INFORMATION TO USERS

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

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

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

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

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

UMI

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

Université d'Ottawa · University of Ottawa

Targeted Disruption of the Sty Dual Specificity Kinase

Simon D. Ginsberg

Thesis submitted to the Department of Biochemistry in partial fulfillment of the requirements for degree of Master of Science

University of Ottawa
Ottawa, Ontario, Canada
September 1996

I certify that this thesis is solely of my own work, with the exception of the following:

Peter Duncan performed the construction of the knockout vectors, and the culturing, electroporation, and selection of the cells. He also performed the DNA extractions from the cell lines and mouse tail clippings.

Manon Dube performed the blastocyst injections, and the subsequent implantations into mice, as well as supervising the mouse breeding and tail clipping.

Abstract

Dual specificity kinases are known to have important roles in cellular regulation. The Sty dual specificity kinase has been shown to segregate with, and phosphorylate, SR family splicing factors, but the consequences of this are not known. Additionally, it is not known if the kinase has any other regulatory functions. In an attempt to determine the function of this protein in a mammalian system, Sty deficient mice were generated. This was accomplished through gene targeting, in ES cells, with a promoterless IRES-β*Geo* based vector. Using this efficient system, we obtained an 82% recombinant frequency. This is in contrast to a PGK-*neo* based vector, targeting the same locus, for which none of 377 clones were homologous recombinants. The knockout mice generated exhibit no overt phenotype, though more detailed analysis is still in progress. A possible explanation for the lack of a phenotype is a potential redundancy between Sty and closely related family members, one of which has recently been identified.

Acknowledgments

I would like to express my thanks to the people that have helped me and provided advice throughout the completion of this work.

Peter Duncan, who was a collaborator on this project, and provided much advice.

Ninan Abraham, who also provided much needed advice.

Dr. John Bell, who supervised and supported me during this work.

The people of the cancer research group, who helped me on numerous occasions.

Table of Contents

Index of Tables & Figures

List of Abbreviations

λ	lambda
μg	microgram
μl	microlitre
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DNA	deoxynucleic acid
dTTP	deoxythymidine triphosphate
EDTA	ethylenediaminetetraacetic acid
ES	embryonic stem
FIAU	(1-(2'-fluoro-2'-deoxy-beta-D-arabinofuranosyl)-5-iodouracil)
GAPDH	glyceraldehyde phospate dehydrogenase
Het	heterozygote
Homo	homozygote
HSV	herpes simplex virus
IRES	internal ribosome entry site
KO	knockout
LIF	leukemia inhibitory factor
ml	millilitre
MOPS	3-N-mopholinopropane-sulfonic acid
NaCl	sodium chloride
NaOH	sodium hydroxide
neo^r	neomycin resistance gene
ng	nanogram
nt	nucleotide
PCR	polymerase chain reaction
PGK	phospho-glycerate kinase
SDS	sodium dodecyl sulfate
SSC	sodium citrate/ sodium chloride
TAE	tris acetate EDTA
TE	tris EDTA
tk	thymidine kinase gene
Tris·HCl	tris hyrdrochloric acid
UV	ultraviolet
WT	wild type

Introduction

Protein phosphorylation is important in the regulation of many cellular processes (1). The proteins that perform these phosphorylations were traditionally divided into two groups: the tyrosine kinases, and the serine/threonine kinases. More recently, a number of kinases that phosphorylate on both tyrosine and serine/threonine have been discovered (2). Some examples of these dual specificity kinases include the yeast wee1 kinase, important in the regulation of cdc2, itself a kinase, important in the progression through the cell cycle(3), and MEK (MAP kinase kinase), a component of mitogen activated pathways, which transmit stimuli into the nucleus, controlling differentiation and proliferation (4). It has also been shown that downstream of MEK is another dual specificity kinase, which activates jun kinases (JNK's) (5).

In an effort to identify other kinases that may be involved in cellular regulation events, the murine Sty kinase was isolated, based on its ability to autophosphorylate on tyrosine. Subsequent analysis showed it to autophosphorylate on both tyrosine and serine/threonine (6), placing it in the dual specificity subclass of kinases.

The human homolog of Sty, Clk, has been shown to phosphorylate serine/arginine rich (SR) splicing factors *in vitro* (7). SR splicing factors have been implicated in splice-site selection and spliceosome assembly (8). Clk has been shown to phosphorylate ASF/SF2 on the same sites that are phosphorylated *in vivo (9)*.

Phosphorylation of SR proteins results in the disassociation of splicosomes; It has been shown that over expression of Clk kinase (or Sty kinase) results in SR splicing

factors acquiring a more diffuse pattern, rather than the normal speckled pattern. It should be noted that another kinase, SRPK-1, has a similar effect on the localization of the SR proteins (8). SRPK-1 has stronger activity on ASF/SF2, and both also phosphorylate different proteins; while SRPK-1 seems to act on other SR proteins, Clk seems to have a broader substrate specificity (9). This suggests wide but overlapping specificities for these proteins. In addition to this overlap in function, two more proteins with a high degree of identity to Clk have recently been identified (10).

The Sty dual specificity kinase has also been shown to phosphorylate members of the SR protein family (7). This implicates Sty as a putative regulator of splicing. The *sty* mRNA is itself alternatively spliced. This results in a frame shift, causing the protein to be truncated upstream of the catalytic domain (11). This further implicates Sty as a regulator of itself.

The analysis of proteins has traditionally been performed by over expression in cultured cells (12). This has drawbacks, however. For example, a protein with much higher than endogenous levels may not act as it would at a physiological level. Also, cell lines are of just one cell type, this does not necessarily indicate the function that the protein may have in a complete organism, with thousands of cell types interacting with one another (13).

A possibly more efficient manner to examine the physiological function of a protein would be functionally to delete it from an animal (14). This might allow for the identification of the processes in which the protein is involved. For example, deletion of

a putative tumor suppressor could result in a specific type of tumor appearing, etc (15). Such mutant mice could also represent custom-made animal models of human disease (16). This disruption of specific genes is possible through the process of gene targeting. This involves disrupting a gene, making it nonfunctional, or unable to be expressed (13).

Very many such targeted disruptions, or knockouts, have been performed in mice, providing a wealth of information (17,18). These knockouts can lead to surprising results. For example, the Src kinase is expressed almost ubiquitously, with very high levels in the neurons and platelets (19), but when it was deleted in mice they suffered from severe osteopetrosis, with no other overt phenotype (20). This was totally unexpected, as there was no previous indication that Src was important in the functioning of osteoblasts. It also raises more questions, why were other tissues not affected, considering Src's ubiquitous expression? This is likely due to a redundancy in some cellular pathways: other kinases may be 'taking over' for the missing protein. Indeed, there are nine Src family members, which are quite similar to each other, and many bind to the same receptors (21), suggesting similar functions. Other knockouts have even more severe phenotypes, Abl deficient mice are runted and die at one to two weeks (22), and Csk mutants die in-utero (23). Still others have no apparent phenotype (19,21).

The targeted event occurs through homologous recombination, which involves the exchange of strands between identical DNA sequences. In order for this to occur, sequences homologous to the gene of interest are inserted on either side of an exogenous sequence, usually a gene for selection of the recombinants. The position of this

intervening sequence determines where the gene will be disrupted. The targeting involves two homologous recombination events, one for each arm of endogenous sequence. Any endogenous intervening sequence will be deleted (13).

Of course, for this knockout vector to be put into cells, they must be in culture, but they must also be able to contribute to the formation of a mouse. The inner mass cells of a blastocyst can be removed and later reintroduced back into another blastocyst (resulting in chimeric mice), allowing for the manipulation necessary to alter their genome (24). Because of their unusual potential, they are called embryonic stem (ES) cells. Great care must be taken to ensure they will retain their potential to contribute to the formation of a mouse. This primarily involves ensuring that they do not differentiate, by growing them in the presence of embryonic fibroblast feeder cells (25). In the absence of feeders, the cells multiply a few times, and then differentiate into primitive tissues. Alternatively, the media may be supplemented with LIF, which prevents the ES cells from dividing, but there is some evidence that this may not satisfy all of the growth requirements of ES cells (13).

The neomycin resistance gene, with a PGK promoter, is often used for the positive selection of homologous recombination events (4). This bacterial gene provides resistance to G418 in mammalian cells, which will ordinarily inhibit translation. This allows for the selection of the recombinants from cells that did not incorporate the vector. Unfortunately, nonhomologous recombination occurs much more readily, in mammalian cells, than homologous recombination (13). Therefore, if the vector carries a promoter to

express the *neo*^{r} gene, the targeted cells must be selected from the surrounding random integrants. This targeting frequency may be as high as one in 50 (19), but may also be less than one in 10^4 (13).

Since it cannot be empirically determined what the targeting frequency will be, it is beneficial to have negative selection, and positive selection. This can be accomplished by adding the HSV thymidine kinase gene, with a strong promoter, to the end of the endogenous sequence on the knockout vector. This gene makes mammalian cells susceptible to gancyclovir and FIAU (13). The idea is that if the vector integrates homologously the *tk* will be lost, since it is not homologous. The *neo*^{r} will remain, because it is between two homologous sequences. In a random integration, there will be no pressure to lose the *tk*. In theory, this should eliminate most of the unwanted integrants, but in practice it may only give a 3 to 10% enrichment (22), probably due to inactivation of the *tk* during integration.

Because of the possible drawbacks in using the PGK-*neo* cassette, it would be beneficial to use *neo* but not include a promoter in the vector. The exogenous gene would be expressed by the promoter of the targeted gene. In the past, this resulted in a fusion protein between the N-terminal end of the endogenous protein and the *neo*^{r} gene product, which might not be functional. The alternative was to have translation stop and then restart, which is also unreliable, since this is not a normal occurrence in mammalian cells. More recently, this problem has been overcome by the use of the internal ribosome entry site (IRES) element (26).

The IRES, a picornoviral element, allows initiation of translation in the absence of a 5' cap, which is normally required (27). This allows for a dicistronic mRNA, so the knockout vector can use the endogenous promoter with little regard for the location or reading frame of the targeted gene's initiating methionine. This element has been used with the β*Geo* gene, which encodes a neomycin resistance/βgalactosidase fusion protein, with 70-80% of the surviving clones scoring positively for homologous recombination (26).

Since the β*Geo* gene carries βgalactosidase activity, and is under control of the targeted gene's promoter, it also acts as a reporter for the expression of that gene. This can be extremely beneficial in the study of the protein; it will identify areas of tissue, and even specific cell types, that the gene is expressed in. This can also aid in developmental studies; it can be easily visualized where the gene is expressed (19,26).

Objective

Relatively little is known about the Sty dual specificity kinase. In order to gain a better understanding of its functions, we used gene targeting to disrupt its gene in ES cells, attempting first with a PGK-*neo*^{r} based vector and subsequently with an IRES-βGeo based vector. Deficient mice were generated, and bred to homozygosity.

During the process of this attempt, two new human cDNAs with a high degree of identity to the human homolog of *sty*, *clk*, were identified (10). The *clk* gene itself has very high identity to *sty* (28); this led us to believe that there may also be murine

homologs of *clk2* and *clk3*. Indeed, *sty2* has been identified in our lab (unpublished). In light of the redundancies in the Src family kinases (19,21), it may not be surprising if the Sty deficient mice have no phenotype. Note, however, that the Src deficient mice had a severe phenotype (20), in spite of the closely related kinases.

Materials and Methods

The pKS NT/Sty KO vector was constructed from pKS NT, which contains the neomycin resistance gene, and the HSV thymidine kinase gene, each under control of a PGK promoter (22). A ~4.5 kb 5' and a ~4 kb 3' fragment that originated from genomic lambda clones of the Sty kinase were added 5' to *neo*^{r}, and between *neo*^{r} and *tk*, respectively (see figure 1, and P. Duncan, PhD Thesis, Univ. Of Ottawa, 1996 for details of construction).

The pBS(KS)/Sty:MTG/IRES-βGeo/Sty vector was constructed from pBS(KS)/βGeo, which harbors the β*Geo* gene, proceeded by an IRES element (26). The same genomic fragments were used as with the pKS NT/Sty KO vector, with the exceptions that the 5' fragment was cut back so that its 5' end was within the cDNA, and a six *myc* tag (29) was added between the 5' fragment and the IRES element (see P. Duncan, PhD Thesis, Univ. Of Ottawa, 1996).

For both vectors 100 μg was electroporated into 50 million J1 ES cells. After 24 hours of growth, for the pKS NT/Sty KO cells, 200 μg/ml of G418 was added for positive selection; the media was also made to 0.2 μM FIAU, for negative selection. The pBS(KS)/Sty:MTG/IRES-βGeo/Sty cells were treated identically, except there was no negative selection. Both cell lines were allowed to grow for 7 to 10 days when individual colonies were picked. Half of each colony was placed into 96 well dishes, with γ-irradiated mouse embryonic fibroblast feeder cells (25), and the other half into 24 well dishes, with 500 U/ml LIF. Both were allowed to grow to confluence, where the 96 well

plates were frozen at -190 °C, and the 24 well plates were used for screening the clones.

The DNA from surviving cell lines was prepared as described (30). Cells from an individual line were mixed overnight at 37°C in extraction buffer (100 mM Tris·HCl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl) with 100 μg/ml of proteinase K. The DNA was then precipitated with an equal volume of isopropanol and then washed with 70% ethanol. The ethanol was removed, and the DNA allowed to dry for 1 to 2 hours. 100 μL of TE (10 mM Tris·HCl pH 8.0, 1 mM EDTA) was then added and incubated at 65°C for ~1 hour. The solution was then left at 4°C until the DNA was resuspended. Tail DNA from the mice (as well as liver DNA) was treated in the same manner, except that they were extracted three times with phenol/chloroform (50:50) before the addition of isopropanol.

Due to the large number of cell lines obtained, the individual DNA samples were not quantitated. Instead a titration was used to determine the optimum amount of DNA to use, with hopes that the different preparations would be similar enough for them all to be visualized (data not shown). For the pKS NT/Sty KO cell lines 40 μl of each sample was used for the Nde I digests, for the Bam HI digests 30 μl of each was used. For the pBS(KS)/Sty:MTG/IRES-βGeo/Sty lines, 15-40 μl was used, depending on the viscosity of the sample. For the mouse tail DNA 5-20 μl was used, again based on viscosity. Attempts were made to standardize later analyses to 3-5 μg per sample, but actual concentration were not generally determined.

Digestion of the DNA was generally carried out with 40-60 units of restriction

enzyme, actual amounts are noted in the figure legends. Later digestions also contained BSA at 100 μg/ml, this is also noted in the figure legends. The digestions were carried out overnight at 37°C. The incubations were then run on 0.8% agarose gels in TAE (0.04 M Tris·Acetate, 0.001 M EDTA), generally overnight. The voltage used is stated in the figure legends.

