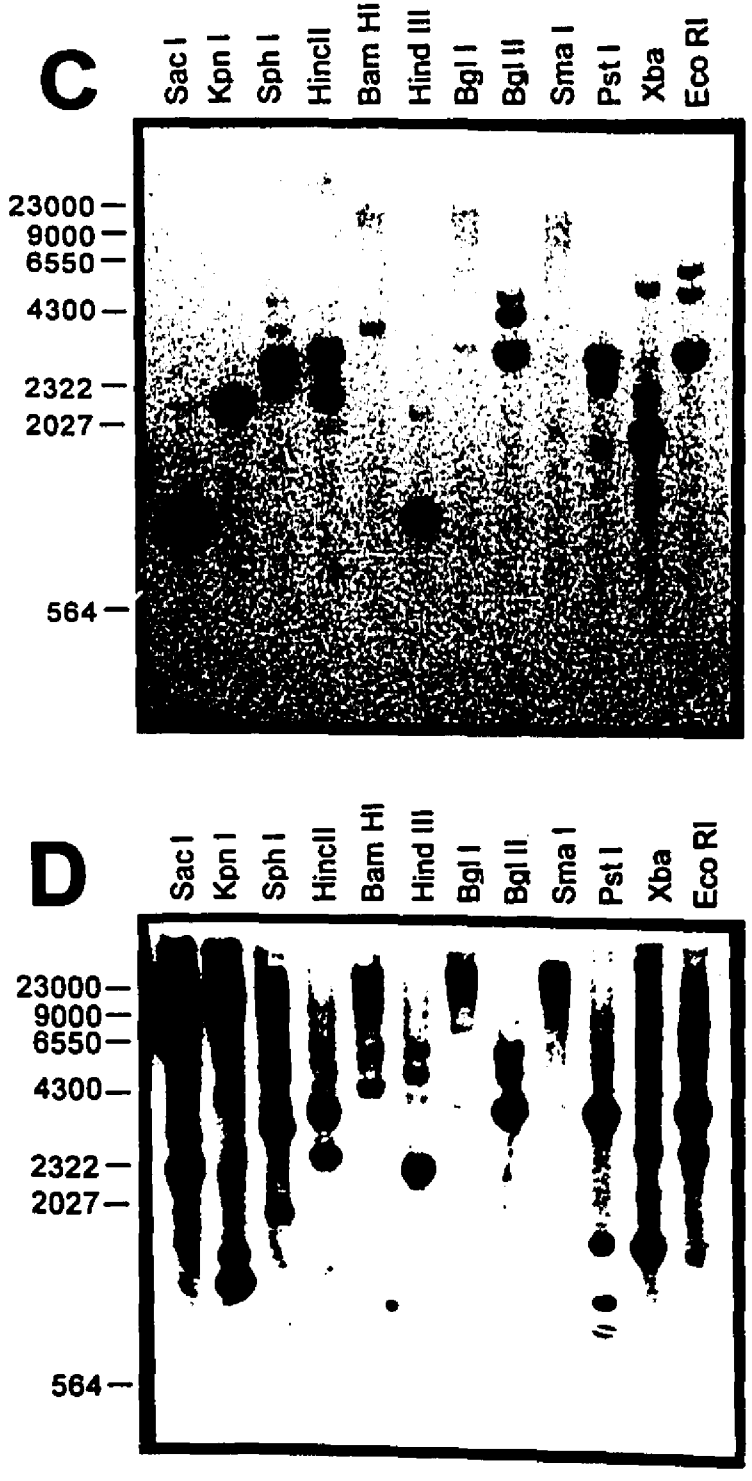


Figure 2

Digested genomic MTPVLT-8 DNA probed with MTPVLT specific probes.

Figure 3

Indentification of 5' and 3' flanking fragments for Nsi I by differential Southern blotting

Total genomic DNA from a transgenic line-8 mouse was digested with Nsi I and southern blotting was performed. In the first panel a probe specific to the 5' end of the transgene was used. The filter was then stripped and reprobed with a 3' end specific probe shown in the second panel. The 5' and 3' end fragments are indicated as well as the characteristic repeat fragment of expected size.

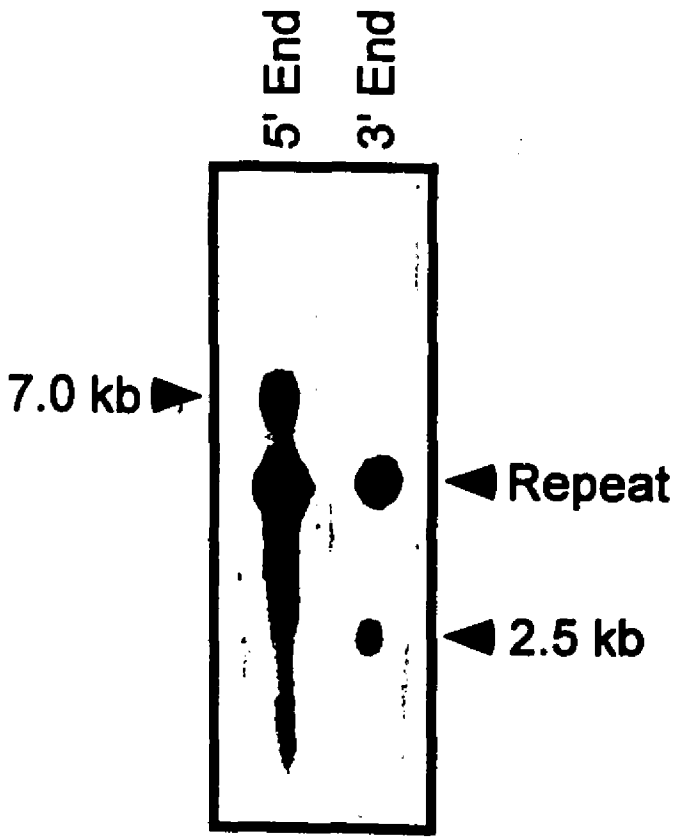


Figure 3

Indentification of 5' and 3' flanking fragments for Nsi I
by differential southern blotting

Figure 4

Transgene orginization and copy number

Panel A: A schemating of the MTPVLT transgene with some several restriction sites marked.

Panel B: A schematic representation of the transgene within the mouse genome.

Panel C: DNA banding pattern as seen by Southern blotting of genomic DNA of a transgenic mouse digested with Eco R1 and probed with radioactive probes specific to the entire transgene. Vol-4 and 3represent DNA flanking fragments, Vol-2 is the expected size of the repeated element, and Vol-1 is the 3' flanking fragment

Panel D: Graphical representation of densotometric measurment of band intensity corresponding to each rectangular region. Rect-3 corresponding to the repeat element appears to be twice as intense as the others, suggesting the presence of two intact copies of the transgene.

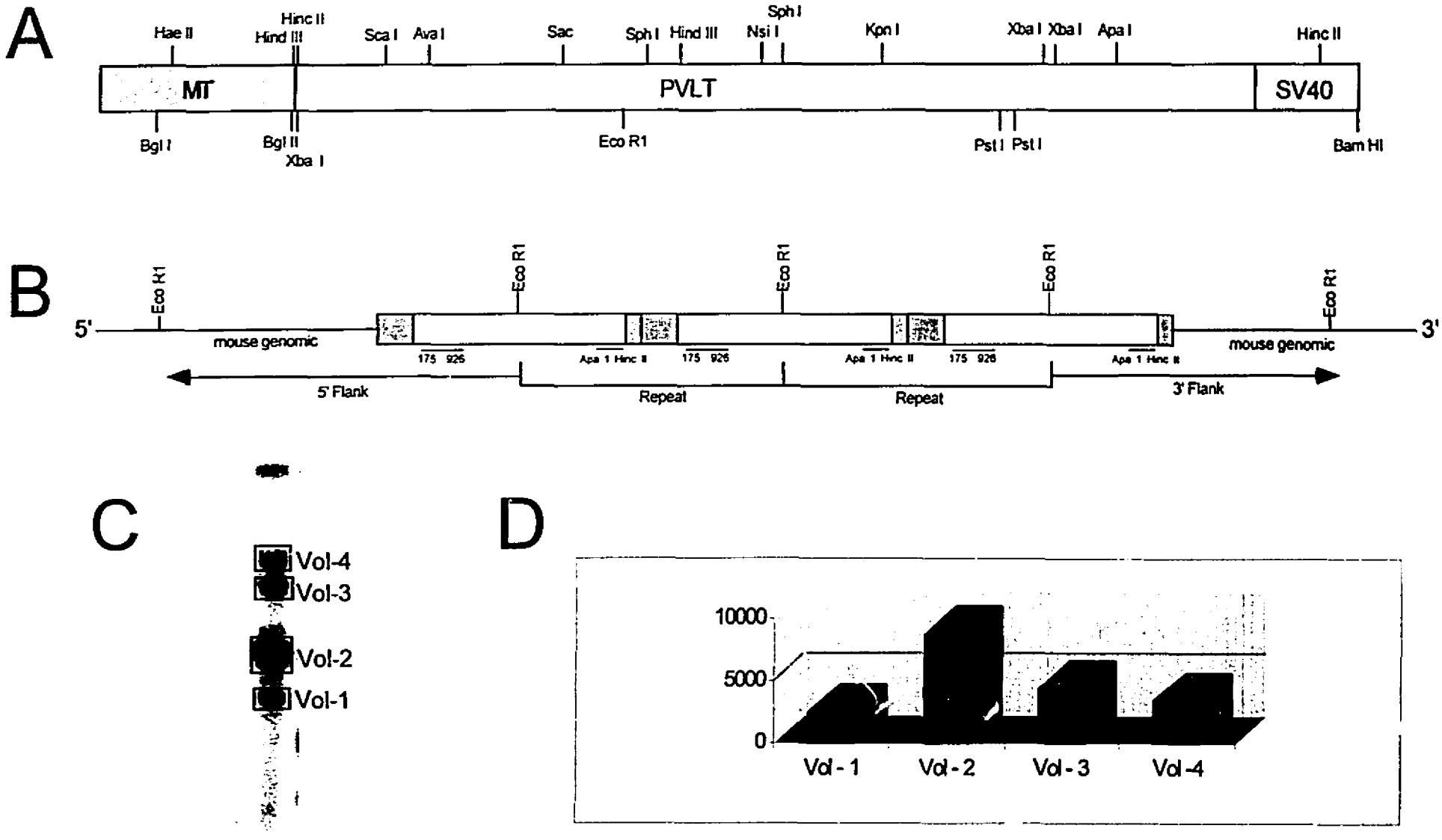


FIGURE 4

Transgene orginization and copy number

4.2: Cloning the flanking regions

4.2.1: By direct PCR

A technique to directly amplify the DNA flanking the transgene by PCR for subsequent cloning was successfully demonastrated by MacGregor and Overbeek (1992). The technique involves digestion of mouse genomic DNA with a restriction enzyme to create a 3' overhang. A small DNA fragment is then ligated to the free ends and serves as a primer template. With one PCR primer specific to the transgene and the other to the primer template it is now possible to PCR amplify the unknown flanking DNA between them (Figure 5).

In the case of line-8 mice, the use of the restriction enzyme NSI I results in a 3' overhang approximately 6-8Kb from the 5' end of the transgene. A 24mer primer template was synthesized with the last 4 bases complementary to the NSI I overhang. Total genomic DNA was digested and separated on an agarose gel. DNA between 4-9Kb was excised (Figure 6A) and purified by spinning through glass wool. The anchor primer template was then ligated to the free ends. A primer specific to a portion of the transgenic MT gene and the complementary sequence primer to the anchor template were used in the PCR reaction.

Figure 6B shows that no PCR amplification between the MT primer and the anchor template was obtained using either size fractionated or total genomic DNA (lanes 2&3). However, DNA amplification was obtained using standard primers for MTPVLT transgene (lane 6) in both the size fractionated and total genomic DNA (lanes 4&5) confirming the presence of the desired DNA fragment.

To test ligation efficiency and the binding of both internal and external primers a control experiment was devised. Plasmid p213, a PML derivative, containing the entire PVLT gene was digested with Nsi I to generate a linear DNA fragment approximately 8.3Kb with 3' overhangs. The external primer was ligated and several PCR reaction were done. The results are shown in figure 7. Lanes 2 & 6 show a PCR amplified fragment of expected size corresponding to the use of a primer specific to PVLT starting at position 175 and the primer complementary to the anchored linker using PFU and Taq polymerases, respectively. Lanes 3 & 7 show the lack of an expected 7.3kb fragment using a PVLT primer starting at position 926 and the same complementary primer to the anchor using PFU and Taq, respectively. Lanes 4&5 results in an expected 751 control fragment and the 160bp fragment respectively. From this data it can be concluded that the ligation of the anchor is efficient and the PCR primers were properly synthesized.

Figure 5

Schematic representation of Direct PCR technique.

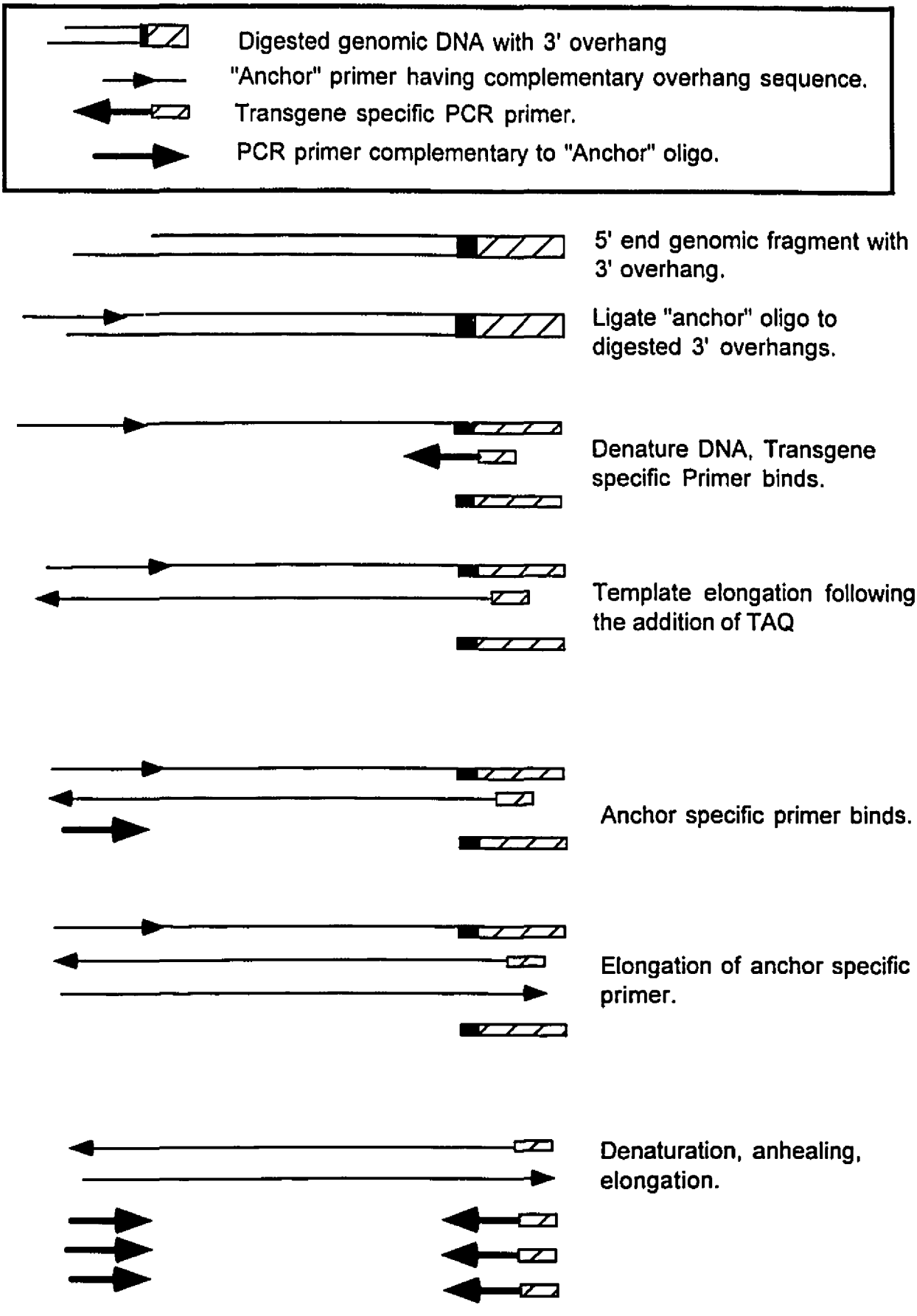


Figure 5

Schematic of direct PCR technique

Figure 6

PCR amplification of size fractionated genomic DNA.

Panel A: Excision of Nsi I digested MTPVLT-8 genomic DNA containing 5' end flanking fragment as determined by southern blotting (see Figure 3)

Panel B: PCR results performed on both size fractionated and total genomic DNA following ligation of the anchor primer. Lane 1 is the lambda/Hind III marker. Lanes 2 and 3 show the absence of any amplification product between the anchor template PCR primer and the transgenic internal primer in the size fractionated and total genomic samples respectively. Lanes 4 & 5 represent PCR products expected using the 175-926 PVLT specific primers in the size fractionated and total genomic samples respectively. Lane 5 is a PCR control using the 175-926 primers on the plasmid pMTPVLT.

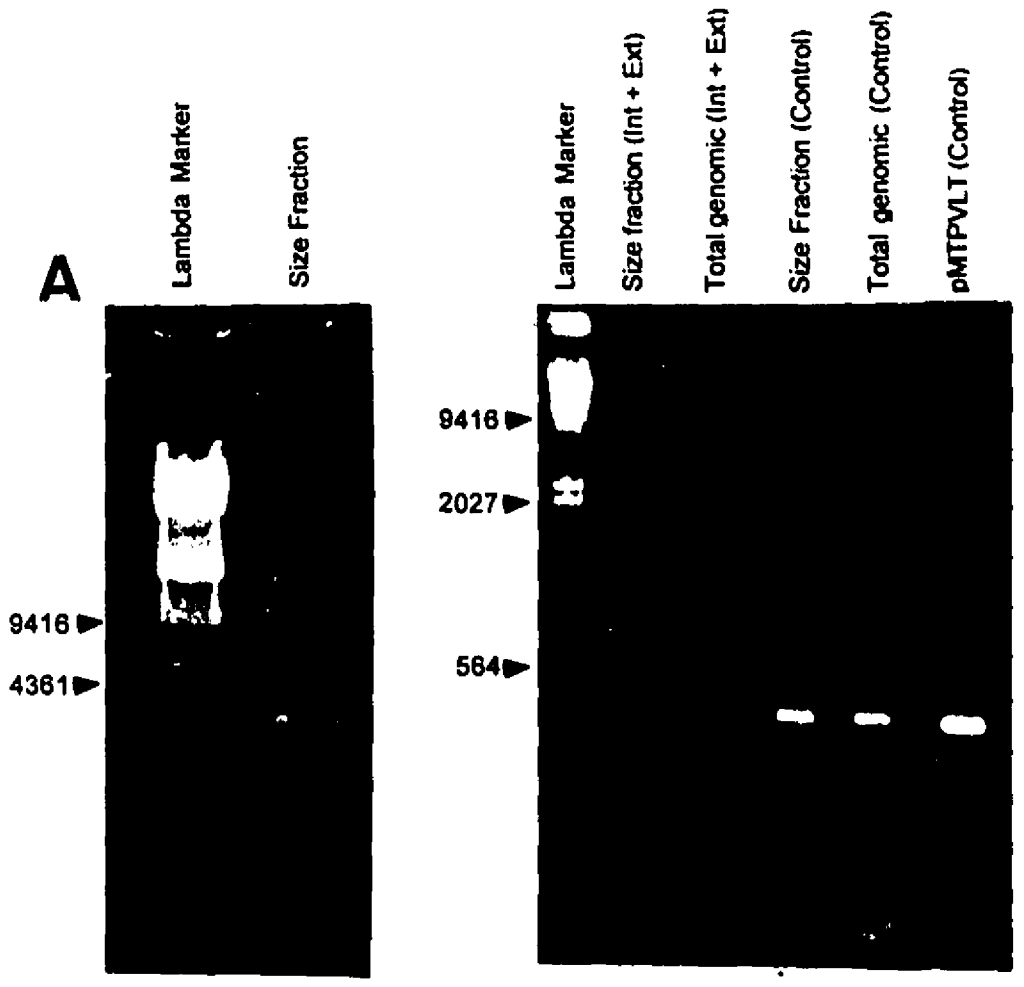


Figure 6

PCR amplification of size fractioned genomic DNA

Figure 7

Determination of anchor template ligation, primer binding and PCR amplification.

PCR amplification was performed on an Nsi I digested p213 plasmid containing the entire PVLT coding region in a PML derivative. Total plasmid size was estimated at 8.3Kb. Internal primers specific to PVLT and starting at the position indicated were used in conjunction with other internal primers or an external primer which is specific to the anchor primer template. Lane 1 is the lambda/Hind III marker. Lanes 2 & 6 represent PCR products resulting in using the 175 internal primer and the external primer using PFU and Taq respectively. Lanes 3 & 7 represent PCR products using a 926 internal primer and the external primer using PFU and Taq respectively. Lane 4 is a PCR reaction using two internal PVLT primers starting at 175 & 926. Lane 5 is a PCR reaction using an internal primer starting at position 2090 and the external primer.