Once the DNA had migrated sufficiently, the excess gel was removed and the genomic smears photographed. The gels were then 'nicked' in a UV crosslinker at 60 mJ (Biorad), to facilitate the transfer of larger fragments. The gels were subsequently denatured for 30 minutes in 0.5 N NaOH/1.5 M NaCl, then neutralized for 30 minutes in 1 M Tris·HCl pH 8.0/1.5 M NaCl.

The gels were then transferred overnight with 10X SSC (from 20X, 300 mM Sodium Citrate pH 7.5, 3 M NaCl), using the capillary method as described (31). The cell line DNA was transferred to Hybond-N (Amersham), the mouse tail DNA was transferred to Nitroplus (MSI), a nylon supported nitrocellulose membrane. The transfers were fixed to the Hybond membrane by UV crosslinking at 100 mJ, and to the nitrocellulose membranes by baking at 80°C under vacuum for 0.5 to 2 hours.

The crosslinked nitrocellulose filters were prehybridized in 50% deionized formamide, 5X SSC, 0.5% SDS, 5X Denhardt's solution (from 50X, 0.01 mg/ml each of Ficoll, Polyvinylpyrrolidone, and Bovine Serum Albumin (Fraction V)) and 0.1 mg/ml denatured herring sperm DNA for about one hour, sealed in a plastic bag which was immersed in a 42°C shaking water bath . While the blots were prehybridizing, the P^{32}

labeled probes were synthesized using a random primed probe kit (Boehringer Mannheim.). Briefly, 25-50 ng of probe template is incubated at 37°C with random hexanucleotides, dATP, dTTP, dGTP, P^{32}dCTP (Amersham) and Klenow enzyme for about one hour. The probe was then spun through a Sephadex G50 (Pharmacia) column at 2000 Xg for 5 minutes to remove unincorporated labeled nucleotides. A 1 µl aliquot was then removed and the activity measured in a scintillation counter (Packard Canberra). The purified probe was then added to the bag containing prehybridization solution and the blot, which was then resealed and returned to the water bath. The hybridization was then allowed to proceed overnight.

The baked nitrocellulose filters were prehybridized in 50% deionized formamide, 50mM sodium phosphate (pH 6.5), 5X Denhardt's solution, 5X SSC, 1% glycine, and 500 µg/ml denatured salmon sperm DNA at 42°C in a hybridization oven (in a hybridization tube). After about an hour the prehybridization solution was poured off, and a hybridization solution containing 50% deionized formamide, 20 mM sodium phosphate (pH 6.5), 1X Denhardt's solution, 1X SSC, 100 µg/ml denatured salmon sperm DNA and 10% dextran sulfate was added, along with the probe. The probes were generated the same way as with the other hybridization buffer, but the probe was not purified away from the unincorporated nucleotides, and therefore the activity was not measured. The hybridization was also allowed to proceed overnight (at 42°C).

In the case of both types of hybridization, the blots were washed twice for fifteen minutes in 0.1% SSC, 0.1% SDS at 65°C in a shaking water bath. The cell line blots

were then put on wet Whatman 3MM paper, wrapped in plastic wrap and exposed to film (Kodak, DuPont) at -70°C with an intensifying screen (Kodak, DuPont). The film was developed after an overnight exposure, and re-exposed for up to seven days if the signal was not sufficient.

The mouse tail blots were put on wet Whatman 3MM paper, sealed in a plastic bag, and then exposed to a phosphor screen (Molecular Dynamics). After an overnight exposure the screen was scanned with a Phosphorimager SI (Molecular Dynamics)

The probe template for the cDNA was the 1.7 kb Eco R1 fragment of pECE/MTG/STY. The 457 bp template for c457 was generated by PCR from 20 ng of pECE/MTG/STY with the oligonucleotides $^{5'}$gatgatgaacaccacagcgccttggtatctacaaga$^{3'}$ (Y330F) and $^{5'}$ggcgcactagtcgtatgctttttaagtgg$^{3'}$ (1539f) using TAQ polymerase (Life Technologies) under standard conditions; the cycle conditions were: 30 seconds at 94°C, 30 seconds at 52°C, and 30 seconds at 72°C, for 30 cycles. The template for c298 was the 298 bp Eco RI /Sal I fragment from pTZ/STY. The c1079 template was the 1079 bp Eco RI/Nde I fragment from pECE/MTG/STY. The GAPDH cDNA was used as the probe template for the GAPDH probe. In all cases the resulting fragment was purified using GeneClean (Bio 101); the concentration was determined by optical density (Biorad) and adjusted to 25 ng/μl in TE.

To generate *sty* deficient mice, 4 day wild type Balb/C blastocysts were injected with ES cells which demonstrated the digestion pattern of homologous recombinants, which were then implanted into pseudo-pregnant foster mothers (32). The resulting

chimeric pups were crossed with wild type Balb/C mice. Tail clippings from the offspring of these crosses were subjected to DNA extraction and southern analysis, as described above.

RNA was prepared using TRIzol reagent (Life Technologies), using approximately 100 mg of lung tissue and 2 ml of TRIzol reagent. 20 μg of the resulting total RNA, in 9 μl of water, was mixed with 4 μl of 10X MOPS buffer (2 M MOPS, 10mM EDTA, and 50 mM Sodium Acetate), 7 μl formaldehyde, and 20 μl deionized formamide and heated at 65°C for five minutes. The samples were then placed on ice and 1 μl of 0.4 mg/ml ethidium bromide, and 2 μl of loading dye (50% glycerol, 1mM EDTA pH 8.0, 0.4% bromophenol blue) added. The RNA was then separated overnight on a 1% formaldehyde gel (1% agarose, 2.2 M formaldehyde, and 1X MOPS buffer) at 1 V/cm. The resulting gel was transferred to Nitroplus membrane, baked, probed, washed, and exposed as the DNA gels.

Results

Part I

Initial Screening of pKS NT/Sty KO clones

The screening for homologous recombinants initially began using the Nde I digest and the *sty* cDNA as the probe template. The *sty* genomic locus contains two Nde I sites, one slightly 5' of the first exon and one which is at nt 1079 of the cDNA. The 3' site is 2.5 kb passed the 3' end of the knockout vector (pKS NT/Sty KO), the 5' site just inside the 5' end of the vector (see figure 1).

Southern analysis of wild type J1 DNA with Nde I and the *sty* cDNA should yield two bands; the fragment between the sites mentioned above, and one from the 3' Nde I site to the next downstream site. The size of the fragment between the known sites is approximately 10 kb. Coincidentally, the other fragment is just slightly smaller, so in a normal analysis one wide band is seen, rather than two discrete ones.

The pKS NT/Sty KO vector, in a homologous recombination event, will insert the 2 kb PGK-NEO cassette into the Sal I site that is seen in the *sty* DNA, with no resulting deletion. This should yield a 12 kb fragment in southern analysis (see figure 1).

This selection has a serious drawback however; as seen in figure 2, all recombination events, whether homologous or not, should yield a band in addition to the WT band. This is because a significant part of the probe template is inside the knockout vector. Wherever the vector incorporates, southern analysis with the *sty* cDNA will show a band of random size, which depends on the distance to the next Nde I site.

Figure 1

Diagram of the: A) pKS NT/Sty KO vector, B) wild type *sty* genomic locus, C) disrupted *sty* genomic locus, demonstrating the strategy for the disruption of the *sty* gene. Also shown is the Bam HI and Nde I screening strategies, with the c457 and cDNA probes, respectively . Sizes expected for southern analysis are also shown. The diagrams are not to scale, and all sizes are approximate. The wide shaded boxes represent exons, the narrower line represent introns, and the lightly shaded exon is alternatively spliced. The exon structure of *sty* has not been fully elucidated, this figure is just an example of a possible configuration. Only the relevent restriction sites are shown. Sty 6-2 and 8-1 are λ genomic clones. cDNA and c457 are probe templates, they are made from the cDNA, there is no intronic sequence in them. I-III, IV, and V-I are subunits of the catalytic domain of the Sty kinase. Met^i is the initiating methionine. NEO is the neomycin resistance gene, with the PGK promoter. TK is the herpes simplex virus thymidine kinase gene with the PGK promoter.

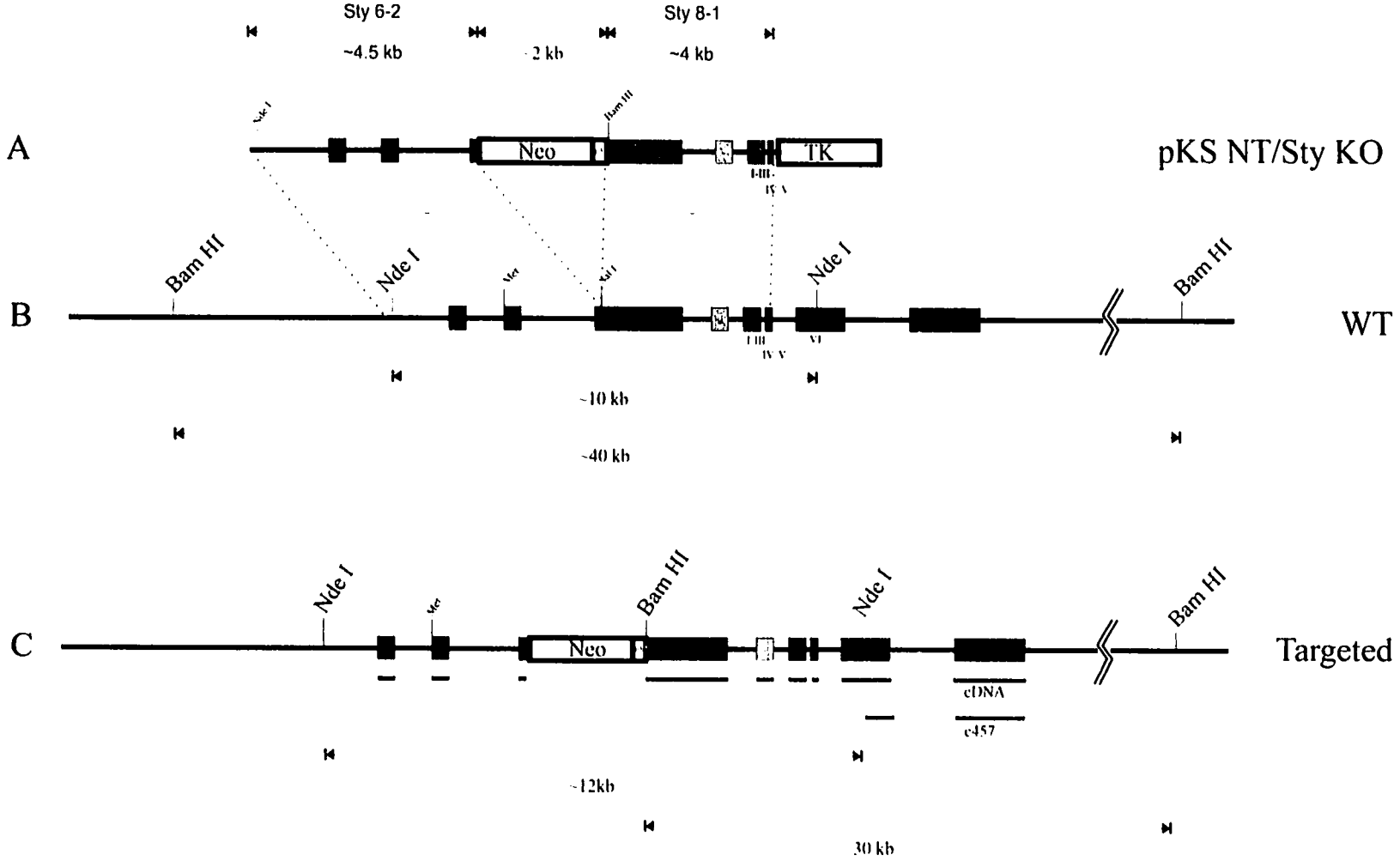

Sty 6-2
~4.5 kb
Sty 8-1
~4 kb
A
Neo
TK
pKS NT/Sty KO
Bam HI
Nde I
Nde I
Bam HI
B
WT
~10 kb
~40 kb
Nde I
Bam HI
Nde I
Bam HI
C
Neo
Targeted
cDNA
c457
~12kb
30 kb

Figure 1

Figure 2
Southern analysis of pKS NT/Sty KO clones, using the Nde I screening strategy, and the cDNA probe.
The respective clones, and wild type DNA (WT) were digested overnight at 37°C with 60 U of Nde I in a total of 100 µl. The samples were then separated on 0.6% agarose gels overnight at 2 V/cm, and southern blotted as described in the materials and methods. 1.3 x 10^8 cpm/µg of cDNA probe was used. The sizes noted are approximate.
The 10 kb band is the wild type, while the other bands are a result of random integrations. A homologous recombinant should have a band at approximately 12 kb, but notice that many clones have a band in this area, while none in fact, are homologous recombinants.