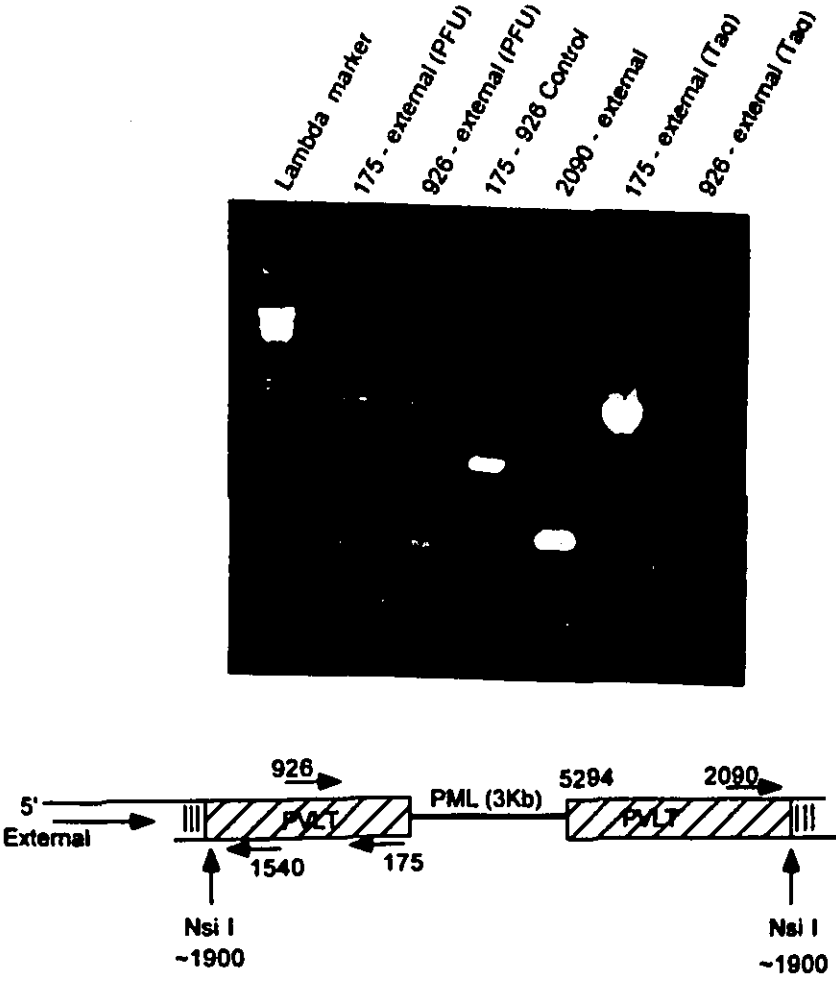


Figure 7

Determination of anchor template ligation, primer binding and PCR amplification

Several attempts were made to amplify the flanking DNA by PCR using various parameters. A range of annealing temperatures, nucleotide and magnesium chloride concentrations were tested. Polymerases such as Taq, PFU, and Vent DNA Polymerase were tested along with PCR helper products such as Taq Extender and Perfect Match DNA polymerase enhancer. However, no amplification of DNA within the expected size range (4-9Kb) was ever observed.

4.2.2: Construction and Screening of MTPVLT-8 mini-library

Cloning DNA fragments into bacterial plasmids is a standard and quick method for amplifying DNA for analysis. This technique can also be used as a mini library where multiple DNA fragments are cloned into bacterial plasmids which are later screened for the presence of specific DNA fragments.

A bacterial library was constructed using the 4.5-9Kb range of Eco RI digested fragments containing the 5' flanking regions as determined by Southern blotting. Figure 8 shows the genomic DNA following extraction and purification. This DNA was cloned into pUC 18 cloning vectors and used in the transformation of competent E.coli cells. The bacteria were plated and colony lifts were performed.

To isolate the 5' flanking region, the filters were probed with a PCR specific probe to the PVLT region 5' of the EcoR1 site. Figure 9A shows a filter with a potential bacterial colony containing the flanking region. Several areas of the colony as well as surrounding growth were replated and rescreened as before (Figure 9B). A single colony from the second screening was used for a final screening (Figure 9C). All but colonies J came out positive following probing. It appears that streak G in the second screening contained a mixture of positive and negative colonies. In the third screening G appeared positive, but in a subsequent trial it was negative (Data not shown).

Minipreps of each colony were prepared and digested with the restriction enzyme Eco R1 and run on an agarose gel (Figure 10A). It is obvious that there was incomplete digestion. It can be postulated that colonies A-F and H, I originated from the same initial colony. Colonies G, J are identical but different from the others. The top band is likely E.coli genomic DNA, the second, third and fifth bands represent the 3 forms of uncut plasmid DNA as compared with clone E which was undigested. Bands 4 and 6 represent the two plasmid fragments.

Figure 10B is a Southern blot using the 5' end PCR probe to PVLT. It shows that clones G, J and the randomly chosen control plasmids came out negative for a PVLT specific insert. The other lanes appear to contain the same cloned plasmid and have homology to the PVLT specific probe indicating a potential insert.

Figure 8

Excision and purification of Eco RI digested DNA

Panel A: Excision of Eco RI digested total MTPVLT-8 mouse genomic DNA between 4.5-9Kb.

Panel B: Sample of DNA recovered either by spinning through Glass Wool, or by using a sephaglass purification kit (Pharmacia).

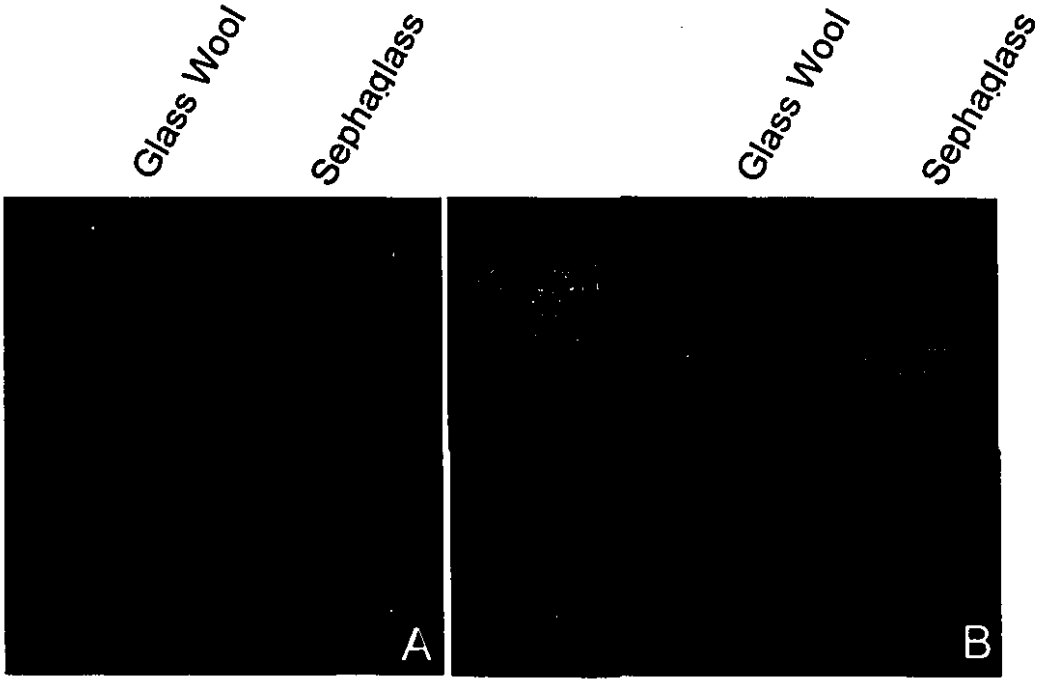


Figure 8

EcoR1 size fraction extraction and purification

Figure 9

First, second and third screening of the mouse mini-library

Panel A: Probing the colony lifts using a specific PVLT probe identifies a potential positive colony.

Panel B: 2nd screening using multiple streaks from the primary colony and its surrounding area.

Panel C: 3rd screening using picks of a single colony from the second screening.

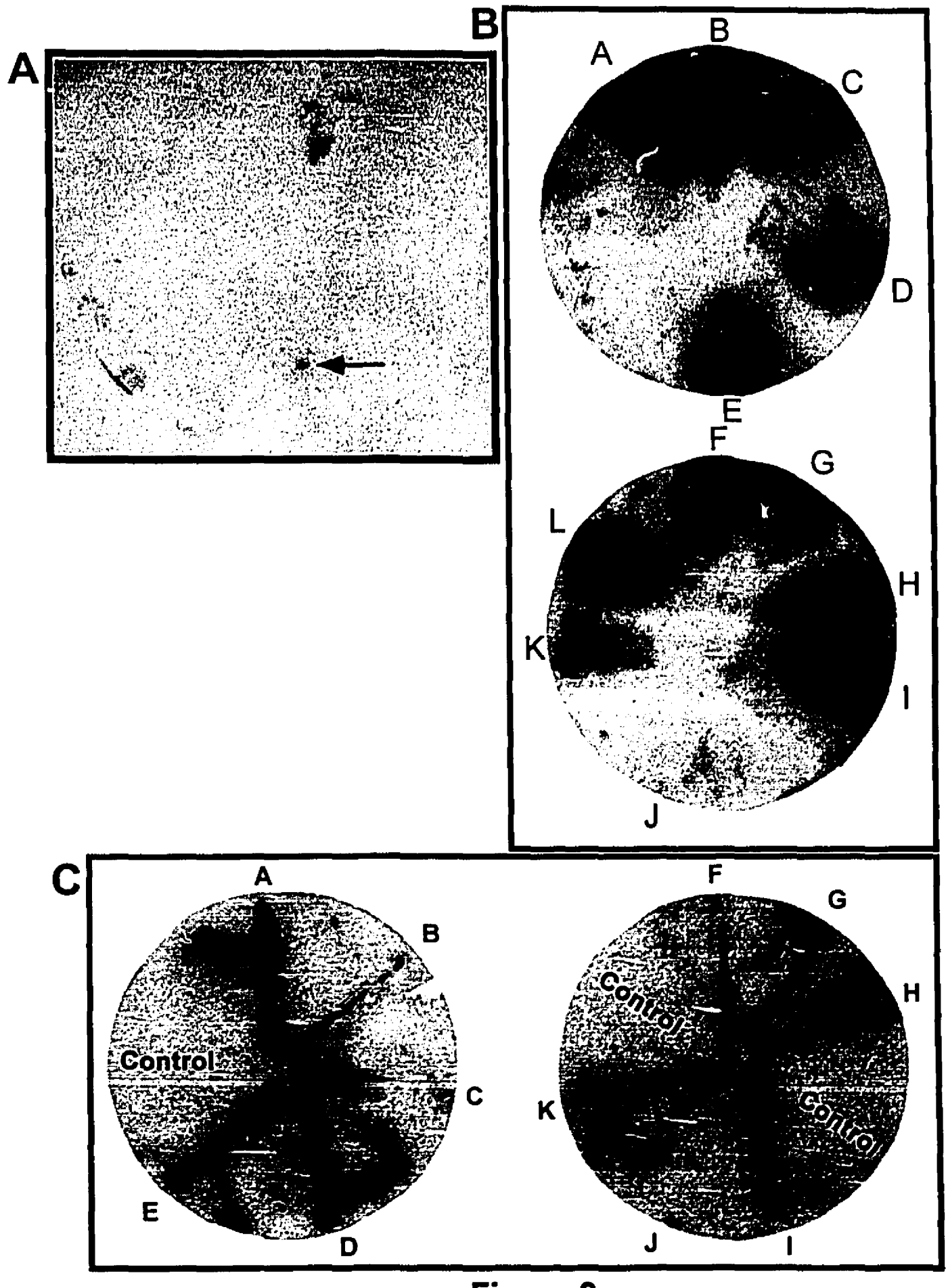


Figure 9
Screening of MTPVLT-8 mini library

Figure 10

Eco RI digestion of plasmid clone and Southern blotting

Panel A: Plasmid preps were grown using a single colony from the 3rd screening. Lanes A-J represent plasmids of the clones digested with Eco RI. Both Control lanes represent other plasmids from the lab also digested with Eco R1. E-Uncut is clone E plasmid that was not digested, and MTPVLT is uncut pMTPVLT run as a hybridization control samples were run on a 1% agarose gel.