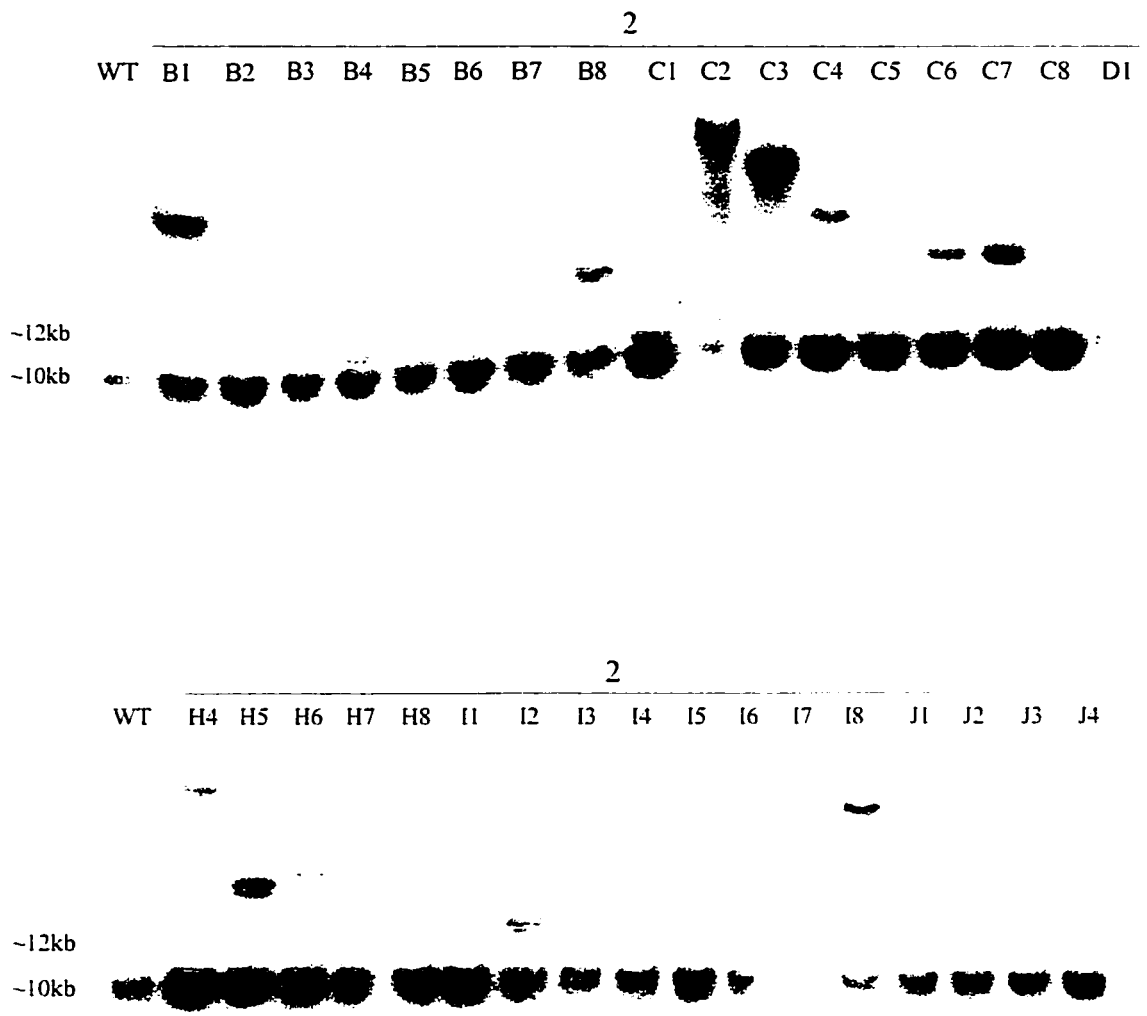

2
WT B1 B2 B3 B4 B5 B6 B7 B8 C1 C2 C3 C4 C5 C6 C7 C8 D1
~12kb
~10kb
2
WT H4 H5 H6 H7 H8 I1 I2 I3 I4 I5 I6 I7 I8 J1 J2 J3 J4
~12kb
~10kb

Figure 2

Since a random integration could produce a band in the 12 kb area, a large number of false positives could result, using this screening strategy. Of the 161 clones screened with this strategy, 59 show a band in this area. For the screening of several hundred clones, this is not acceptable; the screening should be unambiguous.

The solution to this problem would simply be to use a probe template that is not on the knockout vector. However, the only suitable part of the genomic locus not in the vector was a 180 bp cDNA fragment between the end of the knockout vector and the 5' Nde I site. This fragment was too small for efficient labeling, as well as very difficult to isolate. Though there is a ~2 kb intron in this area, it contains repetitive elements and is not suitable as a probe template (data not shown). Therefore, an entirely new screening strategy was needed.

Identification of New Screening Strategy

Since the targeting vector integrates at nt 298 of the cDNA, probes that are entirely to one side of this area would be beneficial. Accordingly, two new probes were produced. c457 is the fragment of the cDNA between nt 1082 and 1539, this is essentially all of the cDNA that is 3' of the knockout vector. c298 is the first 298 bp of the cDNA, up to the Sal I site. This sequence is present in the knockout vector and therefore not suitable for large scale screening, but it is entirely to one side of the vector and therefore would eliminate the doublet that occurs when the cDNA is used as a probe in the Nde I digest. A 5' probe that is outside of the vector would be more beneficial, but probes generated from this area, along with introns from the gene, all seemed to contain

repetitive elements, resulting in nonspecific binding of the probe (data not shown).

An ideal candidate for a screening enzyme would be one whose sites are either side of the genomic locus of interest, with an additional site being introduced by the targeting vector; the fragment should be reduced in size by the targeted disruption of the gene. Since these areas are unknown for the *sty* kinase gene, a southern analysis was performed using 16 enzymes whose sites are introduced by the knockout vector (see figure 3). The Bam HI restriction enzyme produced a single, large sized band, suggesting that these sites might be either side of our genomic locus.

Since the PGK-*neo*^{r} cassette inserts at the Sal I site that is present in the cDNA, a double digestion of wild type J1 DNA by Bam HI and Sal I should give a representation of what a homologous recombinant should look like in a southern analysis, when digested with Bam HI only. This analysis was performed with the 3' probe c457, and with the 5' probe c298 (figure 4).

As seen in figure 4a the double digest results in a small drop in the size of the band, compared to the single digest by Bam HI. The Sal I digest shows a large size smear, since this site is so rare. This confirms that the double digest band is not a result of Sal I alone. Of course, if this were a homologous recombinant, the wild type band would also be present, forming a doublet, since only one allele would be affected by the homologous event. The size of this band has not been accurately determined, but the lower band runs higher than the 24 kb band of a λ-Hind III ladder, suggesting that it is at least 30 kb (data not shown).

Figure 3
Multiple southern analysis of wild type J1 DNA, with the cDNA probe.
Wild type J1 DNA was digested overnight at 37°C (Bst XI at 55°C) with 40 U of the enzymes, as noted, in a total of 50 µl. The samples were then separated on a 0.8% agarose gel overnight at 2.2 V/cm, and southern blotted as described in the materials and methods. 6.8 x 10^7 cpm/µg of cDNA probe was used. M designates 1 kb ladder (Life Technologies). The sizes noted are approximate.
Notice that the Bam HI digest has only one, large sized band, indicating that this fragment may span the *sty* genomic locus.

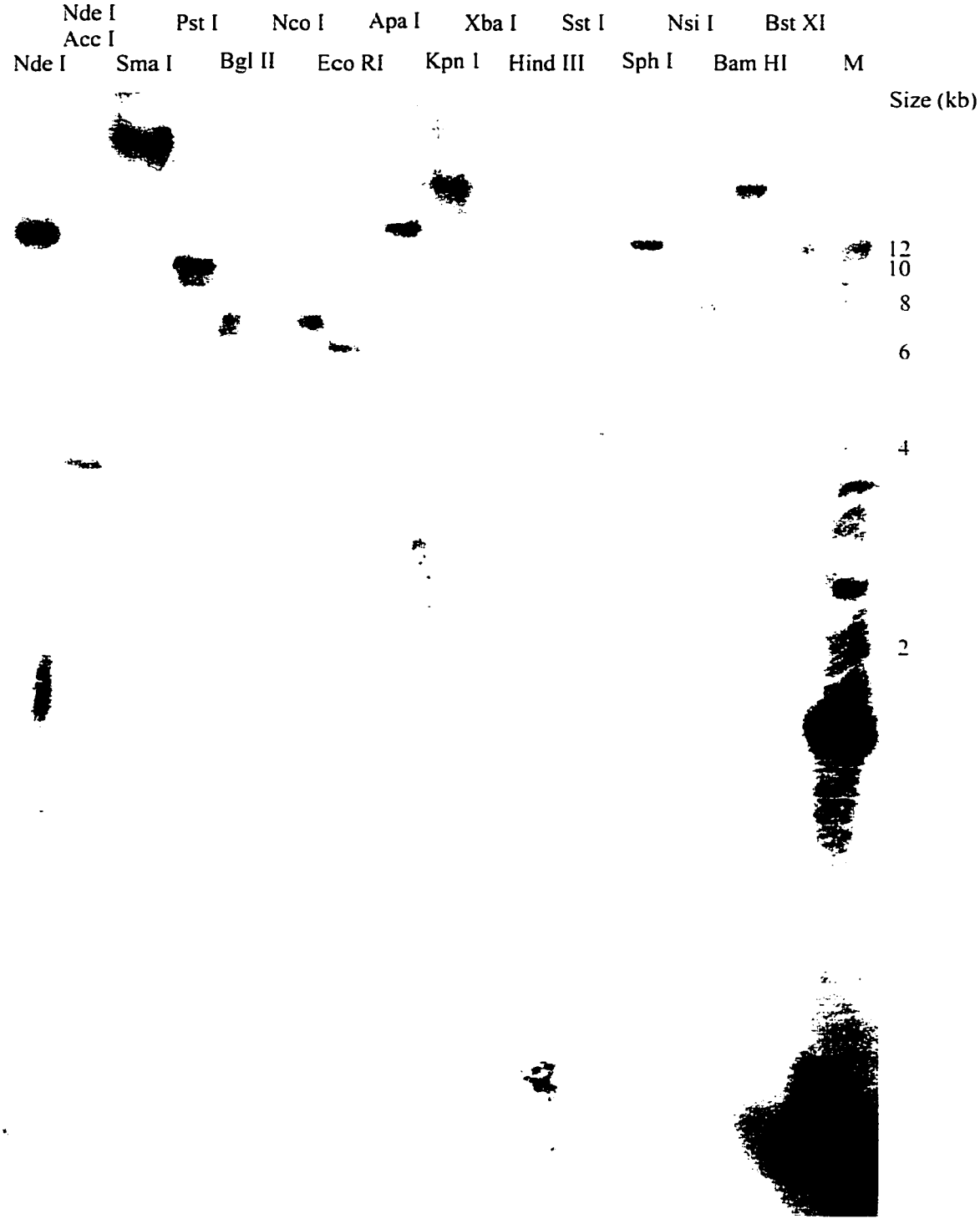

Nde I
Acc I
Pst I
Nco I
Apa I
Xba I
Sst I
Nsi I
Bst XI
Nde I
Sma I
Bgl II
Eco RI
Kpn I
Hind III
Sph I
Bam HI
M
Size (kb)
12
10
8
6
4
2

Figure 3

Figure 4

Southern analysis determining the Bam HI screening strategy with: A) 5' probe c457, B) 3' probe c298.

Wild type J1 DNA was digested overnight at 37°C with 40 U of the noted enzymes, with the Life Technologies digestion buffer noted, in a total of 60 µl, in duplicates. The samples were then separated on 0.8% agarose gel overnight at 1.4 V/cm. The duplicates were then separated and southern blotted as described in the materials and methods. 4.4 x 10^7 cpm/µg of c457 probe was used for A, and 6.4 x 10^7 of c298 for B. M designates 1 kb marker (Life Technologies). The sizes noted are approximate.

The Bam HI/Sal I digest is a representation of what a homologous recombinant, homozygous at the *sty* locus, should look like. This is because the PGK-*neo* cassette inserts at the Sal I site, introducing a Bam HI site at its 3' end. The probes are each entirely on one side of the Sal site, so only one band is seen for each. A heterozygote, such as we would find in our clones, would have both the wildtype and knockout bands.

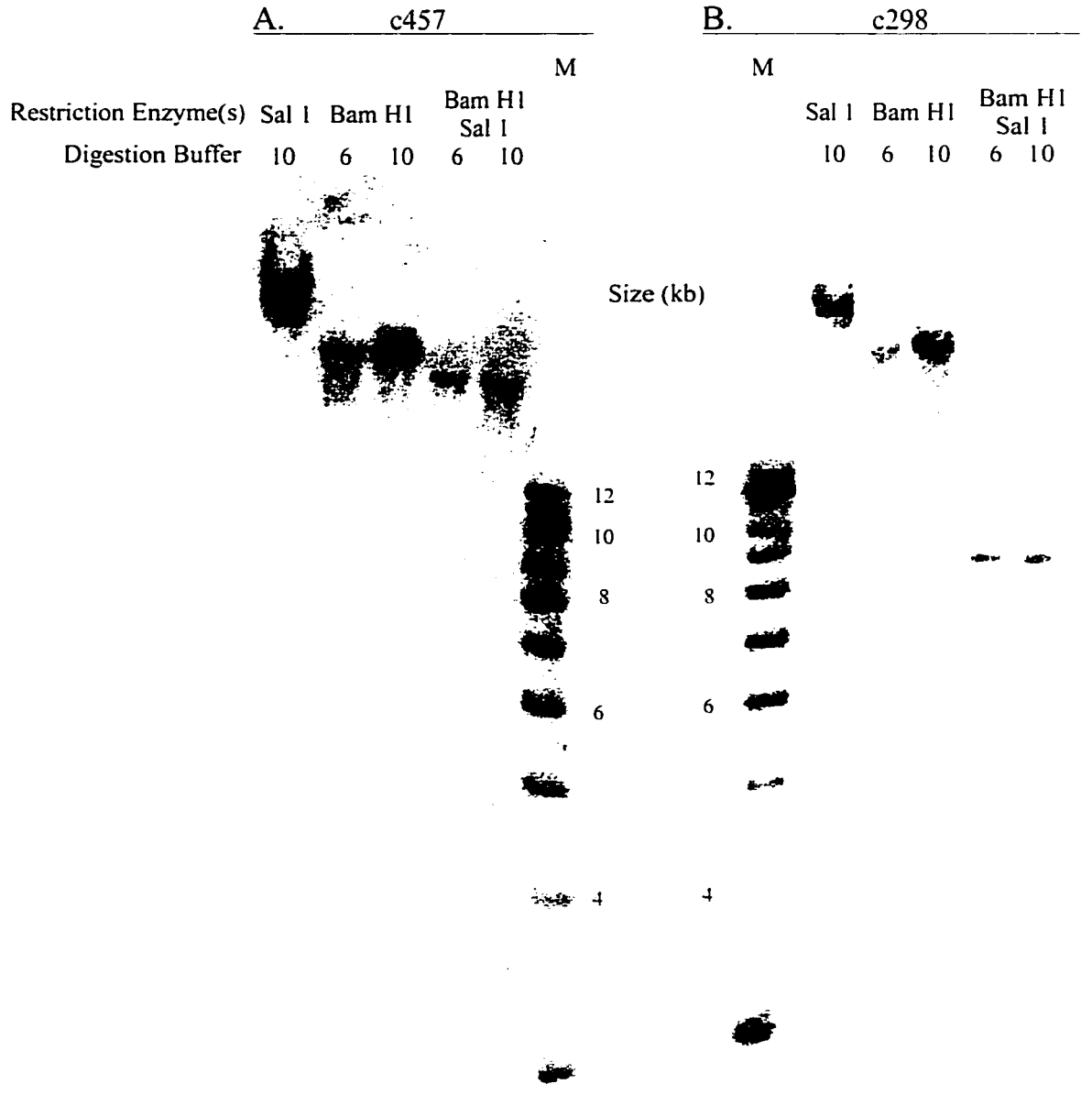

A.
c457
B.
c298
M
M
Restriction Enzyme(s)
Sal I
Bam HI
Bam HI
Sal I
Digestion Buffer
10
6
10
6
10
Sal I
Bam HI
Bam HI
Sal I
10
6
10
6
10
Size (kb)
12
10
8
6
4
12
10
8
6
4

Figure 4

Figure 4b is the same except for the probe used. It shows a similar result, except that the double digest band is now at about 8.8 kb. This suggests that the 5' Bam HI site is 8.8 kb from the Sal I site. With the 3' Bam HI site being (very) approximately 30 kb from the Sal I site, this would indicate the WT Bam H1 band is roughly 40 kb.