Panel B: Southern blot of gel in Panel B following probing with a PVLT specific PCR probe.

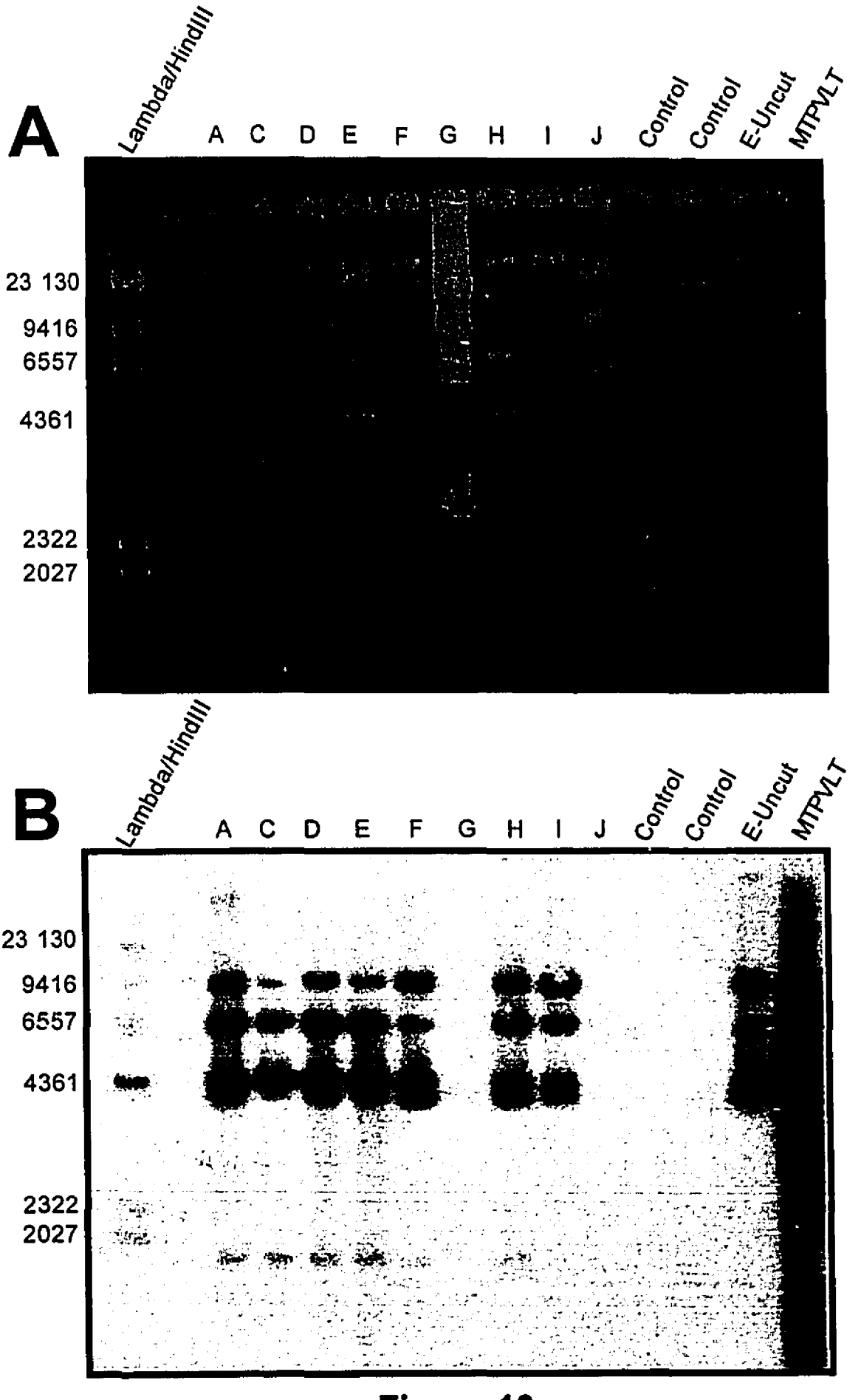


Figure 10
Screening of minipreps

4.3: Isolation and verification of 5' flanking region from a phage library.

A phage clone (designated clone 6) was isolated from a MTPVLT-8 genomic DNA phage library by a colleague using a MTPVLT specific probe. To verify the presence of the transgene within clone-6 sequencing analysis was performed. A primer specific to PVLT was used to sequence 100 bases of clone-6 and the control plasmid pMTPVLT. Sequence data was compared and found to be identical. Southern blot analyis comparing DNA fragments obtained from clone-6 to DNA from transgenic mice were consistant with DNA 5' to the transgene (Figures 2 and 11).

4.4 Use of Clone-6 to identify 5' restriction sites.

Southern blotting using clone-6 digested genomic DNA was carried out to identify restriction sites flanking the transgene. The same filter was probed, stripped and reprobed with four different probes, a random primed probe of MT coding region, a PCR probe using 175-926 primers for LT, a PCR probe using 1540-2090 primers for LT, and a random primed probe of a ApaI/HincII fragment specific to the 3' end of the transgene (Figure 11). Taking Eco RI as an example in panel 1 of figure 11, it can seen that the MT probe detects a 5' flanking fragment of approximately 9 Kb. Similarly the 175-926 probe detects the same fragment, this is expected since the Eco R1 site within the transgene falls 3' to probe sequences (Figure 1 and 4). When probing with the PCR probe 1540-2090 two bands are observed, this is due to the fact that the PCR probe is specific to DNA sequence 5' and 3' to this restriction site (Figure 1). Probing with the ApaI/HincII probe which is 3' to the internal Eco RI site binds the 3' end flanking fragment exclusively. The 3' end fragment in this particular experiment contains lambda vector DNA and is of no importance in this case.

4.5 Use of clone-6 to determine possible gene rearrangements surrounding the transgene integration site.

Major insertional mutations such as gene rearangements, insertions and deletions during transgene integration can greatly influence trangene expression. To identify any major mutations to the mouse genome during trangene integraton clone-6 which represents DNA flanking the integration site of MTPVLT-8 transgenic mice was used. This clone was used as a genetic probe to restriction digests of MTPVLT-8 and wildtype CD1 mouse genomic DNA. Comparisons between digested CD1 and MTPVLT-8 genomic DNA can then be made by Southern blots (Figure 12). We observed identical banding patterns between CD-1 and transgenic mice. These results suggest that the DNA surrounding the trangene is identical to that of wildtype DNA and that no major alteration occured. If major deletions or insertions occured different banding patterns would have been expected.

4.6: Determining whether the flanking DNA codes for an expressed gene.

Expression of the polyomavirus in line-8 heart tissue suggests that the transgene had integrated close to an enhancer or promoter region. Most likely an endogenous gene would be present in the relative area. The possibility of integration near or even within an endogenous gene is possible. To identify such a possibility Northern blotting was performed. Total RNA from various mouse tissues was isolated and run on a denaturing gel. A random primed probe using the entire clone-6 as template was used. A single band was observed in all tissues (Figure 13). No band was observed in lung tissue due to degradation of the RNA in this case. However, this experiment was repeated with similar results including expression in lung tissue (data not shown). It is, therefore, likely that the DNA flanking the 5' end of the transgene does in fact contain part of a gene which is expressed in normal tissues.

4.7: Homology with other species

To determine whether or not the transgene had integrated into a region of the mouse genome that may be conserved within species Southern blotting was performed. Total genomic DNA from non-transgenic Mouse, Human, Monkey, Dog, Chicken, and *Drosophila* was digested with Hind III restriction enzyme and separated on an agarose gel. The DNA was transferred to a membrane and a Southern blot performed. A probe to the entire clone 6 fragment was used in the hybridization. Homology was observed only in the case of the mouse DNA (Figure 14A).

To test the integrety of the DNA, the filter was stripped and reprobed using a β-actin probe (Figure 14B). Actin-specific bands were observed in all DNAs indicating that the DNA was present on the blot and capable of hybridizing in all species.

Figure 11

Southern blots of Clone-6 differentially probed with MTPVLT fragments

Column A: Blots probed with a random primed radioactive label specific to the MT portion of the transgene.

Column B: Blots probed with a PCR primed radioactive label specific to the protion of the PVLT cDNA falling within sequence bp 175-926.

Column C: Blots probed with a PCR primed radioactive label specific to the portion of the PVLT cDNA falling within sequence bp 1540-2090

Column d: Blots probed with a random primed radioactive label specific to a portion of the 3' end of the transgene consisting of the Apa1/HincII fragment.

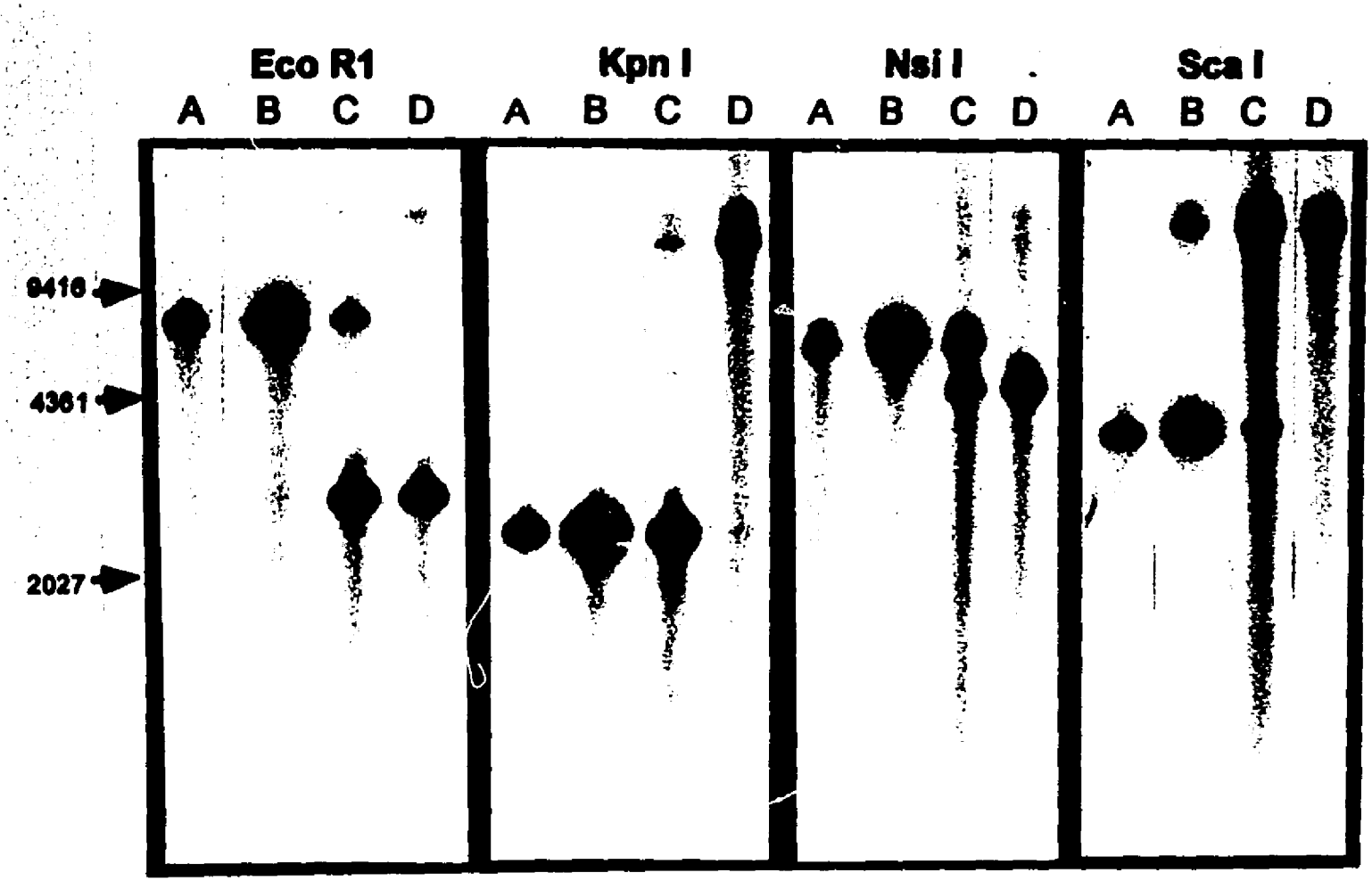


FIGURE 11

Southern blots of Clone-6 differentially probed with MTPVLT fragments

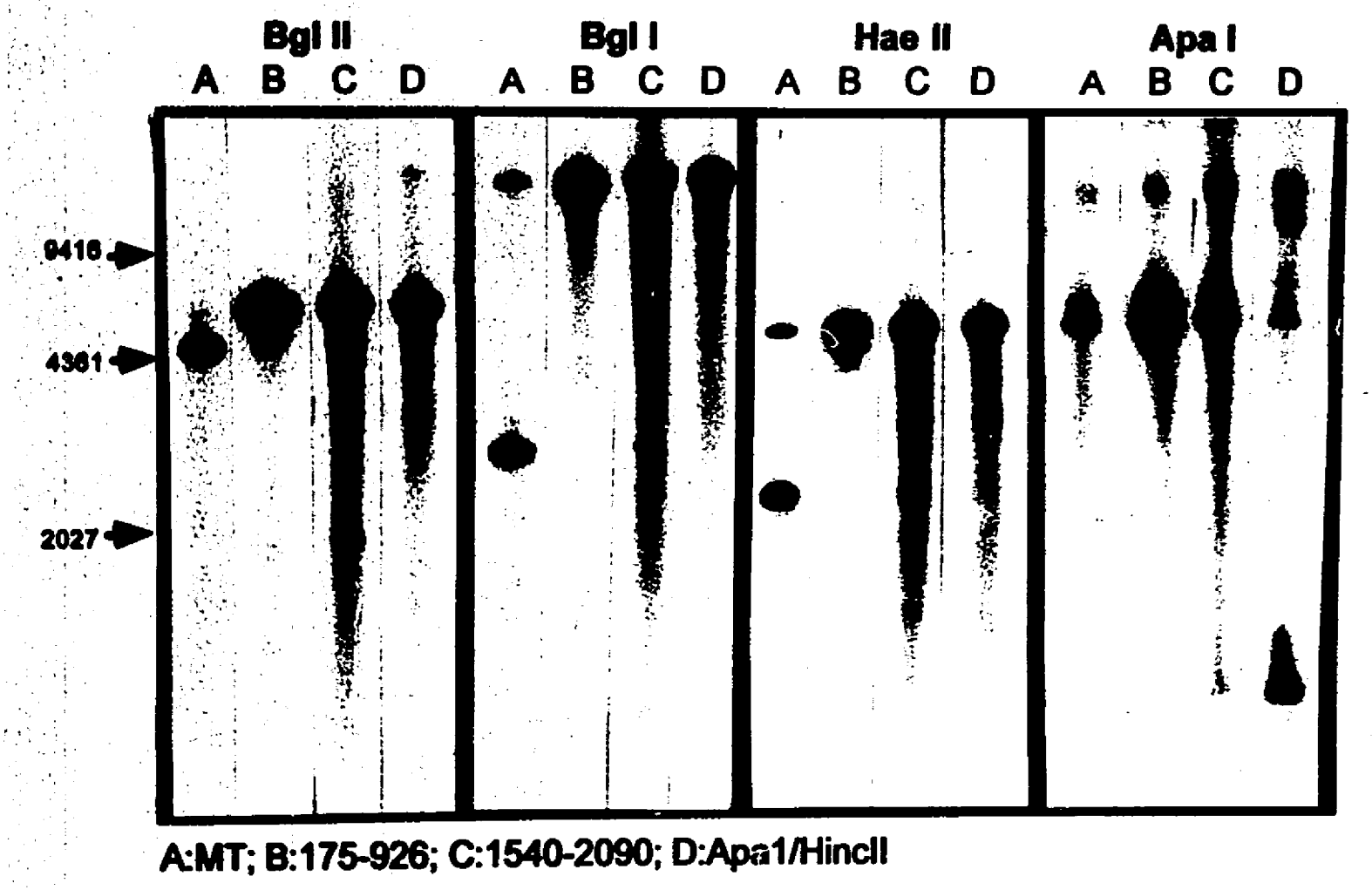


FIGURE 11

Southern blots of Clone-6 differentially probed with MTPVLT fragments

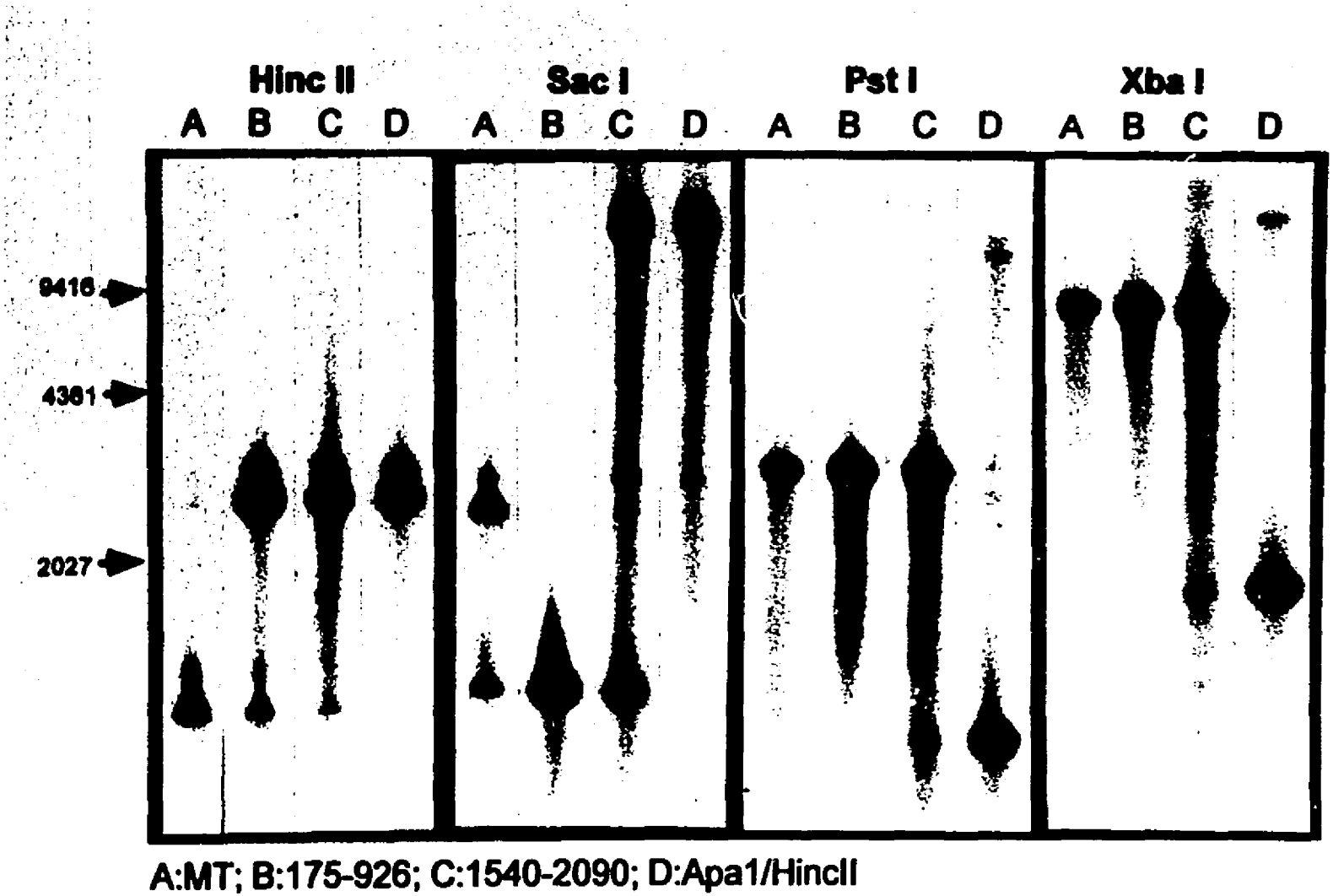


FIGURE 11

Southern blots of Clone-6 differentially probed with MTPVLT fragments

Figure 12

Southern blot comparison of transgenic and non-transgenic mouse DNA surrounding the transgene integration site

Transgenic mouse DNA and non-transgenic CD-1 mouse DNA was digested with Eco RI and Bam HI. The Southern blot was probed with a random primed radioactive labed using Clone-6 as template.

Figure 12

Southern blot comparison of transgenic and non-transgenic mouse DNA surrounding the integration site

Figure 13

Northern blot of mRNA from MTPVLT-8 mice using clone-6 as probe.

Panel A: 28 and 18S bands as seen following electrophoresis on a 1% denaturing gel.

Panel B: 28 and 18S bands as seen following staining of membrane with Methylene Blue.

Panel C: Northern blot following a 3 day exposure using a random probe specific to all of clone-6.