Screening of pKS NT/Sty KO clones with Bam H1

The Bam HI screening strategy, with the c457 probe, was used to screen the remaining 216 of the pKS NT/Sty KO clones as well as the 59 ambiguous clones from the Nde I strategy (see figure 5). The repeated screening showed very well the ambiguity of the Nde I screening strategy; none of the re-analyzed clones were homologous recombinants. In fact, none of the 377 lines analyzed showed a digestion pattern of a homologous recombinant.

As seen with clone 2A2, several of the lines did not digest properly; this is most likely due to the crude extraction process used to isolate the DNA. Repeating digestion on these samples yielded the same result, indicating the DNA was not sufficiently purified. These samples could have been subjected to phenol/chloroform extraction to increase there purity, but after two digestion attempts barely enough DNA was left for one southern analysis. Since some DNA would be lost in the extraction and subsequent precipitation, re-purification was not worthwhile. More DNA could be grown up from the frozen 96 well plates, but the number of passages the ES cells go through must be kept minimized to ensure successful contribution in a chimeric mouse (24). This, compounded with the likelihood that any of these individual clones are a homologous

Figure 5
Southern analysis of selected pKS NT/Sty KO clones, using the Bam HI screening strategy and the c457 probe.
A sample of pKS NT/Sty KO clones with bands in the 12 kb area in the Nde I digest analysis were digested overnight at 37°C with 60 U of Bam HI in a total of 40 µl. The samples were then separated on two 0.8% agarose gels overnight at 4 V/cm. 1.6×10^7 cpm of c457 probe was used.
These are all negative for homologous recombination, since they only have the wildtype band (a doublet would appear in the case of a homologous recombinant). Yet they all had a band in the 12 kb area in the Nde digest (see figure 2). Clone 2A2 did not digest, most likely due to the purity of the DNA, as mentioned in the text.

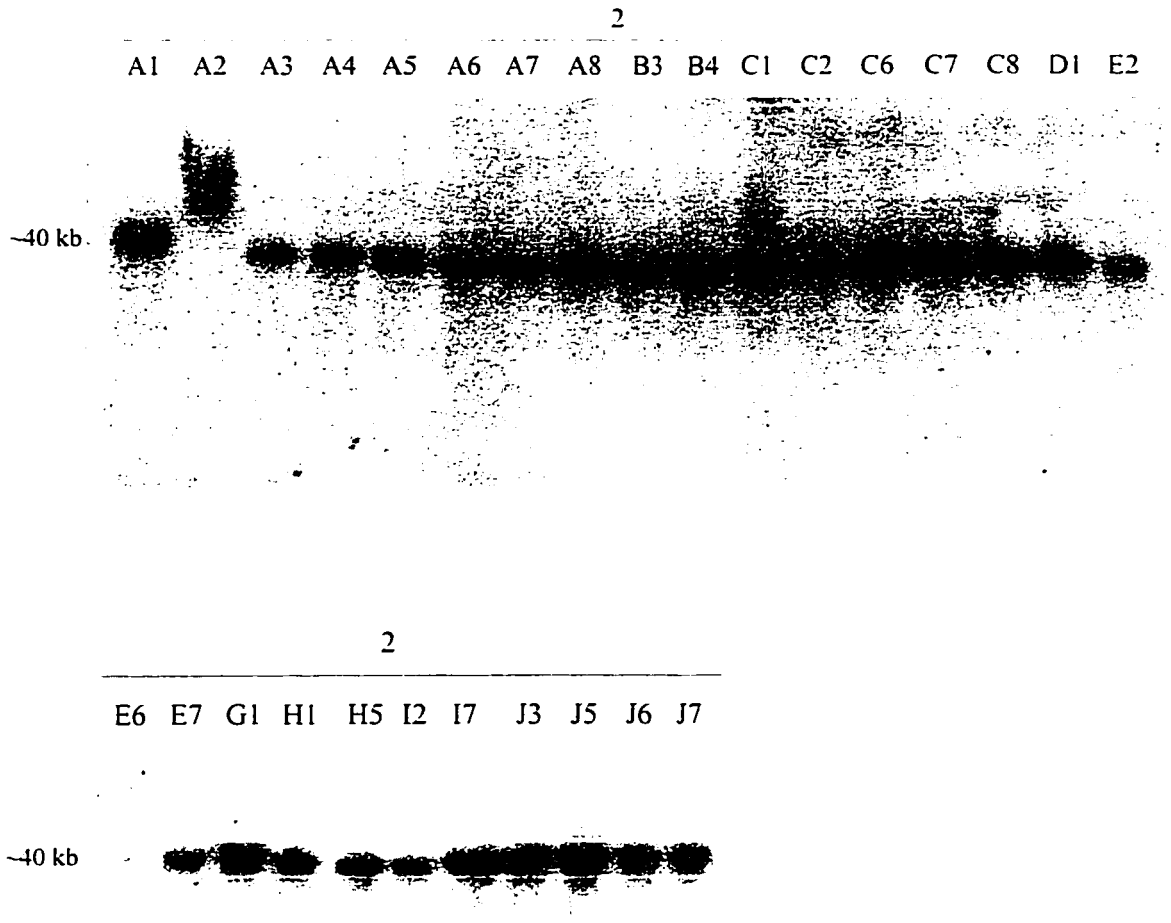

2
A1 A2 A3 A4 A5 A6 A7 A8 B3 B4 C1 C2 C6 C7 C8 D1 E2
~40 kb
2
E6 E7 G1 H1 H5 I2 I7 J3 J5 J6 J7
~40 kb

Figure 5

recombinant suggested that a regrowth would not be worthwhile.

There were actually 960 pKS NT/Sty KO clones, but the remaining 583 lines were not subjected to negative selection. Since negative selection gives perhaps a 10 foldincrease in the selection of homologous recombinants (13,22), and none were identified in the 377 that were subjected to negative selection, it would be quite unlikely that any of these 583 clones would score positive. 163 of these 583 were screened and, as expected, none were identified as a homologous recombinants (data not shown). Screening of these clones was discontinued when the pBS(KS)/Sty:MTG/IRES-βGeo/Sty clones became available.

Part II

pBS(KS)/Sty:MTG/IRES-βGeo/Sty Cell Lines

Due to the disappointing results with the pKS NT/Sty KO vector, and our recent acquisition of the pBS(KS)/βGeo vector, it was decided to attempt the knockout by way of the latter vector. pBS(KS)/Sty:MTG/IRES-βGeo/Sty (see figure 6), being a promoterless construct, was expected to produce much fewer colonies, since the number of nonhomologous recombinants to survive the G418 selection should be small. In our hands, we obtained about as many as with the pKS NT/Sty vector. 96 of the clones were picked, and 86 survived.

Screening of pBS(KS)/Sty:MTG/IRES-βGeo/Sty Cell Lines

The resulting cell lines, produced using the pBS(KS)/Sty:MTG/IRES-βGeo/Sty vector, were subjected to southern analysis using the Bam HI restriction enzyme and the c457 probe (figure 7). The results were surprising, though not unbelievable, owing to the promoterless nature of the vector. Of the 86 colonies screened, 24 produced uninformative results, usually due to insufficient digestion (as with the previous screening attempt). Of the remaining 62 that were not ambiguous 82% showed the doublet that should be indicative of a homologous recombinant.

Confirmation of the Knockout Genotype

This preliminary screen suggested that the identified clones were homologous recombinants, however, it would be conceivably possible for this digestion pattern to appear without a homologous recombination event occurring. So a different analysis is

Figure 6

Diagram of the: A) pBS(KS)/Sty:MTG/IRES-βGeo/Sty vector, B) wild type *sty* genomic locus, C) disrupted *sty* genomic locus, demonstrating the strategy for the disruption of the *sty* gene. Also shown are the Nde I and Bam HI screening strategies, with the various probes. Sizes expected for southern analysis are also shown The diagrams are not to scale, all sizes are approximate. The wide shaded boxes represent exons, the narrower line represent introns. The lightly shaded exon is alternatively spliced, to form sty^f. The exon structure of *sty* has not been fully elucidated, this figure is just an example of a possible configuration. Only the relevant restriction sites are shown. Sty 6-2 and 8-1 are λ genomic clones. c298 c457 and c1079 are probe templates, they are made from the cDNA, there is no intronic sequence in them. I-III, IV, and V-I are subunits of the catalytic domain of the Sty kinase. Met^i is the initiating methionine. NEO is the neomycin resistance gene, with the PGK promoter. TK is the herpes simplex virus thymidine kinase gene with the PGK promoter.

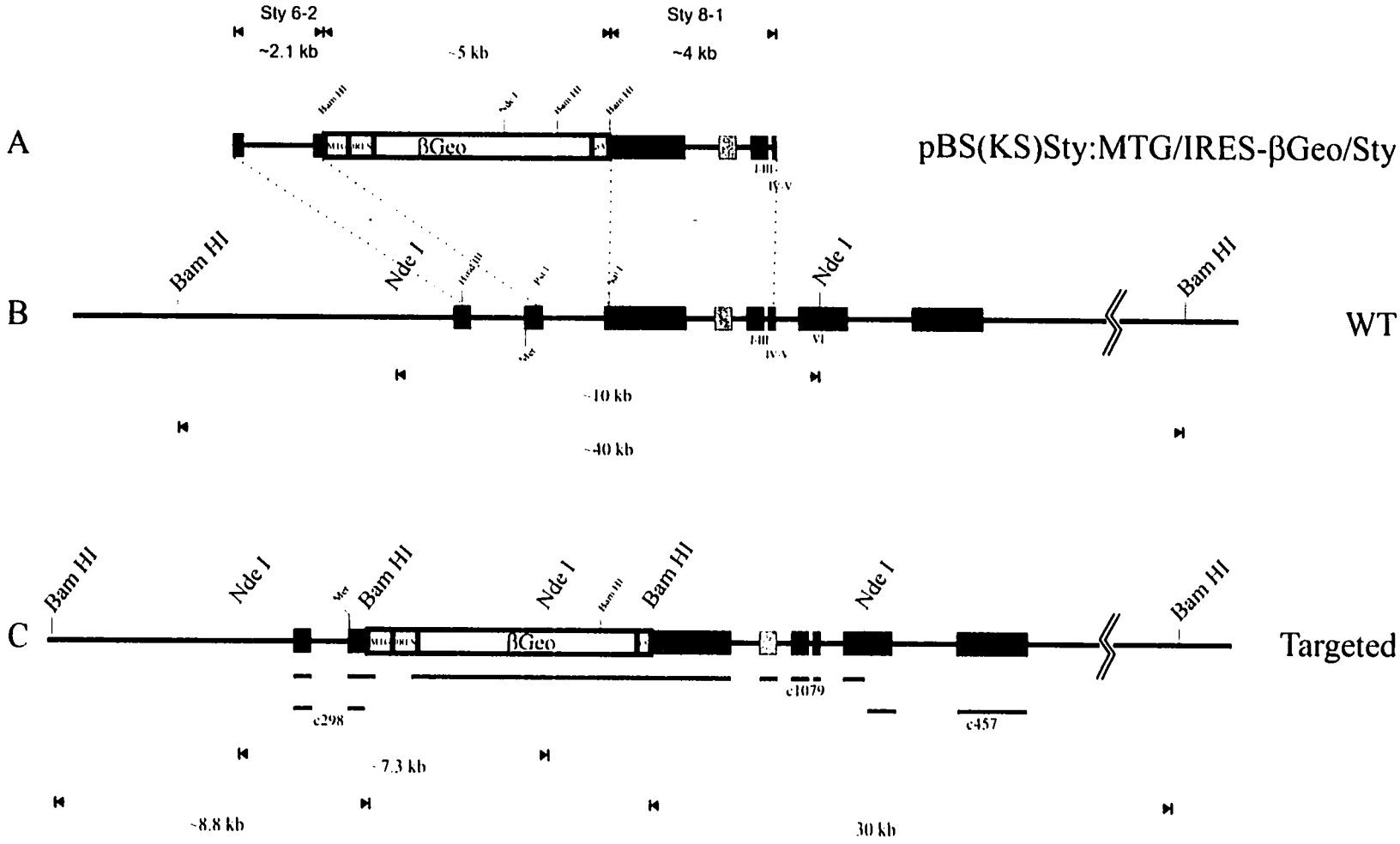

Sty 6-2
~2.1 kb
~5 kb
Sty 8-1
~4 kb
A
βGeo
pBS(KS)Sty:MTG/IRES-βGeo/Sty
Bam HI
Nde I
Nde I
Bam HI
B
WT
~10 kb
~40 kb
Bam HI
Nde I
Bam HI
Nde I
Bam HI
Nde I
Bam HI
C
βGeo
Targeted
c1079
c298
c457
~7.3 kb
~8.8 kb
30 kb

Figure 6

Figure 7

Southern analysis of pBS(KS)/Sty:MTG/IRES-βGeo/Sty clones, using the Bam HI screening strategy and the c457 probe.

pBS(KS)/Sty:MTG/IRES-βGeo/Sty clone DNA was digested overnight at 37°C with 60 U of Bam HI in a total of 30 μl. The samples were then separated on two 0.8% agarose gels overnight at 4 V/cm, and southern blotted as described in the materials and methods. 3×10^7 cpm/μg of c457 probe was used.

These blots demonstrate the high efficiency of the IRES based vector, since most of the clones are homologous recombinants, as seen by the 40/30 kb doublet. Several of the clones DNA samples did not digest (21A6-B1, for example).