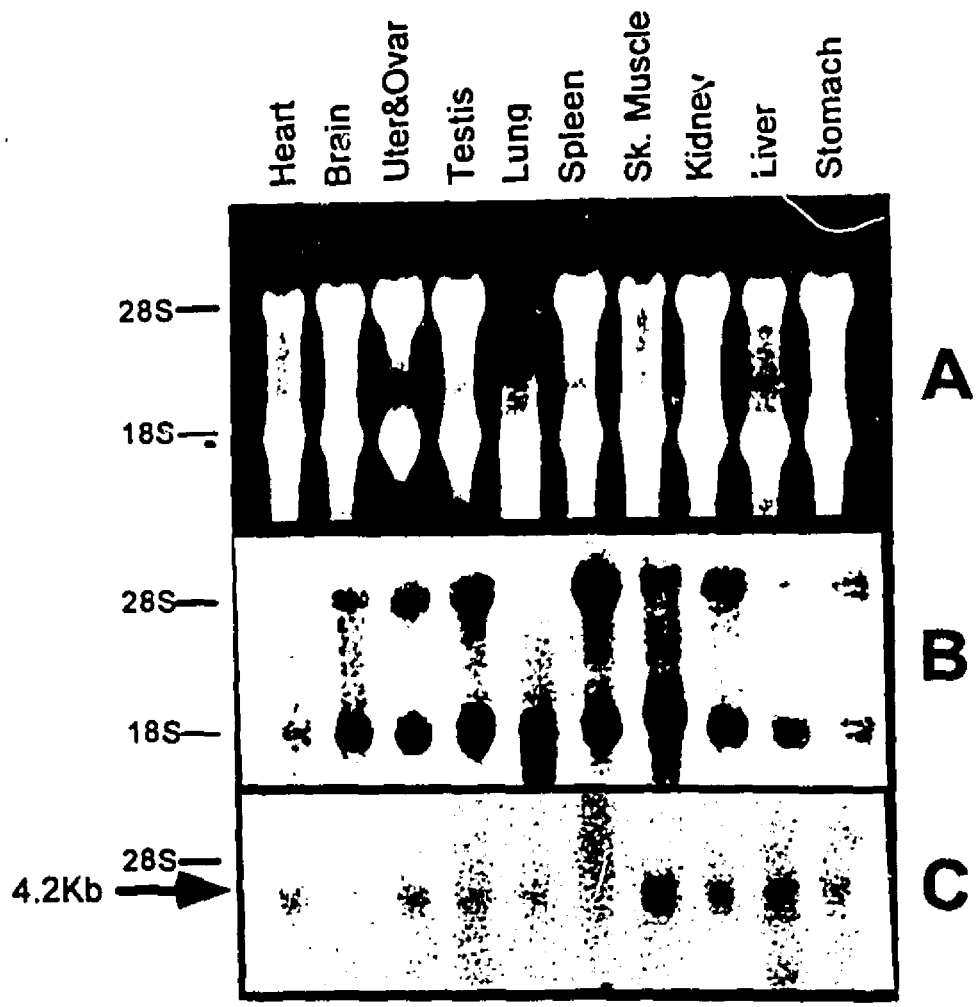


Figure 13

Northern blot of mRNA from MTPVLT-8 mice using clone-6 as probe template

Figure 14

Clone-6 sequence homology with other species

Panel A: Hind III digested Mouse, Rabbit, Chicken, Dog, Monkey, and *Drosophila* genomic DNA run on a 0.8% agarose gel.

Panel B: Southern blot of gel in panel A using random primed probe specific to the entire sequence of clone-6.

Figure 14

Clone-6 sequence homology with other species

4.8: Sequencing

To determine if any DNA markers may identify the integration site, sequencing of approximately 750bp of the DNA fragment contained in clone 6 was performed. Figure 15 contains the entire sequence and sequencing primers used. Searches of the the EMBL/genebank databases indicate that some homology between the human guanalin gene is present (Figure 16).

Figure 15

DNA sequence of clone-6 starting at the T3 promoter (5') end. All subsequent sequence primers have been underlined. The fourth and reverse primer are complementary sequences.

```
  1 CACAGAGATT CAGGAAAATT TCAGGAAACC TCATGGGGCT TCAACATCTT
 51 TGCACCAGAC ACGTGGGCTG CATGCACTCT CATTGATGCC CTGTGGTCTG
101 ATAAACACTG TCATAACCTG CCACTAGTAA GTGTCAGCCA CCTGATAATG
151 GGCTCATCAG TTTATACCTG CACTAGCTAC ATGGGTCCCA GGCAGGAAGC
201 AAGGTGGAGC TGAGTTTCAG AGCATCTGAG TCCAATGACA GATTCCAAGG
251 TCCTTGATTC AAGCTGAATT TCCAGGACAT ACAACCAGCC CAACCTGACA
301 CACCACTATG GGTGGGCACA TGGAGCTAGC CTGAGGGGGT TAGAAGCTCT
351 GGAGTAGAAA TGCCACTCAC TCTATCCTAA ATCCTTGGGT ACTTTCCTTC
401 CAGATGCAGA AAGCCCTCCT GAGCAGTCCA CTGTGGTGGT GGGCAATGGC
451 CATGGCTTGC TGCTGTCGCT CCAGCTGACA CACTGCTGCC TGCTGGTGCC
501 ACACGAACCA GAACCTACGC AAAAGCTGCT GCTCTGCCTG CAGTATAGCC
551 TGAGGCAAGG GTGGGAGGAA ACAGAACAAG TCCAGGAAGG AGAGACAGTT
601 CAGAGGCTGG AGGAAATCTG TGACCAGCAA AAGAGATCCT CCCCGAGCAG
651 CTGCTTATGA ACTGAGCTGA CAACAGACAT CAATGTCTGG ATTTTCCAGG
701 TCTGTCTTCG CAG
```

GGCTCATCAGTTTATACCTGC	Second	Sequencing Primer
CACTCTATCCTAAATCCTTGGGTA	Third	Sequencing Primer
AAAGCTGCTGCTCTGCCTGC	Fourth	Sequencing Primer
TTTCGACGACGAGACGGACG	Reverse	Sequencing Primer

Figure 15

Figure 16

Results of a Blast EMBL/genebank sequence comparison search to the 5' end sequencing data obtained from clone-6

```
gb|M95174|HUMGUANYLA Human guanylin mRNA, complete cds.
           Length = 567

  Plus Strand HSPs:

 Score = 124 (34.3 bits), Expect = 6.1, P = 1.0
 Identities = 28/32 (87%), Positives = 28/32 (87%), Strand = Plus
/ Plus

Query:   592 GAGACAGTTCAGAGGCTGGAGGAAATCTGTGA 623
             |||| | ||||||||||||||||||||  |||
Sbjct:   269 GAGA[illegible]ACTTCAGAGGCTGGAGGAAATCGCTGA 300
```

Figure 16

CHAPTER 5

DISCUSSION

DISCUSSION

Wang et al. (1991) showed that the presence of the complete polyomavirus early region gene causes numerous tumors of mesencymal origin specifically vascular and bone tumors in mice. Although the mice did not develop testicular adenomas, RNAs of Py Large-T antigen (LT), MT, and small T antigen RNA were present in tumors as well as testis. In two of three lineages studied by Wang et al., expression of the transgene, by RNA protection analysis, was limited to unaffected testis and tissues that developed tumors. In both lineages tumor formation included osteasarcamas and rare occurrences of fibrosarcomsa and hemangiomas. The third lineage expressed the transgene in testis and unaffected tissues including spleen, liver, kidney as well as in ostosarcomas, hemangiomas and lymphangiomas.

Of the six lines of transgenic mice harboring a Polyomavirus Large-T antigen transgene produced in our lab, all express the transgene in testis, three develop testicular adenomas, but only one line expresses the gene in both testis and heart. This unique line of mice (MTPVLT-8) develops both testicular adenomas as well as cardiomyopathies which leads to premature death (Chalifour et al. 1990, 1993). The underlying cause of this phenotype is the expression of polyomavirus large-T (PVLT) antigen in cardiac myocytes. Expression of polyomavirus RNA in testis was not surprising as it is very common in transgenic mice harboring a polyomavirus transgene (Wang et al. 1991).

It may be hypothesized that the exclusive expression of Py LT in heart, regarding expression in testis as a common phenomenon, is due to chromosomal position effect. The transgene's restricted expression to the heart is most likely due to the presence of a positively acting regulatory element driving its expression in heart. This element, being either an enhancer or promoter, would have to be cardiac specific, otherwise we would expect to see PVLT expression in other tissues as well.

5.1 Isolation of DNA flanking the transgene in MTPVLT-8 mice

To determine the presence of a possible enhancer element driving the expression of LT in MTPVLT-8 transgenic mice, in a similar fashion used to identify the cTnC cardiac enhancer, the DNA surrounding the integration site must first be isolated and examined. Several techniques to isolate the DNA flanking the transgene were attempted.

5.1.1 Direct PCR:

The technique of directly PCR amplifying DNA flanking the transgene was successfully used by MacGregor and Overbeek (1991) to clone a 2.1Kb fragment. Although simple and straight forward in its application, in our trials it proved problematic in many areas. The requirement of a 3' overhang was essential and the distance between the transgene and the overhang is a variable of the DNA flanking the transgene. In our case the smallest fragment obtained using a restriction enzyme providing a 3' overhang sequence was in the order of 6-8Kb. This size was desirable because of the range of which an enhancer can operate away from the promoter sequence. However, with many modifications to the original protocol to accommodate a larger fragment the process proved unsuccessful.

Control experiments for each step were carefully carried out to determine which process may be problematic. The amplification of a PCR product using transgene specific primers confirmed its presence in the size fraction (Figure 6). Tests for ligation efficiency and primer binding were carried out using a test plasmid. The plasmid was digested with Nsi I and the anchor primer was ligated to the overhangs. The presence of a PCR product using the 175 PVLT specific primer and that to the anchor sequence demonstrates that ligation was effective and that both primers bind (Figure 7). This control experiment demonstrated the plausibility of this technique with amplification of a fragment of approximately 1.8Kb. However, using total genomic or size fractionated genomic DNA has inherent variables that cannot be directly examined. The presence of numerous unwanted genomic DNA fragments competing in all reactions reduce the possibility of

success. There is no means of determining whether or not the anchor template had successfully ligated to the flanking DNA, and even with successful ligation, amplification using the PCR technique was difficult because of the size of the piece being amplified. Even with pure plasmid DNA, PCR trial runs failed to amplify fragments exceeding 6Kb (Figure 7). The use of various polymerases and helper products such as Taq Extender, Perfect Match did not prove useful.

5.1.2 Mini Genomic Library

In order to facilitate restriction mapping, sequencing and deletion analysis it was preferrable to clone the flanking DNA into plasmids. Plasmids were also desirable for their ease of being manipulated and reproduced.

Standard cloning procedures were used where size fractionated EcoR1 digested MTPVLT-8 genomic DNA was ligated to a predigested EcoR1 overhang PUC-18 vector. Prior to ligation the size fractionated DNA was analyzed for the presence of the 5' overhang using PCR amplification on known PVLT sequence. EcoR1 was chosen primarily because of its availability and price as well as the fact that southern analysis demonstrated that the 5' fragment was of relatively large size.