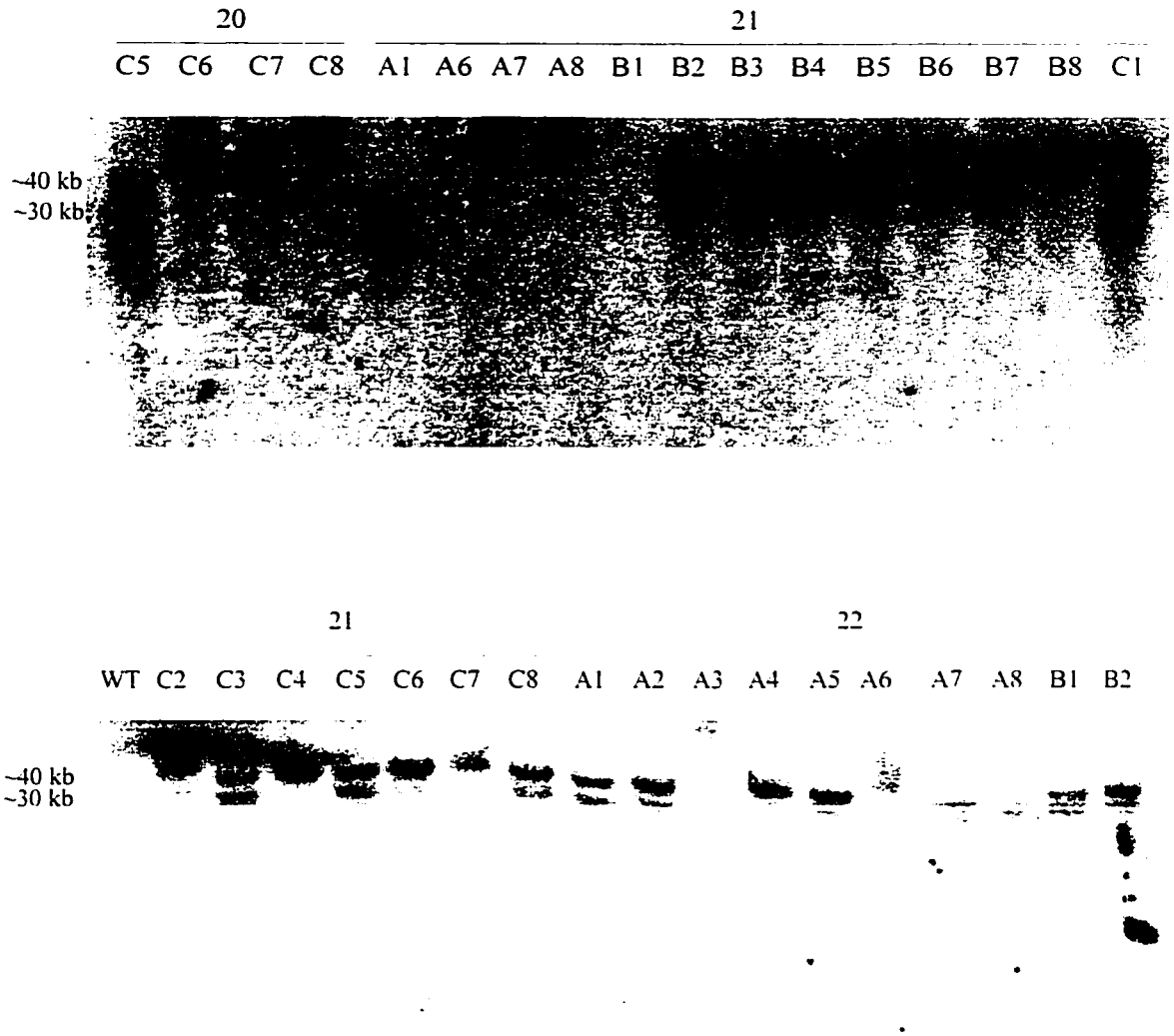

20
21
C5 C6 C7 C8 A1 A6 A7 A8 B1 B2 B3 B4 B5 B6 B7 B8 C1
~40 kb
~30 kb
21
22
WT C2 C3 C4 C5 C6 C7 C8 A1 A2 A3 A4 A5 A6 A7 A8 B1 B2
~40 kb
~30 kb

Figure 7

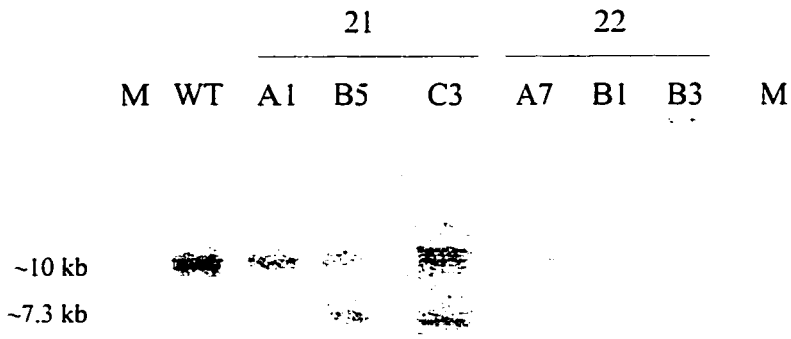
generally required to confirm the first analysis. The Nde I digest is suitable for this purpose. Though it is inefficient as a primary screen, it is valid as a secondary screen because the chances of both digests giving false positive results is very unlikely. Notice that this digest gives a different banding pattern with the pBS(KS)/Sty:MTG/IRES-βGeo/Sty clones compared to the pKS NT/Sty KO clones, due to a Nde I site in the βGeo gene. Clones with wild type and knockout bands of equal intensity were selected and subjected to the Nde I screen, as shown in figure 8. All tested clones were confirmed with the secondary analysis.

Figure 8

Southern analysis of selected pBS(KS)/Sty:MTG/IRES-βGeo/Sty clones, using the Nde I screening strategy and the c298 probe.

DNA from selected pBS(KS)/Sty:MTG/IRES-βGeo/Sty clones was digested overnight at 37°C with 60 U of Nde I in a total of 35 μl. The samples were then separated on a 1% agarose gel overnight at 1 V/cm, and southern blotted as described in the materials and methods. 6.8 x 10^7 cpm/μg of c298 probe was used. M designates 1 kb marker (Life Technologies). The sizes noted are approximate.

This blot confirms the analysis in the previous figure. All have both the wild type 10 kb band, and the knockout 7.3 kb band. The knockout band is 7.3 kb because of the Nde I site in the β*Geo* gene (see figure 6). The full length of the gel was preserved to show that there was no other integrations in these cell lines, which would be identified because the c298 probe is internal to the vector. The mouse line was generated using clone 21A1.



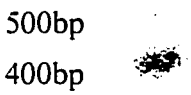


Figure 8

Part III

Generation of Knockout Mouse Line

Cell lines 21A1, 21C5, and 22B1 were thawed from the frozen 96 well dishes and expanded for injection into blastocysts. The degree of chimerism was judged from the coat color of the mice; since mice derived from J1 ES cells are agouti, and Balb/C mice are white, the more agouti the mouse is, the more the ES line contributed to it (24,33). Of the three cell lines tested, only clone 21A1 contributed to the resulting mice, where 3 chimeras were obtained. Mouse 843 was approximately 50% black, whereas 1035 and 1337 were less than 10% black. The chimeras were crossed with wild type Balb/c mice; only chimera 843 transmitted the disrupted allele to it's offspring (data not shown).

Screening of 21A1 Mice

Tail DNA from the mice was subjected to southern analysis using the same screening strategy that was used for the cell lines, the Bam HI digest with the c457 probe. Since the mice were produced in much fewer numbers than the cell lines, and more importantly because all of the genotypes must be determined, the mouse DNA was more highly purified than the cell line DNA. This was accomplished through a triple phenol/chloroform extraction. This, along with the addition of BSA to digestion mixture, resulted in all of the DNA samples being of suitable purity for southern analysis.

It should be mentioned at this time that the southern blotting protocol was altered to increase the sensitivity of detection. This involved using a nitrocellulose membrane, which generally gives lower background (D. Gray, personal communication), rather than

nylon, and improved hybridization and prehybridization solutions (34). This resulted in much higher signal strength, as well as lower background, reducing the amount of DNA necessary to detect a single copy gene to as low as 1 μg.

The male 843 chimera, crossed with wild type Balb/C mice, produced three litters; figure 9 shows a southern analysis of one of the litters. In total there were 11 males and 10 females. Of these mice, 7 were wild type and 14 heterozygote, split evenly between the sexes. Four breeding pairs were set up between 7 heterozygotes from two of the litters; one of the males was in two of the pairs. The breeding pairs were between siblings. Male mouse 1134 produced two litters with female 1129 and four with 1130, 1259 (male) crossed with 1256 (female), and 1258 (male) crossed with 1254 (female), produced a total of four litters each (see figure 10 for an example of the screening of these litters).

Interestingly, the first litters of both the 12xx crosses contained a total of 5 runted pups; they were noticeably smaller and weaker than their litter mates, and most died at three weeks. However, the runted pups were only seen for the 12xx pairs, not the 11xx pairs, and southern analysis revealed no correlation to the Sty genotype. To confirm this, DNA samples from the runts was subjected to southern analysis by the Nde I strategy (figure 11). Later litters from these pairs also demonstrated this phenotype; the second and third litters for the 1259 by 1256 cross had one runted pup each and the third litter from the 1258 by 1254 cross had three. Four of the total of ten runts that were identified died before numbers were assigned to them, this also resulted in their sex not being

Figure 9
Bam HI southern analysis of mice 1254 to 1261, using the c457 probe.
Tail DNA from the above mice digested overnight at 37°C with 50 U of Bam HI in a total of 25 µl. The samples were then separated on a 0.8% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The probe used was c457. WT designates wild type J1 DNA.
This is a litter from the chimera by wild type cross. The doublets identify the heterozygotes. The wild type band is ~40 kb, the knockout band ~30 kb.

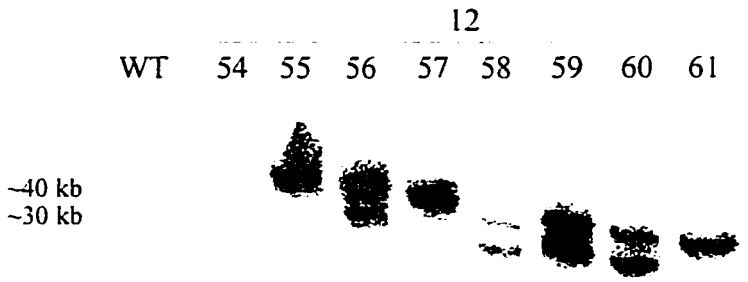


Figure 9

Figure 10
Bam HI southern analysis of mice 1398 to 1409, using the c457 probe.
Tail DNA from the above mice digested for 5 hours at 37°C with 50 U of Bam HI in a total of 25 µl, supplemented with 100 µg/ml BSA . The samples were then separated on a 0.7% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The probe used was c457. WT designates wild type J1 DNA, Balb/C designates wild type Balb/C mouse DNA, and 1254 x 1258 is a mouse that died before number assignment.
These mice are from heterozygote crosses, so homozygotes are expected. These are identified be the single, lower, 30 kb band, as opposed to the wild type with only the upper 40 kb band. Once again, the heterozygotes are identified by the doublet. See the pedigree (figure 17) for the identity of the parental mice.

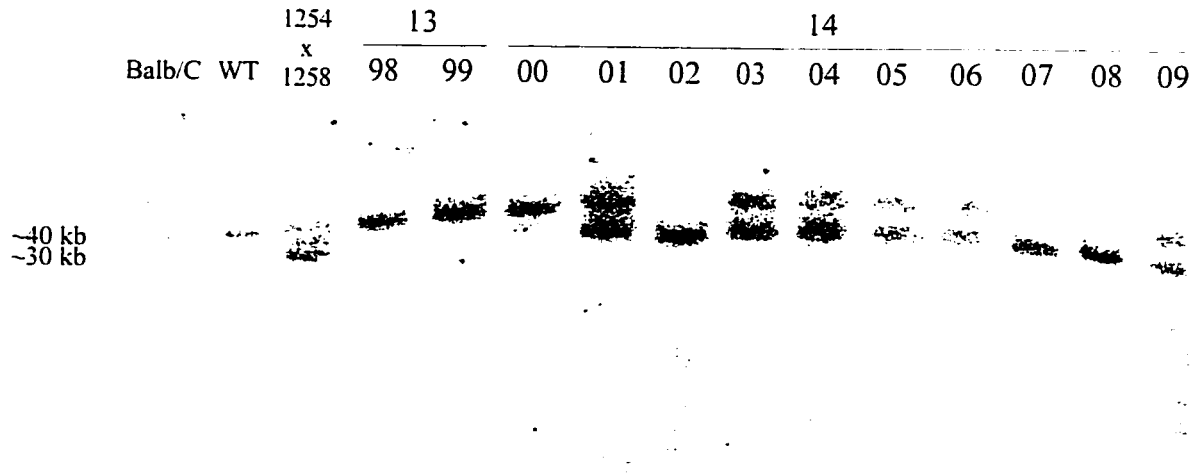

1254
x
1258
13
14
Balb/C
WT
98
99
00
01
02
03
04
05
06
07
08
09
~40 kb
~30 kb

Figure 10

Figure 11
Nde I southern analysis of runted mice up to 1409, using the c298 probe.
Re-isolated tail DNA from 1371, 1375, 1377, 1378, and 1409 mice, which all had the runted phenotype, were digested overnight at 37°C with 50 U of Nde I in a total of 25 µl, supplemented with 100 µg/ml BSA. The samples were then separated on a 0.8% agarose gel overnight at 2 V/cm, and southern blotted as described in the materials and methods. The probe used was c298. WT designates wild type J1 DNA.
This blot was to confirm the genotype of the runts; none are homozygotes, since the 7.3 kb knockout band is not seen without the wildtype band.

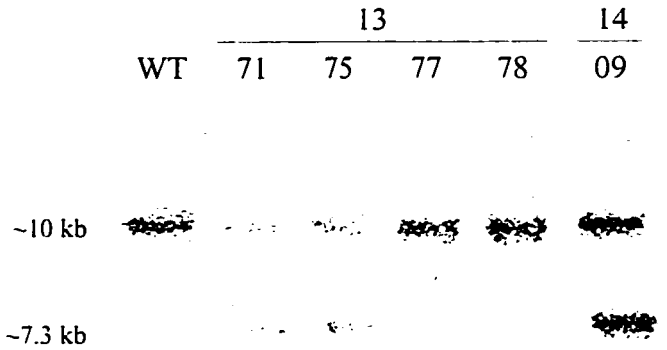


Figure 11

recorded. Subsequent litters, and the one subsequent generation from this lineage did not demonstrate this phenotype.

Of the 46 mice that originated from the first generation of the 12xx crosses, ten had the runt phenotype. Since about one quarter of the mice had this affliction, it is likely to be caused by a mutation in a single gene. This, combined with the appearance of these mutants on only one 'side of the family', and their lack of relation to the genotype with respect to Sty, demonstrates beyond any doubt that this phenotype is not related to the Sty genotype. The Sty homozygotes, as determined by southern analysis, appeared normal, compared to their wild type litter mates.

The mice are actually part Balb/C and part 129/terSv, the mouse line that the J1 cell were derived from, compared to the cell lines being purely 129/terSv in origin (22). As seen with the wild type mice, the *sty* locus, with regard to the Bam HI and Nde I digests, was the same in Balb/C and 129/terSv strains. Additionally, Balb/C DNA was routinely subjected to southern analysis along with knockout mouse DNA (Figures 10,15).

The southern analysis using the Bam HI digest with the c457 probe was very accurate for the screening of the cell lines, however, for the mice it has a drawback. Since the wild type and homozygote both produce a single band, which are very close to each other, it can sometimes be difficult to positively differentiate between them, due to the unevenness of lanes that is produced in some gel electrophoresis apparatuses. Separation can usually be increased by running the genomic smears farther on the gel,

but since the Bam HI fragments are so large, separation between the bands is already near it's maximum.

To overcome this problem, the c298 probe was used instead. This will hybridize to the 8.8 kb fragment, giving a large separation between this and the wild type fragment. This probe is internal to the knockout vector, but this is not as much a concern as with the cell lines since the recombination is already known to be homologous. This solved the problem differentiating between the bands, but with some frequency the probe would not label very efficiently, and often the lower band would be very faint.

To resolve this problem a new probe template was prepared, c1079. This probe is essentially all of the cDNA that is on the knockout vector, both sides of the IRES-βGeo cassette. It should give the same banding pattern as both the c457 and c298 probes combined, when used with the Bam HI digest. More importantly, the template is about 1 kb, which, in this author's experience, produce much more efficient probes than the smaller templates.

Once a suitable number of homozygotes had been obtained, to form breeding pairs, a homozygote knockout line could be propagated. To ensure that the screening process was accurate, some of the Bam HI analyses were repeated with the Nde I strategy, as seen in figure 12. Finally, all of the homozygotes identified up to that point were rescreened (figure 13). Male #1459 and female #1458 were positively identified as homozygotes, and were crossed to produce a homozygote line. All of the offspring were identified as homozygotes (figure 14, note that a different probe was used). Though all

Figure 12
Bam HI (A) and Nde I (B) southern analysis of mice 1449 to 1459, using the c298 probe. Tail DNA from the above mice digested overnight at 37°C with 50 U of A) Bam HI and B) Nde I, in a total of 30 µl, supplemented with 100 µg/ml BSA . The samples were then separated on a 0.8% agarose gels overnight at A) 4 V/cm and B) 2 V/cm, and southern blotted as described in the materials and methods. The c298 POB was used for both. WT designates wild type J1 DNA, DKO designates cell line double knockout DNA and 1256 x 1259 is a mouse that died before number assignment.
The Bam HI blot is using a different probe than previously seen, this results is the knockout band being at 8.8 kb, with the wild type still at 40 kb. Heterozygotes have both bands. The Nde I blot confirms the genotype found in the Bam HI blots.

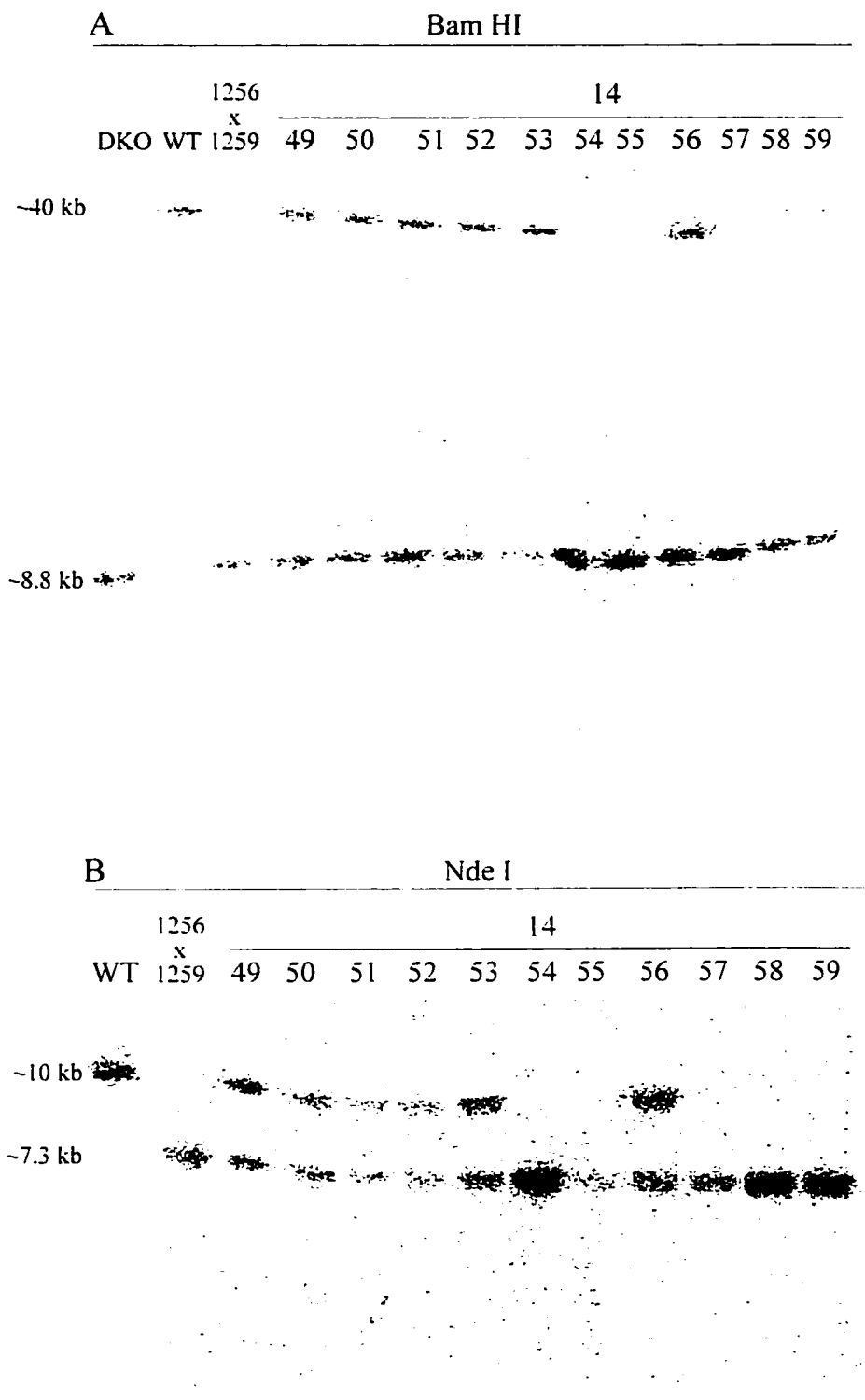

A
Bam HI
1256
x
1259
14
DKO WT 49 50 51 52 53 54 55 56 57 58 59
~10 kb
~8.8 kb
B
Nde I
1256
x
1259
14
WT 49 50 51 52 53 54 55 56 57 58 59
~10 kb
~7.3 kb

Figure 12

Figure 13
Bam HI southern analysis of homozygote mice up to 1459, using probe c298.
Tail DNA from mice 1348, 1402, 1407, 1408, 1454, 1457, 1458, and 1459 digested overnight at 37°C with 60 U of Bam HI in a total of 30 µl, supplemented with 100 µg/ml BSA . The samples were then separated on a 0.8% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The c298 POB was used.
WT designates wild type J1 DNA, DKO designates cell line double knockout DNA.
This blot confirms that all of the previously identified homozygotes are in fact homozygotes, since only the 8.8 kb band is seen. The banding patterns are describe in the legend to figure 12. Homozygote DNA, from a cell line, is included as a reference.

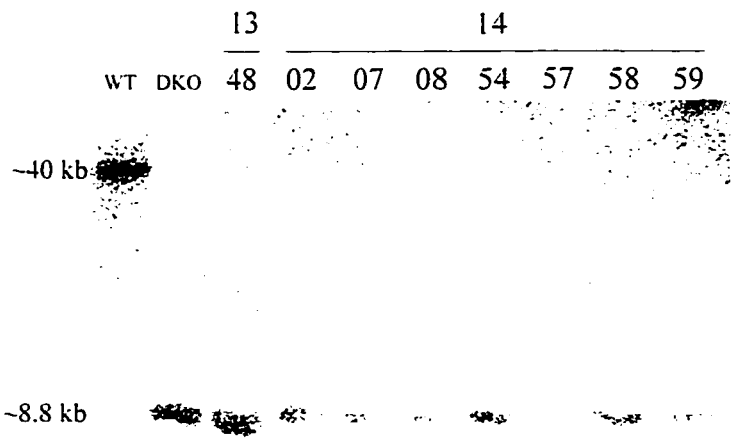


Figure 13

Figure 14
Bam HI southern analysis of mice 1930 to 1934, using probe c1079.
Tail DNA from the above mice, which are the first litter of a homozygote line, was digested overnight at 37°C with 60 U of Bam HI in a total of 30 µl. The samples were then separated on a 0.8% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The c1079 probe was used. WT designates wild type J1 DNA, and DKO designates cell line double knockout DNA.
This blot uses a different probe than previous ones, it hybridizes to both of the knockout fragments seen with the c298 and c457 probes. This results is the homozygotes having ~30 kb and ~8.8 kb bands. Wild type still has a band at 40 kb, heterozygotes would have all three bands.

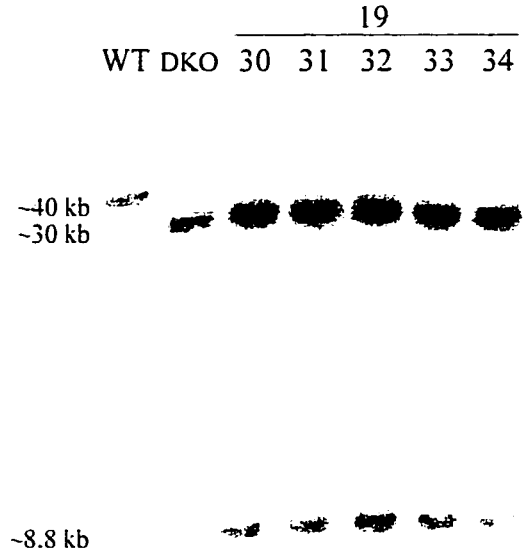


Figure 14

the mice appeared normal, they are later generations of the mice that produced the runted phenotype, so an alternate homozygote cross was set up between offspring of the 11xx crosses. Male 1402 (figure 10) and female 1348 (figure 13) were used. This cross also produced a healthy, but small litter (data not shown).

Analysis of the heterozygote crosses also continued, to gather enough data for our analysis to be statistically accurate (figures 15, 16). Of the 104 mice resulting from the heterozygote crosses, including the four non-numbered nor sexed runts, 27 were wild type, 51 heterozygote, and 25 heterozygotes. Of the 100 whose sex is known, 58 were male and 42 female; the genotypes were split fairly evenly between the sexes. This is all demonstrated graphically, along with the genotype of each mouse, in the pedigree for chimera 843 (figure 17).

Northern Analysis of Mice

RNA was prepared from the lungs of wild type (#1742, figure 16) , heterozygote (#1738, figure 16) and homozygote (#1685) Sty knockout mice to confirm that null alleles had indeed been produced. A southern blot, using DNA from the liver of the same animals, was prepared to demonstrate the relationship between the genotype and the expression of the Sty mRNA, and to ensure there had been no error in the handling of the tail clippings or the subsequent southern analysis (figure 18).

Volumetric analysis of the c1079 northern blot and the GAPDH control, are shown in table 1. It should be noted that the wild type and homozygote mice were 79 and 78 days old, respectively, while the heterozygote mouse was 43 days old.

Figure 15
Bam HI southern analysis of mice 1592 to 1599, using probe c298.
Tail DNA from the above mice digested overnight at 37°C with 60 U of Bam HI in a total of 30 µl, supplemented with 100 µg/ml BSA . The samples were then separated on a 0.8% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The c298 probe was used. WT designates wild type J1 DNA, Balb/C designates wild type Balb/C mouse DNA, DKO designates cell line double knockout DNA, and 1256 x 1258 is a mouse that died before number assignment.
This is once again mice from a heterozygote cross, it utilizes the c298 probe so the meaning of the bands is the same as in figure 12. Balb/C DNA is included to show demonstrates the same wild type band as the J1 DNA.

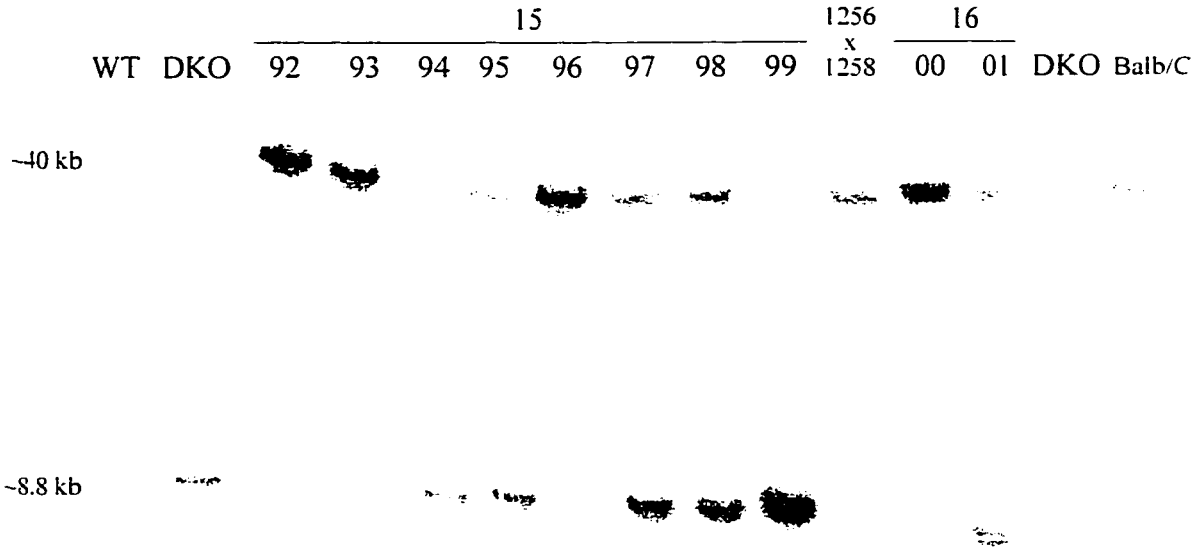

15
1256
x
16
WT
DKO
92
93
94
95
96
97
98
99
1258
00
01
DKO
Balb/C
~40 kb
~8.8 kb

Figure 15

Figure 16
Bam HI southern analysis of mice 1737 to 1744.
Tail DNA from the above mice digested overnight at 37°C with 60 U of Bam HI in a total of 30 μl, supplemented with 100 μg/ml BSA. The samples were then separated on a 0.8% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The c1079 probe was used. WT designates wild type J1 DNA, and DKO designates cell line double knockout DNA.
This is also heterozygote cross DNA, utilizing the c1079 probe. The bands are explained in the legend to figure 14.