Screening using MTPVLT specific probe identified a potential clone which was later analyzed for an insert (Figures 9 & 10). Southern analysis of positive clones showed homology to MTPVLT specific probes. However, after careful observation it was discovered that these clones did not contain a pUC 18 cloned fragment, but in fact they had integrated the MTPVLT plasmid frequently used in the laboratory. Comparing the undigested E clone with MTPVLT used as a binding control, identical binding patterns were observed (Figure 10). Furthermore, the two fragment produced by the partial digestion with Eco R1 when added together result in the same sized fragment approximately 6.4Kb expected from a linear MTPVLT plasmid. It can therefore be concluded that MTPVLT was recloned, and this must have resulted from a contaminant in

one of the solutions. The project was terminated after the successful isolation of the 5' flanking region in a Line-8 phage genomic library by a colleague in the laboratory.

5.1.3 From a MTPVLT-8 genomic library

Total genomic DNA from MTPVLT-8 was digested and a genomic library was made by cloning into Lambda/Zap phage vectors by a co-worker. Screening the library with MTPVLT specific probe identified several potential clones. One clone, designated clone-6, was purified and demonstrated to have strong homology to MTPVLT probes.

5.2 Restriction mapping of 5' flanking DNA

Southern blotting of clone-6 digested DNA suggests that it contains a full copy of the transgene and 6-10Kb of 5' flanking genomic DNA. The 3' end of the transgene is flanked by the lambda vector and so is the far 5' end of the clone. Using Southern blots of digested MTPVL-8 genomic DNA together with the results of Southern blots of similarly treated clone 6 DNA (Figures 2 & 11) it was possible to construct a rough restriction map of the DNA surrounding the transgene (Table 1).

5.3 Potential gene rearrangements surrounding the trangene integration site.

Southern blots comparing digested wildtype CD1 and MTPVLT-8 mouse genomic DNA probed with clone 6 demonstrates identical banding patterns (Figure 12). This observation suggests that no major alterations of the DNA sequence flanking the transgene occurred during integration. Although these observations rule out major genetic changes such as large gene rearrangements, deletions and insertions, they cannot identify point mutations or small gene alteration which can only be resolved by sequencing.

Table 1

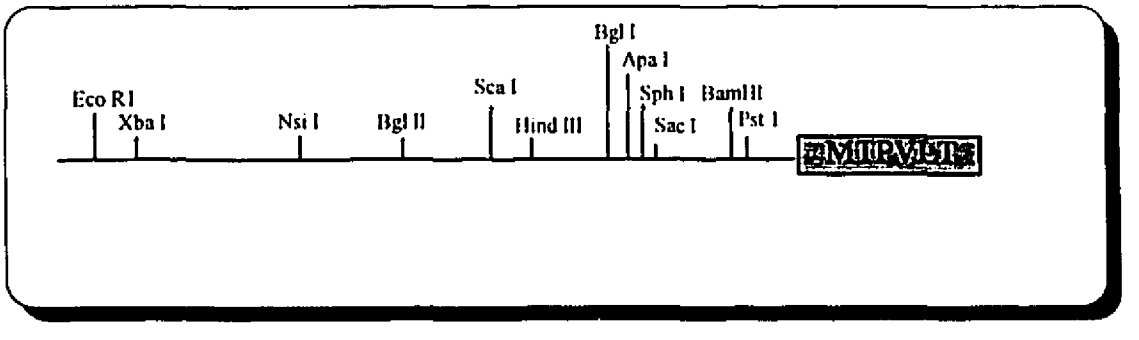
Size fragment determination from Southern blot results

Column 1 represents the restriction enzyme used to determine the size fragment, column 2 is the size of the fragments as observed from the Southern blot resuits, column 3 is the known distance of the restriction site within the transgene from the 5' end of the transgene, column 4 is the actual size of the 5' flanking fragment corresponding to mouse genomic DNA.

A schematic of the order and approximate location of the restriction sites 5' of the transgene is also illustrated.

TABLE 1

Restriction Enzyme	Observed Fragment Kb	Dist. from 5' end of transgene to rest. site within the transgene	Actual 5' flanking fragment corresponding to mouse gemomic DNA.
Eco R1	8.9	1.5	7.4
Nsi I	7.5	1.8	5.7
Sca I	3.8	0.8	3
Bgl II	4.5	0.4	4.1
Bgl I	2.7	0.15	2.55
Hae II	2.2	0.2	2
Apa I	4.6	2.3	2.3
Sac I	2.2	0.4	1.8
Pst I	3	2.2	0.8
Xba I	9.2	2.2	7
Sph I	3.6	1.5	2.1
Bam HI	4	3.1	0.9
Hind III	3.2	0.4	2.8



5.4 Evidence of a potential endogenous gene near the transgene integration site.

On the premise that gene expression in cardiac tissue is due to an endogenous enhancer driving the MT promoter, if follows that an endogenous gene could also be near the integration site. To determine the presence of an endogenous gene, total RNA from various organs of a CD1 mouse was extracted for Northern analysis. Clone-6 which represents about 6-10Kb of the 5' flanking DNA was used as a probe. If the coding area of an endogenous transcribed gene was present within the clone-6 it would hybridize with tissue mRNA.

The results obtained through Northern analysis showed homology to a single band present in all tissues assayed. Taking unequal loading into consideration there appears to be no significant difference in band intensity among the different tissues. The data obtained was consistent and reproducible. This finding was surprising since trangene PVLT RNA was shown to be present only in testis and cardiac tissue (Chalifour et al. 1990). This gene appears to be constitutively expressed in all tissues.

These results do not rule out the possibility of a cardiac specific enhancer. This gene may be located in the vicinity of a cardiac specific gene and transcribed from independent transcriptional machinery. Furthermore, the enhancer may be located 3' of the transgene, a possibility that is not currently being investigated.

5.5 Homology and sequence comparisons

The DNA flanking the 5' end of the transgene, clone 6, was used to probe digests of genomic DNA from different species. Of all the species used, this particular piece of DNA bound only to mouse DNA. This result was consistent in a second screening using a different restriction enzymes. Moreover, in two trials there was no apparent banding, but rather a smear of homologous DNA. (Figure 14). However, when the same probe was

used to compare transgenic and non-transgenic mice banding was evident above that of the background smear (Figure 12). Reprobing the filter using a conserved β-actin shows unique and distinct bands in all animals. This may suggest that Clone-6 is truly specific to mouse DNA. It also implies that this piece of DNA may be homologous to a family of genes that are repeated throughout the mouse genome, leading to the observed smearing.

5.6 Sequence comparisions

A gene-bank sequence comparison using 750bp of the 5' end of clone-6 matched 26bp of the human guanalin gene with 87% homology. The human guanylin gene is believed to be a member of a family of genes constituatively expressed in many tissues which appears to be consistent with the Northern blot data showing the expression of a constitative gene in several mouse tissues. It can be postulated that there is homology between the mouse and human guanylin gene. More experimentation must be done to conclude if in fact a mouse analog of human guanylin or a close relative was found or whether there is only a small sequence of coincidental homology.

5.7 Future experiments

Identifiying enhancer sequences is often a long and tedious undertaking. To classify a sequence element as an enhancer the following criteria must be met. An enhancer sequence must show an ability to increase transcription of cis-linked promotors, operate in an orientation independent manner, and exhibit its effect over large distances independent of position (Atcheson 1988). Standard techniques for identifying enhancer elements has been to insert a sequence of DNA, presumed to contain the enhancer element, downstream of a promoter-reporter gene construct (Parmacek et al. 1992). This construct once introduced into appropriate cells would produce an assayable gene. If an enhancer element is present the amount of product should significantly be higher than that without an upstream insert. To identify the exact region of the enhancer deletions are introduced from one end of the insert until expression is observed to decline. In our

particular case standard techniques are difficult due to the lack of an appropriate cardiac cell line. At present there are no known cardiac cell lines available for this type of analysis. Furthermore, it is often observed that expression in culture is not representative of that in *vivo*.Accordingly, expression must be analyzed in *vivo* directly. To achieve this, it is essential to produce new transgenic mice harboring the DNA insert with a potential enhancer linked to a readily assayable marker in the tissues. A visible marker such as B-gal has been used in similar experiments (Al Moustafa et al. 1992). If an enhancer element is present, precise isolation of the sequence would entail deletions to the original insert and reintroduction of the new constructs into new transgenic mice. Inherent problems in this technique is the time factor for proper growth and development of the mice before assaying and the likelihood of transgene integration expressional differences occuring. Presently there is enough information on the 5' flanking region to attempt such experiments.

REFERENCES

REFERENCES

Acsadi, G., Jiao, S.S., Jani, A., Duke, D., Williams, Pl, Chong, W., Wolff, J.A., 1991. Direct gene transfer and expression into the heart in vivo. *New Biol.* **3**:71-81.

Al Moustafa, A. E., Quatannens, B., Dieterlen-Lievre, F., Saule, S. 1992. Tumorigenic effects mediated in the avian embryo by one or more oncogenes associated with v-myc. *Oncogene* **7(8)**:1667-1670.

Al-Shawi, R., Kinnaird, J., Burke, J., Bishop, J., 1990. Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect. *Mol. Cell. Bio.* **10(3)**:1192-1198.

Atchison, L. M., 1988.Enhancers: Mechanisms of action and cell specificity. *Ann. Rev. Cell Biol.* **4**:127-53.

Ausebel, F M., Brent, R., Kingston, R. E., Moore, D. D., Seidmen, J. G., Struhl, K., (eds), Current Protocols in Molecular Biology, Vol I, Greene Publishing Associates and Wiley-Interscience, NewYork, (1994).

Banerji, J., Rusconi, S., Schaffner, W. 1981. Expression of a beta-gobin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**:299-308.

Banerji, J., Olson. L., Schaffner. W. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**:729-740.

Benoist, C., Chambon, P. 1980. Deletions covering the putative promotor region of early mRNAs of simian virus 40 do not abolish T-antigen expression. *Proc. Natl. Acad. Sci. USA* **77**:3865-3869.

Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Avarbock, m. R. 1980. Translation of globin messenger RNA by the mouse ovum. *Nature* **283**:449-501.

Brinster, R. L., Chen, H. Y., Trumbauer, M. E. 1981a. Mouse oocytes transcribe injected *Xenopus* 5S RNA genes. *Science* **211**:396-398.

Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Senear, A. W., Warren, R., Palmiter, R. D. 1981b. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* **27**:223-31.

Brinster, R. L., Ritchie, K. A., Hammer, R. E., O'Brien, R. L., Arp, B., Storb, U. 1983. Expression of a microinjected immunoglobulin gene in the spleen of transgenic mice. *Nature* **306**:332-336.

Brinster, R. L., Chen, H. Y., Messing, A., van Dyke, T., Levine, A. J., Palmiter, R. D. 1984. Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors. *Cell* **37**:357-379.

Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Yagle, M. K., Palmiter, R. D. 1985. Factors affecting the efficinecy of introducing foreign DNA into mice by microinjecting eggs. *Proc. Natl. Acad. Sci. USA* **82**:4438-4442.

Chada, K., Magram. J., Costantini, F. 1986. An embryonic pattern of expression of human fetal globin gene in trangenic mice. *Nature* **319**:685-688.

Chalifour, L. E., Gomes, M. L., Wang, N. S., Mes-Masson, A. M. 1990. Polyomavirus large T-antigen expression in heart of transgenic mice causes cardiomyopathy. *Oncogene* **5(11)**:1719-1726.