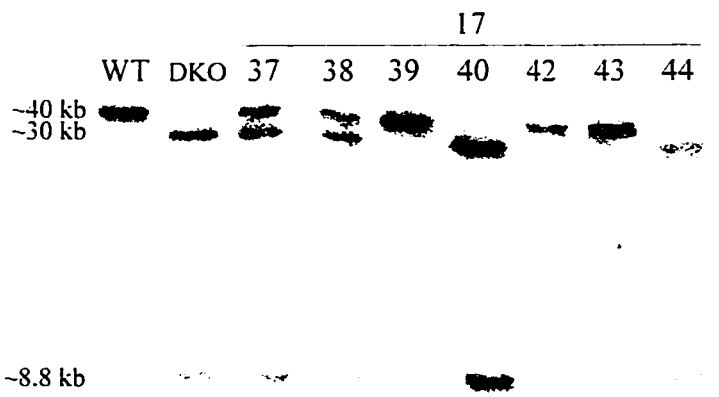


Figure 16

Figure 17
Pedigree of Sty knockout mouse line from chimera 843.
Offspring and subsequent generations from founder 843, indicating the sex and genotype of each mouse. Circles are female, squares are male, diamond are undetermined (died before number assignment and sex recording). The chimera is indicated by the stripes, wild type are empty, heterozygotes are half full, and homozygotes are filled. Mice with the runted phenotype are indicted with *runt*.

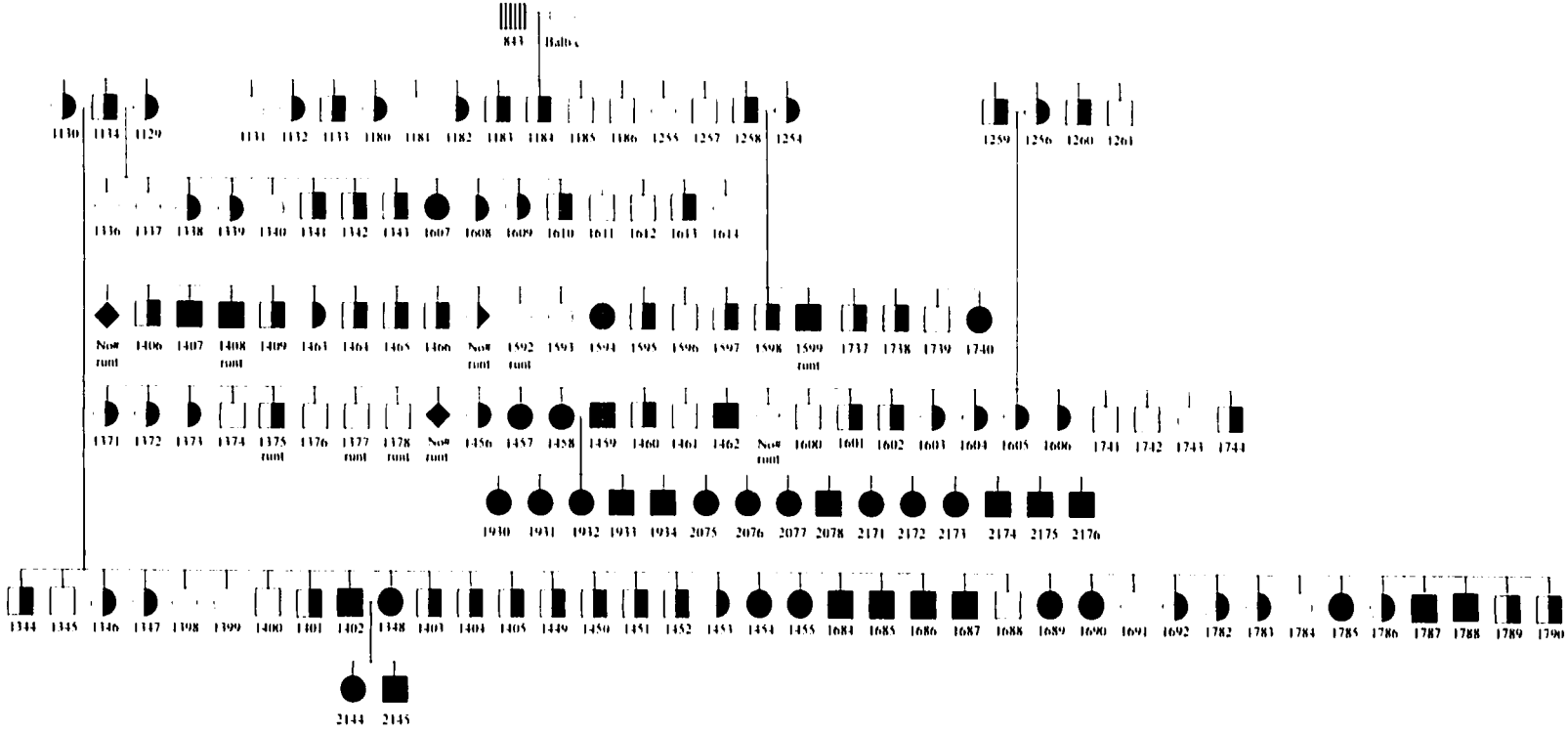


Figure 17

Figure 18

Northern and southern analysis of wild type, heterozygotes and homozygote Sty mice. A) Ethidium Bromide stain of RNA, B) c1079 probing of northern blot, C) GAPDH probing of northern blot, D) southern analysis of DNA from the same mice as were used for the northern analysis. The northern analysis was performed as described in the materials and methods, using the c1079 probe, and subsequently, the GAPDH probe. For the southern analysis, ~5 μg each of mouse liver DNA was digested overnight at 37°C with 60 U of Bam HI in a total of 30 μl, supplemented with 100 μg/ml BSA. The samples were then separated on a 0.8% agarose gel overnight at 4 V/cm, and southern blotted as described in the materials and methods. The c1079 probe was used. WT designates wild type, Het designates heterozygote, and Homo designates homozygote, DNA (for D) or RNA (for A-C). In panel C, the GAPDH bands are at 1200 bp. The samples in A. B. and C are all from mouse. In D. Cell Line designates DNA from cell lines. and Mouse designates DNA from mice.

This analysis demonstrates that *sty* is, in fact, functionally deleted, since its transcript is not seen in the homozygote.

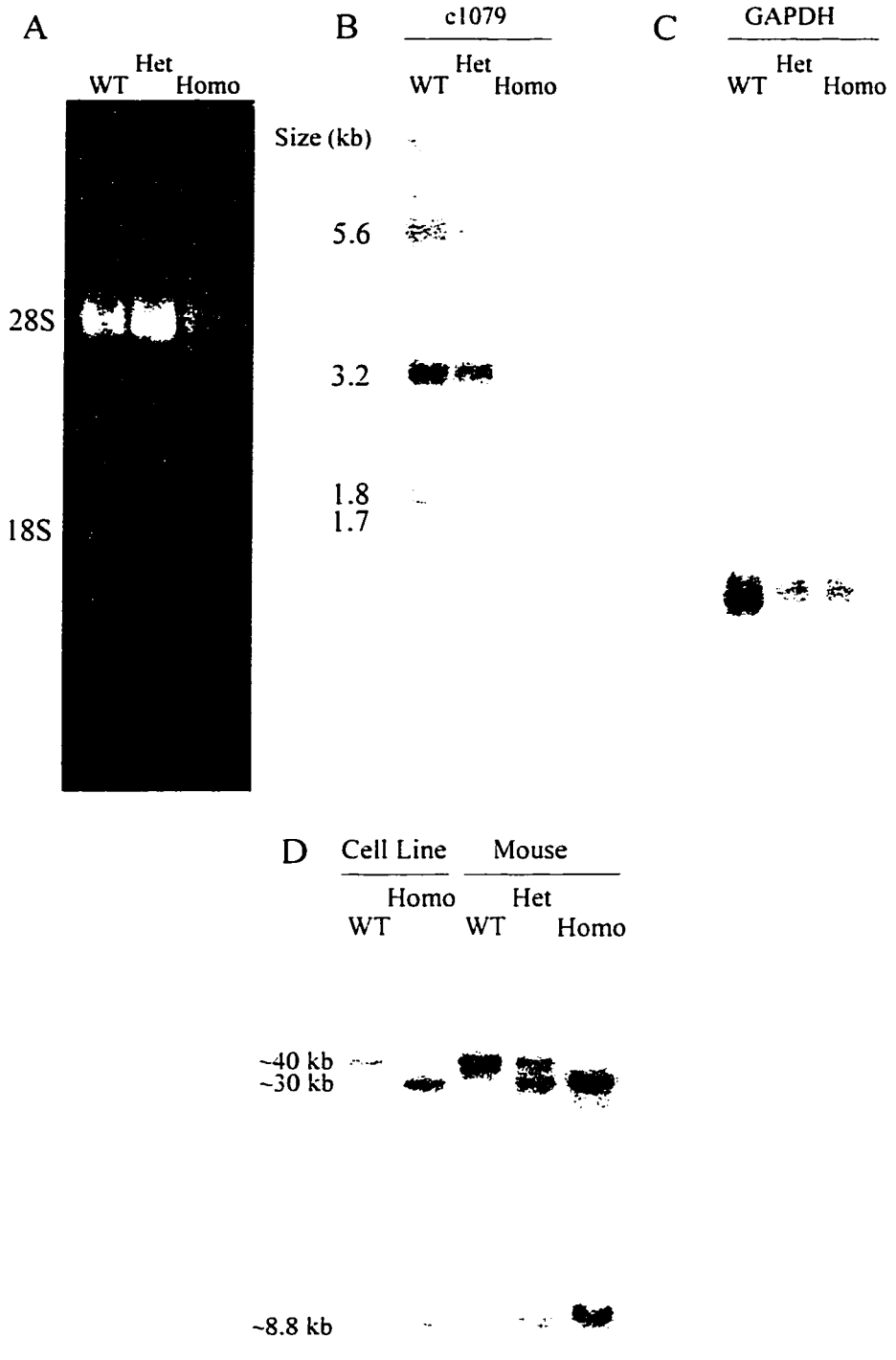

A
WT
Het
Homo
28S
18S
B
c1079
WT
Het
Homo
Size (kb)
5.6
3.2
1.8
1.7
C
GAPDH
WT
Het
Homo
D
Cell Line
Mouse
WT
Homo
WT
Het
Homo
~40 kb
~30 kb
~8.8 kb

Figure 18

	Probe	
Mouse	c1079	GAPDH
Wild Type	1.00	1.00
Heterozygote	0.67	0.42
Homozygote	0.07	0.39

Table 1: Volume analysis of bands from northern blots (figure 18, B & C) shown as a fraction of the intensity of the wild type bands. The 3.2 kb bands were used for the c1079 probing. Volume analysis was performed, with background correction, using ImageQuant software (Molecular Dynamics).

Discussion

From the results of the original screening method, it is obvious that the derivation of an efficient screening strategy is well worth the effort, both in time and resources. Not only does it make the screening process more efficient, by reducing the redundancy necessary to screen the clones, but also makes the analysis more definite. The screening with an internal probe (that is, internal to the knockout vector) will never give a definite answer to whether it is a homologous recombinant or not, even with multiple restriction digests, because it is always possible that the exogenous band is due to a random integration. This possibility can be reduced by using more and more digests, but this is not practical.

The major problem with the Nde I screening is in the design of the pKS NT/Sty KO vector itself; it incorporates almost all of the sequence between the Nde I sites. This problem is reduced with the pBS(KS)/Sty:MTG/IRES-βGeo/Sty vector, since it does not extend as far in the 5' direction. The Bam HI southern analysis strategy, using the c457 probe, is a very efficient screen; since the probe is external to the vector (that is, it is not in the sequence of the vector), it will not identify random integrants at all. Of course, random integrations must be identified if they occur in cells that are also homologous recombinants, since random integrations of the vector could have very detrimental effects (disrupting another gene, for example). A probe that is internal to the vector is therefore a good choice for a subsequent screen, it will help confirm the first analysis, as well as identify any unwanted integrations.

The results of the analysis of the pKS NT/Sty KO clones was quite disappointing, since none of the 377 cell lines (as well as the 163 that were not subjected to negative selection) was a homologous recombinant. This is most likely because of the propensity of mammalian cells to favor nonhomologous recombination over homologous recombination. This ratio has been observed to be in the range of 1:1,000 to 1:10,000 (13). Since it is possible to transfect the vector into millions of cells, it should not be a problem to obtain homologous recombinants, but to identify them amongst the non-homologous recombinants.

This frequency can be increased through the use of negative selection, as was used in this case. However, this generally only results in a ten fold increase in the frequency of homologous recombinants (22). The problem has been overcome in using PCR based sibling selection (13,25). This process, however, requires the vector have one arm quite short in order to facilitate the PCR. The pKS NT/Sty KO has 4.0 and 4.5 kb arms, and though PCR fragments greater than this size have been generated from genomic DNA (35), it is not practical in this case.

Another technique by which the homologous recombinants are identified is by use of a promoterless construct. Since the vector carries no promoter, the resistance gene will only be expressed when incorporated downstream of a promoter. This greatly favors homologous recombinants, since they will be under control of the promoter of the targeted gene. Of course, the gene must be expressed in ES cells for this strategy to work. Random integrations could occur just downstream from promoters, but this is

unlikely, due to the organization of the mammalian genome (1). pBS(KS)/Sty:MTG/IRES-βGeo/Sty is a promoterless knockout vector; since it's construction was nearing completion by the time the pKS NT/Sty KO clones had been screened, it was used to generate more cell lines, rather than repeating the process with pKS NT/Sty KO.

The 5' *sty* genomic fragment used in the pBS(KS)/Sty:MTG/IRES-βGeo/Sty vector was shorter than with the previous vector. This is because the vector must not carry any promoter elements. The 5' end of the vector starts within the 5' UTR of Sty genomic locus. The selection of these IRES-βGeo clones resulted in about as many clones as with the PGK-*neo*^{r} vector. This was more than what was expected, due to the findings of others (26). However, of the 86 clones screen, 82% were identified as homologous recombinants. The reason for the large number of clones is most likely that the electroporation efficiency was higher than with the previous vector, resulting in more cells containing the pBS(KS)/Sty:MTG/IRES-βGeo/Sty vector. The high frequency of homologous recombination is explained by the above reasoning, only a very small fraction of the nonhomologous recombinants integrate behind a promoter. Of course, there is the same frequency of nonhomologous recombinants as with the previous vector, they just don't survive the selection.