Chalifour, L. E., Mes-Masson, A. M., Gomes, M. L., Wang, N. S. 1992. Testicular adenoma and seminal vesicle engorgement in Polyomavirus large-T antigen transgenic mice. *Molecular Carcinogenesis* **5**:178-189

Constantini, F., Lacy, E. 1981. Introduction of a rabbit Beta-globin gene into mouse germ line. *Nature* **294**:92-94.

Covarrubias, L., Nishida, Y., Terao, M., D'Eustachio, P., Mintz, B. 1987. Cellular DNA arrangements and early develomental arrest caused by DNA insertion in transgenic mouse embryos. *Mol. Cell. Biol.* **7**:2243-2247.

Covarrubias, L., Nishida, Y., Terao, Mintz, B. 1986. Early postimplantation embryo lethatlity due to DNA rearrangements in a transgenic mouse strain. *Proc. Natl. Acad. Sci. U.S.A.* **83**:6020-6024.

Dawe, C.J., Freund, R., Mandel, G., Ballmer-Hofer, K., Talmage, D.A., Benjamin, T.L., 1987. Variations in polyoma virus genotypes in relation to tumor induction in mice. *Am. J. Pathol.* **127**:243-261.

Eddy, B.E. 1982. Polyomavirus in: The mouse in Biomedical Research, vol II, Foster, H.L., Smal, J.D., Fox, J.D. (eds), Academic Press, New York, U.S.A. pp 293-311.

Freund, R., Mandel, G., Carmichael, G.G., Barncastle, J.P., Dawe, C.J., Benjamin, T.L., 1987. Polyomavirus tumor induction in mice: influences of viral coding and non-coding sequences on tumor profiles. *J. Virol.* **61**:2232-2239.

Gridley, T., Soriano, P., Jaenisch, R. 1987. Insertional mutagenesis in mice. *Trends in Genetics*. **3(6)**:162-166.

Gross, L., 1970. The parotid tumor (polyoma) virus in: Oncognic Viruses, 2nd ed. Pergammon Press, Oxford, pp 651-750.

Gillies, S. D., Morrison, S. L., Oi, V. T., Tonegawa, S. 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33**:717-728.

Glaichenhaus, N., Mougneau , E., Connan, G., Rassoulzadegan, M., Cuzin, F. 1985. Cooperation between multiple oncogenes in rodent embryo fibroblasts: An experimental model of tumor progression. *Adv. Cancer. Res* **49**:291-305.

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., Ruddle, F. H. 1980. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci USA* **77**:7380-7384.

Gordon, J. W., Ruddle, F. H. 1981. Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science* **214**:1244-1246.

Gossler, A., Joyner, A. L., Rossant, J., Skarnes, W. C. 1989. Mouse embyonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* **244**:463-465.

Grosschedl, R., Birnstiel, M. L. 1980a. Identification of regulatory sequences in the prelude sequence of an H2A histone gene transcription *in vivo*. Proc. *Natl. Acad. Sci. USA* **77**:1432-1436.

Grosschedl, R., Birnstiel, M. L. 1980b. Spacer DNA sequence are essential for promotion of H2A histone gene transcription in vivo. *Proc. Natl. Acad.Sci. USA* **78**:943-947.

Hamada, H. 1986. Random isolation of gene activator elements from the human genome. *Mol. Cell. Bio.* **6**:4185-4194.

Hammer, R. E., Pursel, V. G., Rexroad, C. E. Jr., Wall, R. J., Bolt, D. J., et al. 1985. Production of transgenic rabbits, sheep and pigs by microinjection. *Nature* **315**:680-683.

Hanahan, E. 1985. Heritable formation of pancreatic beta-cell tumors in transgenic mice expressiong recombinant insulin/simian virus 40 oncogenes. *Nature* **315**:115-122.

Holder, E., Mitmaker, B., Alpert, L., Chalifour, L. 1995. Morphometry and muscle gene expression in hypertrophied heats from polyomavirus large T antigen transgenic mice. *Am. J. Physioi.* **269:**H86-H95.

Johnson, J.E., Wold, B.J., Hauschka, S.D. 1989. Muscle Creatin Kinase Sequence Elements Regulating Skeletal and Cardiac Muscle Expression in Transgenic Mice. *Mol. and Cell. Bio.* **9(8):**3393-3399.

Krumlauf, R., Chapman, V. M., Hammer, R. E., Brinster, R., Tilghman, S. M. 1985. Differential experession of -fetoprotein genes on the inactive X chromosome in extraembryonic and somatic tissues of transgenic mice. *Nature* **319:**224-226.

Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D., Costantini, F. 1983. A foreign Beta-globin gene in transgenic mice: integration at abnormal chromosomal positions and expression in inappropriate tissues. *Cell* **34:**343-358.

Lo, c. 1986. Localization of low abundance DNA sequences in tissue sections by in situ hybridization. *J. Cell Sci.* **8:**143-162.

Lin, H., Parmacek, M.S., Morle, G., Bolling, S., Leiden, J.M., 1990. Expression of recombinant genes in myocardium in vivo after direct injection of DNA. *Circulation* **82:**2217-2221.

MacGregor, GR. and Overbeek, PA. 1991. Use of a simplified single-site PCR to facilitate cloning of genomic DNA sequences flanking a transgene integration site. *PCR metods and applications* **1(2):**129-135.

McKnight, G. S. , Hammer, R. E., Kuenzel, E. A., Brinster, R. L. 1983. Expression of the chicken transferrin gene in transgenic mice. *Cell* **34:**335-341.

Mercola, M., Wang, X. F., Olsen, J., Calame, K. 1983. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. *Science* **221:**663-665.

Ornitz, D. M., Palmiter, R. D., Hammer, R. E., Brinster, R. L., Swift, G. H., MacDonald, R. J. 1985a. Specific expression of an elastase-human growth hormone fusion gene in pancreatic acinar cells of transgenic mice. *Nature* **313**:600-603.

Ornitz, D. M., Palmiter, R. D., Messing, A., Hammer, R. E., Pinkert, C. A., Brinster, R. L. 1985b. Elastase 1 promoter directs expression of human growth hormone and SV40 T- antigen genes to pancreatic acinar cells in transgenic mice. *Cold Spring Harb. Symp. Quant. Biol.* **50**:339-409.

Ottesen, J.L., Nilsson P., Jami. J., Weilguny. D., Duhrkop. M., Bucchini. D., Havelund. S., Fogh. J. M. 1994. The potential immunogenicity of human insulin and insulin analogues evaluated in a transgenic mouse model. *Diabetologia.* **37(12)**:1178-1185.

Palmiter, R. D., Chen, H. Y., Brinster, R. L. 1982. Differential regulation of metallothionein-thymidine kinase fusion genes in transgenic mice and their offspring. *Cell* **29**:701-710.

Palmiter, R. D., Wilkie, T. M., Cehn, H. Y., Brinster, R. L. 1983a. Transmission distortion and mosaicism in an unusual transgenic mouse pedigree. *Cell* **36**:869-877.

Palmiter, R. D., Norstedt, G., Gelinas, R. E., Hammer, R. E., Brinster, R. L. 1983b. Metallothionein-human GH fusion genes stimulate growth of mice. *Science* **222**:809-814.

Palmiter, R. D., Chen, H. Y., Messing, A., Brinster, R. L. 1985. SV40 enhancer and Large T-antigen are instumental in development of choroid plexus tumors in transgenic mice. Nature **316**:457-460.

Palmiter, R.D. and Brinster RL. 1986. Germline transformation of mice. *Ann Rev. Genet* **20**:465-500.

Palmiter, R. D. 1986. Molecular biology of metallothioneingene expression. In *2nd Int. meet. metallothionein and other Low Mol. Weight Metal-Binding Proteins*, ed. J. H. R. Kagi, Y. Kojima. Basel: Birkhauser-Verlag.

Parmacek, MS., Vora, AJ., Shen, T., Barr, E., Jung, F., Leiden, JM. 1992. Identification and charactarization of a cardiac-specific transcriptional regulatory element in the slow/cardiac troponin C gene. *Mol. Cell. Biol.* **12(5):**1967-1976.

Rassoulzadegan, M., Nahashfur, Z., Cowie., Carr, A., Grisoni, M., Kamen, R., Cuzin, F. 1983. Expression of the large T protein of polyoma virus promotes the establishement in culture of "normal" rodent fibroblasts cell lines. *Prod. Natl. Acad. Sci. U.S.A.* **80:**4354-4358.

Rochford, R., Campbell, B.A., Villareal, L.P., 1990. Genetic analysis of the enhancer requirements for polyomavirus DNA replication in mice. *J. Virol.* **64:**476-485.

Rosenberg, M., Court, D. 1979. Regulatory sequences involved in the prromotion and termination of RNA transcription. *Ann. Rev. Genet.* **13:**319-353.

Skarnes, W. C., Auerbach, B. A., Joyner, A. L., 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 & Development.* **6:**903-918.

Sternberg, E. A., Spizz, G. S., Perry, M. W., Vizard, D., Weil, T., Olson., E. N., 1988. Identification of upstream and intragenic regulatory elements that confer cell-type-restricted and differentiation-specific expression on the muscle creatin kinase gene. *Mol. Cell. Bio.* **8:**2896-2909

Storb, U., O'Brien, R. L., McMullen, M. D., Gollahon, K. A., Brinster, R. L. 1984. High expression of cloned immunoglobulin k gene in transgenic mice is restricted to B lymphocytes. *Nature* **310:**238-41.

Swift, G. H., Hammer, R. E., MacDonald, R. J., Brinster, R. L., 1984. Tissue-specific expression of the rat pancreatic elastase 1 gene in transgenic mice. *Cell* **38**:639-646.

Townes, T. M., Chen, H. Y., Lingrel, J. B., Palmiter, R. D., Brinster, R. L. 1985. Expression of human beta-globin genes in transgenic mice: effects of a flanking metallothionein-human growth hormone fusion gene. *Mol. Cell. Biol.* **5**:1977-83.

Wagner, T. E., Hoppe, P. C., Jollick, J. D., Scholl, D. R., Hodinka, R. L., Gault, J. B. 1981. Microinjection of a rabbit Beta-globin gene in zygotes and its subsequent expression in adult mice and their offspring. *Proc. Natl. Acad. sci. USA* **78**:6376-6380.

Wilkie, T.W., Palmiter, R.D. 1987. Analysis of the integrant in Myk-103 transgenic mice which fail to transmit the integrant. *Mol. Cell. Biol.* **7**:1646-1655.

Serfling, E., Jasin, M., Schaffner, W. 1985. Enhancers and eukaryotic gene transcription. *Trends Genet.* **1**:224-230.

Stewart, T. A., Wagner, E. F., Mintz, B. 1982. Human Beta-globin gene sequences injected inyo mouse eggs, retained in adults, and transmitted to progeny. *Science* **217**:1046-48.

Weiher, H., Noda, T., Gray, D.A., Sharpe, A.H., Jaenisch, R. 1990. Transgenic mouse model of kidney disese: insertional inactivation of unbiquitously expressed gene leads to nephrotic syndrome. Cell **62(3)**:425-434.