If the cell lines were clonal, that is, arose from one cell, they should all have an identical genetic complement. Therefore, if that line arose from a homologous recombinant, all of the subsequent cells should have one wild type allele, and one

disrupted allele. This should mean that in southern analysis, the two bands in the doublet should have equal intensity. In fact, for many of our clones the lower, or 'knockout' band, is of lower intensity. This would indicate that when the colonies were picked after positive selection, they were not clonal in origin. However, since we had such a large number of positive clones, it was not a problem to pick out a suitable number that showed an equimolar doublet.

Of the three cell lines injected into blastocysts, only one contributed to the formation of mice. This exemplifies the problems with manipulating inner mass cells in culture and then reintroducing them into their native environment. There are many stages where something could compromise their potential. For example, the genome of a ES cell could be damaged during the electroporation, or subsequent handling. This may not affect the cell in culture, but render it unable to contribute to the development of a mouse, or cause a embryonic lethal effect. Since one founder was obtained, the original J1 ES cells must have their full potential, thus ruling out any problems with them. Additional problems could occur with the blastocyst injection, since disruption of the blastocyst may result in it not forming a mouse at all (24). Another problem is that sometimes the ES cells contribute very little, or not at all, to the embryo, though they are known to have this potential. This is seen with our 21A1 cell line, it contributed to three chimeras, but at very different levels, resulting in only one of the chimeras carrying the disrupted allele in its germ line.

Since all of the mice originated from the same chimera, and therefore the same cell line, positive conclusions cannot be made about the effect of this knockout. This would be more of a concern if we had seen a phenotype associated with this disruption, since a phenotype could be caused by the accidental disruption of another gene. But even in this case it is remotely possible that there is a compensatory mutation that obliterates the phenotype that the absence of Sty has. Attempts are currently underway to produce new chimeras, using different cell lines.

From the heterozygote crosses, which produced 104 mice, a normal Mendelian ratio of the genotypes was obtained (1). This is what is expected for a single gene, if there are no embryonic lethal effects (36). This demonstrates that this mouse line has no partial embryonic lethality, which has been observed previously (15). Notice, in the pedigree for chimera 843 (figure 17), that a rather large number of mice must be obtained to determine this; the genotype of the individual litters varies widely. As mentioned previously the runt phenotype that was seen in 10 of the mice has no relationship to the Sty knockout.

It is possible for proteins to still maintain at least part of their function, in spite of being disrupted, by the use of a cryptic initiation codon, or the slicing out of the insertion cassette from the mRNA (37). Therefore, to confirm that the targeting of *sty* had, in fact, produce a null allele, a northern analysis was performed. As seen in figure 18 (A,B,C) and table 1, the *sty* transcript is absent in the homozygote, while the GAPDH control probing shows that the RNA is present, though it is less abundant in this example. It is

quite peculiar that the heterozygote, when corrected for loading with the GAPDH probing, seems to express more *sty* than the wild type. A southern analysis, using liver DNA from the same animals, was performed to confirm the genotypes (figure 18D). It is of note that the wild type and homozygote were 79 and 78 days old, while the heterozygote was only 43 days old. It is possible that the Sty kinase is expressed differentially in the different aged mice. Indeed, Sty has been shown to be somewhat developmentally regulated, though this difference has only been seen between ES and differentiated cells (11). Unfortunately, this northern blot was not repeated, so it is entirely possible that the analysis is not accurate; for example, there could have been unequal hybridization of the GAPDH probe to the RNA. Regardless, the wild type and homozygote lanes show that the mice generated from cell line 21A1 are, in fact, functionally knocked out.

The homozygote mice had no overt phenotype. It may be that these mice have a phenotype that would not be apparent in their ordinary life. Other knockout mice have such phenotypes; for example, mice deficient for the thymocyte isoform of Fyn kinase have impaired signaling in thymocytes (38), and Hck kinase deficient mice have phagocytosis impaired macrophages (21), though both appear normal. Knockout mice may also have learning deficiencies which may not be easily discernable (39). Many other mutants have been shown to be susceptible to pathogens that their wild type counterparts are not (40). The Sty deficient mice need to be further challenged to identify any potential phenotype. Indeed, progress is underway to challenge the

homozygotes with viral infections, to note any effect this might have (unpublished).

It could be possible for phenotypes not to appear until later stages in their life, due to development changes. The oldest homozygote (1348-female) is currently healthy, and is in fact one of the breeders for one of the homozygote lines. The mutation does not appear to effect the fertility of mice.

Future Directions

In light of the discovery of *sty2* gene, and the probability that there is also a *sty3*, it may not be surprising if there is no phenotype associated with this knockout, because of potential redundancy between the similar kinases. This is suggested by the findings for some of the Src family kinases, which appear to be able to compensate for each other (19,21). These compensations became apparent only when deficient mice were crossed to produce doubly deficient mice. Work is currently underway to produce *sty2* deficient mice (unpublished), in order to identify if there is redundancy between the Sty kinases.

It is quite possible that the knockout mice harbour some phenotype that is not readily distinguishable from its wild type counterparts. To identify a possible subtle phenotype a detailed histological examination of the mice should be performed, to identify any tissue abnormalities that may arise in the absence of the Sty protein. Since the β-glactosidase activity of the knockout vector should show where *sty* is expressed, this could be exploited to identify which tissues or cell types should be tested for abnormalities. To test for neurological effects of this knockout, the mice could be tested for their learning ability, compared to wild type mice (38). It is also possible to challenge

the mice with various pathogens to check for any effect on their immunity (21).

Preliminary work was performed to exploit the βgalactosidase activity in the β*Geo* gene (unpublished). This could be very beneficial it the study of the Sty kinase, since it's expression is controlled by the *sty* promoter. It could be used to identify specific cell types that the kinase is active in, as opposed to the traditional northern analysis, which is restricted to identifying tissues. It could also be used identify the expression of Sty during development (19,26); this could prove to be of greater utility than the fact that *sty* is not expressed in the mice.

Conclusions

The IRES-β*Geo* containing vector proved to be far superior to the PGK-*neo*^{r} vector in generating homologous recombinants in the targeting of the *sty* allele. This is in spite of the fact the PGK-*neo*^{r} construct had an even greater degree of homology. The high number of clones that survived the positive selection, coupled with the 82% recombination frequency obtained, may be a result of high *sty* expression in the ES cells. The *sty* mRNA is known to be differentially regulated in embryonic cells, lending support to this theory. However, the IRES based knockout strategy is still fairly new, so more knockouts need to be performed this way so that their frequencies can be compared.

The knockout mice that were generated appeared the same as their wild type litter mates, even when bred to homozygosity. This could be due to a redundancy between Sty and other closely related proteins, as seen with the Src family kinases. The knockout of these other kinases may result in a phenotype that reveals the true function of Sty. Of course, the analysis of the deficient mice is continuing, to identify if there is a more subtle phenotype that will only expose itself in the correct environment.

The βgalactosidase activity that is produced by the now disrupted endogenous allele can be of great use in the study of the distribution of Sty throughout tissues. This will be of even greater virtue if the Sty kinase is involved in developmental processes, as suggested by its developmentally regulated expression. These knockout mice may eventually help lead to the discovery of the function of Sty, even if no phenotype is identified.

Bibliography

1. Lewin, B. (1990) in *Genes IV*, (Oxford University Press, New York).

2. Douville, E. M. J., Duncan, P. I., Abraham, N. & Bell, J. C. (1994) *Cancer Metast. Rev.* **13**, 1-7.

3. Featherstone, C. & Russell, P. (1991) *Nature* **349**, 808-811.

4. Pages, G., Brunet, A., L'Allemain, G. & Pouyssegur, J. (1994) *EMBO J.* **13**, 3003-3010.

5. Lin, A., Minden, A., Martinetto, H., Claret, F. X., Lange-Carter, C., Mercurio, F., Johnson, G. L. & Karin, M. (1995) *Science* **268**, 286-290.

6. Howell, B. W., Afar, D. E. H., Lew, J., Douville, E. M. J., Icely, P. L. E., Gray, D. A. & Bell, J. C. (1991) *Mol. Cell. Biol.* **11**, 568-572.

7. Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C. & Duncan, P. I. (1996) *EMBO J.* **15**, 265-275.

8. Misteli, T., Spector, D. L. (1997) *Trends in Cell Bio.* **7**, 135-138.

9. Colwill, K., Feng, L. L., Yeakley, J. M., Gish, G. D., Cáceres, J. F., Pawson, T. & Fu, X.-D. (1996) *J. Biol. Chem.* **271**, 24569-24575.

10. Hanes, J., von der Kammer, H., Klaudiny, J. & Scheit, K. H. (1994) *J. Mol. Biol.* **244**, 665-672.

11. Duncan, P. I., Howell, B. W., Marius, R. M., Drmanic, S., Douville, E. M. J. & Bell, J. C. (1995) *J. Biol. Chem.* **270**, 21524-21531.

12. Katagiri, T., Urakawa, K., Yamanashi, Y., Semba, K., Takahashi, T., Toyoshima, K., Yamamoto, T. & Kano, K. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 10064-10068.

13. Sedivy, J. M. & Joyner, A. L. (1992) in *Gene Targeting*, (W.H.Freeman & Company, New York)

14. Melton, D. W. (1994) *Bioessays* **16**, 633-638.

15. Bradley, A. (1994) *Mutation Research* **307**, 557-572.

16. Lo, D. (1996) *Clin. Imm. & Immunopathology* **79**, 96-104.

17. Shastry, B. S. (1995) *Experientia* **51**, 1028-1039.

18. Viney, J. L. (1995) *Cancer & Metastasis Reviews* **14**, 77-90.

19. Stein, P. L., Vogel, H. & Soriano, P. (1994) *Genes and Dev.* **8**, 1999-2007.

20. Soriano, P., Montgomery, C., Geske, R. & Bradley, A. (1991) *cell* **64**, 693-702.

21. Lowell, C. A., Soriano, P. & Varmus, H. E. (1994) *Gene and Dev.* **8**, 387-398.

22. Tybulewicz, V. L. J., Crawford, C. E., Bronson, R. T. & Mulligan, R. C. (1991) *Cell* **65**, 1153-1163.

23. Imamoto, A. & Soriano, P. (1993) *Cell* **73**, 1117-1124.

24. Joyner, A. L (1993) in *Gene Targeting*, (IRL Press, New York).

25. Zijlstra, M., Li, E., Sajjadi, F., Subramani, S. & Jaenisch, R. (1989) *Nature* **342**, 435-438.

26. Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M., Chambers, I. & Smith, A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4303-4307.

27. Borman, A. M., Bailly, J. L., Girard, M. & Kean, K. M. (1996) *Nucleic Acids Res.* **23**, 3656-3663.

28. Ben-David, Y., Letwin, K., Tannock, L., Bernstein, A. & Pawson, T. (1991) *EMBO J.* **10**, 317-325.

29. Roth, M. B., Zahler, A. M. & Stolk, J. A. (1991) *J. Cell Biol.* **115**, 587-596.

30. Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R. & Berns, A. (1991) *Nucleic Acids Res.* **19**, 4293-4294.

31. Fritsch, E. F. & Maniatis, T. (1987) in *Molecular Cloning - A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

32. Hogan, B., Constantini, F. & Lacy, E. (1986) in *Manipulating the Mouse Embryo*, (Cold Spring Harbour Laboratory Press, Cold Spring Harbour).

33. Li, E., Bestor, T. H. & Jaenisch, R. (1992) *Cell* **69**, 915-926.
34. Wahl, M. G., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683-3687.

35. Cheng, S., Fockler, C., Barnes, W. M. & Higuchi, R. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5695-5699.

36. Copp, A. J. (1995) *Trends Genet.* **11**, 87-93.

37. Yang, Y., Reis, L. F. L., Pavlovic, J., Aguzzi, A., Schafer, R., Kumar, A., Williams, B. R. G., Aguet, M. & Weissmann, C. (1995) *EMBO J.* **14**, 6095-6106.

38. Stein, P. L., Lee, H.-M., Rich, S. & Soriano, P. (1992) *Cell* **70**, 741-750.

39. Grant, S. G. N., O'Dell, T. J., Karl, K. A., Stein, P. L., Soriano, P. & Kandel, E. R. (1992) *Science* **258**, 1903-1909.

40. Kaufmann, S. H. & Ladel, C. H. (1994) *Trends Microbiol.* **2**, 235-242.

CURRICULUM VITAE

NAME: Simon Ginsberg

DATE OF BIRTH: 12th July, 1967

PLACE OF BIRTH: Bristol, United Kingdom

CITIZENSHIP: Canadian

EDUCATION:

Earl of March High School
Kanata, Ontario (Grades 9-12)

Bell High School
Nepean, Ontario (Grade 13)

University of Ottawa
Ottawa, Ontario
B.Sc. (Hons) Biochemistry, 1994

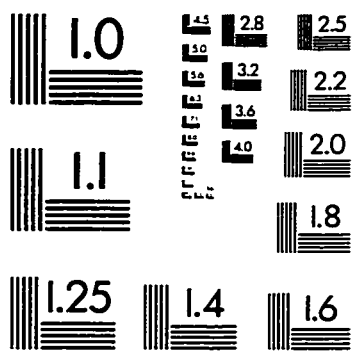
University of Ottawa
Ottawa, Ontario
M.Sc. Biochemistry, registered 1994-present

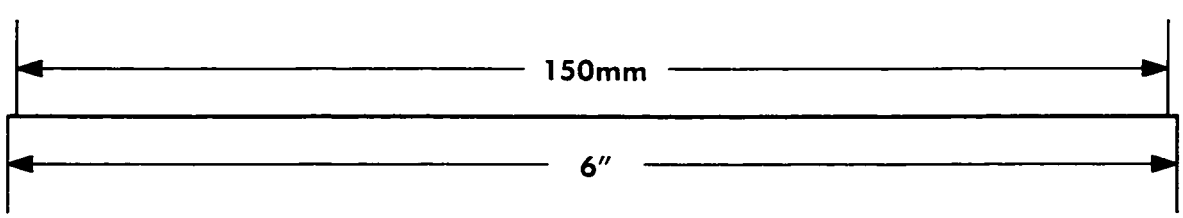
AWARDS: Ontario Scholar, 1990

EXPERIENCE:

Teaching Assistant - 2nd Year Biochemistry Laboratory
University of Ottawa 1995, 1996
Teaching Assistant - 3rd Year Biochemistry Laboratory
University of Ottawa 1995

TEST TARGET (QA-3)





